

## **A incomensurabilidade das teorias científicas\***

Autores: Paul Hoyningen-Huene e Eric Oberheim

Tradutora: Laura Machado do Nascimento

Revisor: Rogério Passos Severo

O termo “incomensurável” significa “sem medida comum”, e tem suas origens na matemática da Grécia Antiga, onde significava nenhuma medida comum entre magnitudes. Por exemplo, não há uma medida comum entre o comprimento do cateto e o comprimento da hipotenusa de um triângulo retângulo isósceles. Relações incomensuráveis como essa são representadas por números irracionais. A aplicação metafórica dessa noção especificamente matemática à relação entre teorias científicas sucessivas tornou-se controversa em 1962, depois de ser popularizada por dois filósofos da ciência influentes: Thomas Kuhn e Paul Feyerabend. Eles pareciam estar desafiando a racionalidade da ciência natural e foram chamados, na revista *Nature*, de “os piores inimigos da ciência” (Theocharis e Psimopoulos, 1987, p. 587; cf. Preston *et al.* 2000). Desde 1962, a incomensurabilidade das teorias científicas é uma ideia controversa que tem sido amplamente discutida, e foi fundamental para o giro histórico na filosofia da ciência e para o estabelecimento da sociologia da ciência como uma disciplina profissional.

### **1. Introdução**

No influente livro *A estrutura das revoluções científicas* (1962), Kuhn afirma, dramaticamente, que a história da ciência revela que os proponentes de paradigmas concorrentes falham em estabelecer contato completo entre suas concepções, de maneira que

---

\* Tradução do verbete "The Incommensurability of Scientific Theories" de Paul Hoyningen-Huene & Eric Oberheim (2009). In: *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.) URL = <<http://plato.stanford.edu/archives/spr2009/entries/incommensurability/>>. Publicada com a autorização do editor.

estas sempre têm, ao menos ligeiramente, propósitos diferentes. Kuhn caracterizou as razões coletivas para estas limitações na comunicação como a incomensurabilidade entre tradições científicas pré e pós-revolucionárias, afirmando que o paradigma newtoniano é incomensurável com os seus predecessores cartesiano e aristotélico, na história da física; assim como os paradigmas de Lavoisier e de Priestley, na química (Kuhn, 1970 [1962], 147-150; Hoyningen-Huene, 2008). Falta aos paradigmas competidores uma medida comum, pois cada um utiliza conceitos e métodos diversos para tratar de problemas igualmente diversos, o que limita a comunicação através da linha divisória revolucionária. Inicialmente, Kuhn utilizou a noção de incomensurabilidade predominantemente para desafiar concepções *cumulativas* de avanço científico, que afirmam que o progresso científico consiste em uma aproximação gradual à verdade, e para desafiar a ideia de que existam padrões metodológicos neutros e invariáveis para a comparação de teorias ao longo do desenvolvimento das ciências naturais. Como na teoria da evolução, o processo não muda *em direção* a uma meta fixa de acordo com regras, métodos ou padrões fixos, mas, em vez disso, tende a *afastar-se* das pressões exercidas por anomalias na teoria predominante (Kuhn 1970 [1962], 170-173). O processo de mudança científica é eliminativo e permissivo, mais do que instrutivo. No processo de confrontar anomalias, certas alternativas são excluídas, mas a natureza não nos guia em direção a uma única teoria correta.

Kuhn desenvolveu e refinou sua ideia inicial ao longo das décadas seguintes, repetidamente enfatizando que incomensurabilidade não significa nem implica incomparabilidade, nem torna a ciência irracional (por exemplo, Kuhn 2000 [1976], 197). Ele concentrou-se cada vez mais na incompatibilidade conceitual manifesta nas diferenças estruturais usadas para classificar os tipos cujas relações são expressas por leis e teorias, tais como elementos químicos e espécies biológicas (Kuhn 2005). Ele utilizou a incomensurabilidade para atacar a ideia, proeminente entre os positivistas e empiristas lógicos, de que comparar teorias exige traduzir suas consequências para uma linguagem observacional neutra (cf. Hoyningen-Huene 1993, 213-214). No final dos anos 90, ele explicou a incomensurabilidade em termos de *inefabilidade*, enfatizando que se torna possível para os

cientistas formular e compreender certos enunciados novos *somente depois* que uma teoria particular for introduzida (no vocabulário antigo, as novas frases são sem-sentido), da mesma forma que só é possível para os historiadores compreenderem certos enunciados antigos colocando de lado concepções atuais que de outro modo causam distorções (Kuhn 2005 [1989], 58-59; 2005 [1993], 244). Tal '*incomensurabilidade taxonômica*' resulta em falha na tradução entre subconjuntos locais de termos interdefinidos devido à classificação cruzada de objetos em taxonomias mutuamente excludentes. Isso pode ser distinguido da '*incomensurabilidade metodológica*', de acordo com a qual não há medida comum entre teorias científicas sucessivas, no sentido de que a comparação de teorias é, às vezes, uma questão de ponderar valores que se desenvolvem historicamente, sem seguir regras fixas e definitivas (Sankey e Hoyningen-Huene 2011, vii-xv). Isso deixa espaço para o desacordo racional na comparação de teorias, já que cientistas podem aplicar valores diferentes (como escopo, simplicidade, fecundidade, precisão) na avaliação e na comparação de teorias particulares, de maneira que a escolha entre teorias não é determinada inequivocamente em toda a comunidade científica.

Paul Feyerabend usou o termo incomensurabilidade pela primeira vez em "Explanation, Reduction and Empiricism" [Explicação, redução e empirismo] (1962), para descrever a ausência de relações lógicas entre os conceitos de teorias fundamentais em sua crítica dos modelos lógico-empiristas de explicação e redução. Ele argumentou que no decorrer do avanço científico, quando teorias fundamentais mudam, significados mudam, e isso pode resultar em uma concepção nova da natureza da realidade. Ao chamar duas teorias fundamentais de incomensuráveis, Feyerabend quis dizer que elas eram incompatíveis conceitualmente: os conceitos principais de uma não poderiam ser definidos com base em termos descritivos primitivos da outra, nem relacionados a eles por meio de um enunciado empírico correto (Feyerabend 1962, 74, 90). Por exemplo, Feyerabend afirmou que os conceitos de temperatura e entropia na teoria cinética são incomensuráveis com aqueles da termodinâmica fenomenológica (1962, 78), enquanto os conceitos newtonianos de massa, distância e tempo são incomensuráveis com os conceitos da mecânica relativística (1962, 80).

Feyerabend utilizou repetidamente a ideia de incomensurabilidade para desafiar uma ampla gama de formas de conservadorismo conceitual, argumentando que elas favorecem injustificadamente conceitos bem sucedidos e entrincheirados, em detrimento de melhorias potenciais e negligenciando a possibilidade de desenvolvimento de alternativas incomensuráveis (Oberheim 2005, 2006).

Embora a aplicação da noção de incomensurabilidade a teorias científicas remonte pelo menos a LeRoy e Ajdukiewicz, na tradição convencionalista de Poincaré (Oberheim 2005), escolhemos focar nas noções mais recentes da incomensurabilidade. Este verbete está centrado e limita-se às popularizações do conceito de incomensurabilidade por Kuhn e Feyerabend em 1962, e àquelas ideias que influenciaram *diretamente* seus desenvolvimentos dessa noção. Primeiramente, caracterizaremos a noção de incomensurabilidade desenvolvida inicialmente por Kuhn, assim como sua origem e suas supostas consequências. Em seguida, examinaremos o caminho que levou Kuhn a essa ideia e ao seu desenvolvimento após 1962. As seções sobre a noção de incomensurabilidade de Feyerabend refletem essa mesma estrutura básica. A elas, segue-se uma breve comparação das concepções de Kuhn e Feyerabend sobre a incomensurabilidade, especialmente em relação à comparação de teorias.

## **2. Paradigmas revolucionários: A incomensurabilidade segundo Thomas Kuhn**

A noção kuhniana de incomensurabilidade em *A estrutura das revoluções científicas* pareceu implicar, erroneamente, que a ciência é, de alguma forma, irracional. Consequentemente, essa noção enfrentou diversas objeções e causou muita confusão. Isso levou a várias clarificações e, finalmente, ao redesenvolvimento substancial para uma versão mais precisa e restrita ao longo das décadas seguintes. Inicialmente, Kuhn usou o termo holisticamente para captar disparidades metodológicas, observacionais e conceituais entre paradigmas científicos sucessivos, que ele encontrara ao longo de suas investigações históricas sobre o desenvolvimento das ciências naturais (Kuhn 1970 [1962], 148-150). Posteriormente, ele refinou a ideia, argumentando que a incomensurabilidade ocorre devido a

diferenças na estrutura taxonômica das teorias científicas sucessivas e de suas subdisciplinas adjacentes contemporâneas. O desenvolvimento da noção de incomensurabilidade por Kuhn recebeu muita atenção, e continua a provocar considerável controvérsia.

## **2.1 A estrutura das revoluções científicas (1962)**

Em *A estrutura das revoluções científicas* (1970 [1962]), Thomas Kuhn utilizou o termo 'incomensurável' para caracterizar a natureza holística das mudanças que ocorrem em uma revolução científica. Suas investigações na história da ciência revelaram um fenômeno agora frequentemente chamado de 'perda kuhniana' [*Kuhn loss*]: problemas cujas soluções eram vitalmente importantes para tradições antigas podem desaparecer temporariamente, tornarem-se obsoletos e até mesmo não-científicos. Por outro lado, problemas que nunca haviam existido, ou cuja solução era considerada trivial, podem ganhar uma importância extraordinária em uma nova tradição. Kuhn concluiu que proponentes de teorias incomensuráveis têm diferentes concepções de suas disciplinas e do que pode ser considerado como boa ciência; e essas diferenças surgem por causa das mudanças no rol de problemas que uma teoria deve resolver e da mudança correspondente nos padrões para a admissibilidade das soluções propostas. Assim, por exemplo, a teoria newtoniana foi inicialmente rejeitada amplamente porque não explicava as forças de atração entre a matéria, algo exigido para qualquer mecânica, da perspectiva dos proponentes das teorias de Aristóteles e Descartes (Kuhn 1970 [1962], 148). De acordo com Kuhn, com a aceitação da teoria newtoniana, essa questão foi banida da ciência como ilegítima, para ressurgir apenas com a solução oferecida pela relatividade geral. Ele concluiu que as revoluções científicas alteram a própria definição de ciência.

Mudanças nos problemas e nos padrões vêm acompanhadas de mudanças conceituais correspondentes, de maneira que, após uma revolução, muitos dos (embora não todos) conceitos antigos ainda são usados, mas de uma maneira ligeiramente modificada. Tais mudanças conceituais têm tanto aspectos intensionais quanto extensionais, o que quer dizer

que os mesmos termos passam a ter significados diferentes e também a referir-se a coisas diferentes quando utilizadas por proponentes de teorias incomensuráveis competidoras. Quanto aos aspectos intensionais dos conceitos, mudanças ocorrem porque os termos utilizados para expressar uma teoria são interdefinidos e seus significados dependem das teorias às quais eles pertencem. Por exemplo, os significados dos termos ‘temperatura’, ‘massa’, ‘elemento químico’, e ‘composto químico’ dependem de quais teorias são usadas para interpretá-los. Mudanças conceituais também resultam na exclusão de alguns elementos antigos da extensão de um conceito, enquanto novos elementos passam a ser subsumidos por ele, de maneira que o mesmo termo passa a referir coisas diferentes. Por exemplo, na teoria ptolemaica, o termo ‘planeta’ referia-se ao Sol, mas não à Terra. Já na teoria copernicana, refere-se à Terra, mas não ao Sol. Teorias incomensuráveis utilizam alguns dos mesmos termos, mas com diferentes significados, para referir-se a diferentes conjuntos de coisas. Dois cientistas que percebem a mesma situação de maneira diferente, mas que ainda assim utilizam o mesmo vocabulário para descrevê-la, falam a partir de pontos de vista incomensuráveis (Kuhn 1970, 201).

Uma das afirmações mais controversas que surgem das asserções de Kuhn sobre incomensurabilidade das teorias científicas é que os proponentes de paradigmas diferentes trabalham em mundos diferentes (Kuhn 1970 [1962], 150; cf. Hoyningen-Huene 1990; 1993). A partir de experimentos da psicologia da percepção, Kuhn argumentou que o treinamento rigoroso exigido para a admissão em um paradigma condiciona as reações, expectativas e crenças do cientista (Kuhn 1970 [1962], 128; 2000 [1989], 66-71), de maneira que aprender a aplicar os conceitos de uma teoria para resolver problemas exemplares determina as experiências dos cientistas. Então, por exemplo, enquanto um proponente da teoria newtoniana vê um pêndulo, um aristotélico veria uma queda livre constrangida; onde Priestley viu ar deflogistificado, Lavoisier viu oxigênio; onde Berthollet viu um composto que poderia variar em proporção, Proust viu apenas uma mistura física. Kuhn (e Feyerabend) utilizaram a analogia da mudança gestáltica para ilustrar essa questão. Desta forma, uma

fonte importante da noção de incomensurabilidade das teorias científicas foi o desenvolvimento da psicologia Gestalt.

De acordo com Kuhn, esses três aspectos inter-relacionados da incomensurabilidade (mudanças nos problemas e padrões que definem uma disciplina, mudanças nos conceitos usados para expressar e resolver esses problemas, e mudança no mundo) conjuntamente restringem a interpretação do progresso científico como cumulativo. Progresso científico, Kuhn argumentou, não é simplesmente a descoberta contínua de novos fatos devidamente explicados. Em vez disso, revoluções mudam o que pode ser considerado como os próprios fatos. Quando teorias predominantes são substituídas por desafiantes incomensuráveis, os supostos fatos são re-descritos de acordo com princípios teóricos novos e incompatíveis com os anteriores. O objetivo principal de Kuhn em *Estrutura* é desafiar a ideia de progresso científico cumulativo, segundo o qual o que é descartado ou corrigido ao longo do avanço científico é aquilo que nunca foi realmente científico em primeiro lugar, e Kuhn utilizou a incomensurabilidade como base para seu desafio. Em vez de compreender o progresso científico como um processo que muda em direção a alguma verdade fixa, Kuhn comparou sua sugestão com a de Darwin: o progresso científico é como a evolução, seu desenvolvimento deve ser entendido sem referência a um objetivo fixo e permanente ([1970] 1962, 173).

## **2.2 A rota kuhniana para a incomensurabilidade**

### **2.2.1 A descoberta da incomensurabilidade por Kuhn**

De acordo com Kuhn, ele descobriu a incomensurabilidade quando era um estudante de pós-graduação, na segunda metade da década de 40, enquanto debatia-se com passagens aparentemente sem sentido na física aristotélica (Kuhn 2005 [1989], 59-60). Ele não conseguia acreditar que alguém tão extraordinário como Aristóteles poderia tê-las escrito. Por fim, começaram a surgir padrões nas passagens desconcertantes, e de repente, de uma vez só, todo o texto fez sentido para ele: uma mudança gestáltica que ocorreu quando ele mudou

os significados de alguns dos termos centrais. Ele viu esse processo de mudança de significado como um método de recuperação histórica. Ele percebeu que em suas leituras anteriores, estava projetando significados contemporâneos nas fontes históricas (história Whiggish), e que ele precisaria desfazer-se delas a fim de remover a distorção e compreender o sistema aristotélico por si mesmo (história hermenêutica). Por exemplo, ao encontrar a palavra “movimento” em Aristóteles (a tradução padrão do grego *kinesis*), ele estava pensando em termos de mudança de posição dos objetos no espaço (como fazemos atualmente). Mas, para aproximar-se do uso original de Aristóteles, teve que expandir o significado de “movimento” para abranger uma gama mais ampla de fenômenos que incluem vários tipos de mudança, como crescimento e diminuição, alternância, e geração e corrupção, tornando o movimento de um objeto no espaço (deslocamento ou ‘locomoção’) apenas um caso especial de movimento. Kuhn percebeu que esses tipos de diferenças conceituais indicavam rupturas entre diferentes modos de pensar, e suspeitou que essas rupturas deveriam ser significativas tanto para a natureza do conhecimento quanto para o sentido em que se pode dizer que o desenvolvimento do conhecimento progride. Ao fazer essa descoberta, Kuhn mudou seus planos profissionais, abandonando a física teórica para investigar esse estranho fenômeno. Aproximadamente quinze anos depois, o termo ‘incomensurabilidade’ apareceria pela primeira vez no seu clássico *A estrutura das revoluções científicas* (1962).

### **2.2.2 Substituição conceitual e impregnação teórica da observação: Ludwik Fleck**

De todas as fontes que influenciaram Kuhn quanto à incomensurabilidade, pelo menos uma merece atenção especial. No prefácio de *A estrutura das revoluções científicas*, Kuhn reconheceu uma dívida profunda com Ludwik Fleck, bacteriologista que desenvolveu a primeira sociologia explícita da ciência natural, e que antecipou muitas das concepções contemporâneas sobre a construção social do conhecimento. Por volta de 1950, Kuhn foi seduzido pela relevância potencial do título paradoxal da obra de Fleck para sua experiência com a incomensurabilidade: *Entstehung und Entwicklung einer wissenschaftlichen Tatsache:*

*Einführung in die Lehre von Denkstil und Denkkollektiv* (1935) (*Genesis and Development of a Scientific Fact*, 1979) [Gênese e desenvolvimento do fato científico]. Ali, e em trabalhos anteriores, Fleck já havia usado 'inkommensurabel' para descrever estilos diferentes de pensamento no interior das ciências naturais assim como para discutir as ramificações das mudanças conceituais radicais na história da ciência. Por exemplo, Fleck utilizou o termo 'inkommensurabel' para descrever as diferenças entre 'pensamento médico' e 'pensamento científico'. O primeiro trata de fenômenos irregulares e temporalmente dinâmicos, como uma doença, enquanto o último trata de fenômenos uniformes (Fleck 1986 [1927], 44-45). Fleck também utilizou o termo 'inkommensurabel' para descrever substituições conceituais em transições teóricas no que ele considerava ser a mais vital das ciências naturais, as ciências médicas. Por exemplo, ele afirmou que um conceito antigo de doença tornara-se incomensurável com um conceito mais novo que não era completamente adequado para substituí-lo (Fleck 1979 [1935], 62). Enquanto o programa de epistemologia comparativa de Fleck antecipou as ideias de Kuhn em muitos aspectos significativos, essas concepções também são completamente diferentes (Harwood 1986; Oberheim 2005). As diferenças mais penetrantes em relação à incomensurabilidade são que Fleck trata significado e mudança de significado como uma função de como conceitos são recebidos e desenvolvidos pelo coletivo, enquanto que, para Kuhn, são os indivíduos que desenvolvem e aplicam os conceitos que fazem a ciência avançar. Além disso, para Fleck, mudanças de significado são características permanentes do desenvolvimento científico, enquanto Kuhn distingue entre desenvolvimento científico normal, que não altera os significados, e o desenvolvimento revolucionário, que altera.

Mesmo assim, Fleck enfatizou os três aspectos inter-relacionados das mudanças que Kuhn chamou de revoluções e que tentou captar ao chamar as teorias de 'incomensuráveis' (mudanças em problemas e padrões (1979 [1935], 75-76; 1979 [1936], 89), mudança conceitual (por exemplo, 1979 [1935]; 1979 [1936], 72, 83) e mudança de mundo (por exemplo, 1986 [1936], 112)). Fleck argumentou que a ciência não se aproxima da verdade porque estilos de pensamento sucessivos originam novos problemas e descartam formas

anteriores de conhecimento (Fleck 1986 [1936], 111–112; 1979 [1935], 19, 51, 137–139; cf. Harwood 1986, 177). Ele enfatizou que os termos científicos adquirem seus significados por meio de sua aplicação em um contexto teórico particular e que eles mudam quando as teorias mudam ao longo do avanço científico, inclusive oferecendo os exemplos de ‘elementos químicos’ e ‘compostos’, repetidos por Kuhn (Fleck 1979 [1935], 25, 39, 40, 53–54). Fleck enfatizou a impregnação teórica da observação com referência explícita às mudanças gestálticas, salientando que um ‘estilo de pensamento’ determina não somente os significados dos conceitos que são utilizados, mas também a percepção dos fenômenos a serem explicados, apresentando exemplos da história da representação anatômica (1979 [1935], 66; 1986 [1947]). Fleck (como Kuhn, Feyerabend e Wittgenstein) reconheceu a obra de Wolfgang Köhler em psicologia da percepção a esse respeito. Ele concluiu que o avanço científico é não-cumulativo, que diferenças conceituais entre membros de comunidades científicas diferentes causam dificuldades na comunicação entre eles (1979 [1936], 109) e para os historiadores que posteriormente tentarem compreender as ideias anteriores (1979 [1936], 83–85, 89). Fleck também enfatizou que a mudança de significado no curso do avanço científico causa falhas na tradução entre teorias, antecipando um aspecto central da noção kuhniana tardia de incomensurabilidade taxonômica (por exemplo, 1986 [1936], 83). Embora Fleck seja frequentemente celebrado como uma figura influente no desenvolvimento da sociologia da ciência, análises recentes vêm criticando muito as suas pesquisas científicas, considerando-as de má qualidade, e suas práticas experimentais duvidosas, que incluem experimentos em seres humanos, com violações éticas óbvias (Hedfors 2007a; 2007b).

### **2.2.3 Psicologia Gestalt e percepção organizada**

Outra fonte primária da ideia de Kuhn de incomensurabilidade de teorias científicas é a psicologia Gestalt, especialmente aquela desenvolvida por Wolfgang Köhler. Ele enfatizou o papel ativo que a organização tem na percepção e argumentou que, em psicologia, começa-se com *Gestalten* (totalidades organizadas e segregadas, tais como os objetos da percepção

humana ou comportamentos humanos identificáveis) e então, procede-se à descoberta de suas partes naturais (e não vice-versa, como na física de partículas). Köhler sugeriu que, enquanto *Gestalten* são comuns em psicologia, sua existência na física violaria os fundamentos da ciência exata, enfatizando que as ideias e a linguagem da física foram estabelecidas de uma perspectiva diferente daquela da psicologia Gestalt (1938 [1920], 17). A estratégia de Köhler para tentar superar o problema entre a relação das totalidades psicológicas e os conceitos reducionistas da física foi tentar encontrar as primeiras na física. Nas frases iniciais da versão condensada em inglês das investigações de Köhler sobre as relações entre os conceitos mentais da psicologia e os conceitos materiais da física, Köhler escreve: “A fim de orientar-se, quando em companhia das ciências naturais, é dever da psicologia descobrir conexões entre seus fenômenos próprios e os de outras disciplinas sempre que possível. Se essa busca falhar, então é dever da psicologia reconhecer que suas categorias e as da ciência natural são incomensuráveis” (1938 [1920], 17).

### **2.3 O desenvolvimento subsequente da noção de incomensurabilidade em Kuhn**

Kuhn continuou esforçando-se em desenvolver e refinar sua compreensão da incomensurabilidade até sua morte, em 1996. Embora o desenvolvimento de sua noção de incomensurabilidade tenha passado por vários estágios (cf. Hoyningne-Huene 1993, 206-222), ele afirmou ter feito “uma série de avanços muito significativos”, a partir de 1987 (Kuhn 2000 [1993], 228). Eles são descritos em vários ensaios e conferências publicadas que foram reunidas em (Kuhn 2000, cf. caps. 3, 4, 5, 10 & 11), e em um último livro ainda não publicado sobre incomensurabilidade (Conant and Haugeland 2000, 2). A natureza desses desenvolvimentos é controversa. Alguns comentadores afirmam que a tese da incomensurabilidade de Kuhn passou por uma ‘grande transformação’ (Sankey, 1993), enquanto outros (incluindo o próprio Kuhn) veem somente uma caracterização mais específica do *insight* central original (Hoyningen-Huene 1993, 212; Kuhn 2000, [1983], 33 ss.; Chen 1997). Distingue-se a caracterização holística original de Kuhn da incomensurabilidade

em duas teses separadas. ‘Incomensurabilidade taxonômica’ envolve mudança conceitual, diferentemente da ‘incomensurabilidade metodológica’, que envolve os valores epistêmicos utilizados para avaliar teorias (Sankey 1991; Sankey e Hoyningen-Huene 2001; ver Seção 2.3.2).

### **2.3.1 Incomensurabilidade taxonômica**

Kuhn continuou a enfatizar a diferença entre o crescimento cumulativo normal, que não altera os conceitos existentes, como a descoberta da lei de Boyle, que deixou os conceitos de gás, pressão e volume e os instrumentos usados para medi-los inalterados, das descobertas revolucionárias que não poderiam ser feitas com base nos conceitos previamente disponíveis, como a descoberta da teoria de Newton (Kuhn 2000 [1981], 14). Tais desenvolvimentos exigem *substituir* conceitos existentes por novos conceitos que são *incompatíveis* com ideias anteriores. Para ajudar a explicar a natureza dessa incompatibilidade conceitual, Kuhn adotou a noção de uma ‘taxonomia lexical’ ou ‘léxico’, da linguística (2000 [1989]). Uma taxonomia lexical armazena e estrutura o vocabulário utilizado para expressar leis e teoria; ou mais especificamente, os tipos de coisas que elas descrevem. Consiste no “módulo mental” em que membros de uma comunidade de fala armazenam os tipos de termos que são pré-requisitos para generalização e descrição (2000 [1991], 92-97; 2000 [1993], 229, 233, 238-239, 242, 244). Para comunicar, os termos que os falantes utilizam não têm que ter os mesmos significados, e os falantes também não precisam usar os mesmos critérios para identificar algo como sendo membro de um tipo particular (um ornitólogo pode identificar patos por meio de seus bicos, e outro, por meio dos seus rabos) (cf. Hoyningen-Huene 1993, 99). Os falantes precisam apenas compartilhar estruturas lexicais homólogas; ou seja, vocabulários que incorporem o mesmo sistema de classificação taxonômica. Isso basta para garantir que eles se refiram aos mesmos objetos com os mesmos termos de espécie. Se duas teorias não compartilham a mesma estrutura lexical taxonômica, então elas são incomensuráveis (2000 [1989], 63; 2000 [1991]; cf. Sankey 1997).

Para ajudar a explicar a incomensurabilidade em termos de classificação taxonômica, Kuhn desenvolveu o *princípio de não-sobreposição* [*no-overlap principle*]. O princípio de não-sobreposição impede a classificação cruzada de objetos em diferentes espécies na taxonomia de uma teoria. De acordo com o princípio da não-sobreposição, não haverá sobreposição entre os referentes de dois termos de espécie a não ser que eles estejam relacionados como espécies ao gênero. Por exemplo, não há cachorros que sejam também gatos; o ouro não é também prata, e é isso que torna os termos ‘cachorros’, ‘gatos’, ‘prata’ e ‘ouro’, termos de espécie (Kuhn 2000 [1991], 92). Tais termos de espécie são utilizados para expressar leis e teorias e devem ser aprendidos *conjuntamente* através de experiências (2000 [1993], 230; cf. Barker *et al.* 2003, 214 ss.). Existem duas possibilidades. A maioria dos termos de espécie deve ser aprendida como membros de um ou outro conjunto de contraste. Por exemplo, para aprender o termo “líquido”, deve-se também dominar o uso dos termos “sólido” e “gás”. Outros tipos de termos de espécie não são aprendidos através de conjuntos de contraste, mas junto com termos intimamente relacionados por meio da sua aplicação conjunta a situações para exemplificar leis naturais. Por exemplo, o termo “força” deve ser aprendido junto com termos como “massa” e “peso” através de uma aplicação da lei de Hooke ou das três leis do movimento de Newton, ou ainda, a primeira e a terceira leis em conjunto com a lei da gravidade (2000 [1993], 231). De acordo com Kuhn, revoluções científicas mudam as relações estruturais entre termos de espécie pré-existentes, violando o princípio da não-sobreposição (2000 [1991], 92-96). Isso quer dizer que teorias separadas por uma revolução classificam de maneira cruzada as mesmas coisas em conjuntos de espécies mutuamente excludentes. Uma espécie de uma taxonomia é mutuamente excludente com outra se ela não puder ser simplesmente introduzida na outra porque os objetos a que ela se refere estão sujeitos a conjuntos diferentes de leis naturais. Isso resultaria em expectativas conflitantes em relação aos mesmos objetos, perda de relações lógicas entre enunciados feitos com esses conceitos, e em última instância, em incoerência e falha na comunicação (Kuhn 2000 [1993], 232, 238). Por exemplo, a teoria ptolemaica classifica o Sol como um planeta, onde planetas orbitam a Terra, enquanto a teoria copernicana classifica o Sol como uma estrela, onde planetas orbitam

estrelas como o Sol. Um enunciado correto, de acordo com a teoria copernicana, tal como “planetas orbitam o Sol” é incoerente com o vocabulário ptolemaico (2000 [1991], 94). Além disso, a afirmação copernicana de que os planetas orbitam o Sol não poderia nem mesmo ser feita sem abandonar os conceitos ptolemaicos e desenvolver novos conceitos para *substituí-los* (e não para suplementá-los).

Além disso, Kuhn (avançando em direção à concepção de Feyerabend) passou a afirmar que os mesmos tipos de dificuldades na comunicação que surgem devido à incomensurabilidade entre membros de comunidades científicas diferentes separadas pela passagem do tempo *também* ocorrem entre membros de subdisciplinas contemporâneas diferentes que resultam das revoluções científicas (Kuhn 2000 [1993], 238). Isso representa uma mudança significativa no seu modelo de períodos de avanços científicos original, e um deslocamento correspondente em sua aplicação da noção de incomensurabilidade. Kuhn não mais representa o avanço científico como uma progressão *linear* da ciência pré-normal à ciência normal, da crise à revolução *que resulta em um novo período de ciência normal*. Em vez disso, revoluções científicas são comparadas ao processo de especiação na biologia, em que ocorrem *ramificações* para subdisciplinas que se assemelham a uma árvore filogenética. A aplicação da relação de incomensurabilidade não é mais restrita a episódios diacrônicos de avanço científico em que duas teorias são separadas por uma revolução. As subdisciplinas contemporâneas que resultam de uma revolução científica podem também ser incomensuráveis entre si. A incomensurabilidade deriva do treinamento diferente exigido para dominar os termos de espécie incompatíveis utilizados para expressar as leis e teorias. Esses termos de espécie compartilhados classificam de maneira cruzada o mesmo conjunto de objetos em diferentes conjuntos de espécies, resultando em taxonomias lexicais mutuamente excludentes que violam o princípio da não-sobreposição. Além disso, agora, ambos os processos (progresso científico e evolução biológica) são não somente similares, no sentido em que não têm fixado previamente nenhum tipo de meta (isto é, a verdade), mas “empurrados” (isto é, para longe das anomalias que desempenham um papel análogo na pressão seletiva), como também no sentido em que a incomensurabilidade das teorias

científicas é apresentada como análoga aos mecanismos de isolamento exigidos para especiação (Kuhn 2000 [1990], 94-99).

Kuhn comparou a função de tais taxonomias lexicais com o *a priori* kantiano quando considerado em um sentido relativizado. Cada léxico torna uma forma de vida correspondente possível no interior do qual a verdade ou a falsidade de proposições podem ser ambas afirmadas e racionalmente justificadas. Por exemplo, com o léxico aristotélico, pode-se falar da verdade ou da falsidade das asserções aristotélicas, mas esses valores de verdade não têm qualquer influência na verdade de asserções aparentemente similares feitas com o léxico newtoniano (Kuhn 2000 [1993], 244). Um léxico é, assim, *constitutivo* dos objetos do conhecimento (Kuhn 2000 [1993], 245); e conseqüentemente, Kuhn rejeitou caracterizações de progresso científico em que a ciência aproxima-se progressivamente da verdade: “*nenhuma métrica compartilhada* está disponível para comparar nossas asserções... e assim, fornecer uma base para uma afirmação de que a nossa afirmação (ou, quanto a isso, a dele) está mais próxima da verdade” (2000 [1993], 244). Em vez disso, o estatuto lógico de uma estrutura lexical, como a dos significados das palavras em geral é o de uma convenção, e a justificação de um léxico ou de uma mudança lexical pode ser apenas pragmática (2000 [1993], 244). Kuhn assim reafirmou sua declaração anterior de que a noção de combinação ou correspondência entre a ontologia de uma teoria e a contrapartida real na natureza é ilusória em princípio (1970, 206; 2000 [1993], 244). As implicações que a incomensurabilidade tem para o realismo científico têm sido discutidas amplamente e continuam a ser controversas (cf. Devitt 2001, Oberheim e Hoyningen-Huene 1997, Sankey no prelo).

Um léxico é pré-requisito não somente para tornar enunciados significativos, mas também para estabelecer limites sobre o que pode ser dito significativamente na comunidade de falantes que o compartilha: “Não há, por exemplo, nenhuma maneira, mesmo em um vocabulário newtoniano enriquecido, de expressar as proposições aristotélicas regularmente mal interpretadas como asserindo a proporcionalidade da força e do movimento ou a impossibilidade de um vazio. Ao usar nosso próprio léxico, essas proposições aristotélicas não podem ser expressadas – são simplesmente inefáveis – e o princípio de não-sobreposição nos

impede de acessar os conceitos exigidos para expressá-las” (Kuhn 2000 [1993], 244; cf. 2000 [1989], 76). Dessa maneira, a noção posterior de incomensurabilidade das teorias científicas de Kuhn está baseada na *efabilidade*. A estrutura do léxico compartilhada por uma comunidade particular determina como o mundo pode ser descrito pelos seus membros, e também como eles interpretarão mal a história de sua própria disciplina; isto é, a não ser que aprendam a compreender os termos antigos conforme a estrutura do léxico antigo. Onde Kuhn havia previamente aproximado o processo por meio do qual historiadores chegam a compreender a ciência obsoleta como um tipo especial de tradução, ele retirou suas declarações, insistindo que o processo é de aprendizagem da linguagem, não de tradução (2000 [1993], 238, 244). Kuhn afirmou frequentemente que teorias incomensuráveis são intraduzíveis (por exemplo, Kuhn 1990, 94). Contudo, ele também enfatizou que a tradução não é necessária nem na comparação de teorias incomensuráveis nem no método histórico hermenêutico para compreender ciências obsoletas (Kuhn 1993, 237, 238, 244). Para superar as barreiras impostas pela incomensurabilidade à compreensão de ciências obsoletas, e para compreender o vocabulário técnico especial utilizado pelas subdisciplinas contemporâneas relacionadas filogeneticamente, não é necessário e nem possível traduzi-las. Na verdade, é necessário tornar-se *bilíngue*, aprendendo a usar (e a manter separadas) as taxonomias lexicais incongruamente estruturadas que subjazem a leis e teorias diferentes.

### **2.3.2 Incomensurabilidade metodológica**

À medida que Kuhn refinou sua noção de incomensurabilidade como um tipo especial de incompatibilidade conceitual, alguns comentadores começaram a distingui-la da ‘incomensurabilidade metodológica’. Incomensurabilidade metodológica é a ideia de que não há padrões objetivos e compartilhados de avaliação de teorias científicas, de maneira que não há padrões neutros ou externos que determinem inequivocamente a avaliação comparativa de teorias competidoras (Sankey e Hoyningen-Huene 2001, xiii). Essa ideia também tem sido discutida em detalhe sob a rubrica “Subdeterminação de Kuhn” [*Kuhn-underdetermination*]

(Carrier 2008, 278). A ideia básica foi desenvolvida a partir de Kuhn e das objeções de Feyerabend à concepção tradicional que afirma que uma característica distintiva da ciência é o seu método científico uniforme e invariante, que permanece fixo ao longo de seu desenvolvimento (Kuhn 1970 [1962], 94, 103; Feyerabend 1975, 23–32; cf. Farrell 2003). Feyerabend apresenta um famoso argumento, segundo o qual toda regra metodológica proposta tem sido fecundamente violada em algum momento do processo de avanço científico, e é somente por meio da violação dessas regras que os cientistas poderiam ter dado os passos progressivos pelos quais são reconhecidos (1975). Ele concluiu que a ideia de um método científico fixo e historicamente invariante é um mito. Não existem regras metodológicas aplicáveis universalmente. A única regra metodológica aplicável universalmente é ‘vale tudo’, que mantém a universalidade, em troca de vacuidade (1970a, 105). Kuhn desafiou a concepção tradicional do método científico como sendo um conjunto de regras, afirmando que os padrões para apreciação de teorias, como simplicidade, precisão, coerência, escopo e fecundidade (1977, 322), dependem e variam de acordo com o paradigma dominante em curso. Ele é frequentemente citado por ter assinalado que, assim como nas revoluções políticas, também na escolha de um paradigma, não há um padrão mais alto do que o assentimento da comunidade relevante (1970 [1962], 94), e por ter argumentado que “não existem algoritmos neutros para a escolha de uma teoria. Nenhum procedimento sistemático de decisão, mesmo quando aplicado adequadamente, deve necessariamente conduzir cada membro de um grupo a uma mesma decisão” (1970, 200). Kuhn desenvolveu a ideia de que tais padrões epistêmicos não funcionam como regras que determinam a escolha racional de teorias, mas meramente como valores que a guiam (1977, 331). Cientistas diferentes aplicam esses valores diferentemente, e eles podem até mesmo inclinar-se em diferentes direções, de maneira que pode haver desacordo racional entre cientistas de paradigmas incomensuráveis, que sustentam teorias diferentes por considerar os mesmos valores de maneira diferente.

### **3. Combatendo o conservadorismo conceitual: A incomensurabilidade segundo Paul Feyerabend**

Um exame da utilização e do desenvolvimento da ideia de incomensurabilidade entre teorias científicas de Feyerabend revela o quão difundida ela estava antes de 1962. Também revela o núcleo comum velado da filosofia da ciência de Feyerabend que, de outra maneira, parece ser uma série de artigos críticos sobre tópicos sem relação (Oberheim 2005, 2006). Esse núcleo comum é a utilização da ideia de incomensurabilidade para desafiar várias formas diferentes de conservadorismo conceitual na ciência e na filosofia. O conservadorismo conceitual favorece injustamente ideias existentes em detrimento de aperfeiçoamentos potenciais oferecidos por alternativas incomensuráveis: a ideia de Heisenberg de teorias fechadas (1951) tenta distinguir descrições filosóficas da ontologia da física (1954), o papel fundacional da tese da estabilidade no positivismo lógico (1951, 1958a), o princípio da complementaridade de Bohr (1958b, 1961, 1962, 1965b, 1970a), e os modelos lógico-empiristas de explicação e redução (1962, 1963a). Ele também utilizou a ideia de incomensurabilidade para atacar formas de conservadorismo conceitual implícitas nos modelos de teste de teorias promovidos por empiristas clássicos, positivistas lógicos, empiristas lógicos, Popper e Kuhn (1958a, 1962, 1963a, 1965b, 1970a), necessidades transcendentais kantianas (1962, 1965b, 1965c), enigmas filosóficos tradicionais como o problema da existência do mundo exterior, o problema da existência de outras mentes, e especialmente o problema mente-corpo (1962, 1965b, 1967, 1969a), certos aspectos da filosofia tardia de Wittgenstein, bem como o movimento linguístico-filosófico que toma a utilidade dos conceitos do cotidiano como dados e simplesmente tenta analisá-los, sem mudar seus significados (ou 'argumentos da sinonímia') (1962, 1963a, 1965b, 1969b, 1970a), a concepção de verossimilhança de Popper (1970b), e caracterizações realistas do conhecimento e do avanço científico (1970b). Feyerabend argumentou que todas essas concepções não podem acomodar o fato de que a história da ciência revela desenvolvimentos

revolucionários em que alternativas incomensuráveis substituíram ideias existentes ao longo do progresso.

### **3.1 ‘Explicação, redução e empirismo’ (1962)**

“Explicação, redução e empirismo” [Explanation, reduction and empiricism] (1962) foi um marco, no qual Paul Feyerabend introduziu pela primeira vez o termo “incomensurável” como parte de uma crítica ao *conservadorismo conceitual* que ele encontrou na teoria da redução de Nagel e na teoria da explicação de Hempel-Oppenheim. Ele argumentou que é impossível obter descrições formais da redução e da explicação para teorias gerais porque elas não podem acomodar o desenvolvimento de conceitos incomensuráveis ao longo do avanço científico. Feyerabend utilizou a noção de incomensurabilidade para caracterizar a relação entre duas teorias científicas fundamentais e universais interpretadas realisticamente, isto é, como supostas descrições da realidade. Ao limitar a incomensurabilidade a teorias universais, Feyerabend tinha em mente apenas aquelas que se aplicam a todos os objetos do universo. Ao limitar a incomensurabilidade a teorias fundamentais, Feyerabend quis dizer somente aquelas que tivessem implicações ontológicas, isto é, aquelas que têm implicações sobre a própria natureza da realidade. De acordo com Feyerabend, teorias científicas fundamentais e universais podem ser incomensuráveis somente se são interpretadas realisticamente porque de outra maneira não poderiam competir como a única teoria verdadeira, e assim não seriam mutuamente excludentes. Ao chamar tais teorias de incomensuráveis, Feyerabend quis dizer que elas eram conceitualmente incompatíveis: os conceitos principais de uma não poderiam ser definidos com base em termos descritivos primitivos da outra, nem relacionados a eles via enunciados empíricos corretos (1962, 74, 90). Como consequência, é impossível reduzir formalmente teorias incomensuráveis.

Em 1962, Feyerabend ilustrou a incomensurabilidade de teorias científicas com uma comparação entre seis pares de conceitos centrais retirados de três episódios de mudança de teoria fundamental ao longo do avanço científico: (1) a caracterização dinâmica do impulso na

teoria impulsionar do movimento e o conceito de força no aparato conceitual da mecânica newtoniana, (2 & 3) os conceitos de temperatura e entropia na transição da termodinâmica fenomenológica para a teoria cinética, e (4, 5 & 6) os conceitos de massa, comprimento e tempo na transição da mecânica newtoniana para a relativística. Feyerabend adicionou subsequentemente o exemplo da ótica geométrica e da ótica de ondas (1965b, 227). Em 1962, Feyerabend também afirmou que a incomensurabilidade ocorre entre teorias fundamentais tradicionais na filosofia, e consequentemente surge nas discussões do problema mente-corpo, da realidade exterior e das outras mentes (1962, 31, 90). Feyerabend argumentou que problemas filosóficos tão antigos não foram resolvidos porque os disputantes resistem ao tipo de mudança de significado necessária para sua dissolução. Conceitos são incomensuráveis, de acordo com Feyerabend, porque pertencem a perspectivas teóricas mutuamente excludentes. Essas perspectivas são mutuamente excludentes porque explicam de maneiras incompatíveis a natureza da realidade. Feyerabend argumentou que, porque os significados dos termos são providos pelas teorias às quais pertencem, quando há mudanças teóricas, ocorrem mudanças de significado (1962, 68, 94). Feyerabend argumentou que essas mudanças nos significados afetam nossos termos teóricos e observacionais assim como nossa concepção da natureza da realidade. Quando isso ocorre, há incomensurabilidade, ou como Feyerabend posteriormente caracterizou sua concepção: “uma teoria é incomensurável com outra se suas consequências ontológicas são incompatíveis com as consequências ontológicas da outra” (1981a, xi). A ideia pretende captar a incompatibilidade conceitual devida a mudanças de significados que ocorrem em transições teóricas que afetam nossas crenças ontológicas. Duas teorias fundamentais são incomensuráveis porque os significados de seus termos são determinados pelos princípios teóricos que governam seu uso, e esses princípios são qualitativamente incompatíveis (1962, 58).

Feyerabend utilizou a ideia de incomensurabilidade para tentar expor um elemento dogmático que os empiristas contemporâneos compartilham com escolas filosóficas tais como o platonismo e o cartesianismo (das quais os empiristas tentaram afastar-se ao insistir em um fundamento empírico para o conhecimento científico). O elemento dogmático deve-se à

suposição de que os termos empíricos permanecem estáveis ao longo das transições teóricas; ou o que Feyerabend chamou o princípio de invariância do significado (1962, 30). Feyerabend argumentou que esse princípio é incoerente com a existência de conceitos incomensuráveis. Feyerabend retirou duas consequências principais do *insight* de que alguns pares de teorias científicas sucessivas são incomensuráveis. Primeiro, teorias científicas sucessivas que são incomensuráveis não têm relações lógicas: “a ordem introduzida em nossas experiências pela teoria de Newton é retida e aperfeiçoada pela relatividade. Isso significa que os conceitos da teoria da relatividade são suficientemente ricos para permitir que enunciemos todos os fatos que foram enunciados anteriormente com a ajuda da física newtoniana. No entanto, esses dois conjuntos de categorias são completamente diferentes e *não têm nenhuma relação lógica entre si*” (1962, 88-89). Esses conceitos têm diferentes significados, por isso não podem ser apreendidos a relações lógicas formais. Segundo, exige-se uma revisão na concepção empirista lógica de avanço científico. Teorias mais antigas, e os conceitos usados para expressá-las, não são absorvidos corretamente, e assim, legitimados. Eles são, na verdade, rejeitados e substituídos, tendo sido falsificados. Dessa maneira, as descrições conceitualmente conservadoras (ou retentivas) de redução e explicação dos empiristas lógicos são minadas. Elas não podem acomodar o desenvolvimento de conceitos incomensuráveis ao longo do avanço científico.

Em 1965, Feyerabend tentou clarificar sua noção, afirmando que duas teorias são incomensuráveis quando os significados de seus principais termos descritivos dependem de princípios mutuamente incompatíveis (1965c, 227; 1975, 269–270, 276). Ele também afirmou que duas teorias são incomensuráveis se uma nova teoria implica que todos os conceitos da teoria precedente têm extensão vazia ou se ela introduz regras que mudam o próprio sistema de classes (1965a, 268). Nessa concepção, para que haja incomensurabilidade, a “situação deve ser ajustada de tal maneira que as condições de formação dos conceitos em uma teoria proíbem a formação dos conceitos básicos da outra” (Feyerabend 1978, 68; cf. 1975, 269; 1981b, 154). Ainda mais tarde, ele afirmou que, com o termo “incomensurabilidade”, ele

sempre quis dizer “desarticulação dedutiva, e nada mais” (1977, 35). Todas essas formulações subsequentes pretendiam apreender a mesma ideia.

### **3.2 A rota feyerabendiana para a incomensurabilidade**

Embora Feyerabend tenha usado o termo ‘incomensurabilidade’ primeiramente para descrever teorias científicas fundamentais sucessivas em 1962, ele desenvolvera sua noção de incomensurabilidade de teorias científicas mais de dez anos antes do aparecimento de *Estrutura das revoluções científicas* (1962) de Kuhn. Feyerabend primeiramente discutiu sua noção de conceitos incomensuráveis com o Círculo de Kraft, em 1949-1951 (Feyerabend 1978, 108 ss.). O Círculo de Kraft foi um grupo de estudantes cujo nome vem de Viktor Kraft (orientador de Feyerabend no Ph.D.) e modelado segundo o Círculo de Viena (do qual Kraft foi um membro). Convidados como Elizabeth Anscombe e Ludwig Wittgenstein frequentaram reuniões (Feyerabend 1978, 115–116; 1993, 259–260; 1995, 92). A rota de Feyerabend para a ideia de incomensurabilidade de teorias científicas foi influenciada fortemente por uma série de indivíduos proeminentes que discutiram uma ampla gama de tópicos relacionados.

Feyerabend traçou as origens de sua noção de incomensurabilidade até sua ‘tese I’ em “Uma tentativa de uma interpretação realista da experiência” [An attempt at a realistic interpretation of experience”] (1958a), uma versão condensada da sua tese de doutorado de 1951 *Zur Theorie der Basissätze* (Sobre a teoria dos enunciados protocolares). Em (1958a), Feyerabend criticou duas concepções de significado. Ele argumentou contra a ideia lógico-positivista de que o significado de um termo observacional é determinado por experiências imediatas, assim como também argumentou contra a ideia wittgensteiniana de que o significado de um termo observacional é determinado pelo seu uso. Em vez disso, Feyerabend argumentou a favor da ideia de que o significado de um termo, mesmo um termo observacional, é determinado pelo seu contexto teórico, ou mais precisamente, pelos princípios teóricos que governam seu uso correto de acordo com nossas melhores teorias. Por exemplo, conforme Feyerabend, o significado correto do termo ‘temperatura’ não é

determinado pelo seu uso cotidiano, mas pelos princípios da termodinâmica estatística. A partir dessa perspectiva, Feyerabend desenvolveu a ‘tese I’, segundo a qual a interpretação de uma linguagem observacional é determinada pelas teorias que usamos para explicar o que observamos, e que muda tão logo as teorias mudem (1958a, 163). Feyerabend continuou a argumentar que, quando teorias antigas são substituídas, os significados dos termos observacionais usados para testar as teorias mudam. Assim como em (1962), o resultado é a incomensurabilidade: a ideia de que teorias científicas sucessivas são incompatíveis conceitualmente e conseqüentemente logicamente disjuntas. A ‘tese I’ de Feyerabend é, assim, uma versão prévia da sua tese da incomensurabilidade. Em seu (1958a), com sua noção de incomensurabilidade, Feyerabend desafiou um conservadorismo conceitual implícito no positivismo lógico: a suposição de que termos teóricos derivam seu significado somente através da sua conexão com a experiência, e que a própria experiência é um fundamento estável (ou invariante) que pode servir de base para o significado teórico. Em vez dessa versão ascendente [*bottom-up*] para a relação entre experiência e conhecimento teórico, de acordo com a qual a experiência determina o significado de nossos termos teóricos, Feyerabend argumentou em favor de uma versão descendente [*top-down*] de acordo com a qual nossas teorias determinam o significado que vinculamos a nossas experiências. Conforme Feyerabend, a experiência não pode ser tomada sem questionamentos como algum tipo de base neutra e fixa para a comparação de teorias. Na verdade, ela assume seu caráter particular à luz das teorias que levamos a ela. Uma investigação da fonte dessas ideias revela o antepassado da noção de incomensurabilidade na história contemporânea e na filosofia da ciência.

### **3.2.1 Progresso através de mudança de significado: Pierre Duhem**

Feyerabend serviu-se profundamente de *The aim and structure of physical theory* [O objetivo e a estrutura da teoria física] (1954 [1906]), de Duhem, em seu desenvolvimento da noção de incomensurabilidade de teorias científicas. Os principais pontos que Feyerabend

ênfatisou ao chamar teorias científicas de incomensuráveis já haviam sido desenvolvidos por Duhem, que argumentara que a lógica é insuficiente para determinar o resultado de disputas teóricas nas ciências naturais, e documentou as dificuldades que os historiadores têm para compreender o desenvolvimento das ciências naturais devido a mudanças de significado. Duhem também já havia ressaltado as dificuldades de comunicação entre proponentes de teorias científicas competidoras por causa das diferenças nos significados. Por exemplo, Duhem afirmara que o que um físico enuncia como resultado de um experimento não é apenas a narração de alguns dos fatos observados. Na verdade, trata-se da interpretação desses fatos com base nas teorias que os cientistas consideram verdadeiras (1954 [1906], 159). Segue-se que, de acordo com Duhem, para compreender os significados que os cientistas atribuem aos seus próprios enunciados, é necessário compreender as teorias que eles usam a fim de interpretar o que eles observam. Assim, Duhem afirmara uma versão prévia da tese da incomensurabilidade de Feyerabend. Além disso, Duhem limitou explicitamente sua discussão a teorias físicas não-instanciais, por oposição a meras leis experimentais. Isso é muito similar aos critérios que marcam as diferenças mais significativas entre os desenvolvimentos de Kuhn e Feyerabend da ideia de incomensurabilidade (ver seção 4), e que também delimita o uso de Einstein de ‘incomensurável’, ao discutir os problemas da comparação de teorias (ver seção 3.2.3).

Após explicar que o significado de um termo depende da teoria à qual ele pertence, e que a consequência de um avanço teórico é a mudança de significado, Duhem continuou: “Se as teorias admitidas por esse físico são aquelas que nós aceitamos, e se nós concordamos em seguir as mesmas regras na interpretação dos mesmos fenômenos, falamos a mesma linguagem e podemos nos entender. Mas isso nem sempre é o caso. Não é assim quando discutimos os experimentos de um físico que não pertence à nossa escola; e especialmente não é assim quando discutimos os experimentos de um físico separado de nós por cinquenta anos, um século ou dois séculos” (1954 [1906], 159). Duhem segue: “Quantas discussões científicas existem em que cada um dos debatedores afirma ter esmagado seu adversário sob o irresistível testemunho dos fatos!... Quantas proposições foram consideradas como erros

monstruosos na escrita daqueles que nos precederam! Talvez devêssemos celebrá-las como grandes verdades se nós quiséssemos realmente investigar as teorias que deram às proposições seus significados verdadeiros” (Duhem 1954 [1906], 160–161). Esses trechos afirmam basicamente as mesmas coisas que tanto Feyerabend e Kuhn pretenderam com suas afirmações sobre a incomensurabilidade de teorias científicas: porque ideias antigas são mal compreendidas, como resultado de serem retiradas de seu contexto teórico, proponentes de teorias científicas incomensuráveis entendem mal uns aos outros, ambos afirmando que os fatos estão do seu lado. Kuhn e Feyerabend ambos afirmaram que, em tal situação, até mesmo argumentos empíricos podem tornar-se circulares (Feyerabend 1965b, 152; Kuhn 1970 [1962], 94).

### **3.2.2 A raiz quadrada de 2 e complementaridade: Niels Bohr**

Em sua autobiografia, Feyerabend reconheceu a influência direta de Niels Bohr no desenvolvimento da sua noção de incomensurabilidade nos anos 1950. Feyerabend recorda uma conversa na qual Bohr falava sobre a descoberta de que a raiz quadrada de dois não pode ser um número inteiro ou uma fração. De acordo com Feyerabend, Bohr apresentou o evento como tendo levado à extensão do conceito de número que reteve algumas propriedades dos inteiros e frações, mas que alterou outras; e afirmou que a transição da mecânica clássica para a quântica foi feita em acordo precisamente com esse princípio (1995, 78). Feyerabend também utilizou a noção de incomensurabilidade em uma publicação sobre a tese da complementaridade de Bohr, em que argumentou que esta é um exemplo de um conservadorismo conceitual injustificado, divergindo da afirmação de Bohr, de que toda a evidência para a mecânica quântica sempre será necessariamente expressa em termos clássicos (Feyerabend 1958b). Ele apresentou a defesa de Bohr do princípio de complementaridade como estando baseada na convicção de que toda experiência deve necessariamente aparecer nos moldes de nossos pontos de vista costumeiros, que atualmente é o da física clássica. No entanto, de acordo com Feyerabend, mesmo que conceitos clássicos

tenham sido bem sucedidos no passado, e mesmo que, no momento, possa ser difícil ou mesmo impossível imaginar como substituí-los, não se segue que o *framework* clássico não possa ser um dia substituído por um rival incomensurável. Consequentemente, não se segue que todas nossas teorias microscópicas futuras terão que tomar a noção de complementaridade como fundamental. Em vez disso, de acordo com Feyerabend, pode-se descobrir uma teoria cujo aparato conceitual, quando aplicado ao domínio de validade da física clássica, seria tão compreensivo e útil como o aparato clássico, sem que coincidam. Ele sustentou que uma situação assim não é de maneira alguma incomum, e utilizou a transição da física newtoniana para a relativística para fortalecer sua afirmação. De acordo com Feyerabend, enquanto os conceitos da teoria da relatividade são suficientemente ricos para enunciar todos os fatos apreendidos pela física newtoniana, os dois conjuntos de conceitos são “completamente diferentes” e não têm “nenhuma relação lógica” entre si (1958b, 83; 1961, 388; 1962, 88-89). Isso é claramente uma versão prévia da sua tese da incomensurabilidade. Em relação à concepção falibilista de Feyerabend sobre o conhecimento empírico, nenhum elemento de nosso conhecimento pode ser considerado necessário ou absolutamente certo. Em nossa busca por explicações satisfatórias, temos a liberdade de mudar quaisquer partes de nosso conhecimento existente, por mais fundamentais que possam ser, incluindo os conceitos da física clássica.

### **3.2.3 ‘Kant sobre rodas’ e teorias universais: Albert Einstein**

Albert Einstein utilizou o termo ‘incomensurável’ para ser aplicado especificamente a dificuldades na seleção e na avaliação de teorias científicas antes de Kuhn e Feyerabend, e há fortes razões para acreditar que o desenvolvimento de Feyerabend da incomensurabilidade foi diretamente inspirado pelo uso do termo por Einstein. Em ‘Notas autobiográficas’ [*Autobiographical notes*] (1949), Einstein tentou explicar que avaliar os méritos relativos das teorias físicas universais envolve fazer juízos difíceis sobre sua ‘naturalidade’ que exige fazer juízos baseados na ponderação recíproca de qualidades incomensuráveis: “O segundo ponto

de vista não está preocupado com a relação com a observação material, mas com as premissas da própria teoria, com o que pode ser breve e vagamente caracterizado como a ‘naturalidade’ ou a ‘simplicidade lógica’ das premissas (dos conceitos básicos e das relações entre aqueles que são tomados como base). Esse ponto de vista, cuja formulação exata encontra grandes dificuldades, tem desempenhado um papel importante na seleção e na avaliação de teorias desde tempos imemoriais. O problema aqui não é simplesmente sobre tipos de enumeração de premissas logicamente independentes (se algo desse tipo fosse inequivocamente possível), mas de um tipo de ponderação recíproca de qualidades incomensuráveis” (1949a, 23).

Há fortes razões para acreditar que esses comentários admitidamente enigmáticos inspiraram diretamente o uso e o desenvolvimento de Feyerabend da ideia de incomensurabilidade. Embora Feyerabend não cite Einstein (1949) em seu (1962), ele de fato cita o artigo de Bohr naquele mesmo volume editado (Schilpp 1949). Além disso, há algumas semelhanças notáveis entre os usos do termo de Feyerabend e Einstein. Primeiro, ambos fazem uma distinção entre teorias universais e teorias que não se aplicam à totalidade das aparências físicas, e depois usam essa distinção para limitar a aplicação da incomensurabilidade da mesma maneira; isto é, o problema da ponderação de qualidades incomensuráveis na seleção e na avaliação de teorias deve explicitamente aplicar-se a tais teorias físicas *universais*, e não a quaisquer pares de teorias científicas (cf. Feyerabend 1962a, 28 e Einstein 1949a, 23). Segundo, a atitude teórica de Einstein é explicitamente uma forma de neo-kantismo muito parecida com a de Feyerabend e Kuhn. De acordo com Einstein, sua atitude teórica é “distinta da de Kant somente pelo fato de que nós não concebemos as ‘categorias’ como inalteráveis (condicionadas pela natureza do entendimento), mas como convenções livres (no sentido lógico). Elas parecem ser *a priori* na medida em que pensar sem postular categorias e conceitos em geral seria impossível como respirar no vácuo” (Einstein 1949b, 374). Essa é a mesma perspectiva básica que tanto Kuhn quanto Feyerabend delinearam quando desenvolveram suas concepções de incomensurabilidade. Por exemplo, Kuhn diz, “Eu costumo explicar minha própria posição dizendo que sou um kantiano com categorias móveis” (Kuhn 2005 [1993], 264), uma ideia desenvolvida em detalhe por

Hoyningen-Huene (1993). Por sua vez, Feyerabend articulou uma abordagem ‘Kant sobre rodas’ – para usar a expressão adequada de Peter Lipton (Lipton 2001) – no começo de ‘Explicação, redução e empirismo’ (1962a). Feyerabend adotou manifestamente uma perspectiva metafísica neo-kantiana com categorias móveis a fim de investigar a questão: se teorias universais determinam todas as nossas experiências do mundo, como a experiência pode ser usada para testar tais teorias? Terceiro, o exemplo principal que Feyerabend utilizou para ilustrar isso e como teorias incomensuráveis podem de fato ser comparadas com base em um ‘experimento crucial’ foi a predição quantitativa de Einstein, e a confirmação subsequente de Perrin do movimento browniano, que serviu como um experimento crucial entre as termodinâmicas clássica e estatística – um dos três únicos exemplos de teorias científicas incomensuráveis de Feyerabend (1962a) (cf. Oberheim 2006, 253 ss.; Sirtes and Oberheim 2006). Feyerabend argumentou que, embora o movimento browniano já fosse um fenômeno bem conhecido, ele tornou-se *evidência* em favor da termodinâmica estatística e contra a termodinâmica clássica *somente depois* do desenvolvimento da teoria cinética, que era necessária para interpretá-la como tal. Mesmo que não houvesse um experimento crucial entre essas duas teorias universais, as teorias não têm nenhuma medida *comum*, pois as observações do movimento browniano que confirmaram a termodinâmica estatística não estavam formuladas em uma linguagem observacional *neutra*. As observações tornaram-se evidência somente depois de serem interpretadas de acordo com a teoria estatística. Einstein discutiu exatamente esse episódio longamente em suas ‘Notas autobiográficas’, em que utilizou o termo ‘incomensurável’ para aplicar-se às dificuldades no julgamento dos méritos das teorias físicas universais. Quarto, posteriormente, Feyerabend frequentemente defendeu sua concepção de que não há nenhum método científico *fixo* (por exemplo, 1975, 10–11) citando Einstein: “As condições externas que estão estabelecidas [para o cientista] pelos fatos da experiência não permitem a ele ser muito restritivo na construção de seu mundo conceitual pela aderência a um sistema epistemológico. Ele, portanto, deve aparecer para o epistemólogo sistemático como um tipo de oportunista inescrupuloso” (Einstein 1949b, 683 ss.). Embora o desenvolvimento da ideia de incomensurabilidade de teorias científicas por

Feyerabend tenha sido recebido pelas comunidades científica e filosófica como propagando ideias radicais e irracionais da ciência, ele, na verdade, estava tentando desenvolver algo que encontrou em Einstein. No prefácio à versão alemã de *Contra o Método* [*Against Method*], Feyerabend escreveu, “Eu quero enfatizar ainda, novamente, que as concepções em meu livro não são novas – para físicos como Mach, Boltzmann, Einstein e Bohr elas eram uma trivialidade. Mas as ideias desses grandes pensadores foram distorcidas [pelos filósofos positivistas] ao ponto de não serem mais reconhecidas” (1983, 12).

### **3.2.4 O mental e o material**

Feyerabend, como Kuhn, serviu-se explicitamente do trabalho de Wolfgang Köhler, na psicologia Gestalt, ideias para o desenvolvimento da sua noção de incomensurabilidade. Além disso, Köhler já havia aplicado a noção de incomensurabilidade a teorias científicas (ver seção 2.2.2). O uso de Köhler aplica-se a diferentes áreas dentro da ciência natural, e não a teorias científicas sucessivas como no Kuhn inicial (1962). O uso do termo ‘incomensurável’ por Köhler para descrever a relação entre conceitos da psicologia e da física foi uma das principais fontes do desenvolvimento de Feyerabend da ideia de incomensurabilidade (Oberheim 2005). A discussão de Feyerabend na formulação do problema mente-corpo deriva diretamente das afirmações de Köhler (Oberheim 2005). Contudo, a estratégia (wittgensteiniana) de Feyerabend para superar a incomensurabilidade na dissolução dos enigmas filosóficos não é inspirada em Köhler (Oberheim 2006, 43-45). Feyerabend sugeriu que a insolubilidade deve-se à má vontade de filósofos de diferentes convicções em alterar os significados profundamente enraizados em seus termos, e que nesse caso, o mental deveria ser reinterpretado a fim de ser compatível com o materialismo (Feyerabend 1963b).

### **3.3 A noção posterior de incomensurabilidade em Feyerabend**

Enquanto a noção de incomensurabilidade de Feyerabend não se alterou significativamente ao longo de seus primeiros escritos na filosofia da ciência, na sua filosofia mais geral posterior, Feyerabend estendeu sua aplicação do termo para além de teorias científicas universais. Por exemplo, ele a utilizou para caracterizar a transição entre a visão de mundo grega arcaica, agregadora de Homero para a visão de mundo substantivista dos Pré-socráticos (1975, 261-269), e subsequentemente, em um sentido muito amplo, ao discutir *frameworks* incomensuráveis de pensamento e ação (1975, 271), incomensurabilidade no domínio das percepções (1975, 225, 271), descobertas e atitudes incomensuráveis (1975, 269) e paradigmas incomensuráveis (1981b [1970], 131-161), em que tais incomensurabilidades envolvem a suspensão de alguns dos princípios universais da concepção tradicional herdada. Feyerabend caracterizou essa noção mais ampla como uma tese histórica e antropológica (1975, 271), mas também a aplicou a diferentes estágios do desenvolvimento do pensamento e da percepção do indivíduo (1975, 274). Ele também desenvolveu adicionalmente sua noção de incomensurabilidade, afirmando que ela envolve mudanças conceituais importantes tanto de classificações ‘manifestas’ [*overt*] e ‘tácitas’ [*covert*] (no sentido de Whorff), de maneira que a incomensurabilidade é difícil de ser definida explicitamente, e pode apenas ser mostrada (1975, 224-225). Ainda mais tarde, Feyerabend argumentou que ‘potencialmente, toda cultura é todas as culturas’, aparentemente retratando-se de uma noção mais ampla, cultural, de incomensurabilidade (Feyerabend 1994).

### **4. Uma comparação entre as noções de incomensurabilidade de Kuhn e Feyerabend**

Inicialmente, a caracterização de Feyerabend sobre a natureza e sobre as origens da incomensurabilidade era mais concreta que a de Kuhn. Na concepção de Feyerabend, porque a natureza dos objetos depende das teorias mais avançadas sobre eles, e porque o significado

dos enunciados de observação depende da natureza desses objetos, a interpretação de uma linguagem observacional é determinada pelas teorias que usamos para explicar o que observamos. Kuhn, por sua vez, estava inicialmente muito menos seguro sobre o significado exato da sua noção de incomensurabilidade, especialmente em relação à mudança de mundo, que ele via como um aspecto fundamental da incomensurabilidade. Kuhn confessou francamente ter estado perdido: “Em um sentido que sou incapaz de explicar melhor, os proponentes dos paradigmas competidores praticam seus ofícios em mundos diferentes” (1970 [1962], 150). Ele sugeriu que “devemos aprender a compreender o sentido de enunciados semelhantes a esses” (1970 [1962], 121), e então empreendeu grande esforço tentando fazê-lo.

Além disso, o conceito de Feyerabend de incomensurabilidade de teorias científicas era muito mais restrito que o de Kuhn. Para Kuhn, a incomensurabilidade tinha três domínios heterogêneos *prima facie*, ligados holisticamente: mudança nos problemas e nos padrões, mudança nos conceitos usados para expressá-los e resolvê-los, e mudança na visão de mundo em que eles surgem. O foco de Feyerabend, por outro lado, estava inicialmente exclusivamente nos conceitos que ocorrem em teorias universais ou fundamentais, junto com suas implicações ontológicas. Ironicamente, contudo, nos desenvolvimentos pós-1962, ambos os autores moveram-se em direções opostas. Kuhn eliminou gradualmente de sua noção de incomensurabilidade tudo que não envolvesse conceitos científicos, e terminou mais ou menos onde Feyerabend começou (ver Carrier 2001; Chen 1997; Hoyningen-Huene 1990, 487-488; Hoyningen-Huene 1993, 212-218; Hoyningen-Huene 2004, Sankey 1993; Sankey 1994, 16-30; Sankey 1997). Feyerabend, por sua vez, enfatizou cada vez mais os aspectos de alteração perceptual (1975, 225–229, 273–274; 1978, 68; 1988, 172–176), e também as alterações nos conjuntos de problemas legítimos que uma disciplina deve tratar (1975, 274–275), e sua filosofia tardia enfatiza um dos pontos originais de Kuhn: o papel de valores epistêmicos não vinculantes na escolha de teorias (cf. Farrell 2003).

Com respeito à gama de teorias que está sujeita à incomensurabilidade, o conceito de Feyerabend é, novamente, mais restrito que o de Kuhn. Para Feyerabend, somente teorias

*universais, fundamentais, e não-instanciais* que são *interpretadas realisticamente* podem ser incomensuráveis (Feyerabend 1962, 44; 1965b, 216; 1975, 114, 271, 284; 1975, 221–222; 1987, 272). Feyerabend estava interessado em teorias universais e fundamentais que se aplicam a tudo, porque acreditava que somente “estruturas de pensamento tão compreensivas” têm implicações ontológicas capazes de sustentar inteiramente uma visão de mundo (Feyerabend 1962, 28; cf. Feyerabend 1954 e Oberheim 2006, 157 ss.). Ele restringiu a incomensurabilidade a teorias não-instanciais, que ele distinguiu de generalizações empíricas, com base nas diferenças em seus procedimentos de teste. Generalizações da forma ‘Todos os A são B’ (independentemente de A ou B serem observáveis) são testadas por meio da inspeção de instâncias. Assim, por exemplo, a primeira lei de Kepler faz afirmações sobre planetas, que podem ser testadas diretamente por inspeção de seu movimento. A fim de testar teorias não-instanciais, tal como a teoria da gravitação de Newton, as primeiras generalizações empíricas devem ser derivadas delas com a ajuda de hipóteses auxiliares, e somente depois podem ser testadas por inspeção (Feyerabend 1962, 28). Para Feyerabend, devido a essas restrições (somente teorias universais, fundamentais e não-instanciais), a incomensurabilidade de teorias científicas é relativamente rara (1987, 272). Kuhn incluiu uma gama mais ampla de teorias candidatas à incomensurabilidade. Para ele, mesmo episódios menores, como descobertas inesperadas, poderiam ser incomensuráveis com a tradição anterior (cf. Hoyningen-Huene 1993, 197-201). Essa diferença no escopo da incomensurabilidade entre as versões de Kuhn e Feyerabend encontra sua expressão mais notável na maneira em que consideram a transição entre as teorias ptolemaica e copernicana. Para Kuhn, as diferenças entre essas duas teorias compreendem uma ilustração exemplar da incomensurabilidade. Para Feyerabend, no entanto, porque a teoria planetária carece da qualidade de universalidade, não há incomensurabilidade (1975, 114). Além disso, em seus escritos posteriores, Kuhn insistiu que a versão de incomensurabilidade que ele defendeu sempre foi a “incomensurabilidade local”, uma noção que restringe a mudança conceitual a poucos conceitos tipicamente interconectados (cf. Hoyningen-Huene 1993, 213, 219). Assim, pode haver consequências empíricas dos pares de teorias incomensuráveis que podem ser

imediatamente comparadas. Por exemplo, as teorias planetárias geocêntrica e heliocêntrica são incomensuráveis no sentido de Kuhn, enquanto que as predições das posições planetárias de ambas as teorias são inteiramente comensuráveis, e podem ser comparadas imediatamente quanto à sua precisão empírica. Por contraste, Feyerabend sempre pensou seu conceito de maneira mais global, como afetando *todos* os enunciados deriváveis de duas teorias fundamentais (1962, 93; 1965c, 117; 1965b, 216; 1975, 275–276; 1981a, xi).

Tanto Kuhn quanto Feyerabend têm sido frequentemente mal interpretados, como tendo defendido a concepção de que a incomensurabilidade implica incomparabilidade (cf. Hoyningen-Huene 1993, 218 ss.; Oberheim 2006, 235). Em resposta a isso, Kuhn enfatizou repetidamente que a incomensurabilidade não implica incomparabilidade (cf. Hoyningen-Huene 1993, 236 ss.). A comparação de teorias é só mais complicada do que imaginam alguns filósofos da ciência. Em particular, não pode ser completamente comparada ‘ponto a ponto’. Não é um procedimento algorítmico (cf. Hoyningen-Huene 1993, 147-154; Feyerabend 1975, 114; 1981a, 238), e nem exige a tradução a uma linguagem observacional neutra. Diferentes valores epistêmicos, como a universalidade, a precisão, a simplicidade e a fecundidade, podem impulsionar a direções diferentes (cf. Hoyningen-Huene 1992, 492–496; 1993, 150–154; Feyerabend 1981a 16, 1981c, 238), permitindo a possibilidade de desacordo racional. Mas, mesmo se uma comparação ponto a ponto completa fosse impossível, e mesmo que a comparação de teorias não tenha o *status* de prova, uma avaliação comparativa de teorias incomensuráveis ainda é possível (cf. Hoyningen-Huene 1993, 236-258; Carrier 2001) e racional em um sentido de meios/fins ou instrumental. Por exemplo, de acordo com Kuhn, é racional escolher teorias que são melhores solucionadoras de problemas, porque elas servem melhor às finalidades da ciência. Essa propriedade da escolha de teorias torna o processo da ciência como um todo racional e progressivo. Com a noção de incomensurabilidade, Kuhn não estava tanto desafiando a racionalidade da escolha de teorias, quanto tentando abrir espaço para a possibilidade de desacordo racional entre proponentes de paradigmas competidores. De fato, de acordo com Kuhn, “a incomensurabilidade está longe de ser aquela ameaça à avaliação racional de asserções de verdade que com frequência tem parecido ser. Ao

contrário, ela é o que é preciso, de uma perspectiva evolucionária, para devolver à noção de avaliação cognitiva um pouco do impacto de que desesperadamente necessita. Ou seja, ela é necessária para defender noções como verdade e conhecimento, por exemplo, dos excessos de movimentos pós-modernistas, como o programa forte” (2000 [1993], 91).

A extensão da má-interpretação da incomensurabilidade como implicando incomparabilidade é ainda mais dramática no caso de Feyerabend. Longe de afirmar que teorias incomensuráveis não podem ser comparadas, Feyerabend argumentou repetida e explicitamente que as alternativas incomensuráveis de fato oferecem meios *melhores* de comparar os méritos de teorias que o mero desenvolvimento de alternativas comensuráveis (Feyerabend 1962, 66; cf. Oberheim 2006, 235 ss.). Ele ilustrou isso com um exemplo de um *experimento crucial* entre teorias incomensuráveis. Argumentou que mesmo que os conceitos centrais da termodinâmica estatística e da termodinâmica fenomenológica clássica sejam incomensuráveis, ainda é possível indiretamente montar um experimento crucial entre eles (a confirmação de Perrin das previsões quantitativas de Einstein do movimento browniano). Essa é a peça central do argumento de Feyerabend a favor do pluralismo. O argumento depende da afirmação de que essa refutação da termodinâmica fenomenológica clássica não poderia ter sido conduzida sem o desenvolvimento da termodinâmica estatística, que é uma alternativa incomensurável. Feyerabend concluiu que algumas observações podem *somente* ser interpretadas como refutações de uma teoria existente depois que uma alternativa incomensurável com a qual interpretá-la fosse desenvolvida (cf. Oberheim 2006, 240–245). Sua imagem não-ortodoxa do progresso científico é delineada como um aumento no conjunto de alternativas incomensuráveis, cada uma forçando as outras a maiores articulações (Feyerabend 1965c, 107; 1975, 30; cf. Oberheim 2005; Oberheim 2006, 260 ss.).

A origem da má-compreensão a respeito da incomensurabilidade parece ser a sugestão que Feyerabend uma vez fez de que “teorias incomensuráveis podem não possuir consequências comparáveis” (1962, 94). No entanto, nessa passagem, Feyerabend estava argumentando que, embora a mesma configuração experimental possa produzir um resultado numérico que pode ser usado para *confirmar* duas teorias incomensuráveis (por exemplo,

medir o tempo de uma pedra em queda livre para testar as predições tanto da mecânica newtoniana quanto da teoria da relatividade), a fim de utilizar o resultado para confirmar uma teoria, isso tem de ser interpretado por aquela teoria. Dado que duas teorias incomensuráveis utilizarão conceitos *qualitativamente* incompatíveis para interpretar resultados *quantitativamente* idênticos, elas interpretarão os mesmo enunciados quantitativos como enunciados qualitativos diferentes. Depois disso, Feyerabend concluiu diretamente: “Assim, pode não existir nenhuma possibilidade de encontrar uma *caracterização* das observações que deveriam confirmar duas teorias incomensuráveis” (1962, 94, *italico inserido*). Isso realmente exclui a possibilidade de usar uma linguagem observacional *neutra* para comparar as consequências empíricas de duas teorias incomensuráveis. No entanto, Feyerabend não concluiu que isso inibe sua comparação, mas, na verdade, que não há necessidade de uma linguagem observacional neutra para compará-las. Feyerabend também menciona outras possibilidades de comparação entre teorias incomensuráveis (Feyerabend 1965b, 217; 1970, 228; 1975, 284; 1978, 68; 1981a, 16). Algumas dessas envolvem fazer julgamentos subjetivos sobre como pesar valores diferentes que podem impulsionar a direções diferentes (incomensurabilidade metodológica).

Por fim, há um ponto substantivo central de concordância entre Kuhn e Feyerabend. Para ambos, a incomensurabilidade obstrui a possibilidade de interpretar o desenvolvimento científico como uma aproximação à verdade (ou como um “aumento de verossimilhança”) (Feyerabend, 1965c, 107; 1970, 220, 222, 227–228; 1975, 30, 284; 1978, 68; Kuhn 1970, 206; 1991, 6; 1993, 330; cf. Oberheim 2006, 180 ss.; Hoyningen-Huene 1993, 262-264). Eles rejeitam tais caracterizações de progresso científico porque eles reconhecem e enfatizam que as revoluções científicas resultam em mudanças na ontologia. Tais mudanças não são apenas refinamentos, ou adições, à ontologia anterior, de maneira que pudessem ser vistas como adições cumulativas a concepções teóricas já estabelecidas. Na verdade, a nova ontologia *substitui* sua predecessora. Consequentemente, nem Kuhn nem Feyerabend podem ser corretamente caracterizados como realistas científicos que acreditam que a ciência progride em direção à verdade.

## 5. Bibliografia

- Barker, P., Chen, X. and Andersen, A., 2003, “Kuhn on Concepts and Categorization”, in T. Nickles (ed.), *Thomas Kuhn*, Cambridge: Cambridge University Press, pp. 212–245.
- Carrier, M., 2001, “Changing Laws and Shifting Concepts: On the Nature and Impact of Incommensurability”, in P. Hoyningen-Huene and H. Sankey (ed.), *Incommensurability and Related Matters*, Dordrecht: Kluwer, pp. 65–90.
- Carrier, M., 2008, “The Aim and Structure of Methodological Theory”, in L. Soler, H. Sankey and P. Hoyningen-Huene (ed.), *Rethinking Scientific Change and Theory Comparison: Stabilities, Ruptures, Incommensurabilities?*, Berlin: Springer, pp. 273–290.
- Chen, X., 1997, “Thomas Kuhn's Latest Notion of Incommensurability”, *Journal for General Philosophy of Science*, 28: 257–273.
- Conant, J. and Haugeland, J., 2000, “Editors' Introduction”, in J. Conant and J. Haugeland (ed.), *The Road Since Structure*, Chicago: University of Chicago Press, pp. 1–9.
- Devitt, M., 2001, “Incommensurability and the Priority of Metaphysics”, in P. Hoyningen-Huene and H. Sankey (ed.), *Incommensurability and Related Matters*, Dordrecht: Kluwer, pp. 143–157.
- Duhem, P., 1906 [1954], *The Aim and Structure of Physical Theory*, Princeton: Princeton University Press.
- Einstein, A., 1949a, “Autobiographical Notes”, in P. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*, La Salle: Open Court, pp. 3–95.
- Einstein, A., 1949b, “Remarks Concerning the Essays Brought Together in this Co-operative Volume”, in P. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist*, La Salle: Open Court, pp. 665–668.
- Farrell, R., 2003, *Feyerabend and Scientific Values. Tightrope-Walking Rationality*, Dordrecht: Kluwer.
- Feyerabend, P., 1951, *Zur Theorie der Basissätze*, Ph.D. thesis, University of Vienna, Universitäts Bibliothek Wien.

- Feyerabend, P., 1954, "Physik und Ontologie", *Wissenschaft und Weltbild: Monatsschrift für alle Gebiete der Forschung*, 7: 464–476.
- Feyerabend, P., 1958a, "An Attempt at a Realistic Interpretation of Experience", *Proceedings of the Aristotelian Society*, 58: 143–170.
- Feyerabend, P., 1958b, "Complementarity", *Proceedings of the Aristotelian Society*, 32 (Supplement): 75–104.
- Feyerabend, P., 1961, "Niels Bohr's Interpretation of the Quantum Theory", in H. Feigl and G. Maxwell (ed.), *Current Issues in the Philosophy of Science*, New York: Rinehart and Winston, pp. 35–39.
- Feyerabend, P., 1962, "Explanation, Reduction and Empiricism", in H. Feigl and G. Maxwell (ed.), *Scientific Explanation, Space, and Time*, (Minnesota Studies in the Philosophy of Science, Volume III), Minneapolis: University of Minneapolis Press, pp. 28–97.
- Feyerabend, P., 1963a, "How to Be a Good Empiricist: A Plea for Tolerance in Matters Epistemological", in B. Baumrin (ed.), *Philosophy of Science: The Delaware Seminar*, New York: Interscience Press (John Wiley).
- Feyerabend, P., 1963b, "Materialism and the Mind-Body Problem", *The Review of Metaphysics*, 17: 49–66.
- Feyerabend, P., 1965a, "On the 'Meaning' of Scientific Terms", *Journal of Philosophy*, 62: 266–274.
- Feyerabend, P., 1965b, "Problems of Empiricism", in R. Colodny (ed.), *Beyond the Edge of Certainty. Essays in Contemporary Science and Philosophy*, Pittsburgh: Pittsburgh Center for Philosophy of Science, pp. 145–260.
- Feyerabend, P., 1965c, "Reply to Criticism. Comments on Smart, Sellars and Putnam", in R. Cohen and M. Wartofsky (ed.), *Proceedings of the Boston Colloquium for the Philosophy of Science 1962–64: In Honor of Philipp Frank* (Boston Studies in the Philosophy of Science, Volume II), New York: Humanities Press, pp. 223–261.
- Feyerabend, P., 1967, "The Mind-Body Problem", *Continuum*, 5: 35–49.

- Feyerabend, P., 1969a, "Materialism and the Mind-Body Problem" in J. O'Connor (ed.), *Modern Materialism: Readings on Mind-Body Identity*, New York: Harcourt, Brace and World, Inc., pp. 82–98.
- Feyerabend, P., 1969b, "Linguistic Arguments and Scientific Method", *Telos*, 2: 43–63.
- Feyerabend, P., 1970a, "Against Method: Outline of an Anarchistic Theory of Knowledge", in M. Radner and S. Winokur (ed.), *Analysis of Theories and Methods of Physics and Psychology*, (Minnesota Studies in the Philosophy of Science, Volume IV), Minneapolis: University of Minneapolis Press, pp. 17–130.
- Feyerabend, P., 1970b, "Consolations for the Specialist", in I. Lakatos and A. Musgrave (ed.), *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press, pp. 197–230.
- Feyerabend, P., 1975, *Against Method. Outline of an Anarchistic Theory of Knowledge*, London: New Left Books.
- Feyerabend, P., 1977, "Changing Patterns of Reconstruction", *British Journal for the Philosophy of Science*, 28: 351–382.
- Feyerabend, P., 1978, *Science in a Free Society*, London: New Left Books.
- Feyerabend, P., 1981a, *Realism, Rationalism and Scientific Method. Philosophical papers*, Cambridge: Cambridge University Press.
- Feyerabend, P., 1981b, "Consolations for the specialist", in *Problems of Empiricism. Philosophical Papers*, Volume 2, Cambridge: Cambridge University Press, pp. 131–167.
- Feyerabend, P., 1981c, "More clothes from the emperor's bargain basement: A review of Laudan's *Progress and its Problems*", *British Journal for the Philosophy of Science*, 32: 57–71.
- Feyerabend, P., 1983, *Wider den Methodenzwang*, 2nd Edition, Frankfurt am Main: Suhrkamp.
- Feyerabend, P., 1987, "Putnam on Incommensurability", *British Journal for the Philosophy of Science*, 38: 75–81.
- Feyerabend, P., 1988, *Against Method*, revised edition, London: Verso.
- Feyerabend, P., 1993, *Against Method*, 3rd edition, London: Verso.
- Feyerabend, P., 1994. "Potentially Every Culture is All Cultures" *Common Knowledge*, 3: 16–22.

- Feyerabend, P., 1995, *Killing Time: The Autobiography of Paul Feyerabend*, Chicago: University of Chicago Press.
- Fleck, L., 1935 [1979], *Genesis and Development of a Scientific Fact*, T. Trenn and R. Merton (ed.), Chicago: University of Chicago Press.
- Fleck, L., 1927 [1986], "Some Specific Features of The Medical Way of Thinking", in R. Cohen and T. Schnelle (ed.), *Cognition and Fact: Materials on Ludwik Fleck*, Dordrecht: D. Reidel, pp. 39–46.
- Fleck, L., 1935 [1986], "Scientific Observation and Perception in General", in R. Cohen and T. Schnelle (ed.), *Cognition and Fact: Materials on Ludwik Fleck*, Dordrecht: D. Reidel, pp. 59–78.
- Fleck, L., 1936 [1986], "The Problem of Epistemology", in R. Cohen and T. Schnelle (ed.), *Cognition and Fact: Materials on Ludwik Fleck*, Dordrecht: D. Reidel, pp. 79–112.
- Fleck, L., 1947 [1986], "To Look, To See, To Know", in R. Cohen and T. Schnelle (ed.), *Cognition and Fact: Materials on Ludwik Fleck*, Dordrecht: D. Reidel, pp. 129–151.
- Harwood, J., 1986, "Fleck and the Sociology of Knowledge", *Social Studies of Science*, 16: 173–187.
- Hedfors, E., 2007a, "Medical ethics in the wake of the Holocaust: departing from a postwar paper by Ludwik Fleck", *Studies in the History and Philosophy of the Biomedical Sciences*, 38: 642–655.
- Hedfors, E., 2007b, "Fleck in Context", *Perspectives on Science*, 15: 49–86.
- Hoyningen-Huene, P., 1993, *Reconstructing Scientific Revolutions. The Philosophy of Science of Thomas S. Kuhn*, Chicago: University of Chicago Press.
- Hoyningen-Huene, P., 1990, "Kuhn's Conception of Incommensurability", *Studies in History and Philosophy of Science*, 21: 481–492.
- Hoyningen-Huene, P., 2004, "Three Biographies: Kuhn, Feyerabend and Incommensurability", in R. Harris (ed.) *Rhetoric and Incommensurability*, West Lafayette, IN: Parlor Press, pp. 150–175.
- Hoyningen-Huene, P., 2008, "Thomas Kuhn and the Chemical Revolution", *Foundations of Chemistry*, 10: 101–115.

- Horwich, P. (ed.) 1993, *World Changes: Thomas Kuhn and the Nature of Science*, Cambridge, MA: MIT Press.
- Köhler, W., 1938 [1920], "Physical Gestalten", in W. Ellis (ed.), *A Source Book of Gestalt Psychology*, London: Kegan Paul, Trench, Trubner, and Co., pp. 17–54.
- Kuhn, T., 1957, *The Copernican Revolution*, John Hopkins University Press.
- Kuhn, T., 1962 [1970], *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press.
- Kuhn, T., 1970, "Postscript—1969", in *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press: 174–210.
- Kuhn, T., 1977, *The Essential Tension*, Chicago: University of Chicago Press.
- Kuhn, T., 2005, *The Road Since Structure*, J. Conant and J. Haugeland (ed.), Chicago: University of Chicago Press.
- Lipton, P., 2001, "Kant on Wheels", *London Review of Books*, 23: 30–31.
- Oberheim, E., 2005, "On the Historical Origins of the Contemporary Notion of Incommensurability: Paul Feyerabend's Assault on Conceptual Conservatism", *Studies in the History and Philosophy of Science*, 36: 363–390.
- Oberheim, E., 2006, *Feyerabend's Philosophy*, Berlin: De Gruyter.
- Oberheim, E. and Hoyningen-Huene, P., 1997, "Incommensurability, Realism and Meta-Incommensurability", *Theoria*, 12: 447–465.
- Preston, J., Munévar, G. and Lamb, D. (ed.), 2000, *The Worst Enemy of Science? Essays in Memory of Paul Feyerabend*, Oxford: Oxford University Press.
- Sankey, H., 1993, "Kuhn's Changing Concept of Incommensurability", *British Journal for the Philosophy of Science*, 44: 759–774.
- Sankey, H., 1994, *The Incommensurability Thesis*, London: Ashgate.
- Sankey, H., 1997, "Taxonomic Incommensurability", in H. Sankey (ed.), *Rationality, Relativism and Incommensurability*, London: Ashgate, pp. 66–80.
- Sankey, H., forthcoming, "Semantic Incommensurability and Scientific Realism", *Studies in the History and Philosophy of Science*.

Sankey, H. and Hoyningen-Huene, P., 2001, "Introduction", in P. Hoyningen-Huene and H. Sankey (ed.), *Incommensurability and Related Matters*, Dordrecht: Kluwer: vii-xxxiv.

Schilpp, P. (ed.), 1949, *Albert Einstein: Philosopher –Scientist*. La Salle: Open Court.

Sirtes, D. and Oberheim, E., 2006, "Einstein, Entropy and Anomaly", in J. Alimi and A. Füzfa (ed.), *Albert Einstein Century International Conference* (AIP Conference Proceedings, Astronomy and Astrophysics, Volume 861), Berlin: Springer, pp. 1147–1154.

Theocharis, T., and Psimopoulos, M., 1987, "Where science has gone wrong", *Nature*, 329: 595–598.

## 6. Outros recursos na Internet

<http://arts.uwaterloo.ca/~incommen/>

<http://www.britannica.com/EBchecked/topic/284869/incommensurable>

## 7. Verbetes relacionados

Ariew, Roger, "Pierre Duhem", *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/duhem/>

Preston, John, "Paul Feyerabend", *The Stanford Encyclopedia of Philosophy* (Winter 2009 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2009/entries/feyerabend/>

Bird, Alexander, "Thomas Kuhn", *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/thomas-kuhn/>

Schmidt, Heinz-Juergen, "Structuralism in Physics", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/physics-structuralism/>

Cat, Jordi, "The Unity of Science", *The Stanford Encyclopedia of Philosophy (Fall 2010 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2010/entries/scientific-unity/>

Niiniluoto, Ilkka, "Scientific Progress", *The Stanford Encyclopedia of Philosophy (Summer 2011 Edition)*, Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/sum2011/entries/scientific-progress/>>

Nickles, Thomas, "Scientific Revolutions", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL=<<http://plato.stanford.edu/archives/spr2011/entries/scientific-revolutions/>>.

# Consciência\*

Autor: Robert Van Gulick

Tradutores: Marco Aurélio Alves & 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 trata 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 aspectos 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 *telefuncionais*, 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 H<sub>2</sub>O 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 assera 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 H<sub>2</sub>O. 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-reducionista, 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-reducionista* 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-reducionista tem sido criticado pela suposta incapacidade de “pagar sua dívida com o fisicalismo” em moeda reducionista. 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-reducionistas 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)<sup>1</sup> 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,

---

<sup>1</sup> 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).<sup>2</sup>

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.

---

<sup>2</sup> 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 representacionistas**

Praticamente todas as teorias da consciência a veem como possuindo características representacionais, mas as assim chamadas teorias representacionistas se definem pela afirmação mais forte de que suas características representacionais exaurem suas características mentais (Harman 1990, Tye 1995, 2000). De acordo com o representacionista, 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 representacionista 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 representacionista 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 sensitivo a meras diferenças em características subjacentes sem conteúdo dos processos no nível de realização. Portanto, a maior parte dos representacionalistas 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 representacionalistas 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 representacionalismo 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 representacionalismo 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 representacionalismo 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 antirrepresentacionalistas, a mera possibilidade lógica de qualia invertidos mostra que estados conscientes podem diferir em um aspecto mental significativo e, ao mesmo tempo, coincidir representacionalmente. Representacionalistas, 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).<sup>3</sup> 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.

---

<sup>3</sup> 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 atividade 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.<sup>4</sup> 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

---

<sup>4</sup> 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 anestesiológico 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.

# Epistemologia\*

Autor: Matthias Steup

Tradutores: Eros Moreira Carvalho, Flavio Williges, Mateus Stein & Paola Oliveira de Camargo

Revisores: Alexandre Meyer Luz & Delvair Moreira

Em sentido estrito, a epistemologia é o estudo do conhecimento e da crença justificada. Enquanto estudo do conhecimento, a epistemologia interessa-se pelas seguintes questões: quais são as condições necessárias e suficientes do conhecimento? Quais são suas fontes? Qual é sua estrutura e quais são seus limites? Enquanto estudo da crença justificada, a epistemologia pretende responder questões como estas: como iremos entender o conceito de justificação? O que torna justificada uma crença? A justificação é interna ou externa à mente de alguém?

Já num sentido mais amplo, a epistemologia trata de questões relacionadas com a criação e disseminação do conhecimento em áreas particulares de investigação. Esse artigo proporcionará um panorama sistemático dos problemas que as questões colocadas acima geram e abordará, com alguma profundidade, questões relativas à estrutura e aos limites do conhecimento e da justificação.

## 1. O que é o conhecimento?

### 1.1 Conhecimento como crença verdadeira e justificada

Existem vários tipos de conhecimento: saber como fazer algo (*know how*) (por exemplo, como andar de bicicleta), conhecer alguém (*knowing someone*) pessoalmente e

---

\* Tradução do verbete "Epistemology" de Matthias Steup. In: *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.) URL = <<http://plato.stanford.edu/archives/win2011/entries/epistemology/>>. Publicada com a autorização do editor.

conhecer um lugar ou uma cidade (*knowing a place or a city*)<sup>1</sup>. Embora tais tipos de conhecimento sejam também de interesse epistêmico, focaremos o conhecimento de *proposições* e faremos referência a tal conhecimento usando o esquema ‘*S* sabe que *p*’, onde ‘*S*’ representa o sujeito que tem o conhecimento e ‘*p*’ a proposição que é conhecida<sup>2</sup>. Nossa questão será: quais são as condições necessárias e suficientes para *S* saber que *p*? Podemos distinguir amplamente entre uma abordagem tradicional e uma abordagem não tradicional de resposta a essa questão. Faremos referência a elas como ‘TK’ (*Traditional Knowledge*) e ‘NTK’ (*Non-Traditional Knowledge*).

De acordo com TK, saber que *p* (*knowledge that p*) é, pelo menos num sentido aproximado, crença verdadeira e justificada (JTB). Não podemos saber proposições falsas. Portanto, o conhecimento requer a verdade. Uma proposição na qual *S* não crê não pode ser uma proposição que *S* sabe. Portanto, o conhecimento requer a crença. Finalmente, o fato de *S* estar certo ao crer que *p* poderia ser uma mera questão de sorte<sup>3</sup>. Portanto, o conhecimento requer um terceiro elemento, tradicionalmente identificado como justificação. Assim, chegamos à análise tripartite do conhecimento como JTB (*Justification True Belief*): *S* sabe que *p* se e somente se *p* é verdadeiro e *S* está justificado em crer que *p*. De acordo com essa análise, as três condições- verdade, crença e justificação- são individualmente necessárias e conjuntamente suficientes para o conhecimento<sup>4</sup>.

<sup>1</sup> Na língua portuguesa, a tradução do verbo “to know” (saber) admite variações que não são comuns em inglês. Em alguns contextos, “to know” é traduzido corretamente por “saber”; noutras, como nos exemplos de tipos de conhecimento indicados acima, a melhor tradução para o português é “conhecer”. (nota do tradutor)

<sup>2</sup> O que são proposições? Proposições devem ser distinguidas de sentenças. Por exemplo, há a proposição que gatos têm quatro pernas. Essa proposição deve ser distinguida da sentença portuguesa ‘Gatos têm quatro pernas’ que expressa a proposição que gatos têm quatro pernas ou tem essa proposição como seu conteúdo. Sentenças em diferentes línguas podem expressar a mesma proposição. Por exemplo, a sentença alemã “Katzen haben vier Beine”, também expressa a mesma proposição que gatos têm quatro pernas. As sentenças são entidades físicas, dado que elas são sons proferidos e, quando impressas, marcas no papel. Proposições, em contrapartida, são (supostamente são) objetos abstratos, não-físicos. O qualificador no parênteses foi acrescentado como uma indicação do fato que, de acordo com alguns, objetos abstratos e, portanto, proposições, não existem. Para mais informações sobre esse tópico, veja-se o artigo de Jeffrey C. King *Structured Propositions* na *Stanford Encyclopedia of Philosophy*.

<sup>3</sup> Por exemplo, se Hal crê que ele tem uma doença fatal, não por que lhe foi dito por seu doutor, mas unicamente porque, como um hipocondríaco, ele não pode deixar de crer nisso, e ocorre que, de fato, ele tem uma doença fatal, o fato de Hal estar certo sobre isso é meramente acidental: uma questão de sorte (má sorte, nesse caso). Para consultar um estudo que trata do fenômeno da sorte epistêmica e das ferramentas conceituais para capturar o sentido em que conhecimento e sorte epistêmica são incompatíveis, veja Pritchard, 2005. Veja também Engel, 1992.

<sup>4</sup> Para mais informações, veja o verbete ‘A análise do conhecimento’ na *Stanford Encyclopedia of Philosophy*, cujo link encontra-se no final deste artigo. Veja também Shope, 1983 e Steup, 1996, capítulos 1 e 2. Para uma abordagem

Inicialmente, podemos dizer que a função da justificação é assegurar que a crença de *S* não é verdadeira por mera sorte (ou acaso). Nisso, TK e NTK estão de acordo. Eles divergem, contudo, à medida que passamos a ser mais específicos sobre como, exatamente, a justificação preenche essa função. De acordo com TK, quando é razoável ou racional assumir, do ponto de vista de *S*, *p* como verdadeira, a crença de *S* que *p* é verdadeira não é uma mera questão de sorte. De acordo com o *evidencialismo*, o que torna uma crença justificada nesse sentido é a posse de evidência. A idéia básica é que a crença é justificada pelo grau em que se ajusta às evidências de *S*. NTK, por outro lado, concebe a função da justificação de modo diferente. O trabalho da justificação é assegurar que a crença de *S* tenha uma alta probabilidade objetiva de verdade e, portanto, se verdadeira, não será por mera sorte. Uma idéia central é que isso é efetivado se e somente se uma crença origina-se de um processo cognitivo confiável ou de faculdades confiáveis. Essa visão é conhecida como *confiabilismo*<sup>5</sup>.

## 1.2 O Problema de Gettier

A análise tripartite do conhecimento como JTB foi mostrada como sendo incompleta. Existem casos de JTB que não se qualificam como casos de conhecimento. JTB, portanto, não é *suficiente* para o conhecimento. Casos como esses- conhecidos como *casos do tipo Gettier (Gettier-cases*<sup>6</sup>)- surgem porque nem a posse de evidência, nem a origem em faculdades confiáveis são suficientes para assegurar que uma crença não é verdadeira meramente por causa da sorte. Considere o conhecido exemplo dos celeiros de fachada: Henry está dirigindo num campo no qual o que parecem ser celeiros são, com apenas uma única exceção, simples celeiros de fachada. Vistas da estrada em que Henry está dirigindo essas fachadas parecem ser celeiros perfeitamente reais. No entanto, Henry está olhando e único celeiro real dessa região

---

inteiramente diferente da análise do conhecimento, ver Williamson, 2000.

<sup>5</sup> Para mais leituras acerca do evidencialismo, ver Conee e Feldman, 1985 e 2004. Para conhecer a literatura que advoga o confiabilismo, ver Armstrong, 1973, Goldman, 1979, 1986, 1991, e Swain, 1981.

<sup>6</sup> Eles são referidos como casos de tipo Gettier pois, em seu artigo de 1963 “Is Justified True belief Knowledge”, Edmund Gettier descreveu dois casos que refutaram decisivamente a análise do conhecimento como crença verdadeira e justificada. Para uma excelente discussão do problema de Gettier, ver o apêndice em Pollock, 1986.

e crê que há um celeiro real lá. A crença de Henry é justificada, de acordo com TK, por que a experiência visual de Henry justifica sua crença. Conforme o NTK, a sua crença é justificada por que a crença de Henry se origina de um processo cognitivo confiável: a visão. No entanto, a crença de Henry pode ainda ser plausivelmente descrita como verdadeira por mera sorte. Se Henry tivesse visto um dos celeiros de fachada em vez dos celeiros reais, ele poderia também ter acreditado que existia um celeiro lá. Há, portanto, amplo acordo entre epistemólogos que a crença de Henry não constitui conhecimento<sup>7</sup>.

Para estabelecer condições que são, quando tomadas conjuntamente, suficientes para o conhecimento, qual elemento ulterior deve ser acrescentado à JTB? Esse problema é conhecido como o *problema de Gettier*. De acordo com TK, resolver o problema requer uma quarta condição. De acordo com alguns teóricos do NTK, ele exige um refinamento no conceito de confiabilidade. Por exemplo, se a confiabilidade pode ser adequadamente indexada ao ambiente do sujeito, os confiabilistas podem dizer que a crença de Henry não está justificada porque, em seu ambiente, a visão não é confiável para distinguir celeiros reais de celeiros de fachada.<sup>8</sup> Alguns teóricos do tipo NTK abandonam a condição de justificação inteiramente. Eles diriam que, se concebemos o conhecimento como crença verdadeira confiavelmente produzida, não há necessidade de justificação. O confiabilismo, então, assume duas formas: como uma teoria da justificação e como uma teoria do conhecimento. Na primeira, ele considera a justificação como sendo um importante ingrediente do conhecimento, mas, diferentemente do TK, alicerçam a justificação unicamente na confiabilidade. Como uma teoria do conhecimento, o confiabilismo sustenta que a justificação não é necessária para o

---

<sup>7</sup> O caso dos celeiros de fachada apareceu primeiramente em Goldman, 1976.

<sup>8</sup> Isso pode não ser fácil de conseguir de uma maneira sistemática e adotando princípios. No ambiente de Henry, a sua visão certamente é confiável quando a questão é distinguir uma vaca de uma ovelha. Assim, apelar para o ambiente de Henry deve ser mais refinado de tal maneira que permita dizer que a visão de Henry não é confiável quando a questão é o reconhecimento de celeiros, sem nos comprometer com o resultado de que ela é também não confiável quando tratamos do reconhecimento de vacas. Um segundo problema surge com relação à determinação dos limites da área de Henry. Suponha que comparemos os dois casos. Cada caso envolve um celeiro que Henry vê. Eles estão a 50 metros um do outro. Um está dentro do ambiente dos celeiros falsos e o outro fora. Dado que uma fronteira deve ser traçada em algum lugar, pode parecer difícil evitar tais casos comparativos. Mas uma vez que traçamos uma fronteira, temos o resultado que, quando Henry olha no celeiro de fora da linha, ele sabe que há um celeiro, enquanto que quando ele olha no celeiro de dentro da linha, ele não sabe. Dado que os celeiros estão justamente 50 m. de distância um do outro, isso pareceria estranho.

conhecimento; antes, a crença verdadeira confiavelmente produzida (considerando que a noção de confiabilidade é adequadamente refinada de modo a excluir os casos do tipo Gettier) é suficiente para isso<sup>9</sup>.

## 2. O que é a justificação?

Quando discutimos a natureza da justificação, devemos distinguir entre duas questões diferentes: primeiro, o que queremos dizer quando usamos a palavra “justificação”? Segundo, o que torna uma crença justificada? É importante manter estas questões separadas, pois um desacordo sobre como responder a segunda questão será uma disputa meramente verbal se as partes discordantes tiverem em mente diferentes conceitos de justificação. Assim, consideremos primeiro o que poderíamos querer dizer por “justificação” e, então, passemos para as questões que não são de definição<sup>10</sup>.

### 2.1 Justificação deontológica e não-deontológica.

Como o termo 'justificação' é usado na linguagem comum? Eis um exemplo: Tom faz uma pergunta a Martha e Martha responde com uma mentira. Ela estava justificada em mentir? Jane pensa que ela estava, pois a pergunta de Tom era inapropriada, uma vez que a resposta da pergunta não era da conta de Tom. O que Jane poderia ter em mente ao pensar que Marta estava justificada em responder com uma mentira? Uma resposta natural é a seguinte: ela tinha em mente que Marta não estava *sob obrigação alguma* de abster-se de

---

<sup>9</sup> Para exemplos de confiabilismo como uma teoria do conhecimento, ver Dretske, 1971, 1981 e Nozick, 1981. Ambos são descrições subjuntivas do conhecimento. Dretske coloca a confiabilidade na caracterização do conhecimento ao impor a condição que, se *S* sabe que *p*, então *S* tem uma razão conclusiva para *p*: uma razão tal que, se *p* fosse falso, *S* não teria essa crença. De modo similar, Nozick assume a confiabilidade por tornar o que tem sido chamado de *sensitividade* uma condição necessária para o conhecimento: Se *S* sabe que *p*, então *S* não acreditaria que *p* se *p* fosse falso. A sensitividade deve ser distinguida da *segurança*, que é mais fraca: Se *S* sabe que *p*, então não seria fácil que *S* viesse acreditar incorretamente ao crer que *p*. Para uma discussão da distinção sensitividade e segurança, ver Sosa, 1999.

<sup>10</sup> Sobre a distinção entre questões de definição e substantivas, veja o ensaio de Alston, “Concepts of Epistemic Justification”, em Alston 1989 pp. 81-114, Goldman 1979, e Steup 1996, capítulo 2.

mentir. Em virtude do caráter inapropriado da questão de Tom, não era um dever de Martha dizer a verdade. Esta compreensão da justificação, comumente denominada deontológica, pode ser definida da seguinte maneira: S está justificada em fazer x se e somente se S não está obrigada a se abster de fazer x<sup>11</sup>.

Suponha, quando aplicamos a palavra 'justificação' não a ações, mas a crenças, que queremos dizer algo análogo. Neste caso, o termo 'justificação' como usado na epistemologia teria de ser definido desta maneira:

**Justificação deontológica (DJ):** S está justificado a crer que *p* se e somente se S crê que *p* quando não é o caso que S está obrigado a abster-se de crer que *p*<sup>12</sup>.

Que tipo de obrigações são relevantes quando queremos avaliar se uma *crença*, ao invés de uma ação, está justificada ou injustificada? Ao avaliar uma ação estamos interessados em avaliar a ação de um ponto de vista moral ou prudencial, mas, quando se trata da crença, o que importa é a busca da *verdade*. Os tipos relevantes de obrigações, assim, são aqueles que surgem quando almejamos ter crenças verdadeiras. Todavia, exatamente o que devemos fazer na busca deste objetivo? De acordo com certo tipo de resposta, a resposta preferida pelos evidencialistas, nós devemos crer em conformidade com a nossa evidência. Para que esta resposta seja útil, precisamos de uma explicação acerca daquilo que constitui a nossa evidência. De acordo com outra resposta, nós devemos seguir as normas epistêmicas corretas. Se esta resposta vai nos ajudar a determinar quais obrigações a verdade-como-objetivo impõe sobre nós, precisamos que nos seja dada uma explicação do que são normas epistêmicas corretas<sup>13</sup>.

---

<sup>11</sup>Esta definição emprega a noção de obrigações. Definições alternativas podem ser dadas empregando outros membros da família de termos deontológicos: requerimento, dever, permissão ou proibição. Definições adicionais são ainda possíveis quando alargamos a classe de conceitos relevantes, empregando noções como responsabilidade, inocência e culpabilidade.

<sup>12</sup>Em relação à literatura sobre a compreensão deontológica da justificação, veja os ensaios 4 e 5 em Alston 1989, pp. 115-152, Ginet 1975, Bonjour 1985 capítulo 2, Feldman 1988, 2001a, Haack 2001, Plantinga 1993, Russell 2001 e Steup 1996, capítulo 4.

<sup>13</sup>Com respeito à verdade como um objetivo epistêmico e a conexão entre a verdade e a justificação, veja Conee 2004, David 2001 e os ensaios de David e Kvanvig em Steup e Sosa 2005.

A compreensão deontológica do conceito de justificação é comum à maneira pela qual filósofos como Descartes, Locke, Moore e Chisholm pensaram a justificação. Hoje, entretanto, a visão dominante é que a compreensão deontológica da justificação é inadequada para os objetivos da epistemologia. Duas objeções principais foram levantadas contra a concepção deontológica da justificação. Primeiro, argumentou-se que a DJ (*Deontological Justification*) pressupõe que podemos ter um controle suficientemente elevado sobre as nossas crenças. Porém, crenças são muito mais semelhantes a coisas como processos digestivos, espirros e piscadas involuntárias dos olhos, do que de ações. A ideia é que a crença simplesmente surge ou é algo que acontece conosco. Portanto, crenças não são apropriadas para a avaliação deontológica<sup>14</sup>. A esta objeção, alguns defensores da DJ responderam que a falta de controle sobre as nossas crenças não é obstáculo para usar o termo “justificação” no seu sentido deontológico<sup>15</sup>. Outros responderam que é um erro pensar que podemos controlar nossas crenças menos do que as nossas ações<sup>16</sup>.

De acordo com a segunda objeção ao DJ, a justificação deontológica não tende a ‘epistemologizar’ crenças verdadeiras: ela não tende a torná-las não-acidentalmente verdadeiras. Esta afirmação é tipicamente sustentada a partir da descrição de casos que envolvem sociedades ignorantes e culturalmente isoladas ou sujeitos que são cognitivamente deficientes. Estes casos envolvem crenças sobre as quais se alega que são epistemicamente deficitárias, embora não parecerá que os sujeitos nestes casos estão sob qualquer obrigação de abster-se de crer do modo como o fazem. O que torna as crenças em questão epistemicamente deficitárias é que elas são formadas usando métodos não confiáveis e defeituosos intelectualmente. A razão pela qual os sujeitos, do seu próprio ponto de vista, não estão obrigados a crer de outro modo é que eles ou são deficientes cognitivos ou vivem em uma comunidade ignorante e isolada. DJ diz que tais crenças estão justificadas. Se elas satisfazem as condições necessárias remanescentes, os teóricos da DJ terão de considerá-las

---

<sup>14</sup>Veja os ensaios de Alston “Concepts of Epistemic Justification” e “The Deontological Conception of Epistemic Justification”, ambos em Alston 1989.

<sup>15</sup>Veja Feldman 2001a.

<sup>16</sup>Veja Ryan 2003 e Steup 2000.

como conhecimento. De acordo com a objeção, contudo, as crenças em questão, mesmo se verdadeiras, não poderiam se qualificar como conhecimento devido à maneira epistemicamente deficitária pela qual foram formadas. Consequentemente, DJ deve ser rejeitada<sup>17</sup>.

Aqueles que rejeitam a DJ usam o termo “justificação” em um sentido técnico que se afasta da maneira como a palavra é comumente usada. O sentido técnico é concebido para fazer o termo satisfazer às necessidades da epistemologia<sup>18</sup>. Mas como, então, podemos conceber a justificação? O que se tem em mente quando se pensa numa crença justificada em um sentido não-deontológico? Lembre que a função atribuída à justificação é assegurar que a crença é verdadeira, mas não por simples acidente. Digamos que este sentido é satisfeito quando uma crença verdadeira é uma instância da propriedade de *probabilização apropriada*. Podemos, então, definir a justificação não-deontológica como se segue:

**Justificação não-deontológica (NDJ):** *S* está justificado a crer que *p* se e somente se *S* crê que *p* a partir de uma base que probabiliza apropriadamente a crença de *S* de que *p*.

Se quisermos determinar no quê, exatamente, consiste a probabilização, teremos de lidar com uma variedade de questões complicadas<sup>19</sup>. Por hora, vamos focar no ponto principal. Aqueles que preferem NDJ à DJ diriam que a probabilização e a justificação deontológica

---

<sup>17</sup>Para uma resposta a esta objeção, veja Steup 1999.

<sup>18</sup>Veja Alston 1989, p. 7f. Alston escreve: “Eu concordo que a ‘justificação’ é a palavra errada para um conceito não-deontológico, mas parecemos estar presos a ela na teoria contemporânea do conhecimento”.

<sup>19</sup>Uma crença pode ser objetivamente provável de uma maneira completamente irrelevante para o fato de a crença ser ou não ser uma instância de conhecimento. Neste caso, ela não seria “probabilizada adequadamente”. Suponha que Jack crê que Meyer vencerá a eleição. Suponha adicionalmente que as crenças de Jack se originam apenas por pensamento ingênuo. Por fim, suponha que Meyer a vitória de Meyer é objetivamente provável porque é um fato que 80% daqueles que votarão irão votar nele. Assim, a *p* em que Jack crê é objetivamente provável. Como resultado, é objetivamente provável que a crença de Jack seja verdadeira. Mas visto que a crença de Jack resulta de pensamento ingênuo, ela não estaria justificada ou seria uma instância de conhecimento. Assim o tipo de probabilidade objetiva que surge de *p* ser objetivamente provável não é do tipo correto. O que precisamos é de probabilidade objetiva não em virtude do que o sujeito crê (o conteúdo da crença), mas em virtude da maneira em que a crença surge. Assim, o que procuramos é que a origem da crença a faça objetivamente provável. Mas então devemos encontrar, de um modo sistemático e por princípio, um meio de determinar qual é a origem da crença. Se pensarmos a origem da crença em termos de processos cognitivos, este empreendimento levanta o que foi chamado de o “problema da generalidade”. Veja o ensaio 6 em Conee e Feldman 2004.

podem divergir: é possível que uma crença esteja justificada deontologicamente sem estar probabilizada apropriadamente. Isto é o que supostamente mostram os casos envolvendo culturas ignorantes ou sujeitos cognitivamente deficientes<sup>20</sup>.

## 2.2 Evidência vs. Confiabilidade

O que torna uma crença justificada justificada? De acordo com os evidencialistas, é a posse da evidência. O que é, todavia, possuir evidência para crer que  $p$ ? Alguns evidencialistas diriam que é estar em um estado mental que representa  $p$  como sendo verdadeira. Por exemplo, se o café na sua xícara tem paladar doce para você, então você tem evidência para crer que o café está doce. Se você sente uma dor latejante na cabeça, você tem evidência para crer que você tem uma dor de cabeça. Se você tem a lembrança de ter ingerido cereal no café da manhã, então você tem evidência para uma crença sobre o passado: a crença sobre o que você comeu quando tomou café da manhã. E quando você claramente “vê” ou “intui” que a proposição “Se Jack tomou mais do que quatro xícaras de café, então Jack tomou mais do que três xícaras de café” é verdadeira, então você tem evidência para crer nesta proposição. Nesta abordagem, a evidência consiste em experiências perceptivas, introspectivas, mnemônicas e intuitivas, e possuir evidência significa ter uma experiência deste tipo. Assim, de acordo com este evidencialismo, o que torna você estar justificado a crer que  $p$  é o fato de você ter uma experiência que representa  $p$  como sendo verdadeira.

Muitos confiabilistas também diriam que as experiências mencionadas no parágrafo anterior importam. Porém, eles negariam que a justificação é apenas uma questão de ter experiências adequadas. Alternativamente, eles sustentam que uma crença está justificada se e somente se ela resulta de uma fonte cognitiva confiável: uma fonte que tende a produzir crenças verdadeiras e, portanto, probabiliza apropriadamente a crença. Confiabilistas, assim, concordariam que as crenças mencionadas no parágrafo anterior estão justificadas. Mas de acordo com a posição padrão do confiabilismo, o que as torna justificadas não é a posse de

---

<sup>20</sup>Para uma elaboração do conceito não-deontológico da justificação, veja os ensaios 4, 5 e 9 em Alston 1989.

evidência, mas o fato de que os tipos de processos – percepção, introspecção, memória e intuição racional – que deram origem às crenças são confiáveis.

### 2.3 Interno vs. Externo

Na epistemologia contemporânea, tem havido um longo debate sobre se a justificação é interna ou externa. Internalistas afirmam que ela é interna; externalistas o negam. Como devemos entender estas afirmações?

Para entender no que consiste a distinção interno-externo, precisamos ter em mente que, quando uma crença está justificada, há algo que a *torna* justificada. De modo semelhante, se uma crença está injustificada, há algo que a *torna* injustificada. Chamemos as coisas que tornam uma crença justificada ou injustificada de fatores-J. A disputa sobre se a justificação é interna ou externa é uma disputa sobre o que são os fatores-J.

Entre aqueles que pensam que a justificação é interna, não há unanimidade sobre como entender o conceito de internalidade. Podemos distinguir duas abordagens. De acordo com a primeira, a justificação é interna porque desfrutamos de um tipo especial de acesso aos fatores-J: eles são *sempre* reconhecíveis por reflexão<sup>21</sup>. Portanto, assumindo certas premissas adicionais (que serão mencionadas logo mais), a própria justificação é sempre reconhecível por reflexão<sup>22</sup>. De acordo com a segunda abordagem, a justificação é interna porque os fatores-J são sempre estados mentais<sup>23</sup>. Vamos chamar o primeiro de *internalismo acessibilista* e o último de *internalismo mentalista*. Externalistas negam que fatores-J satisfazem qualquer uma destas condições.

---

<sup>21</sup>A palavra “sempre” é importante aqui, pois externalistas não precisam, e na verdade não devem, asseverar que a justificação, entendida externamente em termos de confiabilidade, nunca é reconhecível por reflexão. Por exemplo, você escuta e assim vem a crer que há um cachorro latindo do lado de fora. Sem dúvida, em um caso típico como este, a reflexão lhe diz que a sua crença tem uma origem confiável. Se ela tem, então você pode, nesta ocasião, reconhecer por reflexão que a sua crença está justificada mesmo se entendemos a justificação em termos de confiabilidade.

<sup>22</sup>O internalismo de acesso foi defendido por Roderick Chisholm, que pode razoavelmente ser visto como o principal defensor da epistemologia tradicional e internalista na segunda metade do século vinte. Em seu 1977, p. 17, Chisholm escreve: “pressupomos...que as coisas que sabemos estão justificadas para nós no seguinte sentido: podemos saber, *em qualquer ocasião*, o que é que constitui as nossas bases, ou a razão, ou a evidência para pensar que sabemos” (ênfase adicionada).

<sup>23</sup>Veja Conee e Feldman 2001.

O evidencialismo é geralmente associado ao internalismo e o confiabilismo, ao externalismo<sup>24</sup>. Vejamos porquê. O evidencialismo diz, no mínimo, duas coisas:

E1 O fato de alguém estar ou não justificado em crer em  $p$  depende da evidência que ele tem concernente a  $p$ .

E2 A evidência de alguém consiste nos seus estados mentais.

Em virtude de E2, o evidencialismo é obviamente uma instância do internalismo mentalista.

Se o evidencialismo é também uma instância do internalismo acessibilista é uma questão mais complicada. A conjunção de E1 e E2 por si mesmas não implica nada sobre o reconhecimento da justificação. Lembre, contudo, que na seção 1.1, distinguimos entre TK e NTK: a abordagem tradicional e não-tradicional da análise do conhecimento e justificação. Defensores do TK, entre os quais o evidencialismo goza de ampla simpatia, tendem a endossar as duas seguintes afirmações:

**Luminosidade:** O conhecimento da sua própria mente é cognitivamente luminoso: fiando-se na introspecção, alguém pode sempre reconhecer por reflexão os estados mentais em que se encontra<sup>25</sup>.

**Necessidade:** Princípios necessários, reconhecíveis *a priori*, dizem o que é evidência para quê<sup>26</sup>.

---

<sup>24</sup>Sem dúvida, existem versões não-evidencialistas do internalismo. Por exemplo, considere a visão de que a coerência do seu sistema de crenças é um fator-J. De acordo com esta abordagem, o estatuto de justificação das crenças de um sujeito é determinado por algo mais que a sua evidência. Se a coerência de um sistema de crença é algo adequadamente interno, tal visão não-evidencialista contará como internalista. Entre as teorias externalistas, o confiabilismo não é também o único candidato. Por exemplo, o funcionalismo próprio de Plantinga, a teoria do rastreamento de Nozick e teoria das razões conclusivas de Dretske todas elas qualificam como externalistas, mas nenhuma delas é em qualquer sentido estrito uma versão do confiabilismo. Veja Plantinga 1993b, Nozick 1981 e Dretske 1971 e 1981.

<sup>25</sup>Eu tomei emprestado o termo “luminosidade” de Williamson 2000, capítulo 4. Williamson rejeita a afirmação de que estados mentais são luminosos.

<sup>26</sup>A primeira vista, esta afirmação pode parecer estranha. Por exemplo, quando um papel de tornassol imerso em um líquido torna-se vermelho, isto pode ser visto como evidência para o líquido ser uma solução ácida. Certamente, pode-se argumentar, que a mudança da cor do papel ser evidência para a acidez não é o tipo de coisa que pode ser descoberta apenas por reflexão. Contudo, se você não sabe que o papel de tornassol quando imerso em uma solução ácida torna-se

Fiando-se na intuição *a priori* uma pessoa pode sempre reconhecer por reflexão se os seus estados mentais são evidência para  $p$ <sup>27</sup>.

Embora E1 e E2 por si mesmos não impliquem o internalismo de acesso, é perfeitamente plausível manter que o evidencialismo, quando embelezado com a Luminosidade e a Necessidade, torna-se uma instância do internalismo de acesso<sup>28</sup>.

Consideremos agora porque o confiabilismo é uma teoria externalista. O confiabilismo diz que a justificação das crenças de alguém é uma função não da sua evidência, mas da confiabilidade das suas fontes de crença, tais como os processos e estados mnemônicos, perceptivos, e introspectivos. Apesar de as fontes poderem ser qualificadas como mentais, a sua confiabilidade não pode ser qualificada como sendo mental. Portanto, os confiabilistas rejeitam o internalismo mentalista. Além disso, se a justificação das crenças de alguém é determinada pela confiabilidade das suas fontes de crença, a justificação não será sempre reconhecível por reflexão. Portanto, os confiabilistas rejeitam o internalismo de acesso também<sup>29</sup>.

Usemos um exemplo de engano radical para ilustrar a diferença entre evidencialismo como uma teoria internalista e o confiabilismo como uma teoria externalista. Se o evidencialismo é verdadeiro, um sujeito que é enganado radicalmente será induzido ao erro

---

vermelho, a sua observação da mudança de cor do papel não será evidência alguma para você pensar que o líquido em questão é uma solução ácida. Assim, se formos cuidados ao descrever o que a evidência neste caso realmente é, teremos de dizer que ela consiste dos dois seguintes itens:

(1) O princípio geral de que, quando imerso em uma solução ácida, o papel de tornassol torna-se vermelho.

(2) A observação de que a tira de papel de tornassol no líquido ficou vermelha.

O que é evidência para:

(3) O líquido é uma solução ácida.

não é (2) apenas, mas a conjunção de (1) e (2). Quando teóricos da TK argumentam que as conexões evidenciais são reconhecíveis por reflexão, eles sempre têm em mente princípios que identificam a evidência relevante na sua plenitude. Assim, considerando o exemplo presente, defensores da TK argumentariam que o que é reconhecível por reflexão é a proposição de que a conjunção de (1) e (2) é evidência para (3). Esta afirmação pode certamente ser contestada, mas não lhe falta plausibilidade de nenhuma maneira óbvia.

<sup>27</sup>Chisholm mantém que existem princípios de evidência necessariamente verdadeiros e reconhecíveis *a priori* e que estes princípios são internos “tal que o uso apropriado deles a qualquer momento nos habilitará a determinar o estatuto epistêmico de nossas próprias crenças neste momento”. Chisholm 1989, p. 62. Sobre a visão de Chisholm a respeito do estatuto *a priori* destes princípios, veja a p. 72. Para uma explicação excelente do internalismo clássico, o internalismo de Chisholm e o internalismo pós-Chisholm, veja Plantinga 1993.

<sup>28</sup>A combinação do evidencialismo com a Luminosidade e a Necessidade é suficiente para implicar o internalismo de acesso? Esta é uma questão complicada que não pode ser investigada aqui. Para resolver esta questão, teríamos de abordar, entre outras coisas, a questão de o que exatamente é para algo ser reconhecível por reflexão.

<sup>29</sup>Tipicamente, externalistas também rejeitarão a Luminosidade e a Necessidade.

sobre o que é realmente o caso, mas não sobre o que ele está justificado a crer. Se, por outro lado, o confiabilismo é verdadeiro, então o sujeito será induzido ao erro tanto sobre o que é realmente o caso, quanto sobre o que ele está justificado a crer. Vejamos porquê.

**Distinga entre Tim e Tim\*:** uma mesma pessoa que imaginamos em duas situações completamente diferentes. A situação de Tim é normal, como a sua ou a minha. Tim\*, entretanto, é um cérebro numa cuba. Suponha que um cientista maluco abduziu e “encubou” Tim\* removendo o cérebro do seu crânio e colocando-o em uma cuba na qual o seu cérebro é mantido vivo. Em seguida, o cientista maluco conecta os terminais nervosos do cérebro de Tim\* com fios de uma máquina que, controlada por um computador poderoso, começa a estimular o cérebro de Tim\* de uma maneira que Tim\* não note o que realmente aconteceu a ele. Ele terá experiências perfeitamente comuns, assim como Tim. Na verdade, assumamos que os estados mentais de Tim e os estados mentais de Tim\* são idênticos. Contudo, visto que Tim\* é um cérebro numa cuba, ele está, diferentemente de Tim, radicalmente enganado sobre a sua situação real. Por exemplo, quando Tim acredita que ele tem mãos, ele está certo. Quando Tim\* acredita que ele tem mãos, ele está errado (suas mãos foram descartadas, junto com o resto dos seus membros e tronco). Quando Tim acredita que ele está bebendo café, ele está certo. Quando Tim\* acredita que ele está bebendo café, ele está errado (cérebros não bebem café). Suponhamos agora que Tim\* pergunte a si mesmo se ele está justificado a crer que ele tem mãos. O evidencialismo implica que a resposta de Tim\* está correta. Pois ainda que ele esteja enganado sobre a sua situação externa, ele não está enganado a respeito da sua evidência: a maneira que as coisas lhe aparecem em sua experiência. Isto ilustra a internalidade da justificação evidencialista. O confiabilismo, por outro lado, sugere que a resposta de Tim\* é incorreta. A crença de Tim\* de que ele tem mãos surge de processos cognitivos - “ver” e “sentir” suas (não-existentes) mãos – que agora não produz praticamente nenhuma crença verdadeira. Na medida em que isso implique a sua inconfiabilidade, as crenças resultantes estão injustificadas. Consequentemente, ele está enganado não só sobre

sua situação externa (ele não tem mãos), mas também sobre o estatuto justificacional da sua crença de que ele tem mãos. Isto ilustra a externalidade da justificação confiabilista.

O exemplo de Tim e Tim\* pode servir também para ilustrar uma maneira a mais pela qual podemos conceber a diferença entre internalismo e externismo. Alguns internalistas tomam o seguinte princípio como característico do ponto de vista internalista:

**Mentalismo:** Se dois sujeitos, S e S\*, são mentalmente idênticos, então o estatuto justificacional de suas crenças é idêntico também: as mesmas crenças estão justificadas ou injustificadas para eles na mesma medida<sup>30</sup>.

Quando aplicamos este princípio ao exemplo do Tim/Tim\*, ele nos diz que o evidencialismo é uma teoria internalista e o confiabilismo, externista. Embora haja diferenças físicas significativas entre Tim e Tim\*, eles são mentalmente idênticos. O evidencialismo implica que, visto que Tim e Tim\* são mentalmente idênticos, eles têm a mesma evidência e, portanto, estão também na mesma situação justificacional. Por exemplo, eles estão ambos justificados a crer que eles têm mãos. Isto faz do evidencialismo uma teoria internalista. O confiabilismo, por outro lado, permite que, apesar de Tim e Tim\* serem mentalmente indênticos, eles difiram justificacionalmente, visto que as crenças de Tim são produzidas por faculdades cognitivas confiáveis, enquanto que as faculdades que produzem as crenças de Tim\* podem ser consideradas inconfiáveis. Por exemplo, algumas versões do confiabilismo implicam que Tim está justificado a crer que ele tem mãos, enquanto Tim\* não. Isto faz do confiabilismo uma teoria externalista<sup>31</sup>.

---

<sup>30</sup>Veja Conee e Feldman 2004, p. 56.

<sup>31</sup>Com respeito à leitura sobre a disputa internalismo-externalismo, veja Kornblith 2001, o artigo de George Pappas “Internalist vs. Externalist Conceptions of Epistemic Justification”, referido no fim deste artigo e Sosa 2010b.

## **2.4 Por que o Internalismo?**

Por que pensar que a justificação é interna? Um argumento para a internalidade da justificação avança da seguinte maneira: “A justificação é deontológica: ela é uma questão de cumprimento de dever. Mas o cumprimento de dever é interno. Portanto, a justificação é interna”. Outro argumento apela ao cenário do cérebro-numa-cuba que consideramos acima: “A crença de Tim\* de que ele tem mãos está justificada da mesma maneira que a de Tim. Tim\* é internamente idêntico a Tim e externamente bem diferente. Portanto, são os fatores internos que justificam crenças. Por fim, visto que a justificação resultante da posse de evidência é a justificação interna, o internalismo pode ser apoiado pela defesa do evidencialismo. O que, então, pode ser dito em favor do evidencialismo? Os evidencialistas apelariam a casos em que uma crença é formada confiavelmente, mas não é acompanhada por qualquer experiência que qualificaria como evidência. Eles diriam então que não é plausível afirmar, em casos como estes, que a crença do sujeito está justificada. Portanto, estes casos mostram, de acordo com o evidencialismo, que uma crença não pode ser justificada a menos que ela seja apoiada pela evidência<sup>32</sup>.

## **2.5 Por que o Externalismo?**

Por que pensar que a justificação é externa? Para começar, os externalistas sobre a justificação apontariam para o fato de que animais e crianças pequenas têm conhecimento e, assim, possuem crenças justificadas. Porém, suas crenças não podem ser justificadas da maneira como os evidencialistas concebem a justificação. Portanto, devemos concluir que a justificação de que as suas crenças desfrutam é externa: resulta não da posse de evidência, mas do fato de originar-se de processos confiáveis. E, em segundo lugar, externalistas diriam que o que queremos da justificação é o tipo de probabilidade objetiva necessária para o

---

<sup>32</sup>A respeito de literatura em defesa do internalismo, veja Bonjour 1985, a contribuição de Bonjour à Bonjour e Sosa 2003, Conee e Feldman 2001, Feldman 2005 e Steup 1999b, 2001.

conhecimento e apenas condições externas sobre a justificação implicam esta probabilidade. Portanto, a justificação tem condições externas<sup>33</sup>.

### **3 A estrutura do conhecimento e da justificação**

O debate sobre a estrutura do conhecimento e da justificação é primariamente um debate entre aqueles que sustentam que o conhecimento exige justificação. Deste ponto de vista, a estrutura do conhecimento deriva da estrutura da justificação. Iremos, portanto, focar no último.

#### **3. 1 Fundacionalismo**

De acordo com o fundacionalismo, nossas crenças justificadas estruturam-se como numa construção: elas são divididas em uma fundação e uma superestrutura, sendo a última sustentada pela primeira. Crenças pertencentes à fundação são *básicas*. Crenças pertencentes à superestrutura são *não-básicas* e recebem justificação das crenças justificadas na fundação<sup>34</sup>.

Para um relato fundacionalista da justificação ser plausível, ele precisa solucionar dois problemas. Primeiro: em virtude do quê exatamente crenças básicas são justificadas? Segundo: como crenças básicas justificam crenças não-básicas? Antes de resolvermos estas questões, vamos primeiro analisar a questão do que é que torna, a princípio, uma crença justificada básica. Uma vez que tenhamos feito isso, podemos então passar a discutir em virtude do que uma crença básica pode ser justificada, e como tal crença pode justificar uma crença não-básica.

<sup>33</sup>A respeito de literatura defendendo o externismo, veja os ensaios 8 e 9 em Alston 1989, Greco 2005, Goldman 1999a, Kornblith 1999, 2001 e as contribuições de Sosa a Bonjour e Sosa 2003.

<sup>34</sup> Para ver mais literatura acerca da questão do fundacionismo-coerentismo, ver Audi, 1997, BonJour, 1999, 2001, 2002, capítulo 2, BonJour e Sosa, 2003, Chisholm, 1982, capítulos 4, 8, e 9 em Dancy, 1985, DePaul, 2001, capítulo 4 em Feldman, 2003, Fumerton, 2001, Haack, 1993, Pryor, 2005, os ensaios de Sosa 9 e 10, 1991 em Sosa, 1999, Capítulos 5-7 em Steup, 1996, e capítulos 5-8 em Steup e Sosa, 2005, Williams 1999<sup>a</sup>, 2005, e Van Cleve, 1985, 2005.

De acordo com essa abordagem, o que faz de uma crença justificada básica é que ela não recebe a sua justificação de quaisquer outras crenças. A seguinte definição capta esse pensamento:

**Basicidade Doxástica (BD):** A crença justificada de *S* de que *p* é básica se e somente se a crença de *S* de que *p* foi justificada sem extrair sua justificação de qualquer outra das outras crenças de *S*.

Consideremos o que poderia ser, de acordo com BD, qualificado como um exemplo de crença básica. Suponha que você percebe (por qualquer motivo) o chapéu de alguém, e você também percebe que esse chapéu parece azul para você. Então você acredita

(B) Parece-me que esse chapéu é azul.

A menos que algo muito estranho esteja acontecendo, (B) é um exemplo de uma crença justificada. BD nos diz que (B) é básica, se e somente se não deve a sua justificação a qualquer outra crença. Portanto, se (B) é de fato básica, pode haver algum item ou outro ao qual (B) deve a sua justificação, mas esse item não seria outra crença sua. Chamamos esse tipo de basicidade de 'doxástica' porque isso transforma a basicidade numa função do modo pelo qual seu sistema doxástico (seu sistema de crenças) é estruturado.

Voltemo-nos para a questão de onde a justificativa atribuída à (B) pode vir, se pensarmos em basicidade conforme definido pela BD. Note que BD apenas nos diz como (B) não é justificada. Ela não diz nada sobre *como* (B) é justificada. BD, portanto, não responde essa pergunta. O que precisamos, além de BD, é um relato *do que é* que justifica uma crença, como (B). De acordo com uma corrente de pensamento fundacionalista, (B) é justificada porque não pode ser falsa, dubitável, ou corrigida por outros. Então, (B) é justificada porque (B) traz consigo um *privilégio epistêmico* como a infalibilidade, indubitabilidade, ou

incorrigibilidade<sup>35</sup>. A ideia é que (B) é justificada em virtude de sua natureza intrínseca, o que a faz possuir algum tipo de privilégio epistêmico.

Note que (B) não é uma crença sobre o chapéu. Pelo contrário, é uma crença sobre como o chapéu *aparece* para você. Então, (B) é uma crença introspectiva sobre uma experiência perceptiva sua. De acordo com as intuições que estamos considerando aqui, as crenças básicas de um sujeito são compostas de crenças introspectivas sobre seus próprios estados mentais, dos quais experiências perceptivas compõem um subconjunto. Outros estados mentais sobre os quais um sujeito pode ter crenças básicas incluem coisas como ter uma dor de cabeça, cansaço, sensação de prazer, ou de ter um desejo por uma xícara de café. Crenças sobre objetos externos não devem e certamente não podem ser consideradas básicas, pois é impossível para tais crenças possuir o tipo de privilégio epistêmico necessário para o status de serem básicas.

De acordo com uma versão diferente do fundacionalismo, (B) não se justifica em virtude de possuir algum tipo de status privilegiado, mas por algum outro estado mental seu. Esse estado mental, no entanto, não é outra *crença* sua. Pelo contrário, é a própria *experiência perceptiva* acerca da qual (B) trata: o chapéu que lhe parece azul. Considere que '(E)' representa essa experiência. De acordo com esta proposta alternativa, (B) e (E) são estados mentais distintos. A ideia é que o que justifica (B) é (E). Uma vez que (E) é uma experiência, não uma crença sua, (B) é, de acordo com BD, básico.

Vamos chamar as duas versões de fundacionalismo que distinguimos em *fundacionalismo de privilégio* e *fundacionalismo experiencial*. O fundacionalismo de privilégio restringe crenças básicas a crenças sobre os próprios estados mentais de alguém. O fundacionalismo experiencial é menos restritivo. Segundo ele, crenças sobre objetos externos também podem ser básicas. Suponha que em vez de (B), você acredite (H) que esse chapéu é azul.

---

<sup>35</sup> Para uma discussão das várias espécies de privilégios epistêmicos, ver o ensaio 10 em Alston, 1989.

Ao contrário de (B), (H) é sobre o chapéu em si, e não sobre a forma como o chapéu aparece para você. Tal crença não é uma crença acerca da qual somos infalíveis ou, de outra forma, epistemicamente privilegiados. O fundacionalismo de privilégio poderia, portanto, classificar (H) como não-básica. É, no entanto, bastante plausível pensar que (E) justifica não só (B), mas também (H). Se (E) é de fato o que justifica (H), e (H) não recebe qualquer justificação adicional de qualquer outra crença sua, então (H) se qualifica, segundo a BD, como básica.

O fundacionalismo experiencial, então, combina duas ideias cruciais: (i) quando uma crença justificada é básica, a sua justificação não se deve a qualquer outra crença, (ii) o que de fato justifica as crenças básicas são experiências.

Em circunstâncias normais, crenças perceptivas, tais como (H) não estão baseadas em quaisquer *outras crenças* sobre a experiência perceptual de alguém. Não está claro, portanto, como o fundacionalismo de privilégio pode dar conta da justificação de crenças comuns perceptivas como (H). O fundacionalismo experiencial, por outro lado, não tem problema em explicar como crenças comuns perceptivas são justificadas: elas são justificadas pelas experiências perceptivas que dão origem a elas. Isto poderia ser visto como uma razão para preferir o fundacionalismo experiencial ao fundacionalismo de privilégio.

Observamos acima que a forma de pensar a basicidade não é incontroversa. BD define apenas um tipo de basicidade. Eis uma concepção alternativa disto:

**Basicidade Epistêmica (BE):** A crença justificada de  $S$  que  $p$  é básico se, e somente se, a crença justificada de  $S$  que  $p$  não depende de qualquer outra justificação que  $S$  possua para acreditar noutra proposição qualquer,  $q$ .<sup>36</sup>

<sup>36</sup> Uma concepção da basicidade que segue essas linhas é empregada por Huemer, 2000 and Pryor, 2005. Ela também aparece em Van Cleve, 2005. Huemer sustenta a concepção que uma crença que  $p$  pode ser justificada *unicamente* por um parecer que  $p$ . Pryor sustenta que uma crença perceptual que  $p$  pode ser justificada *unicamente* por uma experiência perceptual que  $p$ . Van Cleve argumenta em defesa de uma posição fundacionista caracterizada pela alegação que as crenças da memória podem ser justificadas *unicamente* por memórias ostensivas. Assim, cada uma deles sustenta que uma crença pode ser justificada por um fundamento experiencial apenas, isto é, sem o sujeito ter alguma justificação ulterior para crer algo em acréscimo a crença em questão.

BE torna mais difícil para uma crença ser básica do que BD faz. Para ver por que, nos voltemos para a questão principal (vamos chamá-la de 'questão-J') que os defensores do fundacionalismo experiencial enfrentam:

**A Questão-J:** Porque experiências perceptivas são uma fonte de justificação?

Uma maneira de responder à questão-J pode ser vista como uma *solução de compromisso*, uma vez que isso significa assumir um compromisso entre o fundacionalismo e seu concorrente, o coerentismo. A solução de compromisso será de interesse para nós porque ela ilustra como BD e BE diferem. Se adotarmos a posição de compromisso, crenças, tais como (H) qualificar-se-ão como básicas conforme a BD, mas de acordo com a BE como não-básicas. Vejamos o que diz a solução de compromisso.

Do ponto de vista coerentista, podemos responder à questão-J da seguinte maneira: experiências perceptivas são uma fonte de justificação, porque estamos justificados em *acreditar* que elas sejam confiáveis. Como veremos abaixo, fazer uma justificação perceptiva depender da existência de crenças de confiabilidade-atribuição é bastante problemático. Há, no entanto, uma resposta alternativa à questão-J que apela à confiabilidade sem fazer a justificação perceptiva depender de crenças que atribuem confiabilidade a experiências perceptivas. De acordo com essa segunda resposta à questão-J, experiências perceptivas são uma fonte de justificação, porque temos justificativa para considerá-las confiáveis. Chamemos essa perspectiva de *solução de compromisso*.<sup>37</sup>

Note que a sua justificação para acreditar que *p* não implica que você realmente acredite em *p*. Por exemplo, se você acredita que a pessoa ao seu lado veste um chapéu azul, você tem justificação para crer que a pessoa ao seu lado veste um chapéu azul ou um chapéu vermelho. Mas é claro que é improvável que você acredite na última, apesar de ter justificação.

---

<sup>37</sup> Para artigos advogando as posições de compromisso, ver DeRose, 2004 e Steup, 2004.

Da mesma forma, a sua justificação para confiar nas suas experiências perceptivas não implica que você tem dado atenção ao assunto e realmente formou a crença de que elas são confiáveis. De acordo com o tipo de coerentismo que consideramos acima, se suas experiências perceptivas são uma fonte de justificação para você, deve ser verdadeiro que você tenha considerado o assunto e *acreditado* que elas sejam confiáveis. A posição de compromisso não diz tal coisa. Ela diz apenas que, se suas experiências perceptivas são uma fonte de justificação para você, você deve ter *justificação* para crer que elas sejam confiáveis.

O que a justificação pode nos dar para pensar que nossas experiências perceptivas são confiáveis? Essa é uma questão complicada. Para os nossos presentes propósitos, vamos considerar a seguinte resposta: Lembremos que eles nos serviram bem no passado. Estamos supondo, então, que a justificação para a atribuição de confiabilidade para as suas experiências perceptivas consista em memórias de sucesso perceptual. De acordo com a posição de compromisso, nunca é uma experiência perceptual (E) *por si só* que justifica uma crença perceptual, mas apenas (E) em conjunto com adequadas memórias de registro-controle que lhe dão justificação para considerar (E) confiável. Considere '(E)' novamente como a experiência de ficar olhando para um chapéu que parece azul, e '(H)' para a sua crença de que esse chapéu é azul. De acordo com a posição de compromisso, (E) justifica (H) somente se (E) é acompanhado por memórias de registro-controle (M) que lhe dão justificativa para atribuir confiabilidade a suas experiências visuais. Então, de acordo com a posição de compromisso por nós descrita, o que justifica (H) é a conjunção de (E) e (M).

Podemos ver agora como BD e BE diferem. De acordo com a posição de compromisso, o ter justificação para (H) depende de você dispor de justificação para crer em algo mais, além de (H), ou seja, que suas experiências visuais são confiáveis. Como resultado, (H) não é fundamental no sentido definido pela BE. No entanto, (H) ainda pode ser fundamental no sentido definido por BD. Enquanto a sua justificativa para (H) depender somente de (E) e (M), nenhum dos quais inclui quaisquer *crenças*, BD nos diz que (H) é fundamental. Segue que um fundacionalista experiencial que pretende classificar crenças como (H) como crenças básicas

não poderá adotar a posição de compromisso, assim como o fundacionalista teria de dizer que (E) por si só é suficiente para tornar (H) uma crença justificada.

Como fundacionalistas experienciais que preferem BE à BD podem responder a questão-J? Em função da maneira como eles concebem basicidade eles não podem dizer que experiências perceptivas são uma fonte de justificação, porque você tem uma razão, *R*, simplesmente por acreditar que elas podem, pois *R* seria uma justificação para crer em outra coisa - a mesma coisa que, de acordo com a EB, é um obstáculo à basicidade. Uma opção para fundacionalistas-BE seria endossar o externalismo. Se o fizerem, eles poderiam dizer que experiências perceptivas são uma fonte de justificação se, e somente se, elas são de tipos que estão associadas de forma confiável à crenças verdadeiras crenças resultantes. Nessa abordagem, seria o próprio fato da confiabilidade e não a evidência da confiabilidade que faria das experiências perceptivas uma fonte de justificação.<sup>38</sup> Outra opção internalista seria dizer que as experiências perceptivas são uma fonte de justificação, porque não poderia ser de outra forma: *é uma verdade necessária* que certas experiências perceptivas podem justificar certas crenças perceptivas. Esta seria uma resposta internalista para a questão-J, porque as experiências perceptivas seriam, se confiáveis ou não, uma fonte de justificação.<sup>39</sup>

Para concluir esta seção, vamos considerar brevemente como a justificação é transferida de crenças básicas para crenças não-básicas. Existem duas opções: a relação justificatória entre crenças básicas e não-básicas pode ser dedutiva ou não-dedutiva. Se tomarmos a relação como sendo dedutiva, cada crença não-básica teria que ser de tal forma que ela poderia ser deduzida a partir das crenças básicas de alguém. Isto parece excessivamente exigente. Se considerarmos uma seleção aleatória de crenças típicas que temos, não é fácil ver a partir de quais crenças básicas elas poderiam ser deduzidas.

---

<sup>38</sup> Um problema que surge para essa abordagem é o seguinte: muitos epistemólogos partilham a intuição que, se um gênio maligno engana você e suas experiências perceptuais são, portanto, completamente enganosas, elas são, não obstante isso, uma fonte de justificação para você, pois de seu ponto de vista, do seu ponto de vista interno, tal engano não é detectável. A resposta externalista que consideramos, contudo, implica que elas não podem ser fontes de justificação.

<sup>39</sup> Segue-se dessa resposta internalista que suas experiências perceptivas são uma fonte de justificação para você mesmo se o gênio maligno o engana.

Fundacionalistas, portanto, tipicamente concebem a relação entre a fundação e a superestrutura em termos não-dedutivos. Eles diriam que para uma crença básica, B, justificar uma crença não-básica, B\*, não é necessário que B implique B\*. Pelo contrário, é suficiente que, dado B, B\* é provavelmente verdadeira.

### **3.2 Coerentismo**

O fundacionalismo diz que o conhecimento e a justificação são estruturados como um edifício, composto de uma superestrutura que repousa sobre uma fundação. De acordo com o coerentismo, essa metáfora é desencaminhadora. O conhecimento e justificação são estruturados como uma *teia*, onde a força de uma determinada área depende da força das áreas circundantes. Coerentistas, então, negam que existem crenças básicas. Como vimos na seção anterior, existem duas maneiras diferentes de conceber a basicidade. Consequentemente, há duas maneiras correspondentes de interpretar o coerentismo: como a negação da basicidade doxástica ou como a negação da basicidade epistêmica. Considere primeiramente o coerentismo como a negação da basicidade doxástica:

**Coerentismo Doxástico:** Cada crença justificada recebe sua justificação de outras crenças na sua vizinhança epistêmica.

Apliquemos este raciocínio para o exemplo do chapéu que consideramos na Seção 3.1. Suponha que você observa novamente o chapéu de alguém e acredita que

(H) esse chapéu é azul.

Vamos assumir que (H) está justificada. De acordo com o coerentismo, (H) recebe a sua justificação de outras crenças na vizinhança epistêmica de (H). Elas constituem a sua prova ou

as suas razões para tomar (H) como verdade. Que crenças podem compor este conjunto de crenças circunvizinhas doadoras de justificação?

Vamos considerar duas abordagens para responder a esta pergunta. A primeira é conhecida como *inferência para a melhor explicação*. Tais inferências geram o que é chamado de *coerência explicativa*.<sup>40</sup> De acordo com essa abordagem, devemos supor que você forma uma opinião sobre a maneira como o chapéu aparece para você nas suas experiências perceptivas, e uma segunda crença no sentido de que sua experiência perceptiva que o chapéu parece azul a você é melhor explicada pela suposição de que (H) é verdadeiro. Assim, o conjunto relevante de crenças é o seguinte:

(1) Estou tendo uma experiência visual (E): o chapéu parece azul para mim.

(2) Minha experiência de (E) é melhor explicada assumindo que (H) é verdadeiro.

Existem, naturalmente, explicações alternativas de por que você tem (E). Talvez você esteja alucinando que o chapéu é azul. Talvez um gênio maligno faça o chapéu parecer azul para você quando na verdade ele é vermelho. Talvez você seja o tipo de pessoa a quem chapéus sempre parecem azuis. Um coerentista explicativo diria que, em comparação com estes, a azulidade do chapéu é uma explicação superior. É por isso que você está justificado em acreditar que (H). Note que um coerentista explicativo também pode explicar a *falta* de justificação. Suponha que você lembre que você tomou uma droga alucinante que faz as coisas parecerem azuis a você. Isso impediria você de estar justificado em acreditar que (H). O coerentista explicativo pode descrever isso apontando que, no caso que estamos considerando agora, a verdade de (H) não seria a *melhor* explicação de por que você está tendo a experiência (E). Em vez disso, ter tomado a droga alucinógena seria uma explicação pelo menos tão boa quanto a suposição de que (H) é verdadeira. É por isso que, de acordo com o coerentista explicativo, introduzindo essa variação no exemplo original, você não estaria justificado em acreditar que (H).

---

<sup>40</sup> Ver o capítulo 7 em Harman, 1986.

Um problema para os coerentistas explicativos é nos fazer entender, em termos *não-epistêmicos*, porque a explicação que defendem é realmente melhor do que as explicações concorrentes. Vamos usar a hipótese do gênio maligno para ilustrar essa dificuldade. O que precisamos é uma explicação de por que você está tendo (E). De acordo com a hipótese do gênio maligno, você está tendo (E) porque o gênio maligno o está enganando. O coerentista explicativo diria que esta é uma explicação ruim, porque você está tendo (E). Mas por que seria ruim? O que precisamos para responder a esta pergunta é uma consideração geral e baseada em princípios do que faz uma explicação melhor do que outra. Suponha que apelemos para o fato de você não estar *justificado* em acreditar na existência de gênios malignos. A ideia geral seria esta: Se há duas explicações concorrentes, E1 e E2, e E1 consiste em ou inclui uma proposição que você não está justificado em acreditar, enquanto E2 não, então E2 é melhor do que E1. O problema com esta ideia é que ela coloca a carroça na frente dos bois. Esperamos que o coerentismo explicativo seja capaz de fazer-nos compreender de onde vem a justificação. Ele não consegue fazer isso se explica a diferença entre explicações melhores e piores fazendo uso da diferença entre a crença justificada e injustificada. Se o coerentismo explicativo procedesse desta maneira, esta seria uma consideração circular, e, portanto, pouco informativa acerca da justificação. Portanto, o desafio a que o coerentismo explicativo deve responder é responder, sem usar o conceito de justificação, o que faz uma explicação melhor do que outra.

Vamos passar para a segunda maneira em que a abordagem coerentista pode ser formulada. Lembre-se o que a justificação de um sujeito para acreditar em  $p$  diz respeito a existência de uma ligação entre a crença de que  $p$  e a verdade de  $p$ . Suponha que o sujeito sabe que a origem de sua crença de que  $p$  é confiável. Então, ela sabe que as crenças provenientes dessa fonte tendem a ser verdadeiras. Esse conhecimento lhe daria uma excelente ligação entre a crença e a sua verdade. Podemos, dessa forma, dizer que as crenças circunvizinhas que conferem justificação a (H) são as seguintes:

- (1) Estou tendo uma experiência visual (E): o chapéu parece azul para mim.  
(3) Experiências como (E) são confiáveis.

Chamamos o coerentismo deste tipo de *coerentismo de confiabilidade*. Se você acredita em (1) e (3), você está de posse de uma boa razão para pensar que o chapéu é de fato azul. Assim, você está de posse de uma boa razão para pensar que a crença em questão, (H), é verdadeira. É por isso que, de acordo com o coerentismo de confiabilidade, você está justificado a crer que (H).

Assim como o coerentismo explicativo, esse ponto de vista (o coerentismo de confiabilidade) enfrenta um problema de circularidade. Se (H) recebe a sua justificação em parte porque você também acredita em (3), (3) em si deve estar justificado. Mas de onde viria sua justificativa para (3)? Uma resposta seria: a partir de sua memória perceptual de sucesso no passado. Você lembra que suas experiências visuais tiveram um histórico de bom desempenho. Elas raramente levaram a erro. O problema é que você não pode justificadamente atribuir um histórico de bom desempenho às suas faculdades perceptivas sem fazer uso de suas próprias faculdades perceptivas. Então, se o coerentismo de confiabilidade funcionasse teria que ser legítimo usar uma faculdade com o propósito de estabelecer a confiabilidade dessa mesma faculdade. Alguns epistemólogos pensam que isto não pode ser legítimo<sup>41</sup>.

Vimos que tanto o coerentismo explicativo, quanto o coerentismo de confiabilidade enfrentam seus problemas de circularidade específicos. Uma vez que ambos são versões do coerentismo *doxástico*, ambos enfrentam uma dificuldade adicional: as pessoas, em circunstâncias normais, realmente formam crenças como (1), (2), e (3)? Parece que não. Pode-se objetar, portanto, que estas duas versões do coerentismo fazem exigências intelectuais excessivas em questões ordinárias que tornam improvável ter as crenças de fundo que, de

<sup>41</sup> Assim Richard Fumerton diz o seguinte, no contexto do emprego do raciocínio circular com o propósito de rebater o ceticismo: “você não pode usar a percepção para justificar a confiabilidade da percepção! Você não pode usar a memória para justificar a confiabilidade da memória! Você não pode usar a indução para justificar a confiabilidade da indução! Tais tentativas de responder aos interesses do cético envolvem circularidade flagrante, de fato, patética”. Ver também Alston, 1993 para uma excelente discussão dos problemas envolvidos em argumentar a favor confiabilidade da percepção.

acordo com estas versões do coerentismo, são necessárias para a justificação. Esta objeção poderia ser evitada despojando o coerentismo do seu elemento doxástico. O resultado seria a seguinte versão do coerentismo, que resulta da rejeição da BE (a concepção epistêmica de basicidade):

**Coerentismo de Dependência:** Sempre que alguém está justificado em acreditar numa proposição  $p_1$ , a justificativa de alguém para acreditar em  $p_1$  depende da justificação para crer noutras proposições,  $p_1, p_2, \dots p_n$ .

Um coerentista explicativo poderia dizer que, para você estar justificado em acreditar que (H), não é necessário que você realmente *acredite* em (1) e (2). No entanto, é necessário que você tenha *justificação* para crer em (1) e (2). É a sua justificação para ter (1) e (2) que lhe dá justificação para crer que (H). Um coerentista de confiabilidade poderia assumir uma posição análoga. Ele poderia dizer que, para estar justificado em acreditar em (H), você não precisa acreditar em qualquer coisa sobre a confiabilidade de origem de sua crença. Você deve, entretanto, ter justificação para crer que a origem de sua crença é confiável, ou seja, você deve ter justificativa para (1) e (3). Ambas as versões do coerentismo de dependência, então, repousam na suposição de que é possível ter justificação para uma proposição sem realmente crer nessa proposição.

O coerentismo de dependência representa uma mudança significativa em relação à maneira pela qual o coerentismo foi tipicamente interpretado pelos seus defensores. De acordo com a formulação clássica do coerentismo, se uma determinada crença é justificada, o sujeito deve ter certas *crenças* que constituem razões para a crença dada. O coerentismo rejeita esta dependência. Segundo ele, a justificativa não precisa vir na forma de crenças. Ela pode vir em forma de evidências mnemônicas e introspectivas que dão justificação ao sujeito para crenças sobre coerência de confiabilidade e coerência explicativa. Na verdade, o

coerentismo de dependência permite a possibilidade de uma crença ser justificada não por receber *quaisquer* justificação de outras crenças, mas unicamente por experiências perceptivas e conteúdos da memória adequados. Acima, chamamos este ponto de vista de "solução de compromisso". A solução de compromisso, então, pode ser caracterizada da seguinte maneira:

- i. ela permite basicidade doxástica;
- ii. ela não permite basicidade epistêmica;
- iii. ela é inconsistente com o coerentismo doxástico;
- iv. ela qualifica-se como uma versão do coerentismo, a saber, do coerentismo de dependência.

Note que (iii) segue de (i), e (iv) de (ii). Um fundacionalista descomprometido rejeitaria o coerentismo de dependência. Um fundacionalista desse tipo vê uma crença básica de que *p* como uma crença cuja justificação não depende de ter *qualquer* justificação para crer noutra proposição *q*. O fundacionalismo deste tipo poderia ser chamado de *fundacionalismo de independência*, uma vez que afirma que a justificação de uma crença básica é completamente independente de ter uma justificação para quaisquer outras crenças. A lógica do conflito entre o fundacionalismo e o coerentismo parece sugerir que, em última análise, o conflito entre os dois pontos de vista se transforma no conflito entre o coerentismo de dependência e o fundacionalismo de independência<sup>42</sup>.

A seguir examinaremos as razões a favor e contra o fundacionalismo e o coerentismo.

---

<sup>42</sup> Para mais literatura defendendo várias versões do coerentismo, ver Bonjour, 1985, Elgin, 1996, 2005, Lehrer, 1990, Lycan, 1996.

### 3.3 Por que fundacionalismo?

O principal argumento para o fundacionalismo é chamado de *argumento regresso*. É um argumento de eliminação. Com relação a cada crença justificada,  $B_1$ , surge a questão de onde vem a justificação de  $B_1$ . Se  $B_1$  não é básico, ele teria de vir de outra crença,  $B_2$ . Mas  $B_2$  pode justificar  $B_1$  somente se  $B_2$  se auto-justifica. Se  $B_2$  é básico, a cadeia justificatória terminaria com  $B_2$ . Mas, se  $B_2$  não é básico, precisamos de uma crença a mais,  $B_3$ . Se  $B_3$  não é básico, precisamos de uma quarta crença, e assim por diante. A menos que a regressão que se seguiu termine em uma crença básica, temos duas possibilidades: a regressão ou fará um *loop de volta* a  $B_1$ , ou continuará *ad infinitum*. De acordo com o argumento regresso, qualquer uma dessas possibilidades são inaceitáveis. Portanto, se existem crenças justificadas, deve haver crenças básicas<sup>43</sup>.

Este argumento tem várias fraquezas. Em primeiro lugar, podemos perguntar se as alternativas ao fundacionalismo são realmente inaceitáveis. Na literatura recente sobre este assunto encontramos, de fato, uma defesa bem desenvolvida da posição segundo a qual o infinitismo é a solução correta para o problema do regresso<sup>44</sup>. E nem mesmo a circularidade deveria ser descartada tão rapidamente. A questão não é se um argumento simples da forma "*p, portanto, p*" é aceitável. Claro que não é! Antes disso, a questão consiste em determinar se, na tentativa de mostrar que a confiança em nossas faculdades é razoável, podemos fazer uso do *input* que nossas faculdades fornecem. É uma questão em aberto se tal espécie de circularidade é tão inaceitável como uma inferência "*p, portanto, p*". Além disso, evitar a circularidade não sai barato. Os fundacionalistas experimentais afirmam que a percepção é

<sup>43</sup> Há uma opção posterior: o regresso termina numa crença que não é justificada. É difícil de ver, contudo, como uma crença que não é justificada pode possivelmente justificar outras crenças.

<sup>44</sup> Podemos distinguir entre o *problema* do regresso e os vários *argumentos* do regresso. O problema do regresso é o problema de explicar como a justificação é possível dado que ela gera um aparente regresso infinito da justificação. Um argumento do regresso tem a pretensão de suportar uma solução particular para o problema do regresso tendo por base de uma rejeição das soluções competidoras. Assim, um argumento do regresso para o fundacionalismo rejeita o coerentismo e o infinitismo como opções viáveis. Klein, contudo, rejeita tanto o fundacionalismo e o coerentismo e argumenta pelo infinitismo. (Ver Klein, 1999 e 2005, Ver Ginet, 2005 para uma resposta à defesa de Klein do infinitismo). Do mesmo modo, coerentistas podem argumentar que nem o fundacionalismo e nem o infinitismo são opções viáveis.

uma fonte de justificação. Por isso eles precisam responder à questão-J: *por que* a percepção é uma fonte de justificação? Como vimos acima, se quisermos responder a essa pergunta sem nos comprometermos com o tipo de circularidade que coerentismo de dependência envolve, temos de escolher entre o externalismo e um apelo à necessidade bruta. Nenhuma escolha é isenta de problemas.

O segundo ponto fraco do argumento do regresso é que sua conclusão só diz o seguinte: se há crenças justificadas, deve haver crenças justificadas que não recebem a sua justificação de outras crenças. Sua conclusão não diz que, se há crenças justificadas, deve haver crenças cuja justificação é independente de qualquer justificativa para outras crenças. Assim, o argumento do regresso, se fosse válido (*sound*), iria apenas mostrar que deve haver basicidade doxástica. O coerentismo de dependência, no entanto, permite basicidade *doxástica*. Assim, o argumento do regresso apenas defende o fundacionalismo experiencial contra o coerentismo doxástico. Ele não nos diz por que devemos preferir o fundacionalismo de independência ao coerentismo de dependência.

O fundacionalismo experiencial pode ser apoiado citando casos como o exemplo do chapéu azul. Tais exemplos tornam plausível assumir que as experiências perceptivas são uma fonte de justificação. Mas eles não permitem decidir entre o coerentismo de dependência e o fundacionalismo de independência, uma vez que qualquer uma dessas visões apela para experiências perceptivas para explicar por que as crenças perceptivas são justificadas.

Finalmente, o fundacionismo pode obter sustentação através da formulação de objeções ao coerentismo. Uma objeção importante é que de alguma forma o coerentismo não consegue garantir que um sistema de crença justificada esteja em contato com a realidade. Esta objeção retira sua força do fato de que a ficção pode ser perfeitamente coerente. Por que pensar, portanto, que a coerência de um sistema de crença é uma razão para pensar que a crença nesse sistema tende a ser verdadeira? Os coerentistas poderiam responder a essa objeção dizendo que, se um sistema de crenças contém crenças como "Muitas das minhas convicções têm sua origem em experiências perceptivas" e "Minhas experiências perceptivas

são confiáveis", é razoável para o sujeito pensar que o seu sistema de crença o coloca em contato com a realidade externa. Isto parece uma resposta eficaz à objeção do não-contato-com-a-realidade. Além disso, não é fácil ver por que o fundacionalismo em si deve ser mais bem posicionado do que o coerentismo quando o contato com a realidade é a questão. O que se entende por "garantir" o contato com a realidade? Se os fundacionalistas esperam uma *garantia lógica* de tal contato, crenças básicas devem ser infalíveis. Isso faria o contato com a realidade uma comodidade bastante cara. Dado o seu preço, os fundacionalistas podem querer diminuir suas expectativas. De acordo com uma interpretação alternativa, esperamos apenas a *probabilidade* de contacto com a realidade. Mas se os coerentistas contam com a importância da percepção de uma maneira ou de outra, eles podem atender a essa expectativa, assim como os fundacionalistas.

Dado que o coerentismo pode ser interpretado de maneiras diferentes, é improvável que haja uma única objeção bem-sucedida para refutar todas as versões possíveis do coerentismo. O coerentismo doxástico, no entanto, parece ser particularmente vulnerável a críticas que vem do campo fundacionalista. Uma delas já foi considerada: parece que o coerentismo doxástico faz excessivas exigências intelectuais para os sujeitos que detém crenças. Ao lidar com as tarefas mundanas da vida diária, normalmente não nos preocupamos em formar crenças sobre a coerência explicativa de nossas crenças ou a confiabilidade das fontes de nossas crenças. De acordo com uma segunda objeção, o coerentismo doxástico fracassa por ser insensível à relevância epistêmica de experiências perceptivas. Os fundacionalistas poderiam argumentar da seguinte forma. Suponha que Kim está a observar um camaleão que muda rapidamente suas cores. Uns minutos atrás ele era azul, agora ele está roxo. Kim ainda acredita que ele é azul. Sua crença é agora injustificada porque ela acredita que o camaleão é azul mesmo que *pareça* roxo. Em seguida, o camaleão muda de cor novamente para o azul. Agora a crença de Kim de que o camaleão é azul está novamente justificada porque o camaleão, mais uma vez, *parece* azul para ela. O ponto aqui seria que o que foi responsável por mudar o status justificatório da crença de Kim foi unicamente o modo como o camaleão *parecia* para ela. Dado que o coerentismo doxástico não atribui relevância

epistêmica à experiências perceptivas em si mesmas, ele não pode explicar por que a crença de Kim estava inicialmente justificada e, em seguida, passou a estar injustificada, e, finalmente, voltou a ser justificada<sup>45</sup>.

### 3.4 Por que coerentismo?

O coerentismo normalmente é defendido por atacar o fundacionalismo como uma alternativa viável. Para argumentar contra o fundacionalismo de privilégio, os coerentistas escolhem um privilégio epistêmico que eles acreditam ser essencial para o fundacionalismo, e argumentam que ou nenhuma crença, ou poucas crenças desfrutam de tal privilégio. Contra o fundacionalismo experiencial, diferentes objeções foram lançadas. Uma linhagem de crítica afirma que experiências perceptivas não têm conteúdo proposicional. Portanto, a relação entre uma crença perceptiva e a experiência perceptiva que dá origem a ela só pode ser causal. Entretanto, considere novamente o exemplo do chapéu. Quando você vê o chapéu e ele parece azul para você, a sua experiência visual – que está parecendo azul para você – não tem o conteúdo proposicional *que o chapéu é azul*? Parece que sim. Se for assim, não parece haver nenhuma razão para negar que a sua experiência perceptiva pode desempenhar uma função justificadora.<sup>46</sup>

Outra linhagem de pensamento é que, se experiências perceptivas têm conteúdo proposicional, elas não podem parar o regresso justificatório porque elas mesmas poderiam precisar de justificção. Isso, no entanto, parece ser um pensamento estranho. Em nossa

---

<sup>45</sup> Coerentistas doxásticos poderiam responder que, quando o camaleão muda sua cor para lilás, Kim forma a crença que o camaleão parece lilás para ele. Por causa dessa crença, ele não pode estar justificado em ainda crer que o camaleão é azul. Portanto, o coerentismo doxástico pode explicar no final por que a crença de Kim (o camaleão é azul) é injustificada depois que o camaleão mudou sua cor para lilás. O problema com essa réplica é que os fundacionalistas são livres para descrever o exemplo em qualquer modo que eles quiserem (à medida que ele permanecer concebível). E, obviamente, eles poderiam descrevê-lo por estipular que Kim *não* tem qualquer crença sobre como o camaleão aparece a ele. Em resposta a isso, o coerentista doxástico poderia dizer que a falha de Kim para formar crenças sobre como o camaleão aparece a ele é inconcebível. Essa alegação, contudo, não parece, em si mesma, uma alegação plausível.

<sup>46</sup> Para literatura sobre essa questão, ver Brewer, 1999, Pryor, 2000, 2005, Sellares, 1963, Steup, 2001c, Williams, 2005, e o debate entre Bill Brewer e Alex Byrne em Steup e Sosa, 2005.

prática epistêmica real, nunca exigimos dos outros que justifiquem o modo pelo qual as coisas parecem a eles em suas experiências perceptivas. Na verdade, tal demanda parece absurda. Suponha que pergunto: "Por que você acha que o chapéu é azul" Você responde: "Porque ele parece azul para mim." Há outras questões sensíveis que eu poderia levantar nessa altura. Eu poderia, por exemplo, perguntar: "Por que parecer azul dá a você uma razão para pensar que é azul?" Ou eu poderia perguntar: "Você não poderia estar enganado em acreditar que parece azul para você?" A última pergunta poderia irritá-lo, mas não seria ilegítima. Afinal de contas, nós podemos razoavelmente duvidar que as crenças introspectivas sobre como as coisas aparecem para nós sejam infalíveis. Mas agora suponha que eu lhe pergunto: "Por que você acha que a experiência perceptiva que o chapéu parece azul a você está justificada?". Em resposta a essa pergunta, você poderia acusar-me de mau uso da palavra 'justificação'. Eu poderia muito bem perguntar-lhe o que é que justifica a sua dor de cabeça quando você tem uma, ou o que justifica a coceira no nariz quando você tem uma. As últimas perguntas, você poderia responder, seriam tão absurdas quanto o meu pedido para dar uma razão justificada para a sua experiência perceptiva.<sup>47</sup>

O fundacionalismo experiencial, então, não é facilmente desmontável. Com que base os coerentistas poderiam objetar a ele? Para lançar problemas ao fundacionalismo experiencial, os coerentistas poderiam pressionar a questão-J: por que as experiências perceptivas são uma fonte de justificação? Se os fundacionalistas responderem à questão-J apelando para evidências que garantam a atribuição de confiabilidade às experiências perceptivas, o fundacionalismo experiencial transformar-se-ia em coerentismo de dependência, ou, como o chamamos, a solução de compromisso. Para evitar esse resultado, os fundacionalistas teriam que dar uma resposta alternativa. Uma maneira de fazê-lo seria defender a independência do

---

<sup>47</sup> Poderia ser argumentado que, por atribuir conteúdo proposicional para experiências perceptivas, nós assim as transformamos em estados mentais que são suficientemente *semelhantes a crenças* para serem crenças neste aspecto: elas podem justificar apenas se elas estão elas mesmas justificadas. Essa alegação é mais facilmente feita do que defendida. Se o chapéu parece azul a você, então sua experiência perceptiva apresenta um certo conteúdo proposicional a você, a saber, *que o chapéu é azul*. Não obstante, muito embora tenha esse conteúdo, *é distinto* da crença que o chapéu é azul. Por que? Obviamente, porque é possível que o chapéu parece azul embora você não acredite, por exemplo, que o chapéu é branco e parece azul apenas por que você está usando óculos pintados de azul. Nesse caso, o chapéu poderia parecer azul a você sem que você acredite que parece azul. Não é fácil ver, portanto, em que sentido a posse de conteúdo proposicional pode fazer as experiências perceptuais 'semelhantes a crenças'.

fundacionalismo, o qual adota a concepção epistêmica de basicidade e considera uma questão de necessidade bruta que percepção seja uma fonte de justificação. Então, em última instância, a tarefa de defender o coerentismo pode descer para a tarefa de mostrar que o coerentismo de dependência como uma posição de compromisso é preferível ao fundacionalismo de independência. Para contribuir com tal preferência, poderia ser argumentado que o coerentismo de dependência nos dá uma resposta mais satisfatória para a questão-J do que o fundacionalismo de independência. Mas isso é realmente assim?

Suponha que nós pedimos "Por que a soma de dois mais dois é quatro?" A resposta "Não poderia ser outra coisa" não é perfeitamente satisfatória? Então, às vezes, pelo menos, um pedido para explicar a verdade de  $p$  é cumprido de forma satisfatória salientando que  $p$  é *necessariamente* verdadeiro. Por que, então, não deveríamos estar satisfeitos quando os fundacionalistas de independência respondem à questão-J, dizendo que experiências perceptivas são *necessariamente* uma fonte de justificação? Para saber se devemos ficar satisfeitos, podemos empregar experimentos mentais. Tentaremos descrever um mundo possível em que, para usar o nosso exemplo de novo, alguém vê um objeto que parece azul para ela, mas o objeto que parece azul não lhe dá qualquer justificação para crer que o objeto é realmente azul. Se pudermos conceber um mundo possível como esse, então nós temos razão para pensar que os fundacionalistas de independência estão enganados quando dizem que a experiência perceptiva é necessariamente uma fonte de justificação.

#### **4. Fontes do Conhecimento e da Justificação**

Crenças surgem nas pessoas por uma grande variedade de causas. Dentre elas, devemos listar fatores psicológicos tais como desejos, necessidades emocionais, preconceitos e vieses de vários tipos. Obviamente, quando crenças se originam de fontes como essas, elas não se qualificam como conhecimento mesmo se verdadeiras. Para que crenças verdadeiras

contem como conhecimento é necessário que se originem de fontes que temos boas razões para considerar confiáveis. Essas fontes são a percepção, a introspecção, a memória, a razão e o testemunho. Consideremos brevemente cada uma delas.

#### **4.1 Percepção**

Nossas faculdades perceptivas são nossos cinco sentidos: visão, tato, audição, olfato e paladar. Nós devemos distinguir entre uma experiência que pode ser classificada como *perceber* que  $p$  (por exemplo, ver que há café no copo e sentir que é doce), que implica que  $p$  é verdadeiro, e uma experiência perceptiva na qual nos parece como se fosse o caso que  $p$ , mas onde  $p$  pode ser falso. Refiramo-nos a esse último tipo de experiência como *aparências perceptivas*. A razão para se fazer essa distinção se encontra no fato de que a experiência perceptiva é falível. O mundo não é sempre como nos parece em nossas experiências perceptivas. Precisamos, portanto, de um modo para nos referirmos a experiências perceptivas em que  $p$  parece ser o caso que tolere a possibilidade de  $p$  ser falso. É essa a função atribuída às aparências perceptivas, de modo que algumas aparências perceptivas que  $p$  são casos de perceber que  $p$ , e outras não. Quando lhe parece que existe uma xícara de café sobre a mesa e de fato existe, os dois estados coincidem. Se, porém, você alucina que existe uma xícara de café sobre a mesa, você tem uma aparência perceptiva que  $p$  sem perceber que  $p$ .

Uma série de questões epistemológicas sobre percepção surge quando nos preocupamos com a natureza psicológica dos processos perceptivos através dos quais adquirimos conhecimento de objetos externos. Segundo o *realismo direto*, nós adquirimos tal conhecimento porque podemos perceber diretamente tais objetos. Por exemplo, quando você vê um tomate sobre a mesa, *o que você percebe* é o próprio tomate. De acordo com o *realismo indireto*, nós adquirimos conhecimento de objetos externos em virtude de perceber outra

coisa, a saber, aparências ou dados sensíveis. Um realista indireto diria que, quando você vê e portanto sabe que existe um tomate sobre a mesa, o que você realmente vê não é o próprio tomate, mas um dado sensível semelhante a um tomate ou alguma outra entidade dessa natureza.

Realistas diretos e indiretos possuem diferentes concepções sobre a estrutura do conhecimento perceptivo. Realistas indiretos diriam que adquirimos conhecimento perceptivo de objetos externos em virtude de perceber dados sensíveis que representam objetos externos. Dados sensíveis, que são uma espécie de estado mental, desfrutam de um status especial: sabemos diretamente com o que eles se parecem. Assim, realistas indiretos pensam que, quando o conhecimento perceptivo é fundacional, ele é o conhecimento de dados sensíveis e outros estados mentais. O conhecimento de objetos externos é indireto; ele deriva do nosso conhecimento de dados sensíveis. A ideia básica é que temos conhecimento indireto do mundo exterior porque podemos ter conhecimento fundacional de nossa própria mente. Realistas diretos podem ser mais liberais sobre a fundação do nosso conhecimento de objetos externos. Como eles sustentam que experiências perceptivas proporcionam contato direto com objetos externos, eles podem dizer que tais experiências podem dar conhecimento fundacional de objetos externos.

Nós consideramos nossas faculdades perceptivas confiáveis. Mas como podemos saber que elas são confiáveis? Para os externalistas, isso não parece ser um problema. Se o uso das faculdades confiáveis é suficiente para o conhecimento, e se usando faculdades confiáveis nós adquirimos a crença de que nossas faculdades são confiáveis, então reconhecemos que nossas faculdades são confiáveis. Mas mesmo externalistas podem se perguntar como eles podem, argumentativamente, *mostrar* que nossas faculdades perceptivas são confiáveis. O problema é esse. Parece que a única maneira de adquirir conhecimento sobre nossas faculdades perceptivas é através da memória através do processo de lembrar se elas nos serviram bem no passado. Mas devo confiar na minha memória e devo pensar que os episódios de sucessos perceptivos de que pareço me lembrar foram de fato episódios de sucesso perceptivo? Se eu

tenho o direito de responder a essas questões com “sim”, então eu preciso ter, para começar, razões para considerar minha memória e minhas experiências perceptivas confiáveis. Pareceria, portanto, que não existe um modo não circular de argumentar pela confiabilidade das faculdades perceptivas de alguém<sup>48</sup>.

## 4.2 Introspecção

Introspecção é a capacidade de inspecionar, metaforicamente falando, o “interior” da mente. Através da introspecção, sabe-se em quais estados mentais se está: se se está sedento, cansado, excitado ou deprimido. Comparada com a percepção, a introspecção parece ter um status especial. É fácil de ver como uma aparência perceptiva pode dar errado: o que parece ser uma xícara de café sobre a mesa pode simplesmente ser um hábil holograma visualmente indistinguível de uma verdadeira xícara de café. Mas poderia ser possível que introspectivamente me pareça que eu estou com dor de cabeça quando, de fato, não estou? Não é fácil ver como isso poderia acontecer. Assim concluímos que a introspecção possui um status especial. Comparada com a percepção, a introspecção parece ser privilegiada em virtude de ser menos passível de erro. Como podemos explicar o status especial da introspecção?

Primeiro, se poderia argumentar que, em se tratando de introspecção, não existe diferença entre aparência e realidade; portanto, aparências introspectivas são necessariamente sucessos introspectivos. Conforme essa perspectiva, a introspecção é infalível. Alternativamente, alguém poderia ver a introspecção como uma fonte de certeza. A ideia aqui é que uma experiência introspectiva de  $p$  elimina toda dúvida possível sobre se  $p$  é verdadeiro. Finalmente, alguém poderia tentar explicar o lugar especial da introspecção

---

<sup>48</sup> Para literatura sobre problemas epistêmicos da percepção, ver Alston, 1999, e capítulos 10 e 11 de Dancy, 1985. Mais referências bibliográficas podem ser encontradas na página 442 em Greco e Sosa, 1999. Ver também o artigo de Bonjour “Epistemological Problems of Perception” e o artigo de Crane “The Problema of Perception, na Stanford Encyclopedia of Philosophy, indicada no final desse artigo.

examinando como respondemos a relatos de primeira pessoa: tipicamente, atribuímos uma autoridade especial a tais relatos. De acordo com essa perspectiva, a introspecção é incorrigível. Os outros não estão, ou ao menos não tipicamente, em posição de corrigir os relatos de primeira pessoa dos estados mentais de alguém.

A introspecção revela como o mundo nos parece em nossas experiências perceptivas. Por essa razão, a introspecção tem recebido uma atenção especial por parte dos fundacionalistas. A percepção não é imune ao erro. Se a certeza consiste na ausência de toda dúvida possível, a percepção falha ao gerar tal espécie de certeza. Daí que crenças baseadas em experiências perceptivas não podem ser fundacionais. A introspecção, porém, pode fornecer o que precisamos para encontrar uma fundação firme para nossas crenças sobre objetos externos: na melhor das hipóteses a imunidade absoluta de erro ou a qualquer dúvida possível, ou, talvez mais modestamente, um tipo de imediaticidade epistêmica que não pode ser encontrada na percepção.

É realmente verdade, porém, que, comparada com a percepção, a introspecção é, em algum sentido, especial? Críticos do fundacionalismo argumentaram que a introspecção certamente não é infalível. Não poderia alguém confundir uma coceira desagradável com uma dor? Eu não posso pensar que uma forma diante de mim me parece circular quando de fato ela me parece ligeiramente elíptica? Se é realmente possível à introspecção induzir a erro, então é difícil ver por que a introspecção seria capaz de eliminar toda a dúvida possível. No entanto, não é fácil ver como, ao sentir clara e distintamente uma latejante dor de cabeça, alguém poderia se enganar sobre isso. A introspecção, então, acaba por ser uma faculdade misteriosa. Por um lado, ela não parece ser, no geral, uma faculdade infalível; por outro lado, quando se olha para casos específicos apropriadamente descritos, o erro parece impossível<sup>49</sup>.

---

<sup>49</sup> Para um artigo introdutório e referências bibliográficas, ver o artigo Brie Gertler "Self-Knowledge" na Stanford Encyclopedia of Philosophy cujo link consta no final desse artigo.

### **4.3 Memória**

Memória é a capacidade de reter conhecimento adquirido no passado. O que alguém lembra, porém, não precisa ser um evento passado. Pode ser um fato presente, como um número de telefone, ou um evento futuro, como a data das próximas eleições. A memória é, claro, falível. Nem todo caso em que assumimos que lembramos que  $p$  é realmente um caso em que lembramos que  $p$ . Nós devemos distinguir, portanto, entre lembrar que  $p$  (que implica na verdade de  $p$ ) e *parecer* lembrar que  $p$  (que não implica na verdade de  $p$ ).

Um problema sobre a memória diz respeito à questão do que distingue as aparências memorativas das aparências perceptivas ou a mera imaginação. Alguns filósofos pensaram que ter uma imagem em mente fosse essencial para a memória, mas isso parece ser equivocado. Quando alguém se lembra de um número de telefone é pouco provável que tenha a imagem de um número na mente. As questões distintamente epistemológicas sobre a memória são essas: Primeiro, o que faz das aparências memorativas uma fonte de justificação? É uma verdade necessária que, se alguém tem uma aparência memorativa que  $p$ , tem por isso uma justificação *prima facie* para  $p$ ? Ou seria a memória uma fonte de justificação somente se, como os coerentistas poderiam dizer, alguém tiver razão para pensar que sua memória é confiável? Ou seria a memória uma fonte de justificação somente se, como os externalistas diriam, ela for de fato confiável? Segundo, como podemos responder ao ceticismo sobre o conhecimento do passado? Aparências memorativas do passado não garantem que o passado seja o que pensamos que é. Nós pensamos que temos um pouco mais de idade do que cinco minutos, mas é logicamente possível que o mundo, com nossas disposições a ter aparências memorativas de um passado mais distante e com itens como fósseis que aparentemente sugerem a existência de um passado de milhões de anos, tenha vindo à existir apenas a cinco minutos atrás. Nossa aparente lembrança que o mundo é mais velho do que meros cinco minutos não implica, portanto, que ele realmente o seja. Por que,

então, deveríamos pensar que a memória é uma fonte de conhecimento sobre o passado?<sup>50</sup>  
[49]

#### 4.4 Razão

Algumas crenças parecem ser justificadas somente pelo uso da razão. Uma justificação desse tipo é dita *a priori*: anterior a qualquer tipo de experiência. Um modo padrão de definir justificação *a priori* é como segue:

**Justificação A Priori:** *S* está justificado *a priori* em crer que *p* se e somente se a justificação de *S* para crer que *p* não depende de qualquer experiência.

Crenças que são verdadeiras e justificadas desse modo (e não de alguma forma “gettierizada”) contariam como casos de conhecimento *a priori*.<sup>51</sup> [50]

O que exatamente conta como experiência? Se por 'experiência' nos referimos somente a experiências *perceptivas*, a justificação derivada de experiências introspectivas ou memorativas contariam como *a priori*. Por exemplo, eu poderia então saber *a priori* que estou sedento, ou o que eu comi no café da manhã essa manhã. Embora termo '*a priori*' seja algumas vezes utilizado dessa maneira, o uso estrito do termo restringe justificação *a priori* à justificação derivada *unicamente* do uso da razão. Conforme esse uso, a palavra 'experiência' na definição acima inclui experiências perceptivas, introspectivas e memorativas. Nessa

<sup>50</sup> Para um artigo introdutório e referências bibliográficas, veja o artigo de Tom Senor “Epistemological Problems of Memory” na Stanford Encyclopedia of Philosophy, indicado no final desse artigo.

<sup>51</sup> Não há escapatória dos problemas de Gettier mesmo na área da justificação *a priori*. O que poderia ser um exemplo de crença verdadeira que é justificada *a priori*, mas não é, não obstante isso, um exemplo de conhecimento? Suponha que Carl é um lógico. Ele está tentando provar que *p* (que nós assumimos ser, de antemão, uma proposição complicada) e uma verdade necessária. Ele desenvolve uma longa e complexa prova e conclui que *p* é realmente necessariamente verdadeira. Infelizmente, muito embora Carl esteja certo, ele fez um erro pequeno e muito sutil tão difícil de apontar que deixa a justificação de Carl intacta. Parece que poderíamos julgar que, em virtude de seu erro, que Carl não sabe que *p* é uma verdade necessária. Assim, a crença de Carl que *p* é verdade necessária é uma crença verdadeira e justificada que não é conhecimento.

compreensão mais estreita, exemplos paradigmáticos do que posso saber tendo como base uma justificação *a priori* são verdades conceituais (tais como “todos os solteiros são não-casados”), e verdades da matemática, da geometria e da lógica.

Justificação e conhecimento que não sejam *a priori* são chamados “a posteriori” ou “empíricos”. Por exemplo, no sentido estrito de '*a priori*', se eu estou sedento ou não é algo que sei empiricamente (com base em experiências introspectivas), enquanto eu sei *a priori* que 12 dividido por 3 é 4.

Várias questões importantes surgem sobre o conhecimento *a priori*. Primeiro, ele realmente existe? Céticos sobre a aprioridade negam sua existência. Eles não querem dizer que não tenhamos conhecimento sobre matemática, geometria, lógica e verdades conceituais. Ao contrário, eles alegam que todo esse conhecimento é empírico.

Segundo, se justificação *a priori* é possível, como ela surge exatamente? O que *faz* com que uma crença tal como “Todos os solteiros são não-casados” seja justificada unicamente com base na razão? Existiria uma apreensão imediata da verdade dessa proposição? Ou consistiria isso na apreensão de que a proposição é *necessariamente* verdadeira? Ou seria a experiência puramente intelectual de “ver” (com o “olho da razão”) ou “intuir” que a proposição é verdadeira (ou necessariamente verdadeira)? Ou seria, como sugeririam os externalistas, a confiabilidade do processo cognitivo pelo qual nós viemos a reconhecer a verdade de tal proposição?

Terceiro, se conhecimento *a priori* existe, qual a sua extensão? *Empiristas* argumentaram que conhecimento *a priori* é limitado ao reino da *análise*, consistindo em proposições de um status de alguma forma inferior, pois elas não são realmente “sobre o mundo”. Proposições de um status superior, que transmitem informação genuína sobre o mundo, são classificadas como  *sintéticas*. Conhecimento *a priori* de proposições sintéticas, diriam os empiristas, não é possível. *Racionalistas* negam isso. Eles diriam que uma proposição tal como “Se uma bola é totalmente verde, então ela não possui pontos pretos” é sintética e cognoscível *a priori*.

Uma quarta questão sobre a natureza do conhecimento *a priori* se refere à distinção entre verdades necessárias e contingentes. A concepção que herdamos é a de que qualquer coisa que seja conhecida *a priori* é necessariamente verdadeira, mas existem epistemólogos que discordam disso<sup>52</sup>.

#### 4.5 Testemunho

O testemunho se diferencia das fontes consideradas acima porque ele não se distingue por ter sua própria faculdade cognitiva. Ao contrário, adquirir conhecimento de que *p* através do testemunho é vir a saber que *p* com base a declaração de alguém que *p*. “Declarar que *p*” deve ser entendido de modo amplo, incluindo expressões comum da vida quotidiana, postagens de bloggers em seus blogs, artigos de jornalistas, informações na televisão, rádio, fitas, livros e outras mídias. Então, quando você pergunta à pessoa ao seu lado que horas são, e ela responde, e você assim vem a saber que horas são, esse é um exemplo de vir a saber algo com base no testemunho. E quando você descobre ao ler o *Washington Post* que o ataque terrorista em Sharm el-Sheikh de 22 de julho de 2005 matou pelo menos 88 pessoas, isso, também, é um exemplo de aquisição de conhecimento pelo testemunho.

O *quebra-cabeças* epistemológico que o testemunho traz é este: Por que o testemunho é uma fonte de conhecimento? Um externalista poderia dizer que o testemunho é uma fonte de conhecimento se e somente se ele provém de uma fonte confiável. Mas aqui, ainda mais do que no caso das nossas faculdades, internalistas não acharão a resposta satisfatória. Suponha que você ouça alguém dizer ' *p*'. Suponha, além disso, que essa pessoa é de fato completamente confiável em relação à questão de *p* ser o caso ou não. Finalmente, suponha que você não possui nenhuma evidência acerca da confiabilidade dessa pessoa. Não seria plausível concluir que, já que a confiabilidade dessa pessoa é desconhecida para você, que o fato daquela pessoa declarar '*p*' não te deixa em posição de saber que *p*? Mas se a confiabilidade de uma fonte

---

<sup>52</sup> Acerca da literatura sobre conhecimento *a priori*, ver BonJour, 1998, 2005, Boghossian e Peacocke, 2000, Casullo, 2003 e Devitt, 2005.

testemunhal não é suficiente para fazer dela uma fonte de conhecimento, o que mais é necessário? Thomas Reid sugeriu que, pela nossa própria natureza, nós aceitamos fontes testemunhais como confiáveis e tendemos a atribuir credibilidade a elas a menos que encontremos razões especiais para não fazê-lo. Mas essa é meramente a afirmação da atitude que de fato tomamos acerca do testemunho. O que faz essa atitude razoável? Poder-se-ia argumentar que, nas próprias experiências pessoais com fontes testemunhais, acumula-se um longo histórico que pode ser usado como um sinal de confiabilidade. Entretanto, quando pensamos na amplitude do conhecimento que derivamos do testemunho, pode-se perguntar se as experiências pessoais constituem uma base evidencial rica o bastante para justificar a atribuição de confiabilidade à totalidade das fontes testemunhais em que se tende a acreditar. Uma alternativa para a abordagem histórica seria declarar que uma verdade necessária que confiar em fontes testemunhais é justificado. Essa sugestão, infelizmente, encontra a mesma dificuldade que a abordagem externalista para o testemunho: não parece que podemos adquirir conhecimento de fontes cuja confiabilidade nos seja completamente desconhecida<sup>53</sup>.

## **5. Os limites do conhecimento e da justificação**

### **5.1 O apelo do ceticismo (The case for Skepticism)**

De acordo com os cétricos, os limites do que você sabe são mais estreitos do que você gostaria de imaginar. Existem muitas coisas que você pensa que sabe, mas que, de fato, você não sabe. Por exemplo, você pensa que sabe que você tem mãos, mas, de fato, você não sabe. Como os cétricos esperam que você leve a sério uma conclusão tão estranha? Assim: primeiramente, o cético focará numa outra proposição, sobre a qual você provavelmente concordaria que não sabe. Num segundo passo, eles o fariam concordar que, se você não sabe essa segunda proposição, então você também não sabe a primeira: a proposição que você tem

---

<sup>53</sup> Para mais literatura acerca dessa questão, ver Lackey, 2003. Esse artigo contém referências bibliográficas extensas e está disponível online.

mãos. Quando os céticos apresentam seu argumento com alguma outra proposição sobre a qual você provavelmente concordaria que não sabe, o que eles têm em mente? Eles pedem sua atenção para o que é chamado de *uma hipótese cética*. De acordo com uma hipótese cética, as coisas são radicalmente diferentes do que elas parecem ser. Aqui vão vários exemplos:

- Eu estou sonhando deitado na minha cama.
- Eu estou sendo enganado por um gênio maligno.
- Eu sou um cérebro numa cuba (BIV).
- Eu estou no mundo de *matrix*.

O que os céticos ressaltam e o que eles pensam que você poderia facilmente concordar é isso: para cada uma das hipóteses particulares dessa lista, você não sabe que ela é falsa. Isso funciona melhor para algumas do que outras. Dá realmente certo para a hipótese do cérebro numa cuba (BIV) que discutiremos em seguida na seção 2.2. A idéia é que, se você é um BIV, você se reduz a um mero cérebro que é estimulado de tal maneira a gerar a ilusão de que você tem uma vida normal. Assim, as experiências que você tem como BIV e as experiências que você tem como uma pessoa normal são perfeitamente semelhantes, indistinguíveis, por assim dizer, “a partir de dentro”. Não lhe *parecerá* que você é um BIV. Afinal de contas, você pode ver que você tem um corpo, e que pode movê-lo livremente em seu ambiente. O problema é que será a mesma coisa para um BIV. O resultado disso é que a evidência que você tem como uma pessoa normal e a evidência que você tem como um BIV não difere de maneira relevante. Consequentemente, sua evidência não pode resolver a questão de se você é ou não um BIV. Tendo como base esse pensamento, os céticos afirmam que você *não sabe* que você não é um BIV. Esse é o primeiro passo para o apelo do ceticismo.

Vamos agora nos focar no segundo passo. O pensamento básico é que, se você não sabe que você não é um BIV, você não sabe que você tem mãos. Esse pensamento é extremamente plausível. Afinal, se você for um BIV, você não tem mãos. Assim, se você não pode distinguir entre ser ou não ser um BIV, você não pode distinguir entre ter ou não ter mãos. Mas se você

não pode distinguir entre ter e não ter mãos, realmente você não sabe que você tem mãos. Colocando os dois passos do raciocínio cético juntos, nós obtemos o seguinte argumento:

O argumento BIV

(a) Eu não sei que eu não sou um BIV

(b) Se eu não sei que eu não sou um BIV, então eu não sei que eu tenho mãos.

Portanto

(c) Eu não sei que eu tenho mãos.

Como já vimos, (1) e (2) são premissas muito plausíveis. Poderia parecer, portanto, que o BIV é um argumento cogente. Se for assim, devemos concluir que não sabemos que temos mãos. Mas realmente essa conclusão não pode ser correta. Assim, nós somos colocados diante de uma questão difícil: sob quais fundamentos podemos rejeitar a conclusão de um argumento aparentemente plausível?<sup>54</sup>

## 5.2 Ceticismo e Fechamento

A segunda premissa está intimamente ligada ao princípio de que o conhecimento é fechado (closed) pela implicação conhecida ou, para encurtar, o *princípio do fechamento*. Deixando as complicações de lado, esse princípio diz o seguinte:

**O princípio do fechamento** Se sei que  $p$ , e sei que  $p$  implica  $q$ , então sei que  $q$ <sup>55</sup>.

---

<sup>54</sup> Para uma seleção da literatura sobre ceticismo, o capítulo 1 em Dancy, 1986, ver os capítulos 6 e 7 em Feldman, 2003, o capítulo 10 em Steup, 1996, Stroud, 1984 e Williams, 1999. Ver também a introdução de DeRose para, bem como os artigos em, DeRose e Warfield, 1999, e o debate entre Vogel e Fumerton em Steup e Sosa, 2005. Para mais literatura recente sobre o ceticismo e como responder a ele, ver Fumerton, 1995, Greco, 2000, Huemer, 2000 e Pryor, 2000.

<sup>55</sup> Uma complicação surge do fato que, sabendo  $p$  e que  $p$  implica  $q$ , eu posso, não obstante, simplesmente não me preocupar em formar a crença que  $p$ . Assim, um modo mais cuidadoso de colocar o princípio diria no conseqüente: então estou em condições de saber que  $q$ . Uma articulação cuidadosa do princípio exige a consideração de complicações posteriores. Para uma discussão sistemática das dificuldades de encontrar uma versão aceitável do princípio do fechamento, ver Hawthorne, 2005.

Esse princípio é muitíssimo plausível. Segue-se um exemplo que permite ilustrá-lo. Suponha que você tem exatamente duas cervejas. Ter exatamente duas cervejas implica que você tem menos do que três cervejas. Se você sabe essas duas coisas, então você sabe que você tem menos do que três cervejas. Até esse ponto seguramente não há nada que pareça discutível. Como o princípio do fechamento se vincula ao argumento cético? A conexão pode ser vista quando você substitui ‘*p*’ e ‘*q*’ pelas proposições relevantes:

*p*: eu tenho mãos

*q*: eu não sou um BIV

Fazendo essas substituições, nós obtemos a seguinte aplicação do princípio do fechamento ao argumento do BIV:

**Fechamento em BIV** Se sei que tenho mãos e sei que ter mãos implica não ser um BIV, então sei que não sou um BIV.

De acordo com o argumento cético, você não pode saber que não é um BIV. Assim, o conseqüente do argumento do fechamento em BIV é falso. Portanto, o antecedente do fechamento em BIV deve ser falso. O antecedente do BIV fechamento é uma conjunção. O segundo conjuncto não pode ser defendido. Se você entende o que significa a hipótese do BIV, então você sabe que você não tem mãos se você for um BIV. Segue-se que o antecedente do fechamento em BIV é falso por que seu *primeiro* conjuncto é falso. Assim, assumindo o princípio do fechamento, chegamos numa conclusão cética: você não sabe que tem mãos.<sup>56</sup>

---

<sup>56</sup> Para mais discussão do princípio do fechamento e a função que ele cumpre no raciocínio cético, veja o artigo de Steven Luper “o princípio do fechamento epistêmico” na Stanford Encyclopedia of Philosophy, indicado no final desse artigo. Veja também Pritchard, 2004.

### 5.3 Alternativas relevantes e a negação do fechamento

A seguir examinaremos várias respostas ao argumento do BIV. De acordo com a primeira podemos distinguir entre alternativas relevantes e irrelevantes. Uma alternativa a um estado de coisas ou proposição  $p$  é um estado de coisas ou proposição que é incompatível com  $p$ . Ter mãos e ser um BIV são alternativas: se o primeiro é verdadeiro, o segundo é falso e vice-versa. De acordo com o pensamento que motiva a segunda premissa do argumento do BIV, você sabe que tem mãos somente se você pode discriminar entre ter realmente mãos e a alternativa de ser um BIV (maneta). Mas você não pode discriminar entre esses dois estados de coisas. É por isso que você não pode saber que você tem mãos. Em resposta a tal raciocínio, um teórico das alternativas relevantes poderia dizer que sua falta de habilidade para discriminar esses dois estados de coisas não é um impedimento para saber que você tem mãos, por que ser um BIV *não* é uma alternativa *relevante* para ter mãos. O que seria uma alternativa relevante?

Isso, por exemplo: seus braços terminarem em cotos, em vez de mãos, ou você ter ganchos em vez de mãos ou, ainda, ter próteses de mãos. Mas essas alternativas não impedem você de saber que você tem mãos- não por que elas são irrelevantes, mas, antes disso, por que você pode discriminar entre essas alternativas e ter mãos. O teórico das alternativas relevantes sustenta, portanto, que você sabe que tem mãos.

O argumento BIV é um argumento válido. Os teóricos das alternativas relevantes devem, portanto, negar uma de suas premissas. Dado que eles concordam que você não sabe que você não é um BIV, eles aceitam a primeira premissa. Conseqüentemente, eles rejeitam a segunda premissa. Você sabe que você tem mãos muito embora você não saiba que não é um BIV. Isso significa, com efeito, que os teóricos das alternativas relevantes negam o princípio do fechamento. Vamos considerar os detalhes desse ponto. Os teóricos das alternativas relevantes dizem:

- i. Você sabe que tem mãos.
- ii. Você sabe que ter mãos implica em não ser um BIV.
- iii. Você não sabe que você não é um BIV.

Os teóricos das alternativas relevantes, então, sustentam o antecedente e negam o conseqüente do fechamento em BIV, como estabelecido na seção anterior. Eles estão, portanto, comprometidos com a alegação que o princípio do fechamento é falso.<sup>57</sup>

Existem dois problemas principais nessa abordagem. O primeiro é que denunciar a alternativa do BIV como irrelevante é *ad hoc* a menos que ela seja acompanhada de um tratamento bem sustentado do que torna uma alternativa relevante e outra irrelevante. A segunda é que o princípio do fechamento goza de um alto grau de plausibilidade intrínseca. Negá-lo gera as assim chamadas conjunções abomináveis. Segue-se um exemplo:

**Uma conjunção abominável:** Eu sei que tenho mãos, mas não sei que eu não sou um BIV (maneta).

Muitos epistemólogos poderiam concordar que essa conjunção é realmente abominável, pois ataca flagrantemente a intuição básica e extremamente plausível que você não pode saber que tem mãos se não souber que não é um BIV<sup>58</sup>.

#### 5.4 A resposta mooreana

A seguir consideraremos uma resposta ao argumento do BIV de acordo com a qual a primeira premissa, e não a segunda, deve ser rejeitada. G.E. Moore chamou a atenção que um

---

<sup>57</sup> Uma advertência: isso é correto se a teoria das alternativas relevantes é entendida como construída por Dretske em 1970. Subsequentemente, alguns defensores da abordagem das alternativas relevantes elaboraram uma variação contextualista dessa doutrina com o objetivo explícito de torná-la compatível com o fechamento. Ver Stein, 1976.

<sup>58</sup> Para uma defesa da negação do fechamento e da abordagem das alternativas relevantes ver Dretske 1970. A negação do fechamento como estratégia anti-cética é também defendida por Nozick, 1981. Para a objeção da “conjunção abominável” à negação do fechamento, ver DeRose, 1995. Para um debate dos méritos de evitar o ceticismo através do abandono do fechamento, ver Dretske 2005 e Hawthorne, 2005.

argumento é bem sucedido somente à medida que suas premissas são mais plausíveis que sua conclusão. Assim, se encontramos um argumento cuja conclusão não gostamos e notamos que a negação da conclusão é realmente plausível, realmente mais plausível que a asserção das premissas, então podemos recusar o argumento. De acordo com essa abordagem, podemos responder ao argumento do BIV da seguinte maneira:

### **Argumento Contra o BIV**

(1) Eu sei que tenho mãos

(2) Se eu não sei que eu não sou um BIV, então não sei que tenho mãos.

Portanto:

(3) Eu sei que não sou um BIV

A menos que sejamos céticos ou oponentes do fechamento, teríamos que conceder que esse argumento é plausível. Ele é válido e suas premissas são verdadeiras. No entanto, poucos filósofos concordam que o Contra BIV equivale a uma resposta satisfatória ao argumento do BIV. O que precisa ser acrescentado é mais do que uma mera asserção de (3), baseada no conhecimento que alguém tem de suas mãos. O que precisamos que nos seja explicado é *como* alguém pode saber que não é um BIV. A observação que as premissas do argumento do BIV são menos plausíveis que a negação da sua conclusão não nos ajuda a entender como tal conhecimento é possível. É por isso que a resposta mooreana fracassa como uma refutação bem-sucedida do argumento cético.<sup>59</sup>

---

<sup>59</sup> Para a resposta de G. E Moore ao ceticismo, ver seus ensaios “Four Forms of Scepticism” e “Certainty” em Moore, 1959. Uma objeção para a resposta mooreana que não é discutida aqui é que a resposta é uma petição de princípio. É, contudo, uma questão aberta se é realmente. O que é necessário para realizar tal consideração é uma abordagem precisa de quando um argumento é uma petição de princípio, até tal abordagem ser fornecida a acusação de petição de princípio é *ad hoc*. Mas mesmo que o argumento contra BIV não seja uma petição de princípio podemos ainda imaginar se nós podemos vir a saber que nós não somos BIVs tendo por base o uso deste argumento. O ponto aqui é o da prioridade epistêmica, poderia ser o caso de que eu devesse saber em primeiro lugar que eu não sou um BIV se eu vier a saber que eu tenho mãos. Se isso está correto ninguém poderia adquirir conhecimento sobre não ser um BIV em virtude de empregar o argumento contra BIV. Para discussões contra a resposta mooreana veja Pritchard 2004, Pryor 2004 e Sosa 1999b.

## **5.5 A resposta contextualista**

Nós examinamos duas respostas ao argumento do BIV. A resposta das alternativas relevantes nega a segunda premissa. Em função da plausibilidade da segunda premissa, isso poderia nos parecer um movimento desesperado. A resposta mooreana nega a primeira premissa. O problema com este movimento é o seguinte: a menos que sejamos capazes de proporcionar uma explicação convincente sobre *como* alguém pode saber que não é um BIV, não estaremos fazendo nada além de continuar nos movendo dentro das engrenagens não-céticas. De acordo com o contextualismo, é possível articular uma réplica mais satisfatória ao argumento BIV. O truque é focar em como nós de fato usamos o termo 'saber'. Se o fizermos, notaremos que nosso uso desse termo varia de uma situação- de um contexto- para outro. O que varia aqui é o que nós *queremos dizer* com o termo.

Três questões surgem imediatamente. A primeira é: o que são esses vários sentidos do termo 'saber'? A segunda é: por que e como o que queremos dizer por 'saber' muda de um contexto para o outro? Terceiro, como a *sensibilidade ao contexto* do 'saber' ajuda-nos a responder ao argumento do BIV? Examinemos cada uma dessas questões.

Primeiro, quando o que dizemos por 'saber' muda de um contexto para o outro, o que muda são os padrões que pensamos que devem ser satisfeitos se alguém pretende ter conhecimento de algo. Para simplificar as coisas, vamos distinguir entre dois conjuntos de padrões: os muito altos e os não tão altos. Chamemo-os de padrões 'altos' e 'baixos'. Em alguns contextos, quando usamos o termo 'saber', temos padrões de baixos de conhecimento em mente: padrões que são fáceis de satisfazer. Nós atribuímos conhecimento de maneira generosa. Em outros contextos, nosso uso do termo 'saber' é guiado por padrões altos e mais exigentes. Satisfazer esses padrões é muito difícil. Em tais contextos, atribuímos conhecimento com relutância. Em segundo lugar, o que produz tais mudanças no que queremos dizer por 'saber'? De acordo com alguns contextualistas, é a saliência das possibilidades de erro. Num contexto comum, de padrões baixos, não lhe passa pela cabeça ser um BIV. Não é uma possibilidade de erro que você ignora. Como consequência, seus

padrões de conhecimento permanecem baixos. Num tal contexto, tudo o que é pressuposto para saber que você tem mãos é poder discriminar entre ter mãos e ter cotos, ganchos ou próteses de mãos. Essa é uma condição que podemos facilmente satisfazer. Portanto, você não teria nenhuma relutância em atribuir a si mesmo o conhecimento de suas mãos. Mas suponha que você comece a pensar sobre o problema do ceticismo. Você está intrigado sobre como você pode saber que não é um BIV. Você nota que é muito difícil saber que não é um BIV. Agora, a alternativa BIV tornou-se saliente para você. Isso faz seus padrões de conhecimento aumentarem. Tendo em mente que BIVs não tem mãos, você agora pensa que, para saber que você tem mãos, você deve ser capaz de eliminar a possibilidade de erro de ser um BIV. Uma vez que você percebe que você não pode eliminar essa possibilidade, você não estará mais disposto a atribuir a si mesmo o conhecimento de suas mãos.

Terceiro, como tudo isso nos ajuda a responder ao argumento do BIV? Os contextualistas vêem o argumento do BIV como a apresentação de um paradoxo. Nós pensamos que é loucura negar o conhecimento de nossas mãos. Ao mesmo tempo, nós não pensamos que sabemos que não somos um BIV. Como pode o conflito entre esses pensamentos ser resolvido? Os contextualistas se propõem a resolver isso dizendo o seguinte: em contextos com padrões baixos (quando as hipóteses céticas não são salientes), a primeira premissa e a conclusão do argumento do BIV são falsas. Em tais contextos, o falante que diz: “você não sabe que você tem mãos” ou “Você não sabe que não é um BIV” está enganado. O falante está enganado por que nós, de fato, satisfizemos padrões *baixos* de conhecimento. Assim, relativamente ao que significa ‘saber’ em tais contextos, sabemos que temos mãos e que não somos BIVs. Contudo, em contextos de padrões altos (quando uma possibilidade tal como ser um BIV é saliente), a primeira premissa e a conclusão do argumento BIV são verdadeiras. Agora, quando o falante diz: “Você não sabe que você tem mãos” ou “você não sabe que você não é um BIV”, eles estão certos, pois com relação a ter mãos ou não ser um BIV, nossa posição epistêmica não é forte o suficiente para satisfazer padrões de conhecimento tão elevados. Portanto, relativamente ao que nós significamos por ‘saber’ quando confrontados

com possibilidades de erro salientes tais como BIV, nós não sabemos nem que temos mãos e nem que não somos BIVs.

O contextualismo pretende ser uma resposta ao ceticismo *preservadora do fechamento*. O princípio do fechamento é verdadeiro mesmo em relação às atribuições de ‘conhecimento’ que estão sujeitas a altos padrões. Portanto, de acordo com o contextualismo, tudo se passa do seguinte modo:

2. conhecemos o princípio do fechamento seja o significado de ‘saber’ fixado por padrões altos ou baixos;
3. Quando o significado de ‘saber’ é fixado por padrões baixos, sabemos tanto que temos mãos, quanto que não somos BIVs.
4. Quando o significado de ‘saber’ é fixado por padrões mais altos, não sabemos nem que temos mãos, nem que não somos BIVs.

Como consequência, o fechamento é preservado. O contextualismo pode também ser entendido como um aperfeiçoamento da resposta mooreana. De acordo com essa resposta, a primeira premissa do argumento BIV é falsa. Isso entra em conflito com nossa intuição que não podemos saber que não somos BIVs. O contextualismo resolve esse conflito dizendo que a primeira premissa é falsa apenas em contextos de padrões baixos. Em contextos de padrões elevados, essa premissa é verdadeira.

Naturalmente, o contextualismo tem suscitado muitas objeções. De acordo com alguns, o que está errado com o contextualismo é que ele substitui nosso interesse pelo conhecimento em si mesmo pelo foco no termo ‘saber’. Essa objeção (chamemo-la a *objeção da substituição*) baseia-se num mal-entendido acerca do contextualismo. Na próxima seção, veremos por que.

De acordo com outra objeção, o contextualismo enfatiza em demasia a importância da sensibilidade ao contexto do termo ‘saber’. Podemos distinguir entre dois elementos do contextualismo. O primeiro é a ascensão *semântica*. Se endossarmos o elemento da ascensão semântica, pensaremos que uma resposta satisfatória ao ceticismo em geral e ao argumento

do BIV em particular requer a capacidade de distinguir entre padrões elevados e padrões baixos do sentido de ‘conhecimento’. A tese da ascensão semântica permanece inteiramente dentro dos limites da epistemologia tradicional. Realmente, em qualquer área da filosofia, é sempre uma boa idéia ter consciência da possibilidade de que os problemas em que alguém se acha enredado podem, pelo menos em alguma medida, se dever a mudanças sutis (e às vezes nem tão sutis) no sentido. O outro elemento do contextualismo pode ser chamado de sensibilidade estrita ao contexto, em oposição à sensibilidade frouxa ao contexto. Considere a tese que o sentido da palavra ‘saber’ varia com o contexto. Há uma interpretação inócua dessa tese: as pessoas nem sempre querem dizer o mesmo quando elas usam o termo ‘saber’. Às vezes elas querem dizer uma coisa por ‘saber’ e, outras vezes, elas querem dizer outra coisa. Essa é a sensibilidade frouxa ao contexto. É difícil ver em que bases uma alegação tão fraca poderia ser posta em causa. Os contextualistas, contudo, fazem uma alegação mais forte. Eles sustentam que o que alguém quer dizer por ‘saber’ é determinado, de um modo que é muito difícil de resistir, pela saliência ou não saliência de possibilidades de erro. Essa é a sensibilidade estrita ao contexto. Se endossarmos a sensibilidade estrita ao contexto, há algo importante que salta aos olhos: como se *pretende* usar o termo ‘saber’. Uma alternativa semântica do termo ‘saber’ pode enfraquecer a ênfase na importância da saliência ou da não saliência de possibilidades de erro e atribuir um grau muito mais elevado de independência semântica aos sujeitos que usam o termo saber. A seguir, vamos considerar a resposta ao BIV que retém o elemento da ascensão semântica do contextualismo, mas rejeita a sensibilidade forte ao contexto<sup>60</sup>.

---

<sup>60</sup> Para uma aproximação inicial ao grande corpo de literatura contextualista veja Cohen 1988, 1999, 2001, 2005, Conee 2005, DeRose 1992, 1995, 1999, Feldman 1999b, 2001b, Hawthorne 2004, Lewis 1996, Schiffer 1996, e Sosa 2003. Para referências bibliográficas adicionais acerca do contextualismo veja bibliografia extensa e comentada proposta por Keith De Rose em “Contextualism in Epistemology – A Bibliography”

## 5.6 A resposta da ambiguidade

Qual a *proposição* que uma sentença atribuidora de ‘conhecimento’ expressa depende de qual *conceito de conhecimento* a pessoa que usa essa sentença (de forma escrita ou falada) tem em mente quando usa o termo ‘saber’. Vamos distinguir entre dois conceitos: um de padrões elevados e outro conceito de padrões baixos. Existem vários modos de delinear essa distinção. Podemos entendê-la em termos de evidência falível e infalível. O conhecimento de  $p$  em padrões elevados ou infalíveis exige que a evidência implique  $p$ . Padrões baixos do conhecimento falível de  $p$  requerem evidência adequada para  $p$ , onde evidência para  $p$  pode ser adequada sem implicar  $p$ .

De acordo com a resposta da ambigüidade, uma sentença atribuidora de ‘conhecimento’ é ambígua ao menos à medida que podemos dizer se o termo ‘saber’, como aparece na sentença, faz referência ao conhecimento falível ou infalível. Suponha que pensemos que o conhecimento falível de que alguém tem mãos seja possível, enquanto que o conhecimento infalível que temos mãos não seja. Suponha, ainda que ouvimos Jane dizer: ‘Carl sabe que ele tem mãos’. Finalmente, suponha que não temos nenhuma idéia se Jane usa o termo ‘saber’ no sentido falível ou infalível. Nesse caso, teríamos ter que dizer que o proferimento de Jane é verdade se interpretado como uma alegação sobre conhecimento falível, mas falso se interpretado como uma alegação sobre o conhecimento infalível. Agora, com respeito ao argumento do BIV, estamos numa situação parecida. Nós não fomos informados se devemos tomar o termo ‘saber’ em suas conclusões e premissas num sentido falível ou infalível. Consequentemente, quando avaliamos os méritos do argumento do BIV devemos considerar três versões dele:

**A versão mista:** Nas premissas, o termo ‘saber’ refere-se ao conhecimento infalível, enquanto que na conclusão, ele faz referência ao conhecimento falível.

**A versão dos padrões altos ou exigentes:** O termo 'saber' refere-se ao conhecimento infalível nas premissas e na conclusão.

**A versão dos padrões baixos:** O termo 'saber' refere-se ao conhecimento falível tanto nas premissas quanto na conclusão.

Distinguindo entre essas três versões, proponentes da resposta da ambigüidade podem responder ao BIV do seguinte modo:

- i. A versão mista é uma instância da falácia de equívocidade e, assim, é inválida.
- ii. A versão dos padrões altos é válida, mas não interessante. Sua conclusão afirma que não temos conhecimento infalível de nossas mãos. Isso não é nada preocupante. O que realmente importa para nós é se temos conhecimento falível de nossas mãos. Mas essa questão simplesmente não é abordada pela versão dos padrões elevados.
- iii. A versão dos padrões baixos é interessante, mas não é sólida. Sua conclusão - nós nem mesmo temos conhecimento falível de nossas mãos - é realmente perturbadora. Se essa conclusão fosse verdadeira, então estaríamos enganados de modo radical sobre o que pensamos que sabemos. Contudo, não temos que aceitar essa conclusão, por que a primeira premissa do argumento é falsa. De acordo com essa premissa, não se pode ter conhecimento falível de não ser um BIV. Isso é falso. Há, sobretudo, boa evidência para pensar que não se é BIV. Essa evidência é suficientemente boa para saber que não se é um BIV, muito embora não implique que não se é um BIV.

Suponha que um oponente da resposta da ambigüidade fosse empregar a objeção da substituição , afirmando que a resposta foca no termo 'saber', em vez de focar no conhecimento em si mesmo. Essa objeção erraria o ponto. A resposta da ambigüidade menciona o termo 'saber' somente no estágio inicial, e então imediatamente muda seu foco

para entidades não linguísticas tais como conceitos e proposições. Assim, os defensores da resposta da ambiguidade poderiam ressaltar que, quando distinguimos entre as versões de (i) até (iii), nós estamos interessados em quais *proposições* as premissas e a conclusão do argumento do BIV expressa e, assim, em última instância, preocupados com o conhecimento em si mesmo. O desfecho de sua réplica consiste, então, em distinguir entre as duas seguintes proposições:

(Kif) Eu *seif* que eu tenho mãos

(K-f) Eu *sei-f* que eu tenho mãos.

Onde o termo ‘Saber if’ em (Kif) refere-se ao conhecimento infalível, enquanto que o termo ‘saber-f’ em Kf refere-se ao conhecimento falível. Ambas proposições são sobre o conhecimento em si mesmo, ou, mais precisamente, sobre diferentes *espécies* de conhecimento. A resposta da ambiguidade, portanto, não é vulnerável a objeção da substituição. Nem o é o contextualismo, pois de acordo com o contextualismo, o que o contexto determina é precisamente qual proposição a conclusão do argumento do BIV expressa: (Kif) ou (K-f)<sup>61</sup>. Assim, o contextualismo a despeito das aparências, diz tanto do conhecimento em si mesmo quanto diz a objeção da ambiguidade.

Agora, então, o contextualismo e a resposta da ambiguidade realmente diferem? Ambos partilham do mesmo elemento da ascensão semântica. Uma resposta satisfatória para o ceticismo exige que distingamos entre os vários sentidos do termo ‘saber’. Além disso, eles procedem em diferentes direções. Enquanto, de acordo com o contextualismo, se rejeitamos ou endossamos a conclusão do argumento do BIV depende de qual contexto nos encontramos, a resposta da ambiguidade torna o contexto irrelevante. Ela torna o contexto irrelevante, pois, não importa em qual contexto estamos, nós podemos sempre eliminar a ambiguidade. Assim, quando estamos pensando ou discutindo sobre o argumento BIV e somos, assim, confrontados

---

<sup>61</sup> Ou seja, *se* uma versão particular de contextualismo estabelece a diferença padrões baixos/padrões altos em termos de evidência falível vs. evidência infalível. Se a diferença padrões altos/ padrões baixos é estabelecida de modos diferentes, então proposições diferentes estarão em destaque.

com uma possibilidade de erro saliente, não precisamos adotar o sentido de saber de padrões elevados. Antes, podemos responder o argumento dizendo que, se é sobre o conhecimento infalível, sua conclusão é verdadeira, mas pouco interessante, enquanto que se for sobre o conhecimento falível, sua conclusão é interessante, mas falsa<sup>62</sup>.

### **5.7 Saber que não se é um BIV**

O contextualismo e a resposta da ambiguidade, conforme foram discutidas nas duas seções anteriores, deixam de fora um detalhe importante. Os contextualistas dizem que, relativamente aos padrões operacionais em contextos de padrões-baixos, alguém pode saber que não é um BIV. Os teóricos da ambiguidade dizem que, no sentido falibilista de ‘saber’, alguém pode saber que não é um BIV. Poderia ser objetado que isso é um tanto otimista. Vamos ver a questão do ponto de vista evidencialista. Um evidencialista que emprega a resposta da ambiguidade teria que dizer que a evidência que se tem para pensar que não se é um BIV é suficientemente boa para ter conhecimento. Mas quando a hipótese do BIV foi introduzida, notamos que parte da hipótese levanta o seguinte ponto: se você é uma pessoa normal ou um BIV não faz diferença com respeito a sua evidência: ela é a mesma em ambos os casos. Chame isso de *tese da evidência idêntica*. Essa tese é simplesmente parte da hipótese em questão e deve, portanto, ser assumida. Como, então, poderia alguém possivelmente saber, mesmo no sentido falibilista de ‘saber’, que ele não é um BIV?

Poderia ser um erro pensar que a tese da evidência idêntica implica que, como uma pessoa normal, alguém não tem boa evidência para pensar que não é um BIV. Nem implica que, como um BIV, alguém não tem boa evidência para pensar que não é um BIV. O que se

---

<sup>62</sup> Para uma apresentação clara da resposta da ambiguidade veja Fred Feldman 1986 capítulo 2 especialmente pp. 33-37. Feldman distingue entre conhecimento prático (o tipo de conhecimento que nós atribuímos a nós mesmos na vida comum) e conhecimento metafísico (o tipo que Descartes tinha em mente nas suas meditações). Feldman escreve: “o desfecho [da minha réplica ao ceticismo] é de que se tomamos o argumento [cético] como sendo sobre o conhecimento prático ele tem uma conclusão digna de nota, mas uma premissa não defensável. Se nós o tomamos como sendo como o conhecimento metafísico ele é cogente mas a sua conclusão não é de grande interesse. Se nós tentamos manter a conclusão interessante, mas ao mesmo tempo tornando as suas premissas todas verdadeiras o argumento perderá a sua cogência. De qualquer modo nós não temos provas em favor de qualquer forma surpreendente de ceticismo” (p. 36). Para a literatura defendendo a resposta da ambiguidade veja Engel 2003, Russell 2004 Steup. 2005.

segue é meramente isso: seja qual for a evidência que se tem, como uma pessoa normal considerando a questão de se é um BIV, alguém poderia ter a mesma evidência se fosse um BIV. Isso deixa aberta a possibilidade que, em qualquer caso, como BIV ou pessoa normal, tem-se evidência excelente para pensar que não se é um BIV.

No que poderia consistir a evidência para pensar que não se é um BIV? Por razões de espaço, podemos meramente sugerir, mediante uma analogia, como essa questão poderia ser respondida. Note, antes de tudo, que a hipótese BIV implica em várias proposições problemáticas:

- (1) Ao menos um BIV existe.
- (2) A perícia ou habilidade necessária para encubar pessoas existe.
- (3) A tecnologia necessária para encubar pessoas existe.

Compare:

- (4) Ao menos uma nave espacial existe e pode ser usada para viajar para outra galáxia e voltar dentro de uns poucos meses.
- (5) A perícia necessária para construir tal nave espacial existe.
- (6) A tecnologia necessária para construir tal nave espacial existe.

De acordo com o anti-cético evidencialista sob análise aqui, você sabe, na base do conhecimento de como o mundo funciona, que (d)-(f) são todas falsas<sup>63</sup>. Pelo menos, você

---

<sup>63</sup> O que dizer então sobre viajar até outra galáxia num par de meses? Bem, se consideramos Andrômeda como a galáxia que está mais próxima a nós, viajar até lá na velocidade da luz levaria 2,2 milhões de anos. Voltar outros 2,2 milhões de anos. Viagens espaciais reais seriam um pouco mais lentas do que a velocidade da luz e por conta disso adicione alguns milhões de anos a mais, assim mesmo mantendo uma saudável margem de erro em relação à quão longe Andrômeda está, parece seguro dizer que sabemos que ir até lá e voltar em um par de meses não está no reino das possibilidades físicas. De acordo com uma descoberta recente a Galáxia Anã do Cão Maior é de fato a galáxia que está mais próxima de nós. Ela está somente há 25 mil anos luz de distância. Ainda muito longe, podemos supor, para irmos e voltarmos em um par de meses. Assim saber que nós não podemos viajar para esses lugares e voltar em um par de meses envolve não mais do que saber algumas trivialidades dos livros escolares de física moderna. De modo similar temos livros escolares de neurofisiologia moderna, eles nos dizem dentre outras coisas o que acontece com quem toma LSD ou outras drogas do gênero. Eles também nos dizem o que acontece se a cabeça de alguém é aberta e várias partes do seu cérebro são

pode *vir* a saber isso consultando os especialistas certos. Mas o que dizer acerca de (a) até (c)? Bem, se você sabe ou pode vir a saber que (d)-(f) são todas falsas, não é plausível afirmar que você sabe ou pode vir a saber que (a)-(c) são todas falsas? Se o cético fosse argumentar que você sabe que (d)-(f) são falsas, embora você não saiba que qualquer proposição em (a)-(c) é falsa, o cético poderia incorrer no ônus de ter que abandonar a analogia, de ter que explicar por que, enquanto o conhecimento que (d)-(f) são falsas é facilmente alcançável, o conhecimento, por sua vez, da falsidade de cada um (a)- (c), está além do nosso alcance. Isso não pode ser obtido tão facilmente.

Suponha que você sabe que (a)-(c) são falsas. Então você sabe que qualquer proposição que *implica* (a)-(c) é falsa. A hipótese BIV implica (a)-(c). Portanto, você sabe que a hipótese BIV é falsa. Mas se você sabe que você não é um BIV, então a premissa (1) do argumento do BIV é falsa<sup>64</sup>.

## 6 Questões adicionais

### 6.1 Epistemologia da Virtude

A epistemologia, como é praticada comumente, direciona sua atenção para as crenças do sujeito. Elas estão justificadas? Elas são instâncias de conhecimento? Quando chega o momento de avaliar como o próprio sujeito está agindo com respeito à busca da verdade e a procura do conhecimento, esta avaliação é realizada olhando para a qualidade epistêmica de suas crenças. De acordo com a epistemologia da virtude, a ordem da análise deve ser invertida. Precisamos começar com o próprio sujeito e avaliar as suas virtudes e vícios epistêmicos: suas maneiras “boas” e “ruins” de formar crenças. O raciocínio cuidadoso e atento seria um exemplo de virtude epistêmica; pular para conclusões seria um exemplo de

---

sondadas e cutucadas. O que eles não dizem é como “encubar”: manter um cérebro vivo por um período extenso e criar a ilusão de uma vida normal. Se nós consultarmos especialistas nesse assunto eles nos diriam que isso não pode ser feito. Deste modo pareceria que alguém pode vir a saber que (A) até (C) são todas falsas.

<sup>64</sup> Duas novas respostas ao ceticismo que não serão discutidas aqui são as do externalismo semântico e externalismo epistêmico. Para a discussão destas respostas veja DeRose, e Warfield 1999.

vício epistêmico. Só depois de termos determinado que maneiras de formar crenças contam como virtudes epistêmicas é que podemos, como um segundo passo, determinar a qualidade epistêmica de crenças particulares. Seus defensores entendem a epistemologia da virtude de modo mais ou menos rigoroso. De acordo com a epistemologia pura da virtude, as virtudes epistêmicas são *sui generis*. Elas não podem ser analisadas em termos de conceitos epistêmicos ou não-epistêmicos mais fundamentais. Defensores de uma abordagem menos rigorosa não concordam com isso; eles diriam que as virtudes e vícios epistêmicos podem ser fecundamente analisados empregando outros conceitos. Na verdade, de acordo com uma vertente externalista da epistemologia da virtude, é a própria noção de confiabilidade que devemos empregar para capturar a diferença entre vícios e virtudes epistêmicos. Maneiras estáveis de formar crenças são virtudes epistêmicas se e apenas se elas tendem a resultar em crenças verdadeiras, e vícios epistêmicos se e apenas se elas tendem a resultar em crenças falsas. A epistemologia da virtude, assim concebida, é uma forma de confiabilismo<sup>65</sup>.

## 6.2 Epistemologia Naturalista

De acordo com uma versão extrema da epistemologia naturalista, o projeto da epistemologia tradicional, perseguida de um modo *a priori* a partir da poltrona dos filósofos, é completamente equivocado. Os “frutos” de tal atividade são teorias demonstrativamente falsas, como o fundacionalismo, bem como debates intermináveis e misteriosos na tentativa de enfrentar questões para as quais não há respostas. Para trazer a filosofia ao caminho correto, ela deve fazer parte das ciências naturais e tornar-se psicologia cognitiva. O objetivo da epistemologia naturalista assim entendida é substituir a epistemologia tradicional por um projeto completamente novo e redefinido. De acordo com uma versão moderada da epistemologia naturalista, a tarefa primária da epistemologia é identificar como o conhecimento e a justificação estão ancoradas no mundo natural, assim como é o propósito da

---

<sup>65</sup>De literatura sobre epistemologia da virtude, veja Axtell 1997, Brady e Pritchard 2003, Greco 1993, 1999 e o artigo de Greco “Epistemologia da Virtude” (referido no final deste artigo), Kvanvig 1996a, Montmarquet 1993, ensaios 8, 11, 13 e 16 em Sosa 1991, Sosa 1997 e as contribuições de Sosa a Bonjour e Sosa 2003 e Zagzebski 1996, 1999.

física explicar fenômenos como calor e frio, ou trovão e relâmpago, em termos de propriedades do mundo natural. A realização desta tarefa não requer dos seus proponentes a substituição da epistemologia tradicional. Ao contrário, esta abordagem moderada aceita a necessidade de cooperação entre a análise tradicional de conceitos e métodos empíricos. A primeira é necessária para o propósito de estabelecer a ligação conceitual entre conhecimento e confiabilidade, os últimos para determinar quais processos cognitivos são confiáveis e quais não são<sup>66</sup>.

### **6.3 Epistemologia da Religião**

Na história da filosofia, há vários argumentos célebres em favor da existência de Deus: o argumento ontológico, o argumento cosmológico e o argumento do desígnio. De um ponto de vista epistemológico, a questão é se estes argumentos podem constituir um fundamento racional para fé, ou mesmo se nos fornece conhecimento de Deus. Uma questão adicional é se, Deus existindo, o conhecimento de Deus não poderia também ser possível de outras maneiras, por exemplo, baseado na percepção ou talvez em experiências místicas. Há também um problema famoso que coloca dúvida sobre a existência de Deus: por que, se Deus é um ser onisciente, onipotente e benevolente, há mal no mundo? Aqui a questão epistemológica é se, baseado nesta dificuldade, podemos saber que Deus (assim concebido) não existe. Outra questão central para a epistemologia da religião é levantada pelo evidencialismo. De acordo com o evidencialismo, o conhecimento requer evidência adequada. Entretanto, não parece haver nenhuma evidência adequada para a existência de Deus. É possível, então, que os teístas endossem o evidencialismo?<sup>67</sup>

---

<sup>66</sup>De literatura sobre a questão da epistemologia naturalizada, veja Feldman 1999a, o artigo de Feldman “Naturalized Epistemology” na Stanford Encyclopedia of Philosophy (referido no fim deste artigo), Kornblith 1999 e 2002, Goldman 1986, Quine 1969 e capítulo 9 em Steup 1996.

<sup>67</sup>De literatura sobre epistemologia da religião, veja Alston 1991, Audi 1997a, 2000, Plantinga 2000, e Wolterstorff 1999. Referências bibliográficas úteis podem ser encontradas em Greco e Sosa 1999, p. 445.

## 6.4 Epistemologia Moral

As categorias morais básicas são aquelas de ação certa e errada. Quando fazemos ética teórica, buscamos descobrir o que é que faz uma ação correta correta e uma ação errada errada. Quando fazemos ética aplicada ou prática, almejamos descobrir quais ações são corretas e quais são erradas. A questão epistemológica que essas áreas da filosofia levantam é esta: como podemos saber qualquer uma dessas coisas? Tradicionalmente, os filósofos tentaram responder as questões da ética pela intuição, raciocínio *a priori* e a consideração de casos hipotéticos. Alguns filósofos que se filiam ao campo naturalista consideram esta abordagem equivocada, pois eles pensam que ela é inconfiável e passível de produzir resultados que meramente refletem os nossos preconceitos sociais e culturais. Entre aqueles que pensam que o conhecimento moral pode ser adquirido via intuição e raciocínio *a priori*, a questão primária é se o tipo de justificação que tais métodos podem gerar é coerentista ou fundacionalista. Por fim, uma questão adicional importante é se o conhecimento moral é de qualquer modo possível. O conhecimento requer a verdade e, portanto, realidade objetiva. De acordo com os anti-realistas, não há realidade objetiva e, portanto, nenhuma verdade, nos assuntos morais. Visto que aquilo que é conhecido deve ser verdadeiro, não é fácil ver como, se os anti-realistas estiverem corretos, poderia haver conhecimento de assuntos morais<sup>68</sup>.

## 6.5 Epistemologia Social

Quando concebemos a epistemologia como incluindo o conhecimento e a crença justificada tal como eles estão posicionados em um contexto social e histórico particular, a epistemologia se torna epistemologia social. Como abordar a epistemologia social é uma questão controvertida. De acordo com alguns, ela é uma extensão e reorientação da epistemologia tradicional com o objetivo de corrigir a sua orientação excessivamente

---

<sup>68</sup>De literatura sobre epistemologia moral, veja Audi 1997b, 1999, 2000, 2004 e o artigo de Richmond Campbell “Moral Epistemology” na Enciclopédia Stanford (referido no final deste artigo). Referências bibliográficas adicionais são dadas em Greco e Sosa 1999, p. 444s..

individualista. De acordo com outros, a epistemologia social deveria estar associada em um afastamento radical da epistemologia tradicional, a qual eles vêm, enquanto advogados de uma naturalização radical, como um empreendimento fútil. Aqueles que favorecem a primeira abordagem retêm o pensamento de que o conhecimento e a crença justificada estão essencialmente ligados à verdade como o objetivo de nossas práticas cognitivas. Eles mantêm que há normas objetivas de racionalidade que os epistemólogos sociais devem aspirar articular. Aqueles que preferem a abordagem mais radical rejeitariam a existência de normas objetivas da racionalidade. Além disso, visto que muitos vêm os fatos científicos como construções sociais, eles negariam que o objetivo de nossas atividades intelectuais e científicas é descobrir fatos. Tal construtivismo, se fraco, assevera a afirmação epistemológica de que as teorias científicas são carregadas de preconceitos e pressuposições sociais, culturais e históricas; se forte, ele assevera a afirmação metafísica de que a verdade e a realidade são elas mesmas socialmente construídas<sup>69</sup>.

## **6.6 Epistemologia Feminista**

Quando entendida de uma maneira não controversa, o tema da epistemologia feminista consiste de questões que têm a ver com o acesso justo e igual das mulheres e a sua participação nas instituições e processos pelos quais o conhecimento é gerado e transmitido. Encarada desta maneira, a epistemologia feminista pode ser vista como um ramo da epistemologia social. Quando nos movemos para além desta caracterização inicial, é uma questão controversa no que se torna a epistemologia feminista. De acordo com alguns, ela inclui o projeto de estudar e legitimar maneiras especiais em que apenas as mulheres podem adquirir conhecimento. De acordo com outros, a epistemologia feminista deve ser entendida como mirando o objetivo político de opor e retificar a opressão em geral e a opressão da

---

<sup>69</sup>De literatura sobre epistemologia social, veja o artigo de Alvin Goldman “Social Epistemology” na Stanford Encyclopedia of Philosophy (referido no final deste artigo), Goldman 1999 e Schmitt 1994 e 1999. Para uma lista de referências bibliográficas, veja Greco e Sosa 1999, p. 448.

mulher em particular. No extremo, a epistemologia feminista está estreitamente associada ao pós-modernismo e ao seu ataque radical à verdade e à noção de realidade objetiva<sup>70</sup>.

## **Bibliografia**

- Alston, William. 1989. *Epistemic Justification. Essays in the Theory of Knowledge*. Ithaca: Cornell University Press.
- . 1991. *Perceiving God. The Epistemology of Religious Experience*. Ithaca: Cornell University Press.
- . 1993. *The Reliability of Sense Perception*. Ithaca: Cornell University Press.
- . 1999. "Perceptual Knowledge." In: Greco and Sosa 1999, pp. 223–242.
- Armstrong, D.M. 1973. *Belief, Truth, and Knowledge*. Cambridge: Cambridge University Press.
- Axtell, Guy (ed.). 1997. *Knowledge, Belief, and Character. Readings in Virtue Epistemology*. New York: Rowman and Littlefield.
- Audi, Robert. 1993. *The Structure of Justification*. Cambridge: Cambridge University Press.
- . 1997a. *Religion in the Public Square: The Place of Religious Conviction in Political Debate*. Lanham: Rowman and Littlefield
- . 1997b. *Moral Knowledge and Ethical Character*. Oxford: Oxford University Press.
- . 1998. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. Cambridge: Cambridge University Press.
- . 1999. *Moral Knowledge and Ethical Pluralism*. In: Greco and Sosa 1999, pp. 271–302.
- . 2000. *Religious Commitment and Secular Reason*. Cambridge: Cambridge University Press.

---

<sup>70</sup>De literatura sobre epistemologia feminista, veja Longino 1999 e o artigo de Elizabeth Anderson "Feminist Epistemology and Philosophy of Science" na *Stanford Encyclopedia of Philosophy* (referido no final deste artigo). Para uma longa lista de referências bibliográficas, veja Greco e Sosa 1999, pp. 455ff.

- . 2004. *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton: Princeton University Press.
- Boghossian, Paul and Peacocke, Christopher (eds.). 2000. *New Essays on the A Priori*. Oxford: Oxford University Press.
- BonJour, Laurence. 1985. *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.
- . 1998. *In Defense of Pure Reason*. London: Cambridge University Press.
- . 1999. "The Dialectic of Foundationalism and Coherentism". In Greco and Sosa 1999, pp. 117–142.
- . 2001. "Towards a Defense of Empirical Foundationalism". In DePaul 2001, pp. 21–38.
- . 2002. *Epistemology. Classic Problems and Contemporary Responses*. Lanham: Rowman and Littlefield.
- . 2005. "In Defense of the A Priori". In Steup and Sosa (eds.) 2005, pp. 98–105.
- BonJour, Laurence and Sosa, Ernest. 2003. *Epistemic Justification. Internalism vs. Externalism, Foundations vs. Virtues*. Malden, MA: Blackwell.
- Brady, Michael and Pritchard, Duncan. 2003. *Moral and Epistemic Virtues*. Oxford: Blackwell.
- Brewer, Bill. 1999. *Perception and Reason*. Oxford: Oxford University Press.
- . 2005. "Perceptual Experience Has Perceptual Content." In: Steup and Sosa 2005, pp. 217–230.
- Byrne, Alex. 2005. "Perception and Conceptual Content." In Steup and Sosa 2005, pp. 231–250.
- Casullo, Albert. 2003. *A Priori Justification*. New York: Oxford University Press.
- Chisholm, Roderick. 1982. *The Foundations of Knowing*. Minneapolis: University of Minnesota Press.
- . 1977. *Theory of Knowledge*, 2nd. ed., Englewood Cliffs: Prentice Hall.
- . 1989. *Theory of Knowledge*, 3rd. ed., Englewood Cliffs: Prentice Hall.
- Cohen, Stewart. 1988. "How to be a Fallibilist." *Philosophical Perspectives*, 2: 91–123.

- . 1999. "Contextualism, Skepticism, and the Structure of Reasons." *Philosophical Perspectives*, 13: 57–90.
- . 2001. "Contextualism Defended: Comments on Richard Feldman's 'Skeptical Problems, Contextualists Solutions'." *Philosophical Studies*, 103(1): 87–98.
- . 2005. "Contextualism Defended." In Steup and Sosa (eds.) 2005, pp. 56–62.
- Conee, Earl. 2004. "The Truth Connection". In Conee and Feldman 2004, pp. 242–258.
- . 2005. "Contextualism Contested". In Steup and Sosa (eds.) 2005, pp. 47–56.
- Conee, Earl and Feldman, Richard. 1985. "Evidentialism." *Philosophical Studies*, 48: 15–35.
- . 2001. "Internalism Defended." In: Kornblith (ed.) 2001, pp. 231–60. Reprinted in Conee and Feldman 2004, pp. 53–82.
- . 2004. *Evidentialism. Essays in Epistemology*. Oxford: Oxford University Press.
- Dancy, Jonathan. 1985. *Introduction to Contemporary Epistemology*. Oxford: Blackwell.
- David, Marian. 2001. "Truth and the Epistemic Goal." In: Steup 2001a.
- Devitt, Michael. 2005. "There is no A Priori." In: Steup and Sosa (eds) 2005, pp. 105–115.
- DePaul, Michael (ed.). 2001. *Resurrecting Old-Fashioned Foundationalism*. New York: Rowman and Littlefield.
- DeRose, Keith. 1995. "Solving the Skeptical Problem." *The Philosophical Review*, 104: 1–52.
- . 1992. "Contextualism and Knowledge Attributions." *Philosophy and Phenomenological Research*, 52: 913–929.
- . 1999. "Contextualism: An Explanation and Defense." In: Greco and Sosa 1999, pp. 187.
- DeRose, Keith, and Warfield, Ted. 1999. *Skepticism. A Contemporary Reader*. Oxford: Oxford University Press.
- Devitt, Michael. 2005. "There is No A Priori." In: Steup and Sosa 2005, pp. 105–115.
- Dretske, Fred. 1970. "Epistemic Operators." *The Journal of Philosophy*, 67: 1007–23.
- . 1971. "Conclusive Reasons." *Australasian Journal of Philosophy*, 49: 1–22.
- . 1981. *Knowledge and the Flow of Information*. Oxford: Blackwell.
- . 2005. "The Case Against Closure." In: Steup and Sosa 2005, pp. 1–26.
- Elgin, Catherine. 1996. *Considered Judgement*. Princeton: Princeton University Press.

- . 2005. "Non-Foundationalist Epistemology: Holism, Coherence, and Tenability." In: Steup and Sosa 2005, pp. 156–167.
- Engel, Mylan. 1992. "Is Epistemic Luck Compatible With Knowledge?" *Southern Journal of Philosophy*, 30: 59–75.
- . 2003. "What's Wrong With Contextualism, and a Noncontextualist Resolution of the Skeptical Paradox." *Erkenntnis*, 61: 203–231.
- Feldman, Fred. 1986. *A Cartesian Introduction to Philosophy*. New York: McGraw Hill.
- Feldman, Richard. 1988. "Epistemic Obligations," in *Philosophical Perspectives*, 2: 235–56.
- . 1999a. "Methodological Naturalism in Epistemology." In: Greco 1999.
- . 1999b. "Contextualism and Skepticism." *Philosophical Perspectives*, 13: 91–114.
- . 2001a. "Voluntary Belief and Epistemic Evaluation." In: Steup 2001a, pp. 77–92.
- . 2001b. "Skeptical Problems, Contextualist Solutions." *Philosophical Studies*, 103: 61–85.
- . 2003. *Epistemology*. Upper Saddle River (NJ): Prentice Hall. *Philosophical Studies*, 103: 61–85.
- . 2005. "Justification is Internal." In Steup and Sosa 2005, pp. 270–284.
- Fumerton, Richard. 1995. *Metaepistemology and Skepticism*. Lanham: Rowman and Littlefield.
- . 2001. "Classical Foundationalism." In: DePaul 2001, pp. 3–20.
- Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis*, 23: 121–123. [Independent transcription in HTML available online].
- Ginet, Carl. 1975. *Knowledge, Perception, and Memory*. Dordrecht: Reidel.
- . 2005. "Infinitism is not the Solution to the Regress Problem." In: Steup and Sosa (eds.), pp. 140–149.
- Goldman, Alvin. 1976. "Discrimination and Perceptual Knowledge." *The Journal of Philosophy*, 73: 771–791.
- . 1979. "What is Justified Belief?" In: *Justification and Knowledge*, ed. George S. Pappas. Dordrecht: Reidel.
- . 1986. *Epistemology and Cognition*. Cambridge: Harvard University Press.

- . 1991. "Epistemic Folkways and Scientific Epistemology." In: *Liaisons: Philosophy Meets the Cognitive and Social Sciences*. (Cambridge: MIT Press.)
- . 1999a. "Internalism Exposed." *The Journal of Philosophy*, 96: 271–293.
- . 1999b. *Knowledge in a Social World*. Oxford: Oxford University Press.
- Greco, John. 1993. "Virtues and Vices of Virtue Epistemology," *Canadian Journal of Philosophy*, 23: 413–433.
- . 1999. "Agent Reliabilism," *Philosophical Perspectives*, 19: 273–96.
- . 2000. *Putting Sceptics in Their Place: The Nature of Skeptical Arguments and Their Role in Philosophical Inquiry*. (Cambridge: Cambridge University Press).
- . 2005. "Justification is Not Internal," in Steup and Sosa 2005, pp. 257–270.
- Greco, John and Sosa, Ernest (eds.). 1999. *The Blackwell Guide to Epistemology*. Oxford: Blackwell.
- Haack, Susan. 1993. *Evidence and Inquiry. Towards Reconstruction in Epistemology*. Oxford: Blackwell.
- . 2001. "'The Ethics of Belief' Reconsidered." In Steup 2001a, pp. 21–33.
- Harman, Gilbert. 1986. *Change in View*. Cambridge: MIT Press.
- Hawthorne, John. 2005. "The Case for Closure". In Steup and Sosa (eds.) 2005, pp. 26–43.
- . 2004. *Knowledge and Lotteries*. Oxford: Clarendon Press.
- Huemer, Michael. 2000. *Skepticism and the Veil of Perception*. New York: Rowman and Littlefield.
- Klein, Peter. 1999. "Human Knowledge and the Infinite Regress of Reasons." *Philosophical Perspectives*, 13: 297–332.
- . 2005. "Infinetism is the Solution to the Regress Problem." In Steup and Sosa (eds.) 2005, pp. 131–140.
- Kornblith, Hilary. 1999. "In Defense of a Naturalized Epistemology." In: Greco 1999.
- . 2001. *Epistemology: Internalism and Externalism*. Malden (MA): Blackwell. Oxford University Press.
- . 2002. *Knowledge and its Place in Nature*. Oxford: Oxford University Press.

- Kvanvig, Jonathan. 1996a. *The Intellectual Virtues and the Life of the Mind*. Lanham: Rowman and Littlefield.
- . 1996b. *Warrant in Contemporary Epistemology. Essays in Honor of Plantinga's Theory of Knowledge*. Lanham: Rowman and Littlefield.
- Lackey, Jennifer. 2003. "A Minimal Expression of Non-Reductionism in the Epistemology of Testimony." *Noûs*, 37: 706–723.
- Lewis, David. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy*, 74: 549–567.
- Lehrer, Keith. 1990. *Theory of Knowledge*. Boulder: Westview Press.
- Longino, Helen E. "Feminist Epistemology." In Greco and Sosa 1999, pp. 325–353.
- Lycan, William G. 1996. "Plantinga and Coherentisms." In Kvanvig 1996b, pp. 3–24.
- Moore, G.E.. 1959. *Philosophical Papers*. London: Allen and Unwin.
- Montmarquet, James. 1993. *Epistemic Virtue and Doxastic Responsibility*. Lanham: Rowman and Littlefield.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge: Harvard University Press.
- Plantinga, Alvin. 1993. *Warrant: The Current Debate*. Oxford: Oxford University Press.
- . 2000. *Warranted Christian Belief*. Oxford: Oxford University Press.
- Pollock, John. 1986. *Contemporary Theories of Knowledge*. Totowa: Rowman and Littlefield.
- Pritchard, Duncan. 2004. "Some Recent Work in Epistemology." *The Philosophical Quarterly*, 54: 605–613. [Preprint available from the author].
- . 2005. *Epistemic Luck*. Oxford: Oxford University Press.
- Pryor, James. 2000. "The Skeptic and the Dogmatist", *Noûs*, 34: 517–49. [Preprint available from the author in PDF].
- . 2004. "What's Wrong with Moore's Argument?" *Philosophical Issues*, 15: 349–378. [Preprint available from the author in PDF].
- . 2005. "There is Immediate Justification." In: Steup and Sosa 2005, pp. 181–202.
- Quine, W. V. 1969. "Epistemology Naturalized." In: *Ontological Relativity and Other Essays*. New York: Columbia Press, pp. 69–90.

- Ryan, Sharon. 2003. "Doxastic Compatibilism and the Ethics of Belief." *Philosophical Studies*, 114: 47–79.
- Russell, Bruce. 2001 "Epistemic and Moral Duty." In: Steup (ed.) 2001 a.
- . 2004. "How to be an Anti-Skeptic and a Noncontextualist." *Erkenntnis*, 61: 245–255.
- Schiffer, Stephen. 1996. "Contextualist Solutions to Skepticism." *Proceedings of the Aristotelian Society*, 96: 317–333.
- Schmitt, Frederick (ed.). 1994. *Socializing Epistemology*. Lanham: Rowman and Littlefield.
- . 1999. "Social Epistemology." In: Greco and Sosa 1999, chapter 15.
- Sellars, Wilfrid. 1963. "Empiricism and the Philosophy of Mind." In: *Science, Perception, and Reality*. London: Routledge and Kegan Paul.
- Shope, Robert K. 1983. *The Analysis of Knowing. A Decade of Research*. Princeton: Princeton University Press.
- Sosa, Ernest. 1991. *Knowledge in Perspective. Selected Essays in Epistemology*. Cambridge: Cambridge University Press.
- . 1997. "Reflective Knowledge in the Best Circles." *The Journal of Philosophy*, 96: 410–30.
- . 1999a. "Skepticism and the Internal/External Divide." In: Greco and Sosa (eds.) 1999, pp. 145–157.
- . 1999b. "How to Defeat Opposition to Moore." *Philosophical Perspectives*, 13: 141–153.
- . 2003. "Relevant Alternatives, Contextualism Included." *Philosophical Studies*, 119: 3–15.
- Steup, Matthias. 1996. *An Introduction to Contemporary Epistemology*. Upper Saddle River: Prentice Hall.
- . 1999. "A Defense of Internalism." In: Louis P. Pojman (ed.). *The Theory of Knowledge. Classical and Contemporary Readings*. Belmont: Wadsworth, pp. 373–384.
- . 2000. "Doxastic Voluntarism and Epistemic Deontology." *Acta Analytica*, 15: 25–56. [Preprint available from the author in PDF].
- \_\_\_ (ed). 2001a. *Knowledge, Truth, and Duty. Essays on Epistemic Justification, Responsibility, and Virtue*. Oxford: Oxford University Press.
- . 2001b. "Epistemic Duty, Evidence, and Internality." In: Steup 2001a.

- . 2004. "Internalist Reliabilism." *Philosophical Issues*, 14: 401–425 [Reprint available from the author in PDF].
- . 2005. "Contextualism and Conceptual Disambiguation." *Acta Analytica*, 20: 3–15 [Preprint available from the author in PDF].
- Steup, Matthias and Sosa, Ernest (eds). 2005. *Contemporary Debates in Epistemology*. Malden (MA): Blackwell.
- Stine, Gail. 1976. "Skepticism, Relevant Alternatives, and Deductive Closure." *Philosophical Studies*, 29: 249–61.
- Stroud, Barry. 1984. *The Significance of Skepticism*. Oxford: Clarendon Press.
- Swain, Marshall. 1981. *Reasons and Knowledge*. Ithaca: Cornell University Press.
- Van Cleve, James. 1985. "Epistemic Supervenience and the Circle of Beliefs," *Monist*, 68: 90–104.
- . 2005. "Why Coherence Is Not Enough: A Defense of Moderate Foundationalism." In: Steup and Sosa 2005, pp. 168–180.
- Williams, Michael. 1999a. *Groundless Belief*. Princeton: Princeton University Press (first published 1977).
- . 1999b. "Skepticism." In: Greco and Sosa 1999, pp. 35–69. 2005, pp. 202–216.
- . 2005. "Doing Without Immediate Justification." In: Steup and Sosa 2005, pp. 202–216.
- Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Zagzebski, Linda Trinkaus. 1996. *Virtues of the Mind. An Inquiry Into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.
- . 1999. "What is Knowledge?" In: Greco and Sosa 1999, pp. 92–116.
- Wolterstorff, Nicholas. 1999. *Epistemology of Religion*. In: Greco and Sosa 1999, pp. 303–324.