

Jesse J. Prinz

A CONSTRUÇÃO EMOCIONAL DA MORAL



Editora Fundação Fênix

Jesse J. Prinz é atualmente Distinguished Professor of Philosophy e Director of the Committee for Interdisciplinary Science Studies no Graduate Center da City University of New York (CUNY). O Professor Prinz tem produzido um vasto e consolidado programa de pesquisas interdisciplinares em filosofia experimental, filosofia da psicologia e filosofia moral, sendo o autor de cinco monografias (listadas logo abaixo) e de mais de uma centena de artigos em neurofilosofia e filosofia da neurociência, abordando temas tais como psicologia moral, filosofia da linguagem, estética, emoção e consciência. Jesse Prinz tem viajado e proferido palestras, comunicações e seminários em vários países do mundo, inclusive no Brasil, onde apresentou vários de seus trabalhos seminiais na PUC-Rio, na PUCRS, na Unisinos e num encontro da Sociedade Brasileira de Filosofia Analítica em Fortaleza.

O presente livro foi meticulosamente traduzido por uma equipe de pesquisadores de diferentes instituições brasileiras, todas e todos com sólida formação filosófica, aqui listados em ordem alfabética: Agemir Bavaresco (PUCRS), Aline Vanin (UFRGS), Camila Barbosa (UFSC-PUCRS), Carlos Ferraz (UFPEL), Charles Borges (PUCRS), Clademir Araldi (UFPEL), Eduardo Garcia Lara (PUCRS), Flávia Carvalho (UFPEL), Juliano do Carmo (UFPEL) e Nythamar de Oliveira (PUCRS). A equipe tradutora se beneficiou, ao longo das várias etapas de revisão e editoração, da assistência técnica e da assessoria generosa das seguintes pessoas: Calhandra Pinter Santos, Débora Abreu, Fabricio Pontin, Ibsen Boff e Rogel Esteves de Oliveira.

Agradecemos a Oxford University Press e o Professor Jesse Prinz pela sua generosa atenção e aquiescência em viabilizar a publicação desta importante obra em língua portuguesa. Agradecemos a Editora Fundação Fênix e os revisores que tornaram esta publicação possível.

(Nythamar de Oliveira).



Editora Fundação Fênix



A construção emocional da moral

Jesse J. Prinz

Série Filosofia

Conselho Editorial

Editor

Agemir Bavaresco

Conselho Científico

Agemir Bavaresco – Evandro Pontel
Jair Inácio Tauchen – Nuno Pereira Castanheira

Conselho Editorial

Augusto Jobim do Amaral
Cleide Calgaro
Draiton Gonzaga de Souza
Evandro Pontel
Everton Miguel Maciel
Fabián Ludueña Romandini
Fabio Caprio Leite de Castro
Fábio Caires Coreia
Gabriela Lafetá
Ingo Wolfgang Sarlet
Isis Hochmann de Freitas
Jardel de Carvalho Costa
Jair Inácio Tauchen
Jozivan Guedes

Leno Francisco Danner
Lucio Alvaro Marques
Nelson Costa Fossatti
Norman Roland Madarasz
Nuno Pereira Castanheira
Nythamar de Oliveira
Orci Paulino Bretanha Teixeira
Oneide Perius
Raimundo Rajobac
Renata Guadagnin
Ricardo Timm de Souza
Rosana Pizzatto
Rosalvo Schütz
Rosemary Sadami Arai Shinkai
Sandro Chignola

A construção emocional da moral

Jesse J. Prinz



Editora Fundação Fênix

Porto Alegre, 2022

Agradecemos a Oxford University Press e ao Professor Jesse Prinz pela sua generosa atenção e aquiescência em viabilizar a publicação desta importante obra em língua portuguesa

Tradutores: Nythamar de Oliveira (PUCRS), Camila Barbosa (PUCRS), Juliano do Carmo (UFPEL), Flávia Carvalho (UFPEL), Carlos Ferraz (UFPEL), Agemir Bavaresco (PUCRS), Eduardo Garcia Lara (PUCRS), Aline Vanin (UFRGS), Clademir Araldi (UFPEL), Charles Borges (PUCRS). A equipe tradutora agradece o apoio recebido das seguintes pessoas, durante o processo de revisão da presente tradução: Calhandra Pinter Santos, Ibsen Boff, Débora Abreu, Fabricio Pontin, Rogel Esteves de Oliveira.

Revisão técnica

Marloren Lopes Miranda

*Para os meus pais, Phyllis e Jonathan Prinz,
que me ensinaram a diferença entre o certo e o errado.*

Direção editorial: Agemir Bavaresco
Diagramação: Editora Fundação Fênix
Capa: Editora Fundação Fênix

Disponível em:

<https://commons.wikimedia.org/w/index.php?search=Theodor+de+Bry&title=Special:MediaSearch&type=image>

O padrão ortográfico, o sistema de citações, as referências bibliográficas, o conteúdo e a revisão de cada capítulo são de inteira responsabilidade de seu respectivo autor.

Todas as obras publicadas pela Editora Fundação Fênix estão sob os direitos da Creative Commons 4.0 –

[Http://creativecommons.org/licenses/by/4.0/deed.pt_BR](http://creativecommons.org/licenses/by/4.0/deed.pt_BR)

Obra editada com apoio - CAPES/PROEX - Auxílio N° 0569/2018, Processo N° 23038.001026/2018-81



Série Filosofia – 17

Catálogo na Fonte

P957c PRINZ, Jesse J.
A construção emocional da moral [recurso eletrônico] / Jesse J. Prinz,
tradutores: Nythamar de Oliveira ... [et al.] – Porto Alegre : Editora
Fundação Fênix, 2022.
517 p. (Série Filosofia ; 17)

Título original: The emotional construction of morals.
Demais tradutores: Camila Barbosa, Juliano do Carmo, Flávia
Carvalho, Carlos Ferraz, Agemir Bavaresco, Eduardo Garcia Lara, Aline
Vanin, Clademir Araldi, Charles Borges.
Disponível em: <<http://www.fundarfenix.com.br>>
ISBN 978-65-5460-013-2
DOI <https://doi.org/10.36592/9786554600132>

1. Moral. 2. Ética. 3. Filosofia moral. 4. Psicologia – Filosofia. I.
Oliveira, Nythamar de (trad.). II. Título

CDD: 170

Responsável pela catalogação: Lidiane Corrêa Souza Morschel CRB10/1721

SUMÁRIO

APRESENTAÇÃO

Nythamar de Oliveira 11

PREFÁCIO 15

PREÂMBULO 19

O NATURALISMO E A LEI DE HUME 19

PARTE I

MORALIDADE E EMOÇÃO 33

1. EMOCIONISMO 35

2. EMOÇÕES NÃO-MORAIS E MORAIS 93

3. SALVANDO A SENSIBILIDADE 151

4. CONTRA A OBJETIVIDADE 229

PARTE II

CONSTRUINDO A MORAL 277

5. JANTANDO COM CANIBAIS 279

6. A GENEALOGIA DA MORAL 343

7. OS LIMITES DA ÉTICA EVOLUCIONISTA 387

8. PROGRESSO MORAL 457

REFERÊNCIAS 489

APRESENTAÇÃO

O presente livro aqui traduzido, consagrado *bestseller* mundial de um dos mais impactantes e respeitados filósofos da atualidade, Jesse Prinz, integra uma trilogia que busca reabilitar criticamente o empirismo do *Tratado da Natureza Humana*, de 1739-40 (*A Treatise of Human Nature*), de autoria do pensador escocês David Hume, que o havia concebido em três volumes: *Do Entendimento*, *Das Paixões* e *Da Moral* (*Of the Understanding, Of the Passions, Of Morals*). Como anunciado no próprio subtítulo da publicação original, *Sendo uma tentativa de introduzir o método experimental de raciocínio em assuntos morais* (*Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*), trata-se de elaborar uma teoria empirista dos conceitos, no primeiro tratado, e uma teoria das emoções no segundo, antes de articulá-las no terceiro tratado, de modo a contestar concepções racionalistas que contrapõem de forma dualista a razão aos afetos humanos. Segundo Prinz, trata-se, em última análise, de mostrar que todos os nossos conceitos e princípios morais nos remetem a uma base emocional. Em sua reapropriação crítica do empirismo humeano, Prinz mostra de forma assaz convincente que o naturalismo de Spinoza é reformulado por Hume e Nietzsche através de uma concepção sentimentalista da moral que mantém elementos cognitivistas e não-cognitivistas de estados emotivos, resultando numa inovadora concepção de um empirismo por transformação (*transformation empiricism*), que o filósofo nova-iorquino acertadamente identifica como a mais plausível reformulação metafilosófica (ou metafísica, num sentido pós-kantiano) para a pesquisa ético-moral após a chamada guinada neurocientífica (*neuroscientific turn*) no final do século passado e as novas tecnologias que dominam os cenários científicos e acadêmicos de nosso século. O próprio Jesse Prinz defende nesta obra o seu sentimentalismo construtivo, coerente e organicamente integrado à sua concepção de empirismo por transformação, embora o termo “empirismo” ocorra uma única vez neste terceiro livro de sua trilogia.

Assim como o neurocientista português António Damásio e os neurofenomenólogos Dan Zahavi e Shaun Gallagher, Jesse Prinz desenvolve uma

filosofia da neurociência bastante próxima do que tem sido caracterizada como uma abordagem neuro-fenomenológica da filosofia da mente e da linguagem, na medida em que propõem uma reformulação da cognição humana em termos inseparáveis da perspectiva (fenomenológica) de primeira pessoa, de seu corpo próprio e de sua inserção no mundo, isto é, em termos incorporados, arraigados, estendidos, enativos e afetivos, a chamada "cognição 4EA" (*embodied, embedded, extended, enactive, and affective cognition*), com foco particular em suas respectivas concepções dos papéis co-constitutivos desempenhados pelas emoções, corporeidade, intersubjetividade e socialidade no processo mesmo de individuação autoconsciente. Em sua trilogia, Prinz logra demonstrar em que sentido uma teoria do eu, da consciência e da natureza humana é hoje inseparável de uma teoria das emoções, assim como qualquer teoria hodierna da normatividade – moral, legal, econômica e política – prescinde de pressupostos naturalistas na medida em que o cérebro e o sistema nervoso se integram na construção social de crenças compartilhadas (inclusive crenças epistêmicas e religiosas). Prinz rejeita, destarte, versões metafísicas, reducionistas e naturalistas dogmáticas que podem ser agora sistematicamente revisitadas à luz da neurociência e de descobertas das ciências empíricas do comportamento, reabilitando um naturalismo de transformação, à medida em que mudamos nossos pontos de vista e transformamos nossa inserção no mundo.

Jesse J. Prinz é atualmente Distinguished Professor of Philosophy e Director of the Committee for Interdisciplinary Science Studies no Graduate Center da City University of New York (CUNY). O Professor Prinz tem produzido um vasto e consolidado programa de pesquisas interdisciplinares em filosofia experimental, filosofia da psicologia e filosofia moral, sendo o autor de cinco monografias (listadas logo abaixo) e de mais de uma centena de artigos em neurofilosofia e filosofia da neurociência, abordando temas tais como psicologia moral, filosofia da linguagem, estética, emoção e consciência. Jesse Prinz tem viajado e proferido palestras, comunicações e seminários em vários países do mundo, inclusive no Brasil, onde apresentou vários de seus trabalhos seminiais na PUC-Rio, na PUCRS, na Unisinos e num encontro da Sociedade Brasileira de Filosofia Analítica em Fortaleza.

O presente livro foi meticulosamente traduzido por uma equipe de pesquisadores de diferentes instituições brasileiras, todas e todos com sólida formação filosófica, aqui listados em ordem alfabética: Agemir Bavaresco (PUCRS), Aline Vanin (UFRGS), Camila Barbosa (UFSM-PUCRS), Carlos Ferraz (UFPEl), Charles Borges (PUCRS), Clademir Araldi (UFPEL), Eduardo Garcia Lara (PUCRS), Flávia Carvalho (UFPEl), Juliano do Carmo (UFPEl) e Nythamar de Oliveira (PUCRS). A equipe tradutora se beneficiou, ao longo das várias etapas de revisão e editoração, da assistência técnica e da assessoria generosa das seguintes pessoas: Calhandra Pinter Santos, Débora Abreu, Fabricio Pontin, Ibsen Boff e Rogel Esteves de Oliveira.

Agradecemos a Oxford University Press e o Professor Jesse Prinz pela sua generosa atenção e aquiescência em viabilizar a publicação desta importante obra em língua portuguesa. Agradecemos a Editora Fundação Fênix e os revisores que tornaram esta publicação possível.

Nythamar de Oliveira
PUCRS – CNPq.

Obras de autoria de Jesse Prinz:

Furnishing the Mind: Concepts and Their Perceptual Basis. MIT Press, 2002.

Gut Reactions: A Perceptual Theory of Emotion. Oxford University Press, 2004.

The Emotional Construction of Morals. Oxford University Press, 2004.

Beyond Human Nature. Penguin/Norton: 2012.

The Conscious Brain. Oxford University Press, 2012.

PREFÁCIO

O *Tratado da Natureza Humana*, de David Hume, é dividido em três livros: "Do entendimento", "Das paixões" e "Da moral". Pode-se perguntar como esses tópicos díspares estão relacionados, a não ser em virtude do fato de que eles têm algo a ver com a mente. Mas os *links* ficam claros ao ler o texto. Hume desenvolve uma teoria de conceitos (ou "ideias") no primeiro livro e uma teoria de emoções no segundo livro, e então ele os integra no terceiro ao argumentar que nossos conceitos morais têm uma base emocional. O projeto também é unificado pela fidelidade de Hume ao empirismo. Sua teoria dos conceitos baseia-se na premissa de que as ideias são cópias armazenadas de impressões sensoriais, e sua teoria das emoções é projetada para ser compatível com essa visão empirista (ele define emoções como impressões de impressões). A teoria moral de Hume também é empirista. Conceitos morais parecem especialmente problemáticos para um empirista, porque não pode haver imagem de virtude, nem gosto de bondade, nem cheiro do mal. Ao apelar para os sentimentos, Hume é capaz de argumentar que todos os conceitos estão em impressões, afinal. O conceito de bondade consiste em um sentimento de aprovação e o conceito de maldade consiste em um sentimento de desaprovação. A classe de virtudes não tem aparência comum, mas as coisas boas parecem certas; a classe dos vícios seria impossível expor, mas cada instância provoca uma pontada palpável de culpa. Em suma, o *Tratado* de Hume tem uma estrutura coerente, e a teoria moral culminante pode ser lida como a resolução de um aparente contraexemplo à sua teoria dos conceitos, ou como a recompensa para aqueles que tomam tempo para entender como a mente funciona. Não importa onde você coloque a ênfase, a teoria dos conceitos de Hume e sua teoria da moral estão unidas, e as paixões são a cola.

Os filósofos gostam de reinventar a roda e eu não sou exceção. As visões que eu defendo aqui têm uma tremenda dívida com Hume. Este livro defende uma teoria sentimentalista da moralidade que se baseia nas ideias desenvolvidas por Hume e alguns de seus contemporâneos. Parto de Hume de várias maneiras, mas o impulso básico da teoria é humeano e, a esse respeito, minhas propostas são notas

de rodapé do Livro III do *Tratado*. E esta não é a primeira nota de rodapé humeana que eu escrevi. Meu primeiro livro, *Furnishing the Mind*, defende uma teoria empirista de conceitos, e meu segundo livro, *Gut Reactions*, defende uma teoria empirista das emoções (que é mais jamesiana que humeana, mas, com Hume, meu objetivo é mostrar que as emoções são um tipo de impressão). Então, aqui, no meu terceiro livro, estou simplesmente concluindo uma trilogia que acompanha a estrutura do *Tratado* de Hume. Essas obras são independentes em um sentido - você pode rejeitar uma enquanto aceita as outras -, mas elas se encaixam exatamente da maneira que o *Tratado* de Hume se encaixa. Eu os vejo como partes de um todo, e vejo isso como um tributo e modesta extensão da obra-prima de Hume.

Eu tenho três objetivos principais em estender o projeto de Hume. O primeiro é fornecer suporte empírico para uma teoria que foi desenvolvida a partir de um gabinete. O segundo é adicionar alguns detalhes à teoria de Hume, incluindo um relato dos sentimentos que sustentam nossos julgamentos morais e um relato da ontologia que resulta de levar a sério uma visão sentimentalista. Meu terceiro objetivo é mostrar que essa abordagem leva ao relativismo moral. Hume resistiu ao relativismo e eu argumento que ele não deveria. Também investigo a origem de nossos sentimentos morais e sugiro que a abordagem genealógica de Nietzsche da moralidade tenha muito a contribuir aqui. A história resultante é metade humeana e metade nietzscheana, mas considero a parte nietzscheana adequada à parte humeana.

Eu menciono Hume e Nietzsche por meio de reconhecimento. Dentro do panteão dos filósofos mortos, eles são aqueles a quem devo as maiores dívidas filosóficas. Também devo mencionar Edward Westermarck, porque ele reconheceu a ligação entre o sentimentalismo e o relativismo há cem anos e reconheceu o valor da antropologia e da história na investigação da moral. Este livro continua na tradição de Westermarck. Entre os filósofos vivos, eu fui especialmente inspirado por Gil Harman, Shaun Nichols, David Wiggins e John McDowell. Steve Stich também merece menção especial por seus esforços para promover uma abordagem da filosofia que faz uso liberal de resultados empíricos. Nessa nota, também devo enormes dívidas aos cientistas que forneceram dados para ajudar a

avaliar as teorias filosóficas. Entre os psicólogos, Jon Haidt e James Blair foram especialmente influentes, e eu também destacaria o falecido Marvis Harris, cujo materialismo cultural deixa sua marca na segunda metade deste livro. Esses autores me educaram através de seu trabalho publicado, mas muitos outros ofereceram orientação por meio de discussões e comentários escritos sobre o material deste livro. Eu me beneficieei de dar palestras em numerosos departamentos de filosofia e conferências, abrangendo quatro continentes e o dobro disso de países. Gostaria de poder listar o nome de todos que ofereceram sugestões ou objeções ao longo do caminho. Eu também quero agradecer a todos os membros do Grupo de Pesquisa de Psicologia Moral, que criaram um dos ambientes mais propícios para a troca de ideias filosóficas que eu já vi. Eu também me beneficieei de *feedbacks* por escrito, o que levou a melhorias grandes e pequenas por toda parte. Neste contexto, permitam-me mencionar em primeiro lugar os participantes em seminários ministrados por Steve Stich, Eric Schwitzgebel e John Mikhail, que enfrentaram versões anteriores deste manuscrito ou artigos relacionados. Também recebi correções filosóficas e tipográficas sobre o manuscrito inteiro de Nigel Hope, Mark Jenkins e Jonathan Prinz, bem como comentários úteis sobre peças selecionadas ou materiais relacionados de Ruth Chang, Matthew Chrisman, Justin D'Arms, Karen Jones e Matt Smith, Valerie Tiberius, Teemu Toppinen, Brian Weatherson e outros a quem, sem dúvida, estou esquecendo. Entre os leitores, minha maior dívida vai para Shaun Nichols, Richard Joyce e dois pareceristas anônimos da Oxford University Press, que me forneceram comentários detalhados sobre rascunhos do manuscrito. Cada um deles detectou erros embaraçosos e me pressionou em dezenas de questões filosóficas. O livro é muito melhor por causa deles, e teria sido melhor ainda se eu tivesse mais sucesso em acomodar todas as suas sugestões. Eu ficarei eternamente grato. É claro que eu não teria recebido um *feedback* tão útil se não fosse pelo meu paciente e excelente editor, Peter Momtchiloff. Peter tem sido uma grande fonte de apoio em todas as etapas.

Ao escrever este livro, também me beneficieei de várias instituições. Fui pesquisador no Collegium Budapest e escrevi algumas coisas lá. Tamar Gendler foi fundamental na orquestração dessa visita e na montagem de um grupo

maravilhoso de colegas de verão. Também devo agradecimentos especiais ao Centro de Estudos Avançados em Ciências Comportamentais, em Palo Alto. CASBS é um lugar mágico, e eu terminei este manuscrito lá. Ao fazê-lo, beneficiei-me dos abundantes recursos intelectuais e do excelente pessoal, que contribuem para torná-lo um ambiente ideal para a pesquisa. Eu pude ir para CASBS por causa de uma licença de pesquisa da minha instituição de origem, a Universidade da Carolina do Norte em Chapel Hill. Sou grato à UNC por isso, mas também e especialmente aos meus alunos e colegas. Não há lugar melhor para se trabalhar.

Finalmente, queria mencionar minha família. Sinto-me feliz por ter sido criado por dois pais com fortes convicções morais e cresci ao lado de um irmão mais velho com um senso moral aguçado. Minhas visões sobre o certo e o errado seriam muito diferentes sem eles, e eles continuam a fornecer apoio de muitas maneiras. Como sempre, minha profunda gratidão vai para Rachel, que estava por perto enquanto eu escrevia quase todas as páginas deste livro, e ela tem pacientemente apoiado cada balanço de humor que vem junto com o processo de escrita. O seu apoio foi decisivo.

PREÂMBULO

O NATURALISMO E A LEI DE HUME

Moralidade é um domínio normativo. Refere-se a como o mundo deve ser, e não como é. Uma investigação sobre a moralidade parece precisar de uma metodologia que difere dos métodos utilizados pelas ciências. Ao menos, esse parece ser o caso se o pesquisador tem ambições normativas. Se o pesquisador quer prescrever, descrever não é o suficiente. Como Hume nos ensinou, não há como derivar um *deve* de um *é*. Mais precisamente, não há como deduzir uma sentença que tem força prescritiva (uma sentença que expressa uma obrigação incondicional) de uma sentença que é puramente descritiva. Nenhum fato sobre como o mundo é implica que se deva evitar roubar ou matar ou destruir prédios. A lei de Hume é atraente porque faz a moralidade parecer especial; verdades morais não são iguais às verdades frias da ciência. Mas, em uma leitura, a lei de Hume é uma receita de niilismo moral. Ao isolar verdades morais do método científico, pode-se dizer que a moralidade é supernatural. Se este é o caso, então a moralidade deveria seguir o mesmo caminho que espíritos e fantasias. Esse é um caminho que eu quero resistir.

Defensores da lei de Hume reconhecem a viabilidade de certos tipos de projetos descritivos para a moralidade. Alguém poderia descrever as convicções morais que obteve em uma cultura. Alguém poderia descrever a natureza dos conceitos que dispõem quando fazem julgamentos morais. Alguém poderia dizer algo de descritivo sobre a natureza dos fatos morais e como eles se relacionam aos outros tipos de fatos. Essas são as questões que me nortearão. Mas, quero começar discutindo como as verdades descritivas sobre a moralidade carregam o prescritivo. A teoria metaética e a psicologia moral que irei defender nos capítulos a seguir oferecem uma maneira de cruzar as restrições do *é/deve*.

Meu argumento é que a moralidade deriva de nós. O bem é aquele que nós reconhecemos enquanto bem. As obrigações são aquelas que reconhecemos enquanto obrigatórias. O "nós", aqui, refere-se a uma pessoa apresentando um argumento moral e o grupo cultural com o qual aquele indivíduo se afilia. Se o bem é aquilo que reconhecemos como bem, então conseguimos descobrir quais são

nossas obrigações ao entender com quais crenças morais estamos comprometidos. Descobrir o que acreditamos sobre a moralidade é uma tarefa descritiva por excelência, e que podemos buscar empiricamente com sucesso. Assim, ética normativa pode ser acordada pelas ciências sociais.

Essa sugestão é difícil de enquadrar com a intuição subjacente da lei de Hume. Há uma intuição persistente de que nenhum fato possível de ser descoberto empiricamente sobre nossas crenças possa deduzir que devemos nos comportar de certa forma. Não quero esmagar essa intuição. A lei de Hume é verdadeira em um sentido, e falsa em outro. Isto é o que espero demonstrar aqui. Mais precisamente, quero apresentar como alguém totalmente naturalista – que repele espíritos e fantasias – pode encontrar um lugar para a normatividade. Considero Hume tal naturalista, e defenderei uma visão da moralidade que está profundamente em dívida com Hume. A visão que defendo preserva muitas das nossas intuições sobre o domínio moral, mas não todas. Eu rejeito o niilismo, mas reconheço o subjetivismo, o relativismo e o irracionalismo. Moralidade é uma construção humana que surge das nossas paixões. Mas isso não significa que devemos desistir delas.

1.1 QUATRO TIPOS DE NATURALISMO

O termo “naturalismo” é usado de diversas maneiras; às vezes, com uma entonação depreciativa, às vezes, como um grito de guerra. Quero debater quatro tipos diferentes de naturalismo, todos os quais eu apoio. Não irei argumentar em prol do naturalismo aqui. Vou simplesmente alegar minha aliança.

Um tipo de naturalismo, já sugerido pelas minhas considerações sobre espíritos e fantasias, é mais bem entendido em contraste com o supernaturalismo. É a visão de que o nosso mundo é limitado pelos postulados e leis das ciências naturais. Nada que viola estas leis pode existir, e todas as entidades que existem devem, em alguma medida, ser compostas pelas entidades que nossas melhores teorias científicas exigem. Esta é uma tese metafísica; refere-se à natureza fundamental da realidade. Irei chamá-la de naturalismo metafísico.

Naturalismo metafísico implica em um tipo de naturalismo explanatório. Se tudo que existe é composto por fatores naturais e restritos às leis naturais, então tudo que não é descrito na linguagem de uma ciência natural precisa, em última instância, ser descritível por tais termos. Isto não é equivalente ao reducionismo no sentido forte da palavra. Reducionistas fortes argumentam que a relação entre ciência natural e domínios de "alto-nível" é dedutiva. Deveríamos ser capazes de deduzir fatos de "alto-nível" de substratos de "baixo-nível". Antirreducionistas negam isso. Pensam, por exemplo, que há leis de alto-nível ou generalizações que poderiam ser implementadas em infinitas maneiras. Regularidades capturadas em um nível baixo não teriam generalizações desse tipo. A explicação naturalista pode ser uma explicação antirreducionista. A explicação naturalista não precisa argumentar que explicações de baixo-nível são as únicas. A ideia central é que precisa haver algum tipo de correspondência sistemática entre os níveis. Deve ser possível de mapear qualquer entidade em um nível alto para entidades em um nível inferior, e deve ser capaz de explicar a instanciação de qualquer generalização de alto nível apelar para recursos de nível inferior que realizam essas generalizações.

Um terceiro tipo de naturalismo pode ser denominado metodologicamente. Se todos os fatos são, em algum sentido, fatos naturais (de acordo com o naturalismo metafísico), então os métodos com os quais investigamos fatos devem ser aplicáveis em uma investigação de fatos naturais. Filósofos, às vezes, argumentam ter um método distinto para fazer descobertas: o método da análise conceitual. Se o naturalismo metafísico é verdade, este não pode ser um método supernatural para descobrir verdades supernaturais. Conceitos eles mesmos são entidades naturais, e eles podem ser investigados usando processos naturais. Análise conceitual é, como todas as ferramentas legítimas de investigação, um método empírico. Como todos os métodos empíricos, não é especialmente poderoso. Análises conceituais procedem do acesso a estruturas psicológicas através da primeira pessoa, ou seja, da introspecção. A introspecção é suscetível a erro, e há riscos metodológicos associados a tirar conclusões de investigações usando um único sujeito (uma pessoa). Podemos investigar conceitos usando ferramentas das ciências sociais. Se conceitos são entidades naturais, então eles surgem de maneira natural. Por exemplo, conceitos podem ser adquiridos pela

experiência, e eles podem ser revisados pela experiência. Eles não têm *status* especial quando servem para revelar fatos sobre o mundo.

Naturalismo metodológico, na forma como eu defino, está associado com Quine. Na sua crítica epistemológica (1969), Quine argumenta que a investigação do conhecimento deveria ser buscada usando recursos das ciências sociais. Na sua defesa da confirmação holística (1953), Quine afirma que todos os argumentos estão sujeitos à revisão empírica. Há um terceiro tipo de naturalismo associado ao holismo de Quine. Estamos sempre operando a partir da nossa teoria atual de mundo. Ao fazer revisões teóricas, podemos nos deslocar das nossas teorias e adotar uma instância transcendental. Para fazê-lo, deveríamos supor que temos uma maneira de pensar sobre o mundo que é independente de nossas teorias do mundo. Se teorias do mundo englobam todas as nossas crenças, então essa instância não é possível. Chame isso de naturalismo transformador, porque é a perspectiva sobre como nós mudamos nossa visão.

Cada forma de naturalismo tem implicações para normatividade. Naturalismo metafísico implica que normas morais, se elas existem, não precisem postular qualquer coisa que vá além daquilo que as ciências naturais permitem. Naturalismo explanatório implica que podemos, em última instância, descrever como qualquer norma moral é realizada por entidades naturais. Naturalismo metodológico implica que devemos investigar normas usando todos os recursos empíricos disponíveis. Naturalismo transformador implica que devemos investigar normas usando nosso sistema de crenças atual, e, como resultado, as normas que aceitamos atualmente vão influenciar nossas intuições sobre quais normas devemos defender. Se escolhermos mudar nossas normas, não podemos fazê-lo adotando uma instância transcendental que coloque em parênteses as normas que atualmente aceitamos.

1.2 TRANSGREDINDO A LEI DE HUME

Se o naturalismo está certo, então fatos morais são fatos naturais, ou eles não são sequer fatos. Fatos naturais são fatos que são consistentes com as quatro estruturas do naturalismo recém descritas. O mundo é como é, e não de outra

forma. Se o mundo inclui fatos sobre o que deve ser, esses fatos devem ser explicados em termos de como as coisas são. Todo *deve* precisa seguir de um *é*. Já que o naturalismo não implica reducionismo, naturalismo não implica que fatos prescritivos são outra forma de fatos descritivos. Todo fato prescritivo precisa ser realizado por, ou ser verdade através de, fatos que podem ser descritos sem o uso de um vocabulário prescritivo. Para cada fato prescritivo há algum fato descritivo subjacente que o torna verdade. Acontece que eu acho que o naturalismo nos permite inferir fatos prescritivos de fatos normativos, e, então, há uma forma de transgredir a lei de Hume. Mas o naturalismo não implica que a lei de Hume é violável, por razões que discutirei na próxima sessão.

Primeiro, quero oferecer um argumento rápido e sujo para como derivar um *deve* de um *é*. Uma defesa completa desse argumento precisaria de uma explicação mais elaborada sobre filosofia da linguagem. Meu objetivo aqui é modesto. Quero indicar um caminho pelo qual um naturalista pode reconhecer simultaneamente fatos morais como naturais (uma vez que implique fatos descritivos), mas também irreduzíveis (e, assim, não tão vinculados). Os argumentos nessa seção e na próxima ilustram como essas demandas aparentemente paradoxais podem ser atendidas.

Para ver como um *deve* pode ser derivado de um *é*, nós precisamos primeiro descobrir quais são esses *devem*. A maneira de fazer isso é descobrir o que a palavra 'deve' significa (aqui me restrinjo ao uso moral de 'deve'). Qual conceito essa palavra expressa? Para responder essa questão, nós precisamos da psicologia (introspectiva ou outra). Nós precisamos determinar o que as pessoas têm em mente quando elas dizem que algo é obrigatório. Boa parte desse livro é sobre essa questão. Por enquanto, quero desenhar uma versão muito simplificada do tipo de resposta que quero defender. Na teoria que eu defendo, quando uma pessoa diz que um curso de ação é obrigatório, aquele julgamento expressa o que pode ser chamado de sentimento prescritivo. Um sentimento prescritivo é um completo dispositivo emocional. Se alguém tem esse sentimento sobre uma forma particular de conduta, então está disposto a se engajar naquela conduta, e está disposto a se sentir mal se alguém não age de acordo. Também estará disposto a condenar aqueles que não se engajam nessa forma de conduta. Suponha que Smith

honestamente julgue que se deve doar para a caridade. Smith está expressando um sentimento que o dispõe a se sentir mal se ele não doa para a caridade ou irritado se você não doa para a caridade. Isso se assemelha a visão filosófica denominada emotivismo, mas, como vai ficar claro no capítulo 3, minha perspectiva difere em detalhes importantes.

Muitos aperfeiçoamentos se seguirão nos capítulos a seguir. Quero debater aqui as implicações. Se a palavra 'deve' expressa um sentimento prescritivo, então é isso que a palavra significa. O conceito subjacente à palavra não pode ser nada mais do que aquilo que expressamos ao usá-la. Assim, se esta teoria psicológica simplificada está certa, então nós aprendemos o que significa dizer que alguém deve fazer alguma coisa. Aprendemos que condições satisfazem o julgamento de que algo é obrigatório.

Agora estamos em uma posição de tentar conseguir um *deve* de um *é*. Ofereço o seguinte argumento:

1. Smith tem obrigação de doar à caridade se 'Smith deve doar à caridade' é verdade.
2. 'Smith deve doar à caridade' é verdade, se a palavra 'deve' expressa um conceito que se aplica à relação de Smith com doar à caridade.
3. A palavra 'deve' expressa um sentimento prescritivo.
4. Smith tem um sentimento prescritivo para com quem doa à caridade.
5. Então, a sentença 'Smith deve doar à caridade' é verdade.
6. Então, Smith tem uma obrigação de doar à caridade.

A conclusão desse argumento é um fato prescritivo. As premissas são descritivas. A palavra 'deve' é mencionada, mas nunca usada. A lei de Hume foi violada.

Meu argumento contrasta com aquele defendido por Searle (1964). Searle também buscou uma estratégia metalinguística. Simplificando um pouco, ele argumenta que, quando as pessoas proferem uma sentença na forma 'Eu prometo fazer X', aquela pessoa se coloca sob uma obrigação. Isto faz parte do significado de prometer. Assim, Searle infere que uma pessoa que se coloca sob uma obrigação

está sob aquela obrigação. Não estou convencido do argumento de Searle. Pode haver problemas nos dois passos (para uma crítica mais aprofundada ver Downing, 1972). Prometer é apenas se colocar sob uma obrigação se as pessoas devem cumprir suas promessas. Então, há uma premissa normativa suprimida. O movimento de se colocar sob uma obrigação para estar sob uma obrigação também é suspeito. Se colocar sob uma obrigação pode ser interpretado como uma convenção social. Pode ser uma questão de ser reconhecido sob uma obrigação aos olhos de uma comunidade. Uma comunidade pode reconhecer pessoas como tendo uma obrigação – podem colocá-la sob uma obrigação – mesmo que as pessoas não sejam realmente obrigadas.

Acredito que precisamos de uma metalinguística mais forte do que a proposta por Searle. Precisamos de uma teoria substantiva do significado de termos normativos. A premissa 3 no meu argumento articula essa teoria. É onde toda a ação está. As outras premissas são difíceis de negar. A premissa 3 é controversa, e um dos objetivos do capítulo a seguir é fornecer argumentos que sejam mais convincentes. Mas antecipo que o argumento pode ser modificado para acomodar outras teorias. Se o naturalismo moral é verdadeiro, então conceitos morais são ou vácuos, ou eles expressam propriedades que podem ser, em última instância, descritos sem um vocabulário moral. Se minha análise do dever está incorreta, substitua por outra análise, e troque a premissa 3 pela descrição correspondente do fato moral subjacente a uma obrigação. Agora reformule a premissa 4 de acordo, e o argumento irá se seguir. Se há obrigações, então elas podem ser derivadas nesta forma puramente descritiva em qualquer perspectiva naturalista.

1.3 SALVANDO A LEI DE HUME

Isso tudo é um pouco inquietante. Primeiramente, há uma intuição favorecendo a lei de Hume. Parece haver um salto lógico das premissas sobre como as coisas são para conclusões sobre como as coisas deveriam ser. Segundo, a teoria de normas dada na premissa 3 torna muito fácil derivar obrigações. Uma pessoa sadista pode ter um sentimento prescritivo sobre fazer uma pessoa sofrer.

O argumento apresentado implicaria que um sadista é obrigado a ser cruel. Alguma coisa deve ter dado errado.

Acredito que essa preocupação pode ser resolvida. Em relação à primeira questão, começo apontando que o argumento que ofereci não viola a lei de Hume. O argumento mostra como podemos usar premissas descritivas para derivar fatos prescritivos, mas a frase “fatos prescritivos” parece ser ambígua. Por um lado, um fato prescritivo é apenas um fato sobre o que alguém é obrigado a fazer. Contudo, um fato prescritivo também pode ser interpretado como um julgamento prescritivo ou, mais sucintamente, como uma prescrição. Repare como a conclusão é expressa no argumento acima. Eu disse ‘Smith tem uma obrigação de doar à caridade’. Eu não disse ‘Smith deve doar à caridade’. Realmente, o argumento em si mostra por que essa conclusão não se segue. ‘Deve’ expressa um sentimento prescritivo. Apenas pode ser usado verdadeiramente por um falante que tenha esse sentimento. Nenhuma premissa no argumento implica que eu, o autor do argumento, tenha qualquer disposição para reagir emocionalmente à caridade. Então, nenhuma premissa no meu argumento poder implicar, na minha voz, que Smith deve doar à caridade. Se ‘deveres’ são prescrições, então não demonstrei como derivar um *deve* de um *é*. A premissa 3, que dá significado ao dever, demonstra porque tal derivação não funciona. Aquela premissa não anula a lei de Hume, ela é a chave para defendê-la.

No final da última seção, argumentei que a premissa 3 poderia ser substituída por premissas descrevendo outras teorias naturalistas de termos normativos. Outras teorias não implicam necessariamente o resultado que apresentei. Elas não explicam necessariamente porque não há interferência direta da obrigação para o dever. É uma vantagem da abordagem que defendo que explica por que a lei de Hume é tão convincente. Reivindicações normativas parecem não ser derivadas de reivindicações descritivas, porque não há forma de derivar um sentimento prescritivo. Identificar conceitos normativos com sentimentos prescritivos captura a verdade na lei de Hume.

Alguém pode contestar que minha tentativa de salvar Hume não pode funcionar porque viola princípios básicos da semântica. No argumento acima, o passo final se move da premissa semântica de que ‘Smith deve doar à caridade’ é

verdade para o argumento de que Smith tem uma obrigação de doar à caridade. Poder-se-ia pensar que a premissa semântica implica algo mais forte. Se 'Smith deve doar à caridade é verdade', então Smith deve doar à caridade. Está é apenas uma instância de descitação. Sempre podemos inferir P de 'P' é verdadeiro. Ou será que podemos? Acredito que o argumento que apresentei é um contraexemplo do princípio descitacional. Esta não é uma tarefa difícil de encarar, porque existem outros contra exemplos. Suponha que Smith profere a sentença 'Eu sou Smith'. Essa sentença é verdadeira. Não se segue que eu veja Smith. Descitação não permite que sempre usemos indexicais como 'Eu'. Acredito que 'deve' é como um indexical em que o seu significado não está esgotado pela sua contribuição para uma proposição expressa. Argumentarei essa conclusão no capítulo 5. Por enquanto, o caso "Eu" simplesmente mostra que descitação tem exceções bem conhecidas. Se 'deve' é uma exceção, e funciona como um "Eu", então meu argumento é convincente.

O fato de não podermos derivar deveres pode parecer pouco reconfortante para alguns. Não é suficientemente ruim que nós possamos inferir obrigações? Inferir obrigações de premissas descritivas é um pouco perturbante, mas acredito que agora podemos diagnosticar o porquê. Ficamos desconfortáveis afirmando que as pessoas têm obrigações que nós não endossamos. Não gostaríamos de afirmar que sadistas são obrigados a serem cruéis. Acredito que esse desconforto tem uma origem pragmática. Atribuições de obrigações envolvem implicitamente julgamentos prescritivos. Se eu lhe digo que alguém é obrigado a doar à caridade, tenho provavelmente um interesse em transmitir como me sinto. Afirmar a existência de uma obrigação é uma forma de transmitir que eu penso que uma pessoa deva fazer algo. Mas 'deve' é uma implicação conversacional de uma 'obrigação', não uma implicação semântica. Para ver isso, repare que a inferência de 'obrigação' para 'deve' pode ser cancelada. Parece contraditório dizer 'Smith deve doar à caridade, embora ele não deva doar à caridade'. Mas não parece contraditório dizer 'Smith tem uma obrigação de doar à caridade, mas ele não deve'. Nós falamos esse tipo de coisa com certa frequência quando estamos conversando sobre valores morais de outras pessoas. Podemos dizer que soldados japoneses da Segunda Guerra Mundial tinham uma obrigação de sacrificar sua vida como pilotos

Kamikaze, mas eles não deveriam ter feito isso. Da mesma forma, posso admitir consistentemente que sadistas têm uma obrigação de serem cruéis enquanto insisto que eles devem abster-se da crueldade. Isso aborda a segunda preocupação levantada no início desta sessão. Obrigações podem ser deduzidas de premissas descritivas, mas elas não precisam ser endossadas pelos seus dedutores. Endossos estão meramente implicados. Eles não podem ser deduzidos. Acreditar que Smith deve doar à caridade requer um julgamento prescritivo. Para fazer uma prescrição, precisamos estar em um estado psicológico particular - precisamos prescrever. É nesse sentido que não podemos derivar um *deve* de um *é*.

1.4 DEFENDENDO O SUBJETIVISMO

Tenho argumentado que a Lei de Hume é basicamente verdadeira. Minha defesa depende de uma teoria de conceitos normativos que apresentei na forma de um esboço simples. 'Deve', argumentei, expressa um sentimento prescritivo. Meu principal objetivo nos capítulos a seguir será a defesa desta afirmação, e trazer algumas implicações. Focarei nos conceitos como BOM ou MAU ou CERTO ou ERRADO (letras maiúsculas denotam conceitos). Estes, como o conceito de DEVE, envolvem essencialmente sentimentos. Esses conceitos são fundamentalmente *subjetivos*.

Meu objetivo não será derivar prescrições de descrições. Este é um projeto normativo e, se o argumento precedente está correto, não é um que se pode levar muito longe. Porém, tentarei derivar fatos metafísicos de psicológicos. Certo e errado são referências do nosso conceito de CERTO e ERRADO se eles realmente são alguma coisa. Se a análise dos nossos conceitos desvenda uma forte conexão com respostas subjetivas, então esses termos podem referir a algo subjetivo. Psicologia moral endossa fatos sobre ontologia moral, e psicologia sentimental pode endossar uma ontologia subjetivista.

Se a moralidade é subjetiva, então por que julgamentos morais importam para nós? Uma resposta, inspirada em Hume, é que não conseguimos evitar de nos preocuparmos com a moralidade. Há algo de correto sobre isso, mas isso só traz a questão de volta um nível. Por que não conseguimos evitar de nos preocuparmos

com a moralidade? Essa pergunta parece ser mais difícil de responder do que aquela sobre por que nós nos *preocupamos*. Não há uma resposta única à última questão. Sistemas morais servem a vários fins. Eles regulam comportamentos, eles atribuem à vida um senso de significado, e eles definem o pertencimento a um grupo. A pergunta "Por que a moralidade importa?" é como a questão "Por que leis importam?" ou por que a "Cultura importa?". Pessoas que se sentem desconfortáveis com a ideia de que a moralidade deriva de nós, deveriam considerar qualquer outra coisa que deriva de nós, como medicina, governos e arte. O fato de a arte ser uma construção social não retira dela seu valor. Nós não esperamos que institutos de arte colapsem com a descoberta de que a arte é um produto da invenção humana.

A discussão a seguir está dividida em duas partes, correspondendo a temas que emergiram deste debate. Na primeira parte, argumento que moralidade depende de emoções, e, na segunda parte, discuto o que considero ser as implicações desta perspectiva: a hipótese de que a moralidade varia entre culturas. Se moralidade depende de sentimentos, argumento, então, que é uma construção, e, se é uma construção, pode variar de acordo com o tempo e espaço.

O primeiro capítulo na parte 1 apresenta uma pesquisa de diferentes formas em que emoções podem envolver moralidade. Introduzo o termo "emocionismo" para nomear qualquer visão que considera essencial as emoções, e ofereço algumas razões para pensar que um emocionismo forte é verdadeiro. No capítulo 2, trago fundamentações para uma teoria emocionista ao apresentar uma teoria geral das emoções. Se a moralidade tem uma base emocional, então é melhor começar com uma teoria independentemente motivada do que são emoções. Neste capítulo, também apresento uma revisão das emoções morais, e sugiro que emoções morais derivam de emoções não-morais. No capítulo 3, começo a apresentar minha perspectiva positiva. É o que os especialistas em ética contemporâneos chamam de "teoria da sensibilidade", apesar de minha visão particular se diferenciar de maneira sutil das perspectivas dominantes (especificamente, desenha uma perspectiva de sentimentos morais apresentados no capítulo 2, que não é metacognitiva). Argumento que essa teoria pode lidar com dez das maiores objeções que têm sido feitas às teorias da sensibilidade. O capítulo 4 endossa uma outra objeção não

mencionada no capítulo 3: teorias da sensibilidade são subjetivistas, e muitas pessoas assumem que a moralidade é objetiva. Argumento contra essa premissa ao distinguir diversos tipos de objetividade e avaliando criticamente as principais teorias éticas que pretendem apresentar que a moralidade é objetiva em cada sentido do termo. Concluo que a moralidade é completamente subjetiva.

Chamo a perspectiva desenvolvida na parte I de "construtivismo sentimentalista". O termo sentimentalista se refere ao papel dos sentimentos, e o termo 'construtivismo' refere-se ao fato de que sentimentos literalmente criam moral, e sistemas morais podem ser criados de maneiras diferentes. A parte II foca nesta implicação do sentimentalismo. Mais especificamente, explora o papel da cultura em moldar valores morais. No capítulo 5, abordo as consequências relativistas do meu caso contra o objetivismo, e respondo aos argumentos padrões contra o relativismo. O sexto capítulo diz respeito à genealogia da moral, no sentido nietzscheano. Argumento que a antropologia histórica pode ser usada para explicar por que certos valores morais persistem, e por que outros desaparecem. Também acesso o nível em que tais análises podem ser usadas para criticar moralidade. O Capítulo 7 muda da genealogia para os genes. Mesmo que alguns valores sejam históricos na origem, outros podem ser biológicos. Especialistas da ética evolucionista têm avançado essa linha nos últimos anos. Argumento que a ética evolucionista falha ao explicar qualquer tipo dos nossos valores específicos. As únicas regras morais baseadas biologicamente são muito abstratas para guiar ações, e seu *status* moral é epigenético. Moralidade envolve essencialmente aprendizagem. Essa conclusão carrega a prospecção do progresso moral, que é o tema do capítulo final. Discuto a natureza dos debates morais e argumento que podemos melhorar nossa moralidade. Evolução moral, às vezes, exige olharmos para além das categorias de bom ou de mau, mas não devemos tentar abandonar a moralidade ou substituí-la por outro tipo de empreendimento normativo.

Minha perspectiva ao defender esses argumentos será naturalista em todos os significados caracterizados acima. Meu comprometimento mais óbvio é com a metodologia naturalista, porque irei demonstrar descobertas empíricas, incluindo descobertas da neurociência, psicologia, psiquiatria, antropologia, história cultural e etologia. Acredito que questões filosóficas que persistem podem ser iluminadas por

resultados empíricos, e que, de fato, elas podem não durar tanto tempo se usarmos recursos da ciência. Dito isso, não rejeito a tradicional metodologia filosófica, tal como a análise conceitual. Realmente, acredito que a análise conceitual é, em algum sentido, uma metodologia empírica: um tipo de semântica léxica alcançada por meios de uma cuidadosa introspecção. Penso que o método frequentemente carrega resultados, mas, às vezes, a introspecção colide ou falha em revelar a estrutura real dos nossos conceitos. Então, é útil achar outros métodos para ajudar a julgar entre teorias filosóficas concorrentes. Estes outros métodos não substituem a filosofia. A filosofia coloca os problemas que investigamos, elabora ferramentas úteis para sondar conceitos (como experiências de pensamento), e nos permite passar de dados para teoria sistematizando os resultados em pacotes coerentes que podem orientar pesquisas futuras. Vejo a filosofia como uma continuação da ciência, e acredito que deveríamos estar abertos para usar qualquer metodologia disponível quando fazemos perguntas sobre a natureza da moralidade.

Também sou pluralista sobre o assunto. Este é um livro sobre psicologia moral, sobre metaética, e sobre a origem e a antropologia da moral; até entro em contato com algumas questões normativas no capítulo final. Leitores com um interesse específico, digamos, em metaética, podem achar de pouco interesse a discussão da história cultural, e leitores com uma orientação antropológica podem protelar a discussão sobre ontologia moral. Espero que não seja o caso. Penso que uma perspectiva completa da moralidade deva abordar cada uma dessas dimensões, e acredito que as dimensões são mutuamente elucidativas. Por exemplo, pode-se argumentar a favor do relativismo ao apresentar evidências semânticas, e pode-se argumentar estudando variações culturais. Ambas as formas dão evidências convergentes, e as observações culturais motivam o questionamento semântico e ajudam a revelar porque a tese semântica pode ser profundamente importante.

PARTE I
MORALIDADE E EMOÇÃO

1. EMOCIONISMO

1.1 MORALIDADE AFETIVA

1.1.1 Duas Espécies de Emocionismo

Julgar que alguma coisa é certa ou errada não é como julgar que 3 é um número primo ou que árvores fazem fotossíntese. Nós podemos formar os últimos juízos sem a mínima comoção (agitação, *stirring*) da paixão. Nós podemos ser totalmente indiferentes a eles. Mas juízos morais são qualquer coisa exceto indiferentes. Eles transpiram sentimento. Nós somos passionais a respeito de nossos valores. Considere a questão, "Como você se sente sobre a pena de morte?". Uma resposta apropriada poderia ser "Eu sinto que ela é completamente injustificável." Esta figura de linguagem é estranha ao domínio avaliativo. Nós não perguntaríamos "Como você se sente sobre árvores", e responderíamos, "eu sinto que elas fazem fotossíntese". Certo e errado, ao contrário de números primos e de fotossíntese, são coisas que sentimos.

É claro que muitos teóricos da ética estão preparados para rejeitar este contraste. Ninguém pode negar que nós nos sensibilizamos fortemente a respeito de nossos valores morais, mas se pode razoavelmente duvidar se tais sentimentos fortes são constitutivos do que é valor ou do que é ser valioso. Pode-se concordar que juízos morais agitam nossos sentimentos mesmo negando que o *status* disso enquanto um juízo moral depende de nós termos tais sentimentos. Alguém pode admitir que nós nos sensibilizamos fortemente sobre fatos morais mesmo negando que aqueles fatos dependam de termos tais sentimentos. Pode-se afirmar que há coisas que nós devemos fazer e não fazer, sem considerar o modo como sentimos. A divisão entre aqueles teóricos que pensam que os sentimentos são essenciais à moralidade e aqueles que pensam que emoções são acidentais é, talvez, a divergência mais fundamental em filosofia moral. Eu estou do lado dos membros do primeiro campo. A afirmação que emoções figuram dentro da moralidade pode ser entendida de várias maneiras. Eu irei usar o termo "emocionismo" como um rótulo

abrangente para qualquer teoria que afirma que as emoções são, de certa forma, essenciais. O termo não deve ser confundido com "emotivismo", que é uma versão específica de emocionismo.

Eu quero distinguir duas teses emocionistas dissociáveis. De acordo com a primeira, propriedades morais não poderiam existir sem emoções. Em outras palavras, não há maneira de determinar as condições de identidade de uma propriedade moral como tal sem referência a uma emoção ou classe de emoções. Mais sucintamente, nós podemos dizer:

Emocionismo Metafísico

Propriedades morais estão essencialmente relacionadas com emoções.

Defensores deste ponto de vista estão comprometidos com o realismo moral, se nós definimos realismo moral como a visão que sustenta fatos morais. Quando uma propriedade moral é instanciada, existe um fato que consiste na sua instanciação. Se alguém acredita em propriedades morais, segue-se que existem fatos morais. O emocionismo metafísico abarca fatos morais e sustenta que estes fatos se baseiam em emoções. Algumas formas de utilitarismo tornam-se preeminentes. Considere, especialmente, o utilitarismo clássico de Bentham e Mill. Eles definem o bem como aquilo que maximiza a utilidade, e eles definem utilidade como felicidade. Há fatos morais, nesta visão, porque há ações que maximizam a felicidade. E estes fatos são essencialmente emocionais, porque a felicidade é uma emoção.

O termo "realismo" é algumas vezes reservado para um tipo de independência mental: o fato de que a é F seja real, de acordo com esta interpretação, se a ser F não depende de nós considerarmos a como F . Utilitaristas são realistas neste sentido forte sobre o bem. Chamo a isto de realismo externo. Realismo interno, ao contrário, é a concepção que F sendo a é um fato, mas este fato depende de nós considerarmos a como F . (Ver Putnam, 1980). O realismo interno é um factualismo sem independência-mental. Considere o ponto de vista de

que propriedades morais são qualidades secundárias. Qualidades secundárias são propriedades dependentes de resposta. De acordo com Locke, cores, gostos e cheiros adequam-se nesta categoria. Limões são azedos – isto é um fato - mas eles têm esta propriedade somente na medida em que eles causam uma certa experiência azeda em nós quando os provamos. Exposições [Accounts] que desenvolvem a analogia entre qualidades secundárias e morais têm sido denominadas “teorias da sensibilidade” (Darwall et al., 1992). A versão recente mais influente nós devemos a McDowell (1985) e Wiggins (1987).

Teorias da sensibilidade descendem da teoria “sentimentalista” dos moralistas britânicos, tal como Shaftesbury, Hutcheson, Hume, e Smith. Hutcheson nos diz: “a palavra bondade moral [...] denota nossa ideia de alguma qualidade apreendida em ações, as quais procuram aprovação [...]. Mal moral denota nossa ideia de uma qualidade contrária, a qual suscita desaprovação ou antipatia”. (1738: 67).

Hume vai ainda mais longe, explicitamente extraindo uma analogia entre qualidades morais e secundárias:

Examinemo-la sob todos os pontos de vista, e vejamos se podemos encontrar o fato, ou existência real, que chamamos de *vício*. Como quer que a tomemos, encontraremos somente paixões, motivos, volições e pensamentos. Não há nenhuma outra questão de fato neste caso. O vício nos escapa por completo, enquanto consideramos o objeto. Não o encontraremos até dirigirmos nossa reflexão para nosso próprio íntimo e darmos com um sentimento de desaprovação, que se forma em nós contra essa ação. Aqui há um fato, mas ele é objeto de sentimento, não de razão. Está em nós, não no objeto. Desse modo, quando declaramos que uma ação ou caráter são viciosos, tudo que queremos dizer é que, dada a constituição de nossa natureza, experimentamos uma sensação ou sentimento de censura quando os contemplamos. O vício e a virtude, portanto, podem ser comparados a sons, cores, calor e frio, os quais, segundo a filosofia moderna, não são qualidades nos objetos, mas percepções da mente (1739: III. i.i).

A teoria moral de Hume tem aspectos que a distingue das teorias modernas da sensibilidade. Uma diferença é que, nesta passagem, Hume pode ser lido como implicando que propriedades morais não existem (a questão sobre se esta foi a visão adequada de Hume eu deixo para os estudiosos). É fácil chegar a uma visão cética sobre propriedades morais se nós partirmos de uma concepção antirrealista de qualidades secundárias. A concepção lockeana das qualidades secundárias é diferente. De acordo com Locke, sons e cores são reais, mas relacionais (poderes de causar sensações em nós). Se alguém é um realista sobre qualidades secundárias, ele pode adotar um análogo realista da tese de Hume. Teorias da sensibilidade contemporâneas tendem a ter aquele caráter particular. Na discussão moderna, teorias da sensibilidade estão comprometidas com o perceptivismo mais do que com o projectivismo (D'Arms and Jacobson, 2006). Perceptivistas afirmam que nós percebemos propriedades morais em virtude de ter certas emoções, e projectivistas afirmam que nós não as percebemos, mas, em vez disso, projetamo-as sobre o mundo. Enquanto perceptivistas, os teóricos da sensibilidade estão comprometidos com o emocionismo metafísico.

Eu terei mais a dizer sobre teorias da sensibilidade abaixo, mas eu quero voltar agora a outro aspecto da filosofia moral de Hume. Hume enfatiza a prioridade do caráter sobre a ação. Estar certo ou errado é uma função de causar certas emoções, mas nós diferenciamos as emoções que importam para a avaliação moral daquelas que não importam. A questão sobre quais emoções importam é analisada por Hume como uma questão sobre para quem tais emoções importam. Hume pensa que certo e errado são determinados por uma resposta emocional de uma pessoa de caráter:

É somente quando um caráter é considerado em geral, sem referência a nosso interesse particular, que causa essa sensação ou sentimento em virtude do qual o denominamos moralmente bom ou mau. É verdade que temos naturalmente uma tendência a confundir e misturar os sentimentos relativos ao interesse e os relativos à moral. Raramente deixamos de pensar que um inimigo é vicioso e raramente somos capazes de distinguir entre sua oposição ao nosso interesse e sua vilania ou baixeza reais. Isso não impede, porém, que esses sentimentos

sejam distintos neles mesmos; um homem dotado de serenidade e discernimento pode se proteger dessas ilusões (1739: III, i, ii).

A este respeito, a teoria da sensibilidade de Hume é também um exemplo de outro tipo de teoria: ela é uma ética das virtudes. Algumas versões da ética das virtudes qualificam-se como formas de emocionismo metafísico. Considere o seguinte exemplo. Uma ação é boa se e somente se esta é algo que uma pessoa virtuosa faria. Uma pessoa virtuosa é uma pessoa que tem certos traços de caráter. Traços de caráter virtuosos são ou incluem disposições emocionais. Segue-se que uma ação é boa somente no caso de ela ser realizada por um agente emocional.

Utilitarismo, teorias da sensibilidade e ética das virtudes: todas fazem alegações metafísicas sobre a natureza de propriedades morais. Muitos dos seus defensores estão também comprometidos com alegações epistêmicas. Para reconhecer um fato moral, deve-se compreender os conceitos morais correspondentes. Se conceitos morais se referem a propriedades morais, e propriedades morais estão relacionadas constitutivamente com emoções, então é razoável pensar que compreender conceitos morais envolve emoções de algum modo. Por exemplo, os utilitaristas podem dizer que você não pode entender o que o bem é a menos que você possua o conceito de felicidade. Assim, para os utilitaristas clássicos, conceitos morais podem estar relacionados essencialmente ao conceito de emoção. Teorias da sensibilidade fazem uma exigência ainda mais forte. Elas geralmente afirmam que conceitos morais devem ser definidos, não em termos do *conceito* de emoção, mas em termos das emoções *elas mesmas*. Vou me referir a esta tese como:

Emocionismo epistêmico

Conceitos morais estão essencialmente relacionados com emoções

Para tornar esta tese plausível, é importante traçar uma distinção importante entre modelos adequados de possuir conceitos morais e modelos inadequados.

Considere a analogia com a cor. Há um sentido no qual uma pessoa congenitamente cega pode compreender o conceito de cor (ver Crimmins, 1989). Ela poderia dominar os tipos de sentenças que contêm palavras de cor ou ela poderia até mesmo detectar cores usando um aparato especial que converte informações espectrais em outro formato. Mas este não é o modo como pessoas que têm visão compreendem o conceito de cor. Se cores são qualidades secundárias, nós poderíamos dizer que uma pessoa cega é incapaz de compreender cores pelas suas qualidades essenciais. Uma pessoa cega não pode pensar sobre cores enquanto tais. O emocionismo epistêmico é para ser considerado uma tese sobre nossa capacidade para compreender propriedades morais em um padrão. O emocionista epistêmico não nega que possa haver outros modos de reflexão sobre a moralidade. Um marciano sem emoções poderia ter conceitos morais deferentes [*deferential*], por exemplo ("Errado é aquilo que terrestres chamam 'errado'"). Eu terei mais a dizer sobre conceitos padrão [*standard concepts*] abaixo.

Outro ponto de clarificação é oportuno. Ao definir o emocionismo epistêmico, eu usei a frase "essencialmente relacionado". A forma mais óbvia de relação essencial é uma relação de constituição. Conceitos morais são essencialmente relacionados com emoções se eles são constituídos por emoções. Sob esta perspectiva, exemplos simbólicos de conceitos tais como ERRADO e CERTO são estados emocionais ou têm estados emocionais como partes componentes. Isto é o que emocionistas epistêmicos frequentemente têm em mente. Mas alguns emocionistas epistêmicos gostariam de levar em conta uma relação disposicional entre conceitos morais e emoções. Eles pretendem conceder que, em algumas ocasiões, pessoas podem fazer juízos morais sem qualquer sentimento, mas eles insistirão que, em tais ocasiões, as pessoas que fazem tais juízos estão dispostas a ter respostas emocionais. Por analogia, suponha que você pense o conceito de ENGRAÇADO é essencialmente relacionado com DIVERTIMENTO. Em algumas ocasiões, você pode julgar, a partir da memória, por exemplo, que alguém é engraçado sem estar realmente se divertindo. Nestas ocasiões, entretanto, você está sendo sincero somente se você está disposto a se sentir entretido quando você estiver interagindo com aquela pessoa.

Alguém pode ser um emocionista metafísico sem ser um emocionista epistêmico. Utilitaristas clássicos são um caso neste ponto. Alguém pode ser também um emocionista epistêmico sem ser um emocionista metafísico. Considere aqueles que negam que propriedades morais existem. Se o realismo moral é falso, então o emocionismo metafísico não pode ser verdadeiro. Mas o antirrealismo moral pode defender o emocionismo epistêmico.

O emotivismo é uma teoria deste tipo. Os emotivistas sustentam que juízos morais não descrevem o mundo; antes, eles expressam nossas atitudes. Ayer (1952) afirma que a sentença "roubar é errado" é equivalente a dizer "roubar!" com um tom de horror. Isto não atribui qualquer propriedade a roubar dinheiro; isto meramente comunica um sentimento. Stevenson (1937) defende uma versão ligeiramente diferente de emotivismo. Ele afirma que "roubar é errado" assere algo, a saber, que eu não gosto de roubar, mas não meramente um fato; ele assere aquilo em um modo "dinâmico" que expressa emocionalmente minha aversão e, assim, convida você a compartilhar daquela atitude. Deste modo, mesmo que Stevenson admita que termos morais expressem fatos (gostos/prazeres e aversões), a função primária deles é expressar e recomendar emoções. O emotivismo tem sido algumas vezes chamado de teoria da vaia e do viva, porque os seus defensores, algumas vezes, comparam termos morais com expletivos. Dizer que roubar é errado é um pouco como expressar "vaia para roubar!", porque ambos "errado" e "vaia" são usados principalmente para exprimir e prescrever sentimentos mais do que para relatar fatos.

Autores recentes têm defendido teorias expressivistas mais sofisticadas. Blackburn (1984) está próximo do emotivismo tradicional, mas ele enfatiza a natureza projetiva dos juízos morais. Nós falamos sobre propriedades morais *como* se elas estivessem no mundo, mas não assumimos qualquer comprometimento ontológico sério ao fazer isso. Blackburn e os emotivistas clássicos são emocionistas epistêmicos, mas eles rejeitam o emocionismo metafísico. A exposição de Blackburn é frequentemente comparada à outra teoria, denominada expressivismo da norma, a qual foi antecipada por Gibbard (1990), mas os dois são relevantemente diferentes. Gibbard sustenta que juízos morais expressam a nossa

aceitação de normas emocionais. Dizer que roubar é errado é expressar aceitação de uma norma que exige o sentimento de culpa se eu roubo, e de raiva se alguma outra pessoa rouba. A concepção de Gibbard é diferente do emotivismo porque juízos morais não expressam emoções diretamente; mais exatamente, eles expressam normas que nos comprometem com o caráter apropriado de emoções. Por conseguinte, Gibbard não é estritamente um emocionista epistêmico. Sob esta concepção, alguém poderia dizer que juízos morais se referem a emoções (eles expressam a atitude que eu tenho o direito para estar bravo), mas eles não usam emoções (eles não expressam raiva).

O emocionismo epistêmico é uma tese psicológica. Este é uma tese sobre conceitos morais. O rótulo "epistêmico" adverte para o fato que conceitos são as ferramentas psicológicas pelas quais nós chegamos a entender a moralidade. Mas a psicologia tem outra dimensão. Ela é o *locus* da ação. E é neste domínio que o emocionismo mostra outra face. Para agir, nós precisamos estar motivados. Emoções e motivação estão ligadas. As emoções exercem uma força motivante. Há evidência clínica que, sem emoções, as pessoas não sentem inclinação para agir. Damasio e Van Hoesen (1983) descrevem uma condição chamada mutismo acinético, nos quais pacientes que têm sofrido lesões em áreas emocionais do cérebro ficam imóveis na cama; após a recuperação, eles relatam que eles estavam totalmente conscientes, mas não sentiram emoções e, portanto, nenhuma inclinação para agir. As emoções morais podem ser especialmente importantes para motivar o comportamento decente. Por exemplo, há provas de que a culpa promove a ajuda. Em um estudo, McMillen e Austin (1971) induziram alguns sujeitos a fraudar em um exame, e então eles pediram para aqueles indivíduos ajudar na pontuação de alguns questionários; os indivíduos que não tinham sido induzidos à fraude ajudaram por apenas 2 minutos, mas os trapaceiros ajudaram por 63 minutos. Se você se sente culpado por fazer algo, você vai tentar compensar isso, e se você antecipa o sentimento de culpa por fazer aquilo, há uma boa chance de você resistir à tentação de fazê-lo. Portanto, se os conceitos morais contêm emoções, então os julgamentos morais irão promover um comportamento que se alinha com os julgamentos.

No jargão filosófico, isso significa que o emocionismo epistêmico pode implicar um internalismo motivacional. Internalistas motivacionais acreditam que há uma conexão necessária entre juízos morais e a motivação para agir de acordo com esses julgamentos (Brink, 1989): se se acredita que roubar é errado, é-se, assim, motivado a agir de um modo determinado (por exemplo, a se abster de roubar ou trabalhar para prevenir outros de roubar) mesmo que, sob algumas circunstâncias, essas motivações fiquem "saturadas" por outras demandas motivacionais sobre a ação. Como o nome "internalismo" sugere, internalistas motivacionais pensam que os julgamentos morais carregam uma força motivacional por conta própria, sem necessitar da ajuda de fora. Por exemplo, se você acredita que roubar é errado, você não precisa de um desejo abrangente para evitar o erro, a fim de estar motivado a não roubar. Mas como pode um julgamento moral ser intrinsecamente motivador? A resposta fica clara sob uma visão emocionista epistêmica. Julgamentos morais contêm conceitos morais, e emocionistas epistêmicos alegam que há uma conexão necessária entre conceitos morais e emoções. Suponha que a conexão necessária é tal que "sinalizar" [*tokening*] conceitos morais sempre resultam em um estado emocional. A evidência empírica demonstra que as emoções têm força motivacional. Assim, se esta versão do emocionismo epistêmico está correta, então os julgamentos morais não podem ocorrer sem motivação.

O internalismo motivacional é uma doutrina controversa. Meu ponto aqui é que os emocionistas epistêmicos têm uma explicação de como ele poderia ser verdadeiro. Deve ser notado também que formas diferentes de emocionismo epistêmico implicariam formas diferentes de internalismo motivacional. Sobre a forma que eu sugeri no meu exemplo, internalistas motivacionais afirmam que os julgamentos morais são *sempre* intrinsecamente motivadores. Há também formas mais fracas de internalismo motivacional. Por exemplo, pode-se afirmar que os julgamentos morais são normalmente motivadores, ou capazes de serem intrinsecamente motivadores, ou ligados disposicionalmente à motivação. Da mesma forma, o emocionismo epistêmico pode vir em variedades diferentes. Pode-se ter a visão de que não se pode sinalizar um conceito moral sem sinalizar uma

emoção. Ou pode-se ter a visão de que sinalizar conceitos morais nos predispõe para emoções. E assim por diante. Cada versão do emocionismo epistêmico parece implicar uma forma correspondente ou um internalismo motivacional. Em cada caso, há uma ligação entre conceitos morais e estados que são motivantes.

1.1.2 Relações Essenciais

Nas definições recém apresentadas, eu disse que emocionistas postulam uma "relação essencial" entre emoções e coisas no domínio moral. O que é para ser relacionado essencialmente? Eu escolhi esta frase, em vez de "relacionado necessariamente", porque pode haver algum desvio entre necessidade e essência. Alguma coisa A pertence à essência de outra coisa B se não se pode especificar o que é ser B sem mencionar A. Esta formulação não invoca a necessidade em um sentido modal forte. Ela não diz que todos os Bs são necessariamente como As. *Podem-se* construir relações essenciais neste modo forte. Não é incomum para filósofos pensar as essências como necessárias e suficientes para a participação em uma categoria. O tipo de essencialismo associado com a filosofia moderna da linguagem tem esse tom. Quando Kripke (1980) afirma que "água" se refere a H₂O, ele quer dizer que a água é H₂O em todos os mundos possíveis. Isso pode dar a impressão de que os emocionistas estão comprometidos com a visão de que as emoções estão presentes toda vez em que juízos morais ou propriedades estão presentes, assim como oxigênio está presente em cada amostra de água. Essa impressão é enganosa.

Antes de tudo, há outras maneiras de interpretar as essências. Por exemplo, Boyd (1988) define uma essência como um feixe de propriedades homeostáticas: uma coleção de propriedades que tendem a coocorrer e promover a ocorrência uma da outra. Nesta visão, alguma propriedade particular pode ser parte da essência de algum tipo de coisa, mesmo que nem sempre ocorram em todos os casos desse tipo.

Em segundo lugar, mesmo em uma visão kripkeana das essências, a frase "essencialmente relacionado com as emoções" não implica que as emoções estão

ativas toda vez que há uma propriedade moral ou julgamento instanciado. Suponhamos, por exemplo, que os conceitos morais são constituídos em parte por *disposições* a ter emoções. Suponhamos, além disso, que tais disposições são essenciais para conceitos morais em sentido kripkeano (em cada indício de um conceito moral em todos os mundos, aquele indício é constituído em parte por uma disposição emocional). Segue-se que os conceitos morais são essencialmente relacionados com as emoções, porque estão essencialmente relacionadas com as disposições emocionais, e uma disposição emocional é uma relação com as emoções. Relações essenciais são transitivas. Enquanto as relações em questão não são relações de constituição, o emocionista pode dizer que há uma conexão modal forte entre moralidade e emoção admitindo que as emoções e a moral nem sempre são coinstanciadas.

1.1.3 Emocionismo Forte

As duas espécies de emocionismo que descrevi podem ser aceitas em conjunto ou separadamente. Eu já mencionei algumas das teorias que levam a uma ou outra espécie sem abraçar todas elas. Uma divisão parcial é apresentada na Tabela 1.1.

Tabela 1.1. Espécies de Emocionismo

	Teorias da Sensibilidade	Emotivismo	Utilitarismo Clássico	Ética Kantiana
Emocionismo Metafísico	Sim	Não	Sim	Não
Emocionismo Epistêmico	Sim	Sim	Não	Não

Os kantianos rejeitam ambas as formas de emocionismo. Ações moralmente más são aquelas que eu não poderia querer como uma lei universal. Isto não pretende ser um axioma sobre minhas paixões ou gostos. O mal não é aquilo que eu

detesto. A universalizabilidade é uma exigência racional sobre a moralidade. Certas formas de conduta não podem ser coerentemente universalizáveis. Kant (1785) fornece as promessas de mentira como um exemplo. Se todos mentissem ao prometer, toda a construção de prometer entraria em colapso. Prometer só faz sentido num contexto em que as promessas são geralmente confiáveis e honestas. Kant vê um fundamento racional semelhante a prescrições positivas. Ajudar os necessitados é moralmente necessário porque não se pode universalizar a falta de ajuda para os necessitados. Todos são necessitados ou potencialmente necessitados em algum momento, de modo que seria irracional para todos querer um mundo onde ninguém ajuda os necessitados.

Os kantianos também rejeitam o emocionismo epistêmico porque conceituar algo como certo ou errado é uma questão de formar um julgamento sobre o que é racional. De um modo geral, pode-se fazer isso sem estar em algum estado emocional. Kant pensa que, para fazer julgamentos morais com sucesso, faria-nos bem ignorar nossas paixões.

Os utilitaristas clássicos concordam com kantianos em negar o emocionismo epistêmico. Eles negam que os conceitos morais estejam essencialmente relacionados com as emoções. Poder-se-ia vislumbrar um conceito moral, sem ter qualquer disposição para experienciar uma emoção. Por outro lado, os utilitaristas pensam que o emocionismo metafísico é verdadeiro. O bem é definido em termos de felicidade. O emotivismo é, a este respeito, o inverso do utilitarismo. Os emotivistas afirmam que as emoções são essenciais para conceitos morais, mas rejeitam as teses metafísicas. Eles afirmam que não há fatos morais.

O utilitarismo e o emotivismo podem ser chamados de teorias emocionistas fracas, porque elas implicam uma tese emocionista, e não a outra. Uma teoria emocionista forte implicaria ambas. A teoria da sensibilidade é o exemplo mais saliente. Aqui está uma apresentação esquemática da visão:

(S1) Tese Metafísica: Uma ação tem a propriedade de ser moralmente certa (errada) apenas no caso em que provoca sentimentos de aprovação (desaprovação) em observadores normais sob certas condições.

(S2) Tese Epistêmica: A disposição para sentir as emoções mencionadas em S1 é uma condição de posse sobre o conceito normal de CERTO (ERRADO).

Eu penso que uma teoria deste tipo pode ser defendida. Eu defendo o emocionismo forte. Grande parte deste livro será dedicada a justificar e esclarecer esta pretensão.

1.2 O EMOCIONISMO PODE SER VERDADEIRO?

Evidências de uma variedade de fontes sugere que as emoções são fundamentais para a moralidade. No restante deste capítulo, vou me concentrar nas evidências para o emocionismo epistêmico, embora eu vá oferecer algum apoio para a tese metafísica no final. Vou acrescentar, além disso, argumentos e respostas às objeções nos capítulos que se seguem.

1.2.1. Julgamentos morais são acompanhados por emoções

A razão mais óbvia para levar o emocionismo a sério decorre da observação mundana de que os julgamentos morais são frequentemente acompanhados de emoções. É difícil de se manter imparcial quando você lê notícias de jornal sobre crianças molestadas, atrocidades de guerra, ou racismo institucionalizado. A intensidade de nossas emoções é frequentemente um guia muito confiável para a força de nossos juízos morais. Por exemplo, crimes contra crianças são muitas vezes considerados piores do que crimes contra adultos e eles também parecem provocar fortes reações emocionais.

O impacto emocional do juízo moral decorre do fato de que nós tendemos a evitar o mau comportamento. Violar regras morais é frequentemente vantajoso. Se nós roubamos coisas, podemos tê-las de graça. Se nós traímos nossos parceiros, podemos multiplicar os nossos prazeres. Mesmo matar pode ser vantajoso; se você entrar para um concurso de redação, não há melhor maneira de aumentar suas chances de ganhar do que acabar com os melhores escritores da competição.

Acontece que nós não temos o hábito de fazer essas coisas, mesmo quando nós podemos nos safar delas. Por que não? A resposta óbvia é que fazer coisas más faz com que nos sintamos mal.

Isto é ilustrado comoventemente por um experimento que Stanley Milgram conduziu no início dos anos 1970. Ele pediu aos seus alunos de pós-graduação para embarcar em um trem do metrô da cidade de Nova York e pedir a estranhos para dar a eles os seus lugares. Isto viola uma norma. Nós obedecemos ordinariamente uma regra segundo a qual qualquer um que encontra um lugar vazio primeiro tem direito a esse lugar. Se você encontrou o último assento livre na Rua 14, e eu embarco na Rua 23, eu não tenho direito ao seu lugar; seria errado da minha parte pedir por ele, a menos que eu fosse velho, que estivesse ferido, ou que fosse incapaz de alguma maneira de ficar em pé sem risco. Milgram pediu aos seus alunos para violar esta norma, porque ele queria saber como as pessoas iriam reagir. Ele tinha um interesse geral na obediência. Mas quase todos os seus alunos se recusaram. Ele conseguiu convencer apenas um estudante a realizar o estudo. Esse aluno embarcou devidamente no metrô e pediu às pessoas que desistissem dos seus assentos. Quando ele retornou, ele disse que a experiência foi extremamente difícil e que ele não poderia coletar tantos dados quanto Milgram havia solicitado. Ao invés de pedir para 20 pessoas pelos seus lugares, ele parou em 14. A dificuldade não tinha nada a ver com o fato de que as pessoas não estavam cooperativas. Pelo contrário, a maioria das pessoas voluntariamente deu os seus lugares. A tarefa foi difícil porque era emocionalmente doloroso quebrar uma norma.

Milgram descobriu isso por si mesmo, depois de perder a paciência com os seus relutantes alunos e realizar o estudo por si mesmo. Esta é a forma como ele descreve a experiência em uma entrevista de 1974:

As palavras pareciam depositadas na minha traqueia e simplesmente não saíam. Recuando, eu repreendi a mim mesmo: 'Que tipo de covarde maldito [*crawen coward*] é você"? Finalmente, depois de muitas tentativas frustradas, eu fui até um passageiro e me engasguei no pedido: "Desculpe-me senhor, posso me sentar no seu lugar?" Um momento de completo pânico anômico me

dominou. Mas o homem se levantou e me deu o assento. Um segundo golpe ainda estava por vir, tomando o assento do homem, eu estava sobrecarregado pela necessidade de me comportar de uma maneira que justificaria o meu pedido. Minha cabeça se afundou entre os meus joelhos, e eu podia sentir o meu rosto branqueando. Eu não estava atuando. Eu realmente senti como se eu fosse morrer (Citado em Blass, 2004: 174).

Esta história ilustra um fato importante sobre normas morais. Quando nós fazemos coisas que violam os valores morais, incorremos em custos emocionais.

Há agora evidências empíricas abundantes de que as emoções ocorrem quando nós fazemos juízos morais. É de particular interesse que todo estudo de neuroimagem da cognição moral parece implicar áreas do cérebro associadas com a emoção (Greene e Haidt, 2002). Veja alguns exemplos. Heekeren et al. (2003) pediu para indivíduos avaliarem se sentenças são moralmente incorretas (tal como, "A rouba o carro de B") ou semanticamente incorretas (tal como, "A bebe o jornal"). Na condição do julgamento moral, os indivíduos mostraram significativamente mais ativação em áreas emocionais. Num estudo similar, Moll et al. (2003) pediram que indivíduos emitissem juízos de "certo" ou de "errado" tanto sobre sentenças morais, tal como "Eles enforcaram uma pessoa inocente," quanto sobre sentenças factuais, tal como "Pedras são feitas de água". Mais uma vez, as áreas de emoção foram mais ativas para os juízos morais. Moll et al. (2002) também encontraram ativação emocional quando os indivíduos ouviam frases moralmente ofensivas ao invés de frases neutras (por exemplo, "Os idosos são inúteis" *versus* "Os idosos dormem mais à noite"). Sanfey et al. (2003) pediu para indivíduos jogarem um "jogo de ultimato", no qual um jogador foi solicitado a dividir uma soma monetária com outro jogador. Quando o segundo jogador julgou uma divisão como sendo injusta, regiões emocionais do cérebro se ativaram. Singer et al. (2006) observou indivíduos enquanto choques elétricos foram administrados em pessoas ([que eram] realmente colaboradores do experimento) que jogaram ou justamente ou injustamente em um jogo anterior ao dilema do prisioneiro. Áreas associadas com emoções negativas e sofrimento vicário eram mais ativas quando os voluntários assistiram pessoas justas recebendo choques. Berthoz et al. (2002) deu a

indivíduos histórias em que as regras sociais foram quebradas, e as contrastou com casos de situações que são apenas socialmente embaraçosas. Por exemplo, indivíduos ou ouviram falar de uma pessoa que rudemente cospe comida em um guardanapo em um jantar, ou sobre uma pessoa que inocentemente cospe comida enquanto se asfixia em um jantar. As violações de regras sociais foram associados a uma maior ativação emocional. As estruturas que estão implicadas nestes estudos incluem o lobo da ínsula, o córtex cingulado anterior, o polo temporal, o giro frontal médio, e o córtex orbitofrontal, os quais são todos participantes regulares em estudos das emoções (Phan et al., 2002). Juízos morais e emoções parecem coincidir no cérebro, assim como o emocionismo epistêmico prevê. A explicação natural destas descobertas é que os julgamentos morais são constituídos por respostas emocionais.

Deve se admitir, no entanto, que esta não é a única explicação. A história de Milgrim e estudos de neuroimagem mostram que os juízos morais têm custos emocionais, mas que são consistentes com dois modelos diferentes de como as emoções se relacionam com julgamentos morais: um modelo causal e um modelo de constituição. O modelo causal diz que os juízos morais podem ter efeitos emocionais. Isto é incontroverso. Qualquer um que pensa que nos importamos com a moralidade estaria disposto a dizer que os juízos morais causam emoções. Música, eventos esportivos e tempo ensolarado, todos, causam emoções também, mas eles não são constituídos pelas emoções. Em um modelo causal, os juízos morais ocorrem antes das emoções, e são, portanto, independente das emoções. No modelo de constituição, conceitos como CERTO e ERRADO, literalmente, contêm emoções como partes componentes. Isto é o que emocionistas epistêmicos têm em mente. A evidência até agora não pode decidir entre essas duas possibilidades. Para dar suporte ao modelo de constituição, evidências adicionais são necessárias.

1.2.2. As emoções influenciam os julgamentos morais

O emocionista pode fazer progresso mostrando que as emoções realmente influenciam nossos julgamentos morais. Se julgamentos morais compreendem

emoções, então essa influência pode ser explicada. Se o julgamento de que algo é errado contém indignação, então se tornar indignado promoveria esse julgamento. Por analogia, suponha que o julgamento de que algo é divertido contém divertimento. Mais especificamente, imaginar que quando julgamos algo ser divertido, estamos fazendo um julgamento da forma "que coisas causam este estado", onde "este estado" é um demonstrativo interior apontando para a diversão. Entreter-se promove o julgamento de que algo é divertido, fornecendo-nos uma das suas partes constituintes.

Há várias formas de mostrar que as emoções promovem e influenciam julgamentos morais. Considere, por exemplo, intuições morais sobre matar e deixar morrer. Nós tendemos a pensar que matar é pior. Por que é assim? Uma resposta é que matar desperta emoções negativas mais fortes. Pense sobre isso a partir da perspectiva da primeira pessoa. Se suas ações permitem que alguém morra, e esta não é sua intenção primária, você pode tirar o foco da vítima e se concentrar em qualquer coisa que seja sua intenção primária. Quando você imagina deliberadamente tirar uma vida, você não pode tirar o foco da vítima, pois sentimentos negativos trazidos por simpatia com a vítima são provavelmente mais fortes e inelutáveis. Pode ser que matar pareça pior como resultado destas emoções fortes.

Esta ideia pode explicar intuições sobre o caso do bonde (Thomson, 1976). Nestes experimentos mentais, nós somos normalmente solicitados a comparar dois cenários. Em ambos, um bonde está indo em direção a cinco pessoas que tinham sido amarradas aos trilhos. Você não está próximo o suficiente para libertá-las, mas você pode salvá-las. Em um cenário, você pode fazer isso empurrando uma pessoa para fora de uma passarela no caminho do bonde, matando-a, e fazendo o bonde parar. No outro cenário, você pode salvar as cinco puxando uma alavanca que muda o bonde para outra pista, onde se sabe que uma pessoa está amarrada, em vez de cinco. Em ambos os casos, a sua intervenção resultaria em que haja uma morte em vez de cinco. Muitas pessoas têm a intuição de que é moralmente inaceitável intervir no primeiro caso, e moralmente admissível intervir no segundo (Mikhail, 2000). Empurrar alguém para a frente de um bonde parece errado, mas

parece correto desviar um bonde de cinco pessoas em direção a uma. Por que é assim? Uma resposta popular entre os filósofos é que matar uma pessoa é moralmente pior do que deixar alguém morrer. No caso de empurrar, estamos matando alguém, mas, no caso da alavanca, nós estamos meramente deixando alguém morrer. Outra explicação é que matar apenas desperta emoções mais intensas. Nós não queremos empurrar ninguém para o trilho do bonde, pois fazer isto nos enche de horror, e o sentimento negativo nos faz pensar que a ação é errada.

A resposta filosófica é compatível com a resposta emocional. Na versão filosófica, temos duas regras: uma que diz que não devemos matar, e outra que diz que devemos salvar vidas, e a primeira é mais forte do que a última. Não temos uma regra contra deixar as pessoas morrerem, ou pelo menos não uma regra muito forte. Assim, salvar prevalece sobre deixar morrer, e matar prevalece sobre salvar. Mas quais são exatamente estas regras, psicologicamente falando? Uma resposta é que elas são baseadas em emoções. Nós temos sentimentos negativos sobre matar, e sentimentos positivos sobre salvar vidas, e poucos sentimentos sobre deixar morrer. Quando consideramos dilemas, os sentimentos mais fortes vencem. Esta história faz duas previsões-chave. Uma delas é que as emoções devem entrar em operação quando consideramos dilemas morais, e a outra é que as nossas intuições sobre o que é certo devem ser influenciadas pelas mudanças no conteúdo emocional dos cenários que consideramos. Se as regras morais são baseadas em emoções, então os fatores que alteram nossas emoções devem afetar a nossa aplicação dessas regras.

Greene et.al (2001) utilizam a técnica de tomografia de ressonância magnética funcional para medir a atividade cerebral quando os sujeitos são defrontados com as situações do bonde. Eles mostraram uma ativação significativa em áreas emocionais do cérebro quando os voluntários foram questionados sobre se seria apropriado empurrar alguém para fora de uma passarela no caminho de um bonde. As ativações emocionais foram menores quando os voluntários foram questionados sobre se seria apropriado puxar uma alavanca que desviaria um bonde para longe de cinco pessoas em direção a uma pessoa. Greene et al. notam

também que, nos cenários de puxar a alavanca, os voluntários também mostraram ativações cerebrais em áreas associadas com a memória de trabalho. Com base nisso, os autores sugerem que o raciocínio moral é dirigido por dois processos dissociáveis: um processo racional ponderado e um processo emocional. Eu interpreto os seus dados diferentemente. Eu suspeito que as emoções estejam envolvidas em ambos os casos. A partir da explicação emocionista que eu há pouco esbocei, nós temos uma regra fundamentada na emoção de que é ruim matar, e uma regra fundamentada na emoção um pouco mais fraca de que é bom salvar vidas. No caso do empurrão, nós imaginamos matar de uma forma muito vívida, e a surra [*wallop*] emocional embalada pela regra "não matarás!" esmaga as emoções mais fracas associadas com a regra "salve vidas!". No caso de puxar a alavanca, nós não imaginamos o dano que estamos causando muito vividamente, de maneira que a regra "salve vidas!" pode guiar as nossas ações. Aqui os números importam. Nós calculamos que puxar a alavanca resultará em mais vidas salvas, resultando numa preferência emocional para puxar a alavanca. As ativações nas áreas da memória de trabalho resultam do fato de que nossas decisões dependem da reflexão sobre os números. Nós podemos calcular friamente qual curso de ação vai salvar mais vidas, mas, assim que descobrimos que é moralmente melhor puxar a alavanca, este julgamento pode ser respaldado por uma resposta emocional. Isto é consistente com os dados. Greene et al. descobriram que as emoções são ativas tanto durante o cenário do empurrão, quanto durante o cenário da alavanca. As emoções são mais intensas no caso do empurrão, mas isso não é surpresa: empurrar alguém para a morte é uma atividade muito evocativa.

Se eu estou certo, nós deliberamos sobre dilemas morais comparando emoções com emoções. Regras conflitantes têm forças emocionais diferentes, e as emoções mais fortes vencem. Se isso é correto, então deve ser possível alterar intuições sobre o caso do bonde alterando os cenários de forma emocionalmente significativa. Aqui está uma predição. Quando os voluntários dizem que é moralmente aceitável puxar a alavanca para salvar cinco pessoas e matar uma, eles estão imaginando que a alavanca está muito longe da pista. Agora, suponha que nós contássemos aos voluntários que a alavanca está a poucos centímetros de

distância da pessoa que seria morta se a alavanca fosse puxada. Imagine-se nessa situação. Um homem está amarrado aos trilhos bem perto de você. Você não pode libertá-lo. Ele está se contorcendo e gritando de terror. Você sabe que há cinco pessoas na outra pista, que está a certa distância, e você sabe que o bonde está indo naquela direção. Você sacrificaria a pessoa aos seus pés? Isto seria moralmente aceitável? Aqui, eu penso que as intuições mudariam. Este é quase como o caso do empurrão. Pessoas que não têm sido expostas a muitos destes exemplos teriam, por padrão, sérias dúvidas morais sobre sacrificar a vida de alguém a centímetros de distância. As fortes emoções suscitadas pela proximidade com a vítima influenciariam, eu suponho, o julgamento.

Por outro lado, nós podemos imaginar uma variante emocionalmente atenuada sobre o caso da passarela. Agora você está localizado em uma sala de controle, e descobre que um bonde está indo em direção a cinco pessoas. Ao puxar uma alavanca, você pode abrir um alçapão, fazendo com que uma pessoa que está em pé sobre uma passarela caia no trilho do bonde, descarrilhando-o. Neste cenário, nenhum contato físico com a vítima é exigido. Isto foi recentemente testado por Greene et al. (a ser publicado)¹ e eles descobriram que a maioria das pessoas pensa que é admissível matar a pessoa sobre a passarela nesta variante. Quando os sujeitos foram informados que eles têm que empurrar o homem para fora da passarela, apenas 31 por cento dizem que é admissível, mas quando eles foram informados que só precisavam puxar uma alavanca que abre um alçapão, 63 por cento acham que é admissível. Diminuir a intensidade emocional do método de matar dobra o índice de aprovação.

Esses exemplos sugerem que nós não somos escravos de um princípio que [diz que] matar é pior do que deixar morrer. Nós normalmente aderimos a tal princípio, mas uma mudança na intensidade emocional pode nos levar a endossar claras violações do mesmo. Além disso, o próprio princípio pode ser parcialmente subscrito pelo fato de que matar é geralmente mais carregado emocionalmente do que deixar morrer. Matar geralmente envolve entrar em contato fisicamente com

¹ Já publicado: : Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111 (3): 364–371 [N.T.].

outra pessoa e experienciar perceptualmente o sofrimento dessa pessoa. Podemos deixar alguém morrer sem qualquer contato (como nós muitas vezes fazemos com crises distantes ao redor do mundo). Eu não estou afirmando que não há diferença moral entre esses casos (para esse ponto de vista, ver Kagan, 1989). Meu ponto é que nossas intuições morais sobre tais casos são influenciadas por emoções.

Considere mais uma vez o caso do bonde (para uma pesquisa mais completa, consulte o capítulo 7 a seguir, e Prinz, a ser publicado a)². Quando você se recusa a empurrar uma pessoa na frente de um bonde em alta velocidade para salvar cinco vidas, você está fazendo um julgamento moral deontológico. Você está do lado daqueles filósofos morais que afirmam que matar intencionalmente uma pessoa é errado, independentemente das consequências. Você deve obedecer ao princípio da humanidade: você não pode usar um ser humano como um meio, mais do que como um fim. Mas tais intuições deontológicas podem, reconhecidamente, ser preponderantemente alteradas pelos números. Suponha que, em vez de cinco pessoas ligadas aos trilhos, o bonde está cheio com explosivos potentes e indo em direção a uma aldeia, na qual ele vai detonar, matando 500 pessoas. Agora parece que empurrar a pessoa para o trilho, fazendo com que o bonde descarrilhe, seria moralmente recomendável. Nós deixamos de ser deontologistas para ser consequencialistas. Essa troca em intuições é um embaraço para os filósofos que pensam que as teorias deontológicas e consequencialistas estão em competição. Mas suponha que nenhuma teoria está correta. Suponha que o conceito do bem não é o conceito sobre calcular as melhores consequências ou o conceito de seguir estritamente as regras que obedecem ao princípio da humanidade. Suponha, em vez disso, que o conceito do bem é o conceito daquilo que causa fortes emoções de aprovação. Em alguns casos, a ação condenada por princípios deontológicos causa aprovação, e, em outros casos, nós aprovamos as exigências consequencialistas. No presente exemplo, aquela mudança é explicada pelo fato que imaginar 500 pessoas mortas nos enche de um sentimento intenso de horror. A escala das perdas nos impressiona profundamente. A diferença emocional entre cinco vidas

² Já publicado: W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality*. Cambridge, MA: MIT Press, 2007 [N.T.].

perdidas e uma é grande, mas não enorme. Não é grande o suficiente para compensar a repulsa que nós sentiríamos empurrando uma pessoa para o caminho de um bonde em alta velocidade. Mas a enormidade da perda no caso dos explosivos supera a repulsa de matar um único indivíduo. Eu penso que nossas emoções influenciam nossos julgamentos.

Estes exemplos sugerem que nossos julgamentos morais estão ligados de um modo essencial às emoções. Se as emoções fossem meramente concomitantes aos julgamentos morais, então elas não deveriam influenciar aqueles julgamentos. O fato de que nós somos influenciados por nossas emoções está previsto pela hipótese de que as emoções são a base de nossos julgamentos e, talvez, [suas] partes constitutivas.

Pode-se responder a esta linha de argumentação salientando que, embora as emoções possam orientar julgamentos morais, elas não precisam. Um deontologista dedicado diria "encher-me-ia de uma angústia indizível permitir a dizimação de uma aldeia, mas é ainda errado, para mim, evitar este resultado tirando uma vida humana". Julgamentos morais e emoções parecem ser dissociáveis deste modo. Isto não minaria a exigência emocionista?

Vou adiar a discussão séria deste tipo de objeção até o capítulo 3. Por ora, vou mencionar quatro maneiras nas quais uma teoria emocionista poderia acomodar o deontologista que insiste que não há problema em dizimar a aldeia. Primeiramente, o deontologista pode ser bastante apaixonado pelo princípio da humanidade. O seu investimento emocional no princípio de que é errado usar uma pessoa como um meio para salvar outras pode ser forte o suficiente para superar considerações de compensação. Segundo, julgamentos morais podem depender de determinados tipos de emoções, e outros não. Quando o deontologista diz que é correto não empurrar a pessoa para os trilhos, ele pode estar reconhecendo que iria se sentir culpado se o fizesse. Se ele permitisse que os moradores da aldeia morressem, ele poderia sentir uma tristeza profunda, mas não culpa. A tristeza pode ser mais intensa do que a culpa que ele sentiria se ele tivesse empurrado a pessoa para o caminho do trem, mas seria a emoção errada. Emoções não-morais podem alimentar emoções morais, mas deliberações cuidadosas podem manter estas

separadas. Em terceiro lugar, o deontologista pode estar julgando que as nossas emoções estão equivocadas neste caso. Por analogia, imagine a vítima angustiada por um crime que condena um suspeito acusado falsamente. A angústia provoca a reprovação, mas ela está direcionada à pessoa errada. Da mesma forma, quando nós imaginamos quinhentos moradores morrendo, a angústia faz com que nós procuremos um criminoso, e nós podemos condenar uma pessoa cujas ações ou inações seriam irrepreensíveis se nós considerássemos o cenário em um momento mais calmo, e com todos os detalhes. Finalmente, o deontologista pode estar autoenganado. Suponhamos que ele permite que quinhentas pessoas morram, e então sente uma culpa intensa. Ele poderia continuar a insistir que não acredita que a ação foi errada, mas nós podemos questionar a sua autoavaliação. Podemos dizer, "claramente, você tem dúvidas morais sobre esta ação; claramente, ela parece errada para você". Nós podemos afirmar que ele está meramente falando da boca para fora quando diz que a sua inação foi correta. Ou se pode supor que ele reconhece indefectivelmente que a ação foi correta em um sentido não-moral (ele fez como a razão exigiu), mesmo reconhecendo dolorosamente que a sua inação entrou em conflito com os seus valores morais básicos.

As intuições sobre os casos do bonde não provam que os julgamentos morais envolvem as emoções *necessariamente*, mas eles sugerem que as emoções podem exercer uma grande influência sobre os julgamentos morais. Esta conclusão ganha apoio adicional da pesquisa sobre os efeitos da indução emocional. Em um estudo, Wheatley e Haidt (2005) hipnotizaram indivíduos para sentir uma pontada de desgosto quando ouvem ou a palavra "aceitar" ou a palavra "frequentemente". Eles são, então, convidados a avaliar moralmente o protagonista de várias histórias contendo uma dessas duas palavras. Por exemplo, eles ouvem sobre um congressista que "é frequentemente subornado" ou "aceita subornos". As avaliações de injustiça sobem quando a escolha da palavra corresponde à palavra que desencadeia desgosto no indivíduo. De fato, quando a palavra desencadeadora é usada em histórias neutras, os indivíduos tendem a condenar o protagonista também. Por exemplo, eles ouvem sobre um estudante encarregado de discussões de programação na escola, o qual muitas vezes escolhe temas interessantes. Os

sujeitos que estão enojados quando ouvem a palavra "frequentemente" acham este estudante moralmente suspeito, embora eles não possam dizer por quê ("ele parece estar tramando algo"). Em outro estudo, Schnall et al. (2005) pediram aos sujeitos para fazer avaliações morais de histórias enquanto estavam sentados em um balcão que estava ou limpo ou sujo. O balcão sujo tem uma caixa de pizza velha gordurosa ao lado dele, um lápis mastigado, lenços de papel usados e um copo de bebida sujo. Os indivíduos que são bons na introspecção de suas emoções (enquanto medidas por uma escala do corpo autoconsciente que está correlacionada com a consciência emocional) responderam diferentemente. Aqueles que estavam sentados no balcão sujo julgaram os cenários serem piores do que os indivíduos sentados no balcão limpo. Por exemplo, eles deram classificações mais elevadas de injustiça a um cenário que descreve uma pessoa que acidentalmente mata seu cão de estimação e depois o come. Estes efeitos não dependem de nojo. Lerner et al. (1998) mostraram aos sujeitos cliques de filmes que eram neutros ou evocativos de raiva. Eles foram, então, convidados a considerar algumas vinhetas independentes que descrevem pessoas que cometem transgressões relativamente menores, tal como a venda de um carro usado sem revelar um defeito. Os sujeitos que viram os cliques que induzem à raiva recomendaram penas mais severas para os autores dessas vinhetas. Além disso, alguns estudos demonstraram que a indução de humor triste pode levar a avaliações mais negativas das pessoas (Fogas e Bower, 1987). Por outro lado, a atração física, que é conhecida por induzir afeto positivo, pode promover avaliações positivas de pessoas, incluindo avaliações de honestidade e integridade (Dion et al., 1972). Foi demonstrado também, em estudos de júri, que os réus atraentes ou sorridentes são tratados de forma mais branda (Darby e Jeffers, 1988).

Juntos, esses resultados sustentam o caso do emocionismo epistêmico. Eles sugerem que as emoções podem influenciar as avaliações morais, mesmo quando as emoções são induzidas por fatores moralmente irrelevantes. Isto é apenas o que o emocionismo epistêmico prevê. O emocionismo epistêmico fornece uma explicação natural do fenômeno: se os conceitos morais têm um componente emocional, então a indução das emoções deve influenciar a aplicação desses

conceitos. Compare: se o conceito ENGRAÇADO contém a emoção de prazer, então fazer cócegas nas pessoas deve aumentar a sua tendência para pensar que uma brincadeira é engraçada. Se as cócegas são óbvias demais, eles irão atribuir prazer a elas e não à brincadeira, mas, se a cócega é sutil, eles podem classificar a brincadeira como mais divertida do que ela seria de outro modo.

Eu não estou sugerindo que este é um argumento arrasador [knock-down] para o emocionismo. O fato de que as emoções influenciam os julgamentos morais não implica que os julgamentos morais contêm emoções. As emoções podem influenciar os julgamentos morais de outra maneira. Por exemplo, a empatia com a vítima de um crime poderia incutir o desejo de punição, e isto poderia nos levar a pesar evidências seletivamente ao atribuir a culpa a um suspeito. Deste ponto de vista, a atribuição de culpa não precisa ser um julgamento emocional em seu sentido próprio, apesar das emoções desempenharem um papel ao realizá-lo. Emoções desempenham um papel causal, aqui, mas elas não são constitutivas.

Eu penso que o emocionismo epistêmico oferece uma melhor explicação de como as emoções influenciam julgamentos morais. Nos experimentos controlados, não há dúvidas sobre se os personagens descritos em várias vinhetas são culpados. O rapaz que come o seu cão de estimação é claramente responsável por suas ações. Assim, as emoções não podem estar afetando julgamentos sobre injustiças influenciando a forma como as pessoas pesam provas. Além disso, há dissociações entre o desejo de punição e juízos sobre injustiças. Compare um assassino louco com um machado a um assassino calculista que usa uma arma. Por causa de sua insanidade, nós podemos pensar que o assassino do machado deve ser punido menos severamente do que o assassino da arma, mas, porque assassinatos com machado são mais horríveis, nós podemos julgar que seus crimes são mais errados. Esta é apenas a minha intuição, mas seria fácil testá-la.

Claro, o adversário do emocionismo poderia conceber outras explicações para por que emoções influenciam julgamentos morais se o desejo de punição não acomodar todos os dados. Eu penso que a principal razão para preferir a explicação emocionista é que ela se encaixa melhor com os resultados que estou prestes a descrever. Se as emoções não eram constituintes dos julgamentos morais, mas

apenas exerceram uma influência causal, então as emoções não seriam nem necessárias nem suficientes para considerar algo como errado. O conceito de incorreção seria algo acima e além das emoções e, portanto, independente delas. Sobre a explicação da influência causal, deve haver casos em que as pessoas moralizam sem ter emoções, e não deve haver casos em que só as emoções são, pensando bem, a única base de julgamento moral. Como veremos, estas previsões da explicação da influência causal estão incorretas. A visão de que as emoções constituem os nossos juízos morais se encaixa melhor com os dados.

1.2.3 Perplexidade

Os elementos apresentados até agora mostram que as emoções podem influenciar o nosso julgamento moral. Mas emocionistas epistêmicos fazem uma reivindicação mais forte. Eles dizem que ter uma atitude moral é uma questão de ter uma disposição emocional. Se isso está correto, então alguém deve ser capaz de ter uma atitude moral na ausência de qualquer justificação racional. As atitudes emocionais devem ser suficientes para as atitudes morais. Há evidências empíricas apoiando esta previsão. Os julgamentos morais de pessoas reflexivas parecem ter uma base emocional. Se perguntarmos às pessoas por que elas têm uma visão moral particular, elas podem oferecer algumas razões, mas essas razões são frequentemente superficiais e ulteriores. Se as razões são desafiadas com sucesso, o juízo moral frequentemente permanece. Quando pressionadas, os valores morais mais profundos das pessoas não estão baseados em argumentos decisivos que elas descobriram enquanto ponderavam sobre questões morais, mas em sentimentos profundamente inculcados.

Para sustentar esse modelo, Murphy et al. (2000) estudaram as atitudes morais em relação ao incesto consensual. Eles solicitaram a estudantes universitários americanos considerar um caso em que um irmão e uma irmã fazem sexo. No cenário, os irmãos consentem em ter relações sexuais, usar contraceptivo, desfrutar da experiência, e mantê-la em segredo. Oitenta por cento dos indivíduos considerou que o comportamento é moralmente errado, mas tiveram grande

dificuldade em explicar o porquê. A cada vez que eles vieram com um argumento para mostrar que os irmãos tinham feito algo imoral, os pesquisadores explicaram por que o argumento falha. Muitos indivíduos se preocuparam com o fato de que o casal teria crianças deformadas, mas os pesquisadores lembraram a eles que a contracepção foi usada. Alguns estavam preocupados com os efeitos sobre a comunidade, mas tal preocupação é inaplicável, porque o casal no cenário não contou a ninguém o que eles tinham feito. Alguns indivíduos teriam se queixado de que o casal poderia ficar traumatizado, mas o cenário especifica que eles realmente gostaram da experiência e que ela fortaleceu o relacionamento deles. Poucos indivíduos sugeriram que o incesto é condenado na Bíblia, mas nenhum conseguia se lembrar de onde (certamente não na história de Lot e suas filhas!). Aos indivíduos foram apresentados contra-argumentos decisivos para cada argumento que eles deram contra o incesto consensual. Eles tenderam a admitir que os contra-argumentos foram bem-sucedidos, mas apenas 17 por cento mudou os seus julgamentos morais iniciais. Os outros se nivelaram tipicamente por baixo em declarações insustentáveis e exclamações emocionais. Incesto é perverso! O incesto é simplesmente errado: é rude! As razões caíram no esquecimento, mas convicções morais e emoções morais eram recalcitrantes.

Murphy et al. (2000) encontraram o mesmo padrão de respostas quando eles apresentaram aos indivíduos um cenário envolvendo canibalismo. Uma mulher trabalhando sozinha tarde da noite em um laboratório de patologia médica decide cozinhar e comer um pedaço descartado de um cadáver que foi doado para o laboratório para a pesquisa médica. Mais uma vez, os indivíduos dizem que isso é errado, mas eles não conseguem articular as razões suficientes para apoiar essa conclusão. Eles dizem que sua avaliação moral do caso é baseada em um sentimento inquestionável.

Estes resultados inquietantes podem ser interpretados de várias formas. Uma possibilidade é que os sujeitos têm boas razões para seus pontos de vista sobre incesto e canibalismo, mas estas razões operam inconscientemente. Afinal, muitas das soluções de problemas são feitas inconscientemente, e as pessoas têm uma compreensão limitada de como elas chegam a julgamentos em outros

domínios (Nisbett e Wilson, 1977; Moscovitch, 1995). Chamo a isto de interpretação de *razões ocultas*.

Esta interpretação me parece altamente improvável. Nós podemos chegar à nossa avaliação moral sobre o incesto inconscientemente, mas não há evidência de que muito raciocínio está ocorrendo. É muito difícil saber que razões seriam essas. É uma coisa dizer que as razões não são acessíveis à consciência quando chegamos inicialmente aos nossos julgamentos, e outra dizer que a reflexão amplamente cuidadosa não pode ter acesso a eles. Propor argumentos contra o incesto consensual é difícil, e uma vez que há pouca discussão pública sobre o incesto, é difícil acreditar que os indivíduos internalizaram argumentos a partir de reflexão anterior ou da educação. Além disso, há uma explicação clara de como as pessoas chegam aos seus julgamentos morais que não requerem a postulação de razões ocultas. Incesto e canibalismo têm sido marcados emocionalmente como repulsivos e como tabus. Isto explica detalhadamente as condenações morais instintivas de seus autores. E a nossa tenacidade em denunciar o incesto e o canibalismo deriva do fato de que as razões não são facilmente substituídas pelas respostas emocionais profundamente arraigadas.

De acordo com uma segunda interpretação dos resultados perplexos, as pessoas sempre baseiam os seus julgamentos morais sobre razões, mas estas razões são, por vezes, ruins. De fato, as pessoas oferecem argumentos contra o incesto e o canibalismo. Só que os seus argumentos se provam falhos. Os resultados de Murphy et al. são consistentes com a hipótese de que juízos morais derivam da razão, em vez da paixão.

O problema com esta interpretação é que as pessoas geralmente não reveem a sua avaliação moral quando as suas razões são desmascaradas. Elas reconhecem que os motivos são falhos, mas se recusam a mudar de ideia sobre o erro do incesto consensual e do canibalismo. Isso sugere que as razões que elas oferecem não desempenhou um papel central na formação ou manutenção de seus julgamentos morais. Outra possibilidade é que os indivíduos não consideram realmente o incesto e canibalismo como imorais. Talvez eles estejam somente dizendo que estas coisas são erradas, porque eles reconhecem que essa é a visão

predominante. Endossar um comportamento tabu em público tem graves consequências sociais, de modo que os indivíduos têm boas razões para fazer parecer com que eles achem o incesto e o canibalismo ruins.

Esta interpretação também não é convincente. Se tabus são poderosos o suficiente para fazer as pessoas dizerem que se opõem categoricamente ao incesto e ao canibalismo, então eles devem ser fortes o suficiente para incutir as crenças correspondentes. A questão poderia ser testada fazendo as pessoas responderem perguntas sobre o incesto consensual e o canibalismo em um questionário anônimo. Eu poderia prever que os indivíduos continuariam a condená-los.

Uma quarta possibilidade é que os indivíduos não têm razões para os seus julgamentos morais. Eles simplesmente têm uma reação instintiva de que o incesto consensual e o canibalismo em laboratório estão errados, e poucas racionalizações *post hoc*, as quais não desempenham nenhum papel importante na condução dessas reações.

Eu acho que esta proposta é quase correta, mas ela é um pouco enganadora. Geralmente, se você não tem nenhuma razão para uma crença, você está racionalmente obrigado a desistir dela, mas eu não acho que as pessoas consideram as suas atitudes morais como sujeitas a esta exigência. Os valores podem ser básicos de uma maneira que os coloca fora do jogo da razão. As pessoas tendem a expressar seus pontos de vista sobre o incesto e o canibalismo dizendo: "É simplesmente errado!". O meu palpite é que elas diriam a mesma coisa sobre matar ou infligir dano em uma pessoa inocente. Considere a questão "Por que é errado estuprar uma criança que nunca vai se lembrar do incidente?". Esta é uma pergunta estranha. É difícil de responder. É simplesmente errado fazer isso. Muito errado. Fundamentalmente errado. E moralmente monstruoso. Quando dizemos: "É simplesmente errado", não estamos antecipando a razão; nós estamos implicitamente dando uma. O "simples" em "simplesmente errado" designa que isto é um valor básico. Estamos no fundo do poço. Alguém que sinceramente afirma que ele não considera errado estuprar uma criança não entende o que nós queremos dizer com "errado". Ele está usando a palavra de forma diferente. Compare alguém que insiste que os morangos não são vermelhos.

Isso revela algo sobre a prática de dar razões na moralidade. Quando fornecemos uma razão para pensar que determinado comportamento é errado, implica que a sua incorreção consiste no fato de que ele tem uma propriedade particular que o torna errado. Mas suponha que iteramos a pergunta sobre o porquê. Por que é errado dirigir bêbado? A resposta é que isto põe em risco vidas inocentes. Por que é errado as pôr em perigo? Porque o perigo é o risco de dano, e prejudicar uma pessoa inocente é errado. Por que é errado prejudicar uma pessoa inocente? Aqui a questão se torna estranha. Filósofos treinados podem ter pontos de vista sobre isto, e outros podem ser capazes de chegar a ter razões, mas é pouco provável que essas razões sejam a fonte da intuição moral. Se se pudesse chegar a algum aspecto que torna errado matar, poderíamos perguntar o que faz desse aspecto errado. Em algum momento, vamos nos tomar dispersos. Em algum momento a pergunta sobre o porquê parece descabida, bizarra, ou até mesmo depravada. Podemos dizer que as pessoas não têm nenhuma razão para os seus valores básicos, mas seria melhor dizer que os valores básicos são implantados em nossa psicologia de um modo que se colocam fora de certas práticas de justificação. Os valores básicos fornecem razões, mas eles não são baseados em razões.

Volto aos valores básicos no capítulo 3 sob o rótulo "normas fundantes". Eu os apresento aqui como uma forma de explicar os resultados inquietantes de Murphy et al. As pessoas ficam confusas quando pedimos para explicarem a sua condenação do incesto, porque este é um valor básico. Além disso, os valores básicos parecem ser implementados de um modo emocional. Quando nós descemos aos valores básicos, as paixões regulam. As pessoas dizem que o incesto e o canibalismo são repugnantes. O assassinato é abominável. Roubar é inconcebível. Um membro típico dessa cultura suportaria uma considerável punição emocional por cometer qualquer um desses atos.

1.2.4. Desenvolvimento moral

Se esta interpretação dos resultados de Haidt está correta, os adultos normais têm valores que não são mantidos por uma rede de razões cuidadosamente pensada. Eles são implementados por sentimentos viscerais. Esta imagem ganha mais apoio a partir da pesquisa sobre o desenvolvimento moral.

A teoria mais amplamente discutida do desenvolvimento moral foi proposta por Laurence Kohlberg (1984). Kohlberg pediu aos sujeitos para resolverem dilemas morais. Por exemplo, ele lhes conta sobre um homem chamado Heinz que não pode pagar por um medicamento que ele necessita para salvar a sua esposa de um câncer. Depois das tentativas sem sucesso para conseguir a droga pelos meios legais, Heinz deve entrar no laboratório e roubá-la? Kohlberg avalia a compreensão moral olhando como os sujeitos justificam as suas respostas em tais casos. Com base nesta pesquisa, ele concluiu que as crianças passam por uma progressão de estágios no desenvolvimento moral. Kohlberg identifica seis estágios, agrupados em três níveis. Na primeira fase, as crianças se concentram na obediência e na punição. Elas justificam julgamentos morais apelando para as respostas punitivas das autoridades. Depois disso, as crianças começam a pensar instrumentalmente sobre a moralidade; elas pensam sobre os benefícios para o agente moral. Kohlberg chama esses dois estágios de moralidade pré-convencional, sendo caracterizados por uma orientação egoísta. Esta é seguida por um nível convencional de duas fases do pensamento moral, no qual as crianças começam a pensar em conformidade a um grupo. No primeiro estágio da moralidade convencional – estágio três da sequência total – as crianças adotam uma orientação de "bom menino/boa menina". Eles começam a se concentrar em como eles serão vistos pelos outros. Na fase quatro, há um foco na lei e na ordem. Nesta fase, há um foco no dever de fixar regras e na manutenção da ordem social. Kohlberg pensa que a moral convencional pode ser seguida pela moralidade pós-convencional, mas reconheceu, empiricamente, que este nível final de desenvolvimento é raramente atingido, mesmo entre os adultos. A moralidade pós-convencional começa com um quinto estágio de desenvolvimento moral em que as pessoas se concentram em

contratos sociais. Nesta fase, as pessoas continuam a pensar em termos da lei e da ordem, mas agora eles justificam as leis pelo apelo a princípios amplamente utilitários. Há um sexto estágio potencial, o qual Kohlberg caracteriza em termos de princípios morais universais. Estas regras são abstratas e categóricas (como a Regra de Ouro), em vez de concretas e particulares (como os Dez Mandamentos). Em outras palavras, Kohlberg pensa que o desenvolvimento moral deve, mas raramente o faz, em última análise, levar-nos a uma concepção kantiana da moralidade. Em um estudo longitudinal, nos Estados Unidos, Colby et al. (1983) encontraram pouca evidência para o raciocínio nas etapas cinco e seis. A maioria dos adultos raciocinam na fase quatro a maior parte do tempo. Em uma revisão de pesquisa transcultural, Snarey (1985) descobriu que o estágio quatro foi o estágio mais alto exibido nas sociedades rurais e nas aldeias.

As descobertas de Kohlberg são consistentes com a visão de que as emoções são essenciais para o julgamento moral. Primeiramente, a ausência relativa de raciocínio nos níveis quinto e sexto sugere que pessoas comuns não são nem utilitaristas nem kantianas. Os conceitos morais padrão não parecem ser fundamentados nos tipos de princípios que dominam a ética filosófica. Isto levanta algumas dúvidas sobre as explicações filosóficas nas tradições utilitaristas e kantianas. Estes programas podem ser melhor interpretados como revisionistas em vez de análises precisas de como as pessoas comuns compreendem conceitos morais (ver capítulos 3 e 4). Em segundo lugar, os primeiros três estágios de desenvolvimento moral de Kohlberg implicam emoções de forma bastante explícita. No primeiro estágio, as pessoas expressam o medo de punição. Na segunda fase, as pessoas apelam para ganhos hedônicos. Na terceira fase, as pessoas expressam o desejo de ser querido pelos outros. Simplificando, pode-se dizer que o conceito ERRADO é sequencialmente ligado a medo, frustração e, finalmente, a tristeza antecipada de rejeição social.

O que devemos dizer sobre o quarto nível de Kohlberg, que é dominante nas sociedades ocidentais industrializadas? O raciocínio neste nível não é explicitamente emocional, mas isso não significa que ele não tenha uma base emocional. No quarto nível as pessoas apelam para a lei e para a ordem. Apelos à

ordem podem ter um tom emocional. As pessoas dizem que as regras morais são justificáveis pelo fato de que a sociedade iria desmoronar sem elas. Esta justificação perde força, se se é neutro sobre o colapso da sociedade. É natural supor que pensamentos sobre o colapso social evocam medo e preocupação pelos entes queridos. Aqueles que comprometem a estabilidade social representam uma ameaça ao bem-estar e são, portanto, vistos com desprezo ou com raiva. Se assim for, apelar para a ordem pode refletir numa atitude emocional. Agora, voltemos da ordem para a lei. Às vezes, em vez de levantar preocupações sobre o colapso da sociedade, as pessoas tentam justificar os seus julgamentos morais simplesmente citando a existência de uma lei ou de uma política. Com efeito, elas dizem que algo está errado, porque há uma regra contra isso. Essas pessoas tratam as regras como se elas tivessem um valor intrínseco. Esse tipo de justificação é de fato previsto pelo emocionismo epistêmico. Se nós fundamentamos normas em reações emocionais, então as nossas convicções morais carecem de uma fundamentação racional. Agora, suponha que você dê aos sujeitos uma tarefa de raciocínio, em que eles são convidados para justificar sua crença de que a ação- ψ é errada, de modo que eles se veem incapazes de articular quaisquer razões porque a regra é baseada em emoções. Neste ponto, as justificações começarão a soar circulares. Por que a ação- ψ é errada? Bem, ela simplesmente é. Em outras palavras, a ação- ψ é errada, porque essa é a regra.

O emocionista pode explicar este tipo de fetichismo de regra. No curso do desenvolvimento moral, estamos condicionados a ter uma forte reação emocional à violação de certas regras. Apenas pensar em alguém violando essas regras provoca sentimentos negativos. Assim, as regras assumem uma espécie de obviedade. Elas são imediatamente obrigantes para nós e nós assumimos que elas são óbvias para os outros. A evocação de uma regra que foi condicionada desta forma parece suficiente para fins de justificação. Assim, o estágio quatro é como os três estágios anteriores do desenvolvimento moral, na medida em que todos eles fazem um apelo implícito às emoções. A principal mudança no desenvolvimento é que, no momento em que atingimos a quarta fase, nós atribuímos um significado emocional diretamente às regras, em vez de forma derivada. Nos níveis anteriores, a ação- ψ é

errada porque causa consequências emocionais negativas para mim. Acima do nível quatro, a ação- φ é errada porque eu a considero negativa em si mesma. Há uma transferência de emoções das consequências egocêntricas da transgressão para a transgressão em si mesma. A mudança do desenvolvimento se encaixa perfeitamente dentro de uma estrutura emocionista.

Na visão de que acabo de descrever, as pessoas no nível quatro de Kohlberg justificam julgamentos morais apelando a regras porque as regras são consideradas intrinsecamente valiosas; as regras alcançam esse *status* porque são fundamentadas na emoção, mais do que a razão. Para testar essa interpretação, as pessoas, nos experimentos do estilo de Kohlberg, devem ser pressionadas um pouco mais. Quando dizem que a ação- φ é errada por causa de uma regra contra a ação- φ , elas devem ser questionadas "Por que devemos fazer o que as regras exigem?". Eu prevejo que a maioria das pessoas teria dificuldade em articular uma resposta. Como a pesquisa inquietante de Murphy et al. (2000), eu esperaria que as pessoas ficassem atordoadas ou expressassem emoções. Eles deveriam dizer: "a ação- φ é horrível", ou algo nesse sentido. Referindo-se ao quarto estágio como "convencionalista", Kohlberg parece ter uma interpretação muito diferente em mente. O rótulo implica que as pessoas no estágio quatro pensam a moralidade em termos convencionais. Se questionado: " Por que devemos fazer o que as regras exigem?", uma pessoa com uma concepção convencional de moralidade deveria responder dizendo: "Bem, isso é o que os membros da minha comunidade fazem".

Em suma, o fato de que a maioria das pessoas progride até o nível quatro pode ser interpretado como evidência para emocionismo. Kohlberg chama este de estágio convencional, mas eu acho que esse rótulo é mal entendido. Eu acho que as regras morais são essencialmente convencionais, mas eu não acho que as pessoas as veem dessa forma. O fato de que as pessoas justificam suas atitudes morais apelando para a lei e para a ordem não implica que elas as consideram como convenções. Na interpretação alternativa que eu recém ofereci, as pessoas nesta fase realmente consideram as regras como tendo um valor intrínseco. Violar uma regra moral só faz mal. O apelo para a lei e para a ordem no raciocínio moral não é um apelo à convenção, e sim um apelo a normas emocionalmente fundamentadas.

A minha interpretação emocionista dos resultados de Kohlberg faz duas previsões empíricas. Ela prevê que o amadurecimento moral é alcançado através de um processo de treinamento emocional, e prevê que as pessoas na assim chamada fase convencional realmente não pensam que as regras morais se mantêm simplesmente em virtude das convenções sociais. Colocando essas duas predições juntas, a explicação emocionista prevê que as pessoas passam a considerar as regras morais como diferentes das regras convencionais, atribuindo significado emocional para as regras morais. Evidências da psicologia do desenvolvimento apoiam todas essas predições.

Antes de tudo, há evidências que sugerem que a educação moral é uma questão de treinamento emocional. As crianças recebem instrução moral através da manipulação cuidadosa de emoções. Psicólogos enfatizam três métodos principais utilizados pelos pais para promover a boa conduta (Hoffman, 1983; Eisenberg, 2000). Um método é o poder da afirmação. Os pais punem ou ameaçam punir os seus filhos. A punição promove medo, e, por imitação, as crianças que foram punidas por fazer algo ruim provavelmente terão raiva de outras pessoas que se comportarão mal no futuro. Outro método é a retirada do amor. Quando as crianças fazem algo ruim, os seus pais, às vezes, expressam desapontamento e se recusam a sinalizar carinho. Isso faz com que as crianças se sintam tristes, o que pode ser a fonte de arrependimento. Os pais também usam uma técnica chamada de indução. Eles chamam a atenção das crianças para os malefícios que causam os seus delitos. Quando as crianças reconhecem que as suas ações fizeram alguém sofrer, elas naturalmente sentem simpatia e sacrifício vicário pela vítima. Cada um destes métodos foram associados com o desenvolvimento de culpa e de vergonha, e com o desenvolvimento de um comportamento pró-social.

Em segundo lugar, há evidências de que as crianças comuns não consideram as regras morais como meramente convencionais. Smetana (1981), Turiel (1983) e Nucci (2001) têm pesquisas pioneiras sobre esta questão. Elas demonstraram que as crianças distinguem entre regras morais e convencionais. Por exemplo, as crianças fazem uma distinção entre as regras que proíbem bater e roubar, por um lado, e regras que proíbem falar sem levantar a mão na classe e regras que

proscurem códigos de vestimenta, por outro. Para estabelecer que esta distinção seja entendida, Smetana, Turiel e Nucci apresentam às crianças exemplos de regras de violações, não indicando quais são morais e quais são convencionais. Elas, então, fazem perguntas às crianças sobre gravidade (por exemplo, quão errado foi o comportamento no exemplo?), dependência de autoridade (por exemplo, seria errado o comportamento se os professores o permitissem?), e de justificação (por exemplo, por que é errado?). As crianças respondem a estas perguntas de forma diferente para as regras morais e convencionais. Elas tratam as violações de regras morais como mais graves e menos dependentes de autoridades. Bater seria errado, não importa o que o professor dissesse, mas falar em sala de aula seria bom, se o professor permitisse. As crianças tendem a justificar as regras morais apelando aos danos infligidos a outros, enquanto eles justificam as regras convencionais apelando para convenções.

Aos resultados de Smetana, Turiel, e Nucci pode ser dada uma explicação emocionista. Por que as crianças acham as transgressões morais mais graves? Talvez elas tenham maiores consequências emocionais. A gravidade pode ser uma avaliação emocional. Por que as crianças apelam para danos explicando o que há de errado com transgressões morais? Talvez os danos causem simpatia e angústia. Por que as crianças consideram as regras morais independentes da autoridade? Talvez pensamentos sobre transgressões morais despertem emoções negativas, e estes permanecem mesmo quando as crianças imaginam as proibições sendo suspensas.

O apoio para essa interpretação pode ser encontrada nas respostas que as crianças dão a perguntas sobre as regras morais. Por exemplo, em uma série de estudos, Nucci (2001) perguntou às crianças se roubar ainda seria errado se Deus dissesse que o roubo foi permitido. A esmagadora maioria das crianças respondeu afirmativamente. Elas insistiram que o roubo seria errado, independentemente do que Deus diz. Quando Nucci pediu às crianças para justificar as suas respostas, elas tenderam a apelar para as emoções. Aqui está o que um menino de 11 anos de idade nos diz: "se as pessoas roubassem, então o mundo não seria um lugar feliz. . . ainda faria todo mundo infeliz. . . Como quando minha irmã roubou minhas pilhas,

isto realmente me irritou. Se as coisas de todo mundo continuassem sendo roubadas, todos ficariam loucos" (Nucci, 2001: 36). Esta resposta é típica. Uma menina de nove anos de idade explica que roubar é errado, porque "aquele que foi roubado ficaria com uma raiva real" (p. 38). Uma menina de 10 anos de idade ressalta essa justificativa: "você está levando as coisas de outra pessoa e eles provavelmente ficariam irritados" (p. 47). Há evidências consideráveis de que, quando as pessoas atribuem emoções, elas também as experienciam (Goldman e Sripada, 2005). É provável, então, que as crianças do estudo de Nucci sintam uma raiva semelhante por parte das vítimas do crime. A própria ideia de roubar as faz sentir raiva porque elas imaginam a raiva dos outros.

Há evidências de que as crianças começam a apreciar as consequências emocionais do mau comportamento consideravelmente mais cedo. No momento que elas têm dois anos de idade, as crianças mostram sinais de culpa e de vergonha quando elas fazem algo errado (Barrett et al, 1993;. Zahn-Waxler e Robinson, 1995;. Kochanska et al, 1995). Em um paradigma experimental, os pesquisadores dão às crianças uma boneca de brinquedo que está preparada para se desfazer quando brincam com ela. Duas crianças de dois anos de idade mostraram sinais de angústia autoconsciente quando tais "acidentes" ocorreram. Por exemplo, elas evitam o contato visual, elas contorcem a cabeça para baixo, elas cobrem os seus rostos. Tais sentimentos negativos podem ajudar as crianças a adquirir o domínio da distinção moral convencional, que começa a aparecer um pouco antes do terceiro aniversário. Smetana e Braeges (1990) descobriram que as crianças nessa idade consideram regras morais diferentemente de regras convencionais, julgando que as primeiras são mais generalizáveis do que as últimas (por exemplo, mais propensas a serem seguidas em outras escolas). No início da infância, a distinção moral-convencional é muito melhor entrincheirada. As crianças chegam a reconhecer que os erros morais são mais independentes de autoridade e mais graves do que os erros convencionais.

Quando estão um pouco mais velhas, as crianças tornam-se muito sensíveis aos efeitos que o mau comportamento tem sobre os outros. Arsenio e Lover (1995) descreveram uma série de comportamentos ruins nos jardins da infância, e, então,

pediu-lhes para escolher expressões faciais que capturassem o que as vítimas e os observadores desses comportamentos sentiriam. As crianças confiantemente escolheram expressões faciais de emoções negativas, como medo, raiva e tristeza. Curiosamente, as crianças esperaram que os autores de transgressões morais se sentissem felizes. Este é o chamado efeito feliz das vítimas. Presumivelmente, elas pensam que as vítimas estão felizes porque seus crimes atingiram os objetivos pretendidos. Por exemplo, uma pessoa que rouba um doce ficará feliz em ter obtido o doce. As crianças do jardim da infância não julgam as vítimas felizes como sendo mais censuráveis do que vítimas que mostram remorso. Eles começam a considerar as vítimas felizes como piores do que as que mostram remorso quando eles estão com oito anos de idade. Este é um marco importante do desenvolvimento. Isto indica que as crianças de oito anos de idade simplesmente não têm normas sobre como as pessoas devem se comportar; elas têm normas sobre como as pessoas devem se sentir. As vítimas felizes cometem dois erros: um em sua conduta e outro em sua atitude.

Essa trajetória de desenvolvimento é, provavelmente, produto tanto da natureza, como da criação. As crianças estão naturalmente propensas à angústia empática, medo e vínculos pessoais. Quando elas fazem coisas ruins, elas descobrem os custos emocionais. Os custos tendem a ser especialmente piores no caso de erros morais, uma vez que violações das convenções sociais tendem a produzir reações mais fracas em outros, e os educadores respondem com formas mais severas de disciplina (Smetana, 1989; Grusec e Goodnow, 1994; Nucci e Weber, 1995). Desta forma, as crianças são condicionadas a associar os erros morais com fortes emoções negativas. Isso explica porque as regras morais são psicologicamente distintas das regras convencionais, e isso explica porque as regras morais são consideradas mais graves e menos dependentes de autoridade. A relação sistemática entre as emoções e a compreensão moral em crianças sustenta a hipótese de que os conceitos morais são emocionalmente implementados. Essa história contradiz afirmação de Kohlberg de que as pessoas tendem a entender a moral convencionalmente. O fato de que as pessoas apelam

para regras na justificação de seus julgamentos morais é mais bem explicado em termos emocionistas.

1.2.5 Maria Moral

Após revisar a literatura sobre o desenvolvimento moral, pode-se admitir que as emoções desempenham um papel central na aquisição de competência moral, mesmo insistindo que este papel é meramente contingente. As crianças não são tão sofisticadas para entender os complexos argumentos filosóficos, assim, elas devem ser emocionalmente condicionadas a se comportar de acordo com as regras morais. Mas adultos são diferentes. Nós somos receptivos à razão. Para explorar essa possibilidade, eu quero construir um experimento de pensamento que descreve um indivíduo que aprende sobre a moralidade sem o benefício do condicionamento emocional.

Imagine uma mulher chamada Maria, a qual nunca foi exposta a qualquer educação moral enquanto ela estava crescendo, mas as suas outras capacidades cognitivas se desenvolveram normalmente. Ela é agora uma adulta inteligente. Imagine que Maria não tem nenhuma atitude moral inata intacta. Ela não se sente culpada ou indignada sobre qualquer coisa. Mas ela decide que quer aprender sobre o que trata, afinal, a moralidade, de modo que ela se encerra em uma sala com as obras-primas de Kant, de Mill e de outros eticistas normativos. Ela aprende as suas teorias, tornando-se muito hábil em identificar os tipos de considerações que eles trazem para fundamentá-las. Para qualquer ação que ela considera, Maria é capaz de determinar (a) se esta maximizaria a utilidade e (b) se ela levaria a contradições práticas se fosse perseguida por todos os agentes. De fato, ela pode discernir quaisquer dos fatos ressaltados se guiando por teorias normativas. Agora, aqui está a questão crucial. Suponha que Maria descobre que fazer X, de fato, maximizará a utilidade. É suficiente para ela saber que fazer X é moralmente certo? Ela pode se perguntar se X é moralmente necessário mesmo que ela saiba que isto maximiza a utilidade? A resposta é óbvia. Maria pode se perguntar. Ela pode estar totalmente insegura sobre X ser uma ação que exige a *moralidade*. Suponha que

Maria também contemple um outro curso de ação Y. Ela sabe que fazer Y levaria a uma contradição prática, se todos o fizessem; talvez a ação exija usar outra pessoa como um meio e não como um fim. Kant diria que Y é moralmente errado, mas Maria pode suspeitar. Ela sabe que Y é praticamente irracional, mas ela não sabe se é *imoral*. Suponha que as recomendações de Kant e de Mill para a ação entrem em conflito. Maria pode decidir qual opção é moralmente superior? Certamente que não. Maria começou a sua formação se perguntando o que as pessoas estão falando quando fazem julgamentos morais, e ela ainda está no escuro; ler livros de ética normativa se provou infrutífero. Intuitivamente, Maria pode ser uma detentora perfeita dos aspectos que os eticistas normativos identificam como a base da moralidade, mas ela pode não ter nenhuma ideia se esses aspectos têm qualquer significado moral.

De fato, a Maria moral pode mesmo ser capaz de colocar as questões relevantes. Ela não poderia se perguntar se a maximização da utilidade é moralmente necessária, porque o conceito de exigência moral fugiria dela. Ela não teria nenhuma compreensão, dado o seu desenvolvimento desviante, do que as pessoas querem dizer quando se referem a algo como "certo" ou "errado". Na melhor das hipóteses, ela poderia perguntar se estas palavras, usadas por membros da comunidade dela, designam as propriedades discutidas por Kant e por Mill. Ela pode falar da boca para fora as palavras "certo" e "errado", mas ela não pode entendê-las. Isto sugere fortemente que os conceitos de certo e de errado (ou bom e mau) não são explicáveis em termos dos conceitos introduzidos por Kant, Mill e outros eticistas normativos.

Este experimento de pensamento é apenas uma versão do "argumento da questão aberta" de G. E. Moore. Moore (1903) tentou refutar a tese de que a moralidade pode ser analisada de modo reducionista, mostrando que, para qualquer propriedade "natural" P, é uma questão aberta se a posse de P é bom (embora o que Moore quis dizer com "natural" não esteja exatamente claro, ver Dreier, 2006). Com o mesmo espírito, estou usando a Maria Moral para argumentar que é improvável que os conceitos morais possam ser analisados nos termos das construções usadas nas teorias éticas normativas familiares. Apesar dessa semelhança, eu

quero distinguir o meu argumento do de Moore, porque o argumento dele explora um erro. Moore estava defendendo uma conclusão metafísica. Ele considerou o fato de que era uma questão aberta se qualquer propriedade P era boa como fundamento para a conclusão de que a propriedade de ser boa não é constituída por qualquer propriedade. Isto é certamente um erro. Desde o tempo de Moore, os filósofos têm chegado a reconhecer a existência de identidades *a posteriori* – fatos metafísicos que não podem ser descobertos através da análise conceitual. Antes da química moderna, ninguém sabia que o álcool é um composto de hidroxila, com uma estrutura molecular $C_nH_{2n+1}OH$. A alegação de que o álcool é um composto de hidroxila certamente não é uma verdade óbvia, porque se pode possuir o conceito de álcool sem saber química. Mas seria um erro desafiar essa identidade, ao indicar que se possa questionar inteligivelmente saber se isto é verdade. Da mesma forma, foi um erro de Moore inferir que as propriedades morais são não-naturais, e seria um erro inferir, a partir do experimento de pensamento sobre a Maria Moral, que as propriedades morais não se referem aos tipos de propriedades discutidas nas teorias normativas éticas de Kant e de Mill.

Os argumentos da questão em aberto não podem estabelecer conclusões metafísicas, mas eles podem estabelecer conclusões conceituais. Se é uma questão em aberto se alguma propriedade natural N é boa, então o conceito de N não pode ser parte do conceito de bom [*goodness*]. Se houvesse uma ligação conceitual entre as duas, a questão estaria fechada. Por exemplo, pode ser parte do conceito ordinário de álcool (ou seja, o conceito utilizado pela maioria dos adultos ao pensar sobre o álcool na maioria das vezes) que o consumo de álcool pode resultar em intoxicação. Se isso é parte do conceito comum, então ninguém que capta esse conceito poderia inteligivelmente se perguntar se beber pode resultar em intoxicação. Onde há um *link* conceitual, não há questão em aberto. Da mesma forma, no caso da Maria Moral. Se a Maria moral não pode adquirir os conceitos ordinários de bom e de mau ou de certo e de errado ao ler sobre as propriedades descritas em teorias éticas normativas, então não pode ser o caso que essas propriedades sejam constitutivas dos nossos conceitos morais comuns. Os conceitos morais podem se referir, a partir de tudo o que se tem argumentado até

agora, a propriedades kantianas ou millianas, mas não é uma verdade conceitual que o certo é o que maximiza a utilidade ou que é errado fazer o que não pode ser universalizado.

O experimento de pensamento da Maria Moral também pode ser comparado e contrastado com o argumento célebre de Frank Jackson contra o materialismo. Jackson nos faz imaginar uma neurocientista brilhante, chamada Maria, que está presa em um quarto preto e branco. A mente de Maria sabe tudo sobre o que acontece no cérebro quando as pessoas veem vermelho, mas isso não a ajuda a compreender como seria ter uma experiência do vermelho. Como Moore, Jackson tenta tirar uma conclusão metafísica. Ele diz que a experiência de ver o vermelho não pode ser idêntica a um estado cerebral. Esta conclusão é muito controversa. Muitos comentadores pensam que é falacioso inferir uma conclusão metafísica a respeito da base da experiência do vermelho a partir de premissas epistemológicas ou semânticas sobre se nós podemos inferir o que o vermelho é a partir do conhecimento do cérebro. Mas quase todo mundo está disposto a conceder que o argumento de Jackson estabelece uma conclusão epistemológica. Os conceitos envolvidos em saber quais experiências são semelhantes diferem dos conceitos envolvidos em saber sobre os estados cerebrais como tais: não se pode inferir uma classe de conceitos a partir do outro.

Da mesma forma, o argumento da Maria Moral estabelece firmemente uma conclusão epistemológica. Os conceitos invocados em teorias éticas normativas diferem quanto aos conceitos de certo e de errado. Não se pode inferir que uma ação é moralmente certa ou errada a partir de uma descrição kantiana ou milliana dessa ação. Isso é suficiente para estabelecer que Kant, Mill, e outros eticistas normativos não conseguem explicar os conceitos morais. Isso não prova que o emocionismo metafísico está certo. Nós não podemos inferir que Kant e Mill estão errados sobre as propriedades morais a partir do fato de que eles estão errados sobre os conceitos morais. É possível (embora, por razões dadas nos capítulos 3 e 4, implausível) que os conceitos morais designam propriedades kantianas ou millianas. Mas conceber essas propriedades não é suficiente para compreender conceitos morais. Maria entende o que é maximizar a utilidade e o que é

universalizar uma máxima comportamental, mas ela não entende o que as pessoas ordinariamente entendem quando usam as palavras "certo" e "errado".

Ao contrário de Jackson e de Moore, eu não estou tentando estabelecer uma conclusão metafísica; o que deve ser defendido de forma independente. Até agora, eu tenho tentado apenas estabelecer uma tese negativa sobre os conceitos morais ordinários: esses conceitos não estão vinculados às propriedades enfatizadas pelos eticistas normativos. Se os conceitos morais se referem a tais propriedades, isto seria uma descoberta *a posteriori*. Eu, agora, quero estender o experimento de pensamento da Maria Moral para estabelecer uma tese positiva. Moore achava que nenhuma análise dos conceitos morais levaria a uma questão fechada, e concluiu que os conceitos morais são conceitualmente primitivos. Eu acho que Moore também estava enganado aqui. Suponha que Maria desenvolve a capacidade para a empatia humana. Suponha que ela comece a se sentir espontaneamente feliz quando os outros estão felizes, e triste quando os outros estão tristes. Suponha, agora, que ela tem um prazer especial em maximizar o bem, porque ela sabe que ele vai trazer felicidade aos outros. Suponha que ela se sente muito culpada quando ela não maximizar o bem, porque ela sabe que ela não conseguiu dar prazer a alguém que poderia ter sido feito mais feliz. Suponha que ela experimente a raiva em outras pessoas quando não conseguem maximizar a utilidade. Neste ponto, parece perfeitamente razoável dizer que Maria acredita que é errado não maximizar a utilidade, e que é certo maximizar a utilidade. Podemos atribuir a crença moral a ela, uma vez que ela começa a ter uma atitude emocional pela maximização da utilidade. Se, ao invés disso, ela tivesse desenvolvido atitudes emocionais pelas ações que a ética kantiana considera moralmente obrigatórias, nós atribuiríamos atitudes morais por essas ações. Esta é uma importante tese positiva. Nós atribuímos atitudes morais a Maria desde que ela tenha atitudes emocionais. Uma vez que certas ações a levam a se sentir culpada e ultrajada, dizemos que ela tem um ponto de vista moral. As intuições sobre o caso de Maria Moral sugerem que há uma ligação conceitual entre juízos morais e respostas emocionais. Nós atribuímos atitudes morais se e somente se uma pessoa tem certas respostas emocionais.

Esta tese positiva pode ser expressa dizendo que, uma vez que Maria adquire atitudes emocionais, algumas questões se encerram. Mas é preciso ter cuidado em afirmar que a questão está encerrada. Suponha que Maria se sentiria culpada se ela fosse roubar e ficasse indignada se alguém fosse roubar. Não parece haver dúvida sobre se Maria tem uma atitude moral em relação a roubar. Com certeza, ela considera o roubo errado. A questão da atitude de Maria está encerrada. Se estou certo, Moore estava errado quando afirmou que os conceitos morais não admitem análise. Conceitos morais ordinários estão conceitualmente ligados às emoções, e isso nos permite dizer que alguém que nutre culpa e raiva está tomando uma postura moral.

Em resposta, pode-se objetar que a ligação entre os conceitos morais e as emoções nos deixa com questões em aberto. Suponha que Maria sente indignação para com aqueles que roubam. Ela ainda pode se perguntar se roubar é *realmente* errado; ela pode se perguntar se a atitude está justificada. Não é *esta* questão em aberto que mina a minha afirmação de que há uma ligação conceitual entre erro e certas emoções? Absolutamente não. A razão pela qual Maria pode se perguntar se roubar é realmente errado tem a ver com uma lacuna entre sentido e referência. Muitos de nossos conceitos são apreendidos por meio de características que não são nem necessários nem suficientes para pertencer a uma categoria (ver, por exemplo, Prinz, 2002). Nós pensamos as aves como criaturas com asas e penas, mas poderia haver um pássaro sem asas nem penas. Como resultado, pode-se geralmente perguntar se um conceito que nós estamos aplicando *realmente* se aplica em alguma dada ocasião. Maria sabe que ela *considera* o roubo errado, mas ela pode se perguntar se roubar realmente cai sob o seu conceito. Porque sentir e poder se destroem, muitas vezes há questões em aberto sobre a que um conceito dado se refere, mesmo se sabemos que o sentido pelo qual compreendemos o conceito tenha sido satisfeito. A questão em aberto, no presente caso, é uma questão sobre a que conceitos morais se referem, e não sobre o seu sentido. O fato de podermos dizer que Maria está fazendo julgamentos morais simplesmente em virtude de suas atitudes emocionais sugere que há uma ligação conceitual entre os conceitos morais e as emoções.

Nesta etapa, eu não quero descartar a possibilidade de que se pode pensar sobre a moralidade sem possuir emoções. Se os conceitos morais se referem a propriedades kantianas ou a propriedade millianas, então certamente se pode. Mas eu acho que o argumento estabelece que o nosso modo ordinário de pensar sobre a moralidade depende das emoções. Acima, eu introduzi o termo conceito “padrão” para referir a um modo ordinário de conceituar algo. O conceito padrão para uma propriedade é o conceito que é mais comumente usado dentro de uma comunidade; o qual, geralmente, é o conceito que aprendemos em primeiro lugar, e é o conceito que tentaríamos transmitir se estivéssemos ensinando a alguém um termo correspondente. Nós interpretamos normalmente aves com penas, nós interpretamos normalmente álcool como intoxicante, e nós normalmente concebemos as cores imaginando as qualidades fenomenais que elas causam. As intuições por trás do experimento de pensamento aqui sugerem que Maria não tem um padrão de conceitos morais, até que ela desenvolva emoções morais. Sem emoções morais, ela não pode formar julgamentos morais no sentido ordinário. Ela não pode tomar uma atitude moral.

1.2.6 Psicopatas

O experimento de pensamento da Maria Moral sugere que as emoções são necessárias para fazer julgamentos morais de uma forma padrão. Este é um argumento conceitual para o emocionismo epistêmico. Conclusões paralelas podem ser retiradas a partir de pesquisas empíricas. Maria é alguém que não consegue apreender conceitos morais, apesar de seu alto nível de inteligência. Há pessoas que têm esse perfil no mundo real. Há pessoas que podem aprender tudo o que há para saber sobre as consequências de uma ação, sem entender que é imoral. As pessoas que eu tenho em mente são psicopatas.

Os psicopatas apresentam um teste importante para a teoria ética. Os emocionistas epistêmicos estão comprometidos com o internalismo motivacional: a visão de que os julgamentos morais são intrinsecamente motivantes. O internalismo enfrenta um sério desafio. Os críticos acreditam que uma pessoa

poderia formar um julgamento moral sem estar motivado a agir. Eles chamam tais indivíduos hipotéticos de "amorais" (Brink, 1989). Quando este desafio é oferecido, internalistas e externalistas tendem a ficar envolvidos em uma batalha de intuição. Internalistas insistem que uma pessoa que não se motivou com os julgamentos morais não compreendeu realmente a moralidade. Externalistas discordam. Se usarmos critérios emocionais ou motivacionais como teste decisivo para a competência moral, isto dirigiria a questão contra os externalistas. Se rejeitarmos esses critérios, os internalistas dirão que isto é injusto. Eu acho que os psicopatas oferecem uma maneira muito promissora para sair deste impasse.

Os psicopatas parecem estar pouca coisa mais próxima que estamos dos amorais reais. Eles são perfeitamente inteligentes e articulados. Eles parecem compreender os valores morais, mas eles são totalmente indiferentes a eles. Eles se envolvem em comportamentos antissociais crônicos, desde mentir e roubar até torturar e matar, e eles cometem esses crimes sem custo emocional. Quando assassinos psicopatas ouvem as palavras relacionadas à violência, eles não têm uma resposta emocional normal (Gray et al. 2003). Eles também mostram pouca empatia, culpa, vergonha, ou remorso (Hare, 1998). Hare (1993) e Cleckley (1941) classificam muitos exemplos de psicopatas que mostram uma indiferença surpreendente em relação aos danos que eles causam. Por exemplo, um assassino psicopata insistiu que ele era inocente, porque ele fez uma ferida limpa quando ele matou a vítima.

Em face disso, a psicopatia parece demonstrar que o amoralismo não é só possível, mas real. Uma análise mais atenta, porém, revela uma história muito diferente. Os psicopatas *parecem* compreender a moralidade, mas eles, de fato, não compreendem. Eles usam os termos morais de um modo que se desvia surpreendentemente da maneira que os não-psicopatas usam aqueles termos. Estes desvios sugerem que eles não possuem conceitos morais; ou pelo menos que os conceitos morais deles são fundamentalmente diferentes dos nossos. Aqui está um trecho de Cleckley:

O [psicopata] não está familiarizado com os fatos ou dados primários daquilo que poderia ser chamado de valores pessoais. [...] Beleza e feiura, exceto em um sentido muito superficial, bem [*goodness*], mal, amor, horror e humor não têm significado real, não tem poder para movê-lo. [...] É como se ele fosse daltônico, apesar de sua inteligência aguçada, a este aspecto da existência humana. Isto não pode ser explicado a ele porque não há nada em sua órbita da consciência que pudesse fazer a ponte com a comparação. Ele pode repetir as palavras e dizer levianamente que ele entende, mas não há maneira de ele perceber que ele não entende. (Cleckley 1941: 40)

Externalistas podem resistir à avaliação de Cleckley, porque ele parece *pressupor* que os estados emocionais são necessários para a compreensão. Cleckley diz que os valores não têm "nenhum significado real, nenhum poder" para mover o psicopata. Se Cleckley está usando as observações sobre a motivação moral para apoiar a tese de que os psicopatas não podem compreender os conceitos morais, então ele está dirigindo a questão contra os externalistas. Felizmente para o internalista, a evidência do atraso moral em psicopatas não é restrita a provas relativas dos seus estados motivacionais. Fundamentalmente, a incompreensão dos psicopatas de conceitos morais pode ser estabelecida usando critérios que *tanto* externalistas quanto internalistas devem concordar. Deixe-me explicar.

Tanto externalistas quanto internalistas devem concordar que as regras morais diferem de outros tipos de regras. "Moralmente certo" e "moralmente errado" significam algo diferente do "convencionalmente certo" e "convencionalmente errado" (para uma discussão de como isso pode ser conciliado com o relativismo, veja o capítulo 5). Qualquer pessoa que apreende os conceitos morais deve ser capaz de fazer essa distinção. Como nós já vimos, as crianças saudáveis já começam a dominar a distinção moral/convencional no momento em que têm três anos de idade. Se os psicopatas não conseguem compreender a distinção, então não pode ser dito que entendem o que as pessoas saudáveis querem dizer com moralidade. Conceitos são individuados pelo sentido e pela referência. Se os psicopatas colocam os erros morais e os erros convencionais na

mesma categoria, então os seus conceitos de "erros morais" têm referência diferente e, presumivelmente, sentido diferente do conceito que nós normalmente expressamos quando falamos de erros morais.

Blair (1995) demonstra que os psicopatas não conseguem entender a distinção moral/convencional. Usando os métodos de Turiel, Smetana, e Nucci, ele pediu aos criminosos que tinham sido diagnosticados como psicopatas para considerar vários cenários em que as regras haviam sido violadas. Algumas das regras eram morais e algumas eram convencionais, mas os psicopatas não foram alertados sobre este fato. Eles foram simplesmente solicitados a classificar o erro e a gravidade das violações, e a justificar as suas respostas. Eles também foram questionados sobre se o comportamento descrito teria sido errado se uma autoridade tivesse permitido. Os resultados foram surpreendentes. Os psicopatas não tratam os erros morais e convencionais significativamente de forma diferente. Ao contrário de um grupo de controle de criminosos não-psicopatas, eles tendiam a ignorar o bem-estar da vítima quando justificavam as suas respostas sobre os erros morais.

Havia um resultado inesperado no estudo de Blair. Os psicopatas tenderam a tratar ambos os erros morais e os convencionais como se eles fossem independentes de autoridade. É errado falar alto em sala de aula não importando o que o professor diz. Pelo jeito, isto parece sugerir que os psicopatas estão interpretando as convenções como morais, e não o contrário. Mas Blair oferece outra explicação mais plausível. Ele especula que os seus sujeitos estavam tentando sondar se eles eram sensíveis a regras morais. Os presos têm grande motivação para fazer isso porque a sensibilidade moral pode apressar a liberação. Os psicopatas tendem a tratar todas as regras como invioláveis em um esforço para convencer o experimentador que eles eram mentalmente saudáveis. O plano deu errado. Os criminosos não-psicopatas no grupo de controle que foram igualmente motivados a impressionar o pesquisador responderam dentro dos padrões normais. Eles trataram as transgressões convencionais como dependentes de autoridade. Se os psicopatas entenderam a diferença entre erros morais e convencionais, eles teriam que ter tratado os tipos de transgressão de forma

diferente, a fim de transmitir a sua saúde moral. A estratégia compensatória de fingir retidão moral realmente revelou a profundidade do seu problema. Esta interpretação encontra confirmação em outro estudo. Em vez de olhar para adultos presos, que têm forte motivação para mentir, Blair (1997) administrou o teste moral/convencional em crianças com tendências psicopatas, descobrindo que essas crianças tendem a tratar transgressões morais como se fossem convencionais. Blair conclui que os psicopatas não conseguem entender a diferença entre erros morais e convencionais, pois eles tendem a considerar todos os erros como meramente convencionais. Para eles, a moralidade é como uma etiqueta, como convenções sobre que lado da estrada se dirige, ou como o xadrez: um grupo de convenções mais ou menos arbitrárias que fazem exigências sobre nós só porque elas foram adotadas por um grupo social. Os psicopatas podem beirar a moralidade, mas a sua compreensão é superficial na melhor das hipóteses.

Os psicopatas não são amoralistas do mundo real, de modo que eles não podem ser usados como prova contra o internalismo. Na verdade, eles parecem fornecer aos internalistas um elemento comprovativo útil. Na psicopatia, um déficit na motivação moral coocorre com um déficit na competência moral. Isto sugere que as duas estão ligadas. Na verdade, as explicações principais de psicopatia sustentam que o déficit na compreensão moral é um resultado direto do déficit emocional.

Primeiro, considere a explicação dada por Blair (1995). Blair pensa que os psicopatas sofrem de uma anomalia no sistema psicológico designado para inibir a violência. A ideia de um mecanismo de inibição da violência (VIM³) é inspirada pelo trabalho em etologia. Cães e outros animais recuam durante os atos de agressão quando a vítima apresenta uma sugestão de submissão, como no caso da garganta. Esta resposta pode ser mediada por um VIM. Um mecanismo homólogo pode ser responsável pelo fato bem estabelecido de que os seres humanos saudáveis ficam emocionalmente perturbados ao ver outros em perigo. A resposta do sofrimento semelhante está presente mesmo em crianças, mas está seriamente

³ A sigla se refere ao termo original "*violence-inhibition mechanism*". [N.T.]

prejudicada em psicopatas (Casa e Milligan, 1976; Blair et al 1997). Blair toma isso como evidência de uma disfunção do VIM. Ele acha que psicopatas nunca dominam as emoções morais, porque eles não têm o mecanismo que os torna sensível aos perigos dos outros. O comportamento antissocial é uma consequência da inibição debilitada da violência e portanto também um subproduto secundário de uma falta de preocupação emocional mais geral pelos outros oriunda dessa deficiência.

Eu acho que a explicação de Blair está parcialmente certa e parcialmente errada. A psicopatia envolve a inibição do debilitado, mas não se limita à inibição da violência. A história do VIM explica por que em psicopatas podem faltar emoções morais, mas deixa outros sintomas inexplicáveis. Primeiro, o foco na inibição da violência parece um pouco estreito, uma vez que os psicopatas se envolvem em todos os tipos de crime, e não apenas em violência. Isto é confirmado por duas ferramentas de diagnóstico que são mais frequentemente usadas para identificar psicopatas: *Psychopathy Checklist* de Hare (1991) inclui um item rotulado de versatilidade criminal, e os critérios do DSM-IV para o transtorno de personalidade antissocial incluem fatores como falsidade e não cumprimento das obrigações financeiras (Associação Psiquiátrica Americana, 1994). Mais seriamente, a exposição de Blair não explica um padrão geral de anormalidades *cognitivas* gerais encontradas nos psicopatas. Por exemplo, os psicopatas cometem muitos erros quando completam labirintos de dificuldade crescente (Schalling e Rosen, 1968). Eles ficam tentados a ir por caminhos que não terão sucesso. E, finalmente, a história do VIM não explica toda a gama de déficits emocionais em psicopatia. Outro critério de diagnóstico é o afeto restrito (Hare, 1991). Os psicopatas não têm emoções morais, mas eles também são assustadoramente deficientes em emoções não-morais. Psicopatas mostram déficits em medo como prova de uma capacidade diminuída para o condicionamento do choque elétrico e aumentos induzidos pelo estresse em resposta a surpresas (Davies e Maliphant, 1971; Patrick et al, 1993). Eles também mostram um déficit na tristeza. Cleckley (1941) fala sobre um psicopata adolescente que mostrou apenas uma preocupação superficial quando o seu cão de estimação foi atropelado por um carro (p. 71), e ele também descreve um homem que não mostra nenhuma tristeza depois de matar a sua mãe e muitos

outros sem provocação (p. 266). Também foi demonstrado que os psicopatas têm dificuldade em reconhecer expressões faciais e vocais de tristeza (Blair et al., 2001; Stevens et al., 2001).

O padrão de déficits cognitivos e emocionais em psicopatia sugere que a disfunção não pode estar restrita à inibição da violência. Em vez disso, parece haver uma disfunção em um sistema afetivo mais global e rudimentar que subjaz às diversas emoções e processos inibitórios na cognição. Gray (1987) postulou um sistema de inibição comportamental (BIS⁴) que se encaixa perfeitamente nesse perfil. É um componente essencial de emoções negativas tais como medo e tristeza, bem como das emoções morais, mas também desempenha um papel nas tarefas cognitivas que exigem uma mudança ou interrupção [*cessation*] de comportamento. Alguns autores sugerem que as deficiências BIS poderiam explicar os sintomas essenciais da psicopatia (Fowles, 1980).

A principal diferença entre a teoria VIM de Blair e a teoria BIS tem a ver com a extensão da deficiência inicial. Blair atribui à psicopatia um tipo específico de deficiência de inibição, em vez de uma deficiência de inibição geral. Ambas as teorias concordam sobre o ponto de interesse central aqui. Elas sustentam que os atrasos morais resultam do atraso emocional. O desempenho fraco na distinção moral/convencional deriva de um déficit de emoções morais causado por uma deficiência na inibição. O ponto-chave é que ambas as exposições veem a psicopatia como um transtorno emocional. A cegueira moral dos psicopatas ocorre a partir de uma cegueira emocional. Se isso está correto, a psicopatia fornece evidências positivas para o internalismo. O internalismo, como eu mencionei, está implicado pelo emocionismo epistêmico. Se os julgamentos morais são intrinsecamente motivantes, pode ser devido ao fato de que os conceitos morais padrão são essencialmente carregados de emoção. Isso é precisamente o que a pesquisa sobre a psicopatia parece confirmar.

Esta explicação da psicopatia se destina, principalmente, a uma explicação dos conceitos morais, mas ela pode ter um resultado metafísico. Se as

⁴ A sigla se refere ao termo original: "*Behavioral Inhibition System*" [N.T.].

propriedades morais não envolvessem essencialmente a emoção, então deveria haver uma maneira de traçar a distinção moral/convencional sem o apelo às emoções. Os psicopatas devem ser capazes de aprender a diferença. Afinal, os psicopatas geralmente têm quocientes de inteligência dentro do padrão normal. Alguns são extremamente brilhantes. Eles são, muitas vezes, articulados e astuciosos. Eles têm um grande interesse em aprender a distinguir o certo do errado. A maioria dos psicopatas têm uma longa história de má conduta (o que, na verdade, é um critério de diagnóstico do DSM-IV). Estas histórias os colocam em contato com pessoas que têm um cuidado especial para ajudá-los a compreender a moralidade. Eles muitas vezes encontram pais preocupados, professores, advogados e agentes da lei. Dada esta combinação de disposição e motivação para aprender, os psicopatas devem estar mais propensos do que outros a desenvolver maneiras de identificar de forma confiável as propriedades morais. O fato de que eles não conseguem dominar a distinção moral/convencional sugere que não pode haver nenhuma maneira de fazer essa distinção sem se referir ou experienciar respostas emocionais. A distinção pode ser emocional no cerne da questão. Certo e errado podem ser constituídos por reações emocionais em nós. Subtraia essas reações e a distinção se torna tão invisível quanto o espectro de cores é para o cego. Os psicopatas podem acompanhar atentamente os julgamentos morais de indivíduos saudáveis. Eles podem classificar exemplos familiares de boa conduta em um compartimento, e de má conduta em outro, mas eles não têm uma compreensão sobre a essência própria desta divisão. Eles teriam dificuldade no raciocínio sobre a moralidade em contextos modais ("Continuaria sendo errado se ...?"), e eles teriam dificuldade em estender a categoria para casos incomuns. O emocionismo epistêmico prevê que indivíduos emocionalmente deficientes não desenvolverão conceitos morais normais. O emocionismo metafísico prevê que será difícil de compensar esta deficiência pela aquisição de conceitos que têm as mesmas referências, mas sentidos diferentes. Como os conceitos de cor, os conceitos morais pode ser um caso em que sentido e referência estão intimamente ligados.

1.2.7. A Desunião da Moralidade

Eu vou concluir com mais um argumento a favor do emocionismo metafísico, já que a maioria dos meus argumentos pertence à tese epistêmica. O emocionismo epistêmico não implica o emocionismo metafísico. Os julgamentos morais poderiam estar essencialmente ligados às respostas emocionais, mesmo que as propriedades morais não estivessem. O bem poderia ter uma essência que é independente de nossas emoções. Comparemos com o papel de tornassol. Ele testa ácidos mudando de cor. O papel de tornassol detecta ácido pelo efeito que o ácido tem sobre ele. Mas o ácido não é definido pelo seu efeito sobre o papel de tornassol. O ácido tem uma essência que pode ser caracterizada de forma independente. Por paridade de raciocínio, nós podemos detectar o bem e o mal medindo as nossas respostas emocionais, mesmo se o bem e o mal tem alguma essência independente.

O emocionismo epistêmico corroboraria o emocionismo metafísico se não houvesse, de nenhum modo, a independência do mental para caracterizar a essência do bem e do mal. Suponha que as coisas boas não têm nada em comum com as outras além do fato de que elas causam emoções positivas em nós. Suponha que coisas ruins não compartilhem nada além da sua capacidade de nos estimular emoções negativas. Se é assim que as coisas são, então o emocionismo metafísico é uma tese plausível. O bem e o mal são essencialmente dependentes de resposta, pois eles são definidos por seus efeitos sobre nós. Comparemos com os venenos. Não há uma unidade intrínseca na classe de coisas que são venenosas. Eles têm constituições químicas extremamente diversas, e os seus reais efeitos sobre o corpo são variados. Os venenos são definidos pelo fato de que eles nos causam doenças e a morte. Propriedades morais poderiam ser assim.

Como decidir entre estas duas opções? A evidência da psicopatia é sugestiva, porque os psicopatas não conseguem encontrar uma maneira de identificar regras morais independente da emoção, mesmo que seja do interesse deles fazer isso. Mas o caso do emocionismo metafísico não deve se apoiar nos psicopatas. O seu fracasso em encontrar uma essência do bem e do mal

independente da mente pode derivar do fato de que eles não se importam o suficiente com os valores, ou que eles não têm formação filosófica suficiente. Para provar que a moral não tem uma essência independente da mente, seria preciso realmente examinar as teorias filosóficas. Seria necessário mostrar que os filósofos têm sido incapazes de identificar uma essência independente da mente para a moralidade. Eu examino algumas dessas tentativas nos capítulos 3 e 4. Aqui eu quero avançar no caso do emocionismo metafísico fazendo um deslocamento do ônus [da prova]. Eu quero argumentar que, *em face disso*, a moralidade é mais como o caso de veneno do que o caso do ácido. A classe de comportamentos morais parece estar dissociada. Se boas e más ações parecem dissociadas, então os oponentes do emocionismo metafísico têm o ônus da prova, pois eles devem mostrar que as aparências enganam. Eles devem estabelecer que o bem e o mal têm essências unificadoras que são independentes das reações emocionais humanas.

Vou me concentrar na classe dos males. Atos imorais têm algo em comum? Aparentemente, a resposta é não. Atos imorais compreendem uma mistura: mentir, roubar, acumular [*hoarding*], ferir, matar, negligenciar, assediar, poluir, insultar, molestar, vandalizar, desrespeitar, e assim por diante. O que essas coisas têm em comum *além do fato de que desaprovamos todas elas*? É tentador responder que todas elas são prejudiciais. Talvez o mal seja a essência da iniquidade. Há três objeções fatais a esta proposta. Em primeiro lugar, a noção de mal é, ela própria, dependente da resposta. Há muitas coisas que são difíceis ou dolorosas, mas não consideradas como más: escrever artigos, exercitar-se, passar oito horas no trabalho, aventurar-se no frio, dizer "adeus" a um ente querido antes de uma viagem de negócios, comer comida picante, ir ao dentista, e assim por diante. Danos parecem ser aquelas provações físicas e psicológicas que consideramos como candidatos para o interesse moral. Em segundo lugar, alguns danos não são considerados moralmente maus. Muitas culturas têm ritos de iniciação danosos, e muitos consideram moralmente aceitável causar danos às pessoas em autodefesa ou retaliação. Em terceiro lugar, e mais importante ainda, algumas das coisas que condenamos moralmente não causam danos; alguns erros não têm vítimas. O

incesto consensual parece, a muitas pessoas, como imoral, embora ninguém pareça ser prejudicado de qualquer maneira óbvia. Considere o cenário de Murphy et al. (2000) em que um irmão e uma irmã fazem sexo consensual usando contraceptivo. Eles fazem isso apenas uma vez, eles gostam disso, e mantêm em segredo. As pessoas tendem a considerar isto como moralmente errado, embora ninguém seja prejudicado.

Males sem vítimas apresentam um sério problema para as outras tentativas de encontrar a unidade na classe das iniquidades. Quando não há vítimas, não há violação de direitos, e não houve redução na felicidade pura. Crimes sem vítimas também são um problema para Kant, o qual tentou mostrar que todos os males são unificados por um determinado tipo de irracionalidade. Em particular, Kant sugeriu que os males são aquelas ações que não podem ser coerentemente universalizáveis. A dificuldade é que não há contradição em querer que o incesto consensual seja uma lei universal. Nenhuma contradição surge em um mundo onde todos os irmãos fazem sexo um com o outro com proteção quando eles querem. Em um mundo assim, o sexo entre irmãos é como jogo de luta entre irmãos. O jogo de luta é moralmente aceitável porque é uma boa prática e ninguém se fere gravemente. O sexo é uma atividade humana central e um pouco de prática faz algum bem. Ainda assim, o incesto consensual nos parece muito errado. Esta não é uma refutação de Kant. Kant poderia argumentar que há uma irracionalidade escondida no incesto consensual, ou ele poderia dizer que ele não é realmente errado. Mas, quanto a isso, a proposta de Kant e de outras teorias éticas normativas parecem não obter o direito de taxonomia pré-teórica. Há coisas que as pessoas comuns consideram errado, mas não são erradas nas principais explicações filosóficas. Essas explicações tendem a se concentrar em categorias que são mais estreitas do que a classe de coisas que são consideradas erradas pré-teoricamente. Quando se considera a classe total de coisas que nós normalmente consideramos erradas, as perspectivas de encontrar uma teoria unificada começam a parecer sombrias.

Vou discutir Kant de modo breve novamente no capítulo 3, e nos capítulos 4 e 5 eu vou me deter mais sobre o tema da dissociação. O meu presente objetivo é

mostrar que não há nenhuma propriedade óbvia intrínseca que unifica a classe de males e que sugere que a incorreção de um ato *pode* ter mais a ver com a nossa reação do que com o ato em si. O incesto consensual e o roubo *pode* pertencer à mesma categoria em virtude de ofender a nossa sensibilidade. O sentimento *pode* ser o fio condutor que mantém juntos todos os vícios e virtudes. Aqueles que acreditam na unidade intrínseca dos erros têm o ônus da prova. Pelo jeito, nós julgamos moralmente uma classe heterogênea de coisas. A única coisa que os itens dessa classe parecem ter em comum são as nossas atitudes morais para com eles.

1.2.8 O resultado emocionista

Eu ofereci uma bateria de argumentos em favor do emocionismo. Eu misturei argumentos filosóficos com evidências da psicologia experimental, da neurociência, do desenvolvimento infantil, e da psicopatologia. O emocionismo conta com o apoio convergente das intuições remotas e do trabalho empírico. Os conceitos morais parecem estar ligados com as respostas emocionais. Aqueles que não têm emoções morais também não têm conceitos morais. Isso apoia o emocionismo epistêmico. Além disso, há pouca evidência de que qualquer um *pode* conseguir distinguir propriedades morais sem a ajuda de emoções, e às propriedades morais parecem faltar quaisquer características unificadoras independentes das nossas reações a elas. Isso dá suporte ao emocionismo metafísico. Pode haver outras maneiras de dar sentido a todas as evidências que eu tenho discutido, mas o emocionismo surge como a explicação mais óbvia.

O caso que eu tenho apresentado não pretende ser uma prova demonstrativa. Pode ser demais esperar por um argumento arrasador para o emocionismo. Por enquanto, vou aceitar parcialmente um apoio *prima facie*. Se o emocionismo tem fundamento *prima facie*, então vale a pena explorar. Nos capítulos seguintes, vou apresentar uma teoria emocionista com mais detalhes, e argumentar que ela *pode* suportar objeções. Eu vou também apresentar objeções às teorias éticas não-emocionistas, tais como a de Kant, e a teorias emocionistas

fracas, como o utilitarismo e o expressivismo. Eu vou concluir que o caso do emocionismo forte é atraente. As emoções estão implicadas, essencialmente, tanto na ontologia como na epistemologia da moral.

2. EMOÇÕES NÃO-MORAIS E MORAIS

No capítulo 1, apresentei o suporte preliminar para um modelo forte de emocionismo. Existem razões para pensar que tanto a ontologia moral quanto os juízos morais envolvem essencialmente emoções. A natureza desse envolvimento continua a ser elaborada, mas existe um obstáculo importante a ser superado antes que mesmo que o projeto possa começar. Historicamente, os filósofos que invocavam as emoções em suas teorias morais também forneceram considerações detalhadas e independentemente motivadas acerca das emoções. Spinoza desenvolveu uma influente teoria das emoções em seu livro *Ética*, e Hume dedicou o Livro Dois de seu *Tratado* para as paixões antes mesmo de propor sua teoria emocionista da moral no Livro Três. Os emocionistas do século XX, contudo, foram menos profundos. Muitas vezes, dizem-nos que as emoções são fundamentais para a moralidade, mesmo sem sermos informados a respeito do que as emoções realmente são. Os emocionistas não podem se dar ao luxo de ficar em silêncio sobre essa questão, pois algumas teorias da emoção são realmente difíceis de conciliar com suas abordagens da moralidade. Neste capítulo, eu defendo uma teoria não-cognitiva das emoções que pode ser conciliada com o emocionismo. Eu apresentarei, portanto, uma visão geral das emoções morais e argumentarei que emoções morais diferentes desempenham papéis funcionais diferentes. Argumentarei também que as emoções morais não são básicas, mas que, ao invés disso, elas são derivadas de emoções não-morais.

2.1 UMA TEORIA DA EMOÇÃO

2.1.1 Dois pontos de vista

A história da pesquisa sobre a emoção pode ser considerada como uma batalha entre dois lados opostos. Alguns autores argumentam que as emoções são cognitivas, pois elas envolvem essencialmente juízos ou pensamentos. Outros autores negam isso, argumentando que as emoções são fundamentalmente não-

cognitivas. A divisão entre teorias cognitivas e não-cognitivas é a fratura mais evidente e volátil dentro da pesquisa da emoção, e a maior parte das teorias alinham-se cuidadosamente de um lado ou de outro. O problema é que nenhuma abordagem se encaixa perfeitamente bem com o emocionismo. Ambas as visões levantam um problema para a tese epistêmica de que os juízos morais envolvem necessariamente emoções. Para tornar os problemas explícitos, será útil descrever as teorias cognitivas e as não-cognitivas em mais detalhes.

Considere, primeiramente, as teorias cognitivas. Spinoza (1677) diz que as emoções são juízos acompanhados por prazer ou por dor. Dor e prazer são sentimentos que, em si mesmos, são desprovidos de significado. Assim, a teoria de Spinoza é realmente parte cognitivista e parte não-cognitivista. A maior parte das teorias cognitivistas são assim, pois, sem um elemento não-cognitivo, é difícil explicar o calor das paixões. Ainda que alguns teóricos cognitivistas pensem que as emoções não são mais do que simples juízos (Solomon, 1976; Nussbaum, 2001), a maioria contudo não é tão ousada. O que define os diferentes teóricos cognitivistas é a alegação de que as emoções contêm elementos essencialmente cognitivos, caso em que um mero sentimento, por exemplo, não seria suficiente.

Na psicologia contemporânea, as teorias cognitivistas dominantes enfatizam uma classe particular de estados cognitivos chamados de avaliações (por exemplo, Arnold, 1960; Lazarus, 1991; Scherer, 1993). Uma avaliação é uma representação de uma relação organismo/ambiente que leva ao bem-estar. Elas chamam essa relação de "interesse". A raiva, por exemplo, envolve uma avaliação de ameaça ou ofensa. O medo envolve uma avaliação do perigo. A tristeza envolve uma avaliação da perda. Ofensas, perigos e perdas são matérias de interesse. Nas teorias cognitivas usuais, as emoções são respostas sentidas com relação aos juízos de avaliação, e os juízos de avaliação, ainda que sejam inconscientes, são avaliações explícitas da relação de alguém com o mundo. Tais juízos são pensados geralmente como altamente estruturados e complexos, e caem ao longo de várias dimensões fixadas. Estamos constantemente nos perguntando sobre como as coisas estão se saindo; meus objetivos estão sendo atingidos? Quem, se alguém o for, é responsável por minha situação atual? Quais opções estão disponíveis para serem

enfrentadas? Scherer (1993) chama esses juízos de verificação de estímulos de avaliação. Cada emoção corresponde a um diferente conjunto de respostas. A raiva não envolve apenas um juízo de que fui ofendido, mas também um juízo de que meus objetivos foram ameaçados, que a fonte da ameaça é outra pessoa, e que a agressão é uma opção disponível para enfrentar a situação. Se a verificação de meus estímulos de avaliação leva a esse padrão de respostas, então o estado resultante é a raiva. Esse estado pode ser um sentimento; pode ser uma dor, ou um prazer, ou uma sensação de excitação. Tivesse o mesmo sentimento surgido de alguma outra forma, então ele não teria sido raiva. Avaliações são essenciais para a identidade da emoção.

Os detratores consideram essa abordagem, e também os precursores filosóficos delas, como excessivamente intelectuais. Eles insistem que emoções podem ocorrer e manter sua identidade sem quaisquer juízos ou pensamentos. Consideram implausível a demanda de que as emoções exigem o domínio de conceitos ou atitudes proposicionais. Ao invés disso, eles preferem teorias não-cognitivistas. As teorias não-cognitivistas mais simples identificam as emoções com sentimentos puros. Um sentimento puro seria um sentimento cuja identidade é exaurida por seu aspecto sentido ou por seu caráter fenomenal. A psicologia popular tem alguma simpatia por essa sugestão, pois parece intuitivamente plausível que uma pessoa diga "eu sei que eu estou com raiva, porque eu sinto raiva – eu reconheço esse sentimento". Mas é preciso cautela para inferir que as emoções são sentimentos puros, principalmente a partir do fato de que elas são conhecidas por seus sentimentos. Por analogia, considere o caso da dor. Eu sei que estou com dor por sentir minha dor, mas a dor é mais do que um sentimento. A dor tem uma função particular, pois ocorre, ordinariamente, quando o corpo é acometido por algum acidente ou doença. É coerente dizer que dores são sentimentos e que elas são reconhecidas pelo modo que são sentidas enquanto insistem que existe algo mais para sua identidade. Suponha que encontrássemos uma criatura que experimentasse estados que fossem qualitativamente iguais às nossas dores, mas que não tivesse nenhuma correlação com uma lesão. Poderíamos chamá-las de dores? Ou, para utilizar outra analogia, suponha que

encontrássemos uma criatura que tivesse experiências qualitativamente iguais às nossas sensações de cores, mas que não tivesse nenhuma conexão com a luz. Poderíamos chamá-las de sensações de cores? Tais casos são extremamente fantasiosos e não pretendo oferecer intuições decisivas sobre eles. Mas eles destacam uma diferença importante entre a identificação de procedimentos e as condições de identidade: estados que podem ser identificados por seu caráter fenomenal podem ter atributos adicionais ainda não sentidos que são necessários para sua identidade.

Uma razão para duvidar que as emoções sejam sentimentos puros é que não existem sentimentos suficientes para contorná-las. A raiva é sentida de uma certa maneira, e esse sentimento é diferente da tristeza ou medo, mas algumas emoções são sentidas de uma mesma forma. Posso falar de raiva a partir da indignação unicamente pelo sentimento? Posso distinguir a alegria do orgulho? Ou o luto do remorso? É provável que as emoções obtenham a sua identidade em parte pelo contexto a partir do qual elas surgem.

Esse ponto já estava implícito na apresentação das teorias da avaliação. Sob tal consideração, as emoções são diferenciadas por suas causas cognitivas, e são, portanto, inerentemente ligadas à cognição. Um teórico não-cognitivista poderia admitir que as emoções obtêm sua identidade, em parte, a partir das condições em que surgem e, ao mesmo tempo, negar que essas cognições sejam necessariamente cognitivas.

Hume pode ser interpretado como o defensor de uma teoria semelhante. Ele diz que as emoções são impressões e que elas surgem de duas formas. Existem impressões de sensação e impressões de reflexão. Sensações são causadas quando os sentidos são estimulados, e reflexões são causadas por sensações ou outros estados mentais. Reflexões são como que sensações internamente dirigidas: são respostas imediatas aos eventos no mundo mental. Cada impressão de reflexão tem uma causa mental, e muitas têm efeitos mentais. Emoções obtêm sua identidade a partir dessas causas e efeitos. Elas são impressões causadas por outras impressões ou por ideias. O medo é uma impressão que surge quando temos uma impressão de dor que se conectou a uma impressão de incerteza. A alegria

pode surgir, digamos, ao reconhecer um amigo. Algumas emoções obtêm a sua identidade a partir de impressões ou de ideias que elas causam, ao invés de obtê-las através das impressões ou ideias que as causam. Orgulho, por exemplo, faz com que alguém pense sobre si mesmo. Em alguns casos, a posição de Hume é indistinguível de uma consideração cognitivista moderna, pois as emoções são individualizadas pelos estados mentais com os quais elas interagem, e tais estados incluem as ideias pertinentes ao modo como estamos nos saindo na vida. Mas a consideração de Hume é, contudo, não-cognitivista, pois as emoções podem ser causadas diretamente pelas experiências sensoriais sem qualquer juízo avaliativo explícito. Quando a alegria é causada pelo prazer sensorial ou por vislumbrar o semblante de um amigo querido, não existe mediação cognitiva, pois é a percepção que desencadeia o afeto. Hume diz também que as emoções são construídas por sentimentos, independentemente de suas causas e efeitos. As emoções não são nem construídas por juízos, nem necessariamente causadas por eles. A história é, contudo, um pouco mais complicada pelo fato de que Hume identifica uma classe de emoções, chamada de paixões indiretas, que essencialmente envolvem ideias de um objeto: orgulho e humildade são direcionadas a si mesmo, enquanto o amor e o ódio são direcionados ao outro. Hume afirma que essas emoções trazem à mente a ideia do eu ou de outra pessoa, e isso implica que elas possuem um componente conceitual, ao contrário de outras emoções, as quais ele chama de paixões diretas. Pode ser que Hume tenha uma teoria cognitiva das paixões indiretas e uma teoria não-cognitiva das paixões diretas. Alternativamente, Hume pode ser interpretado como afirmando que as paixões indiretas são sentimentos de estados não-cognitivos que, como uma questão de fato contingente, apenas acontecem para nos levar a pensar de forma segura no eu ou no outro. De qualquer modo, Hume está comprometido com a visão de que muitas emoções (as paixões diretas) podem surgir sem quaisquer pensamentos ou conceitos.

A abordagem de Hume é também não-cognitiva em outro sentido. Ele defende que as emoções carecem de significado ou de intencionalidade. Elas não são representações. Impressões sensoriais representam características do mundo, ao passo que as emoções não representam nada. Elas são apenas sensações de

um tipo particular. Pode-se ter uma ideia de por que Hume faz essa afirmação se se considerar o fato de que a representação no século XVIII estava intimamente relacionada à ideia de semelhança. A sensação causada por ver o rosto de um amigo se assemelha com esse rosto, podendo-se argumentar que: existe um isomorfismo entre a imagem interna e a causa externa. A alegria desencadeada por essa sensação, contudo, não se assemelha à sensação de modo algum.

A afirmação de que as emoções carecem de intencionalidade está intimamente associada com as teorias não-cognitivas, mas isso não é essencial para elas. Considere a teoria desenvolvida por William James (1884) e Carl Lange (1885). Eles argumentam que as emoções são sensações de mudanças estampadas no corpo. Ainda que Lange enfatize a mudança na pressão sanguínea, a posição de James, contudo, é muito mais inclusiva. Ele menciona as expressões faciais, as mudanças musculoesqueléticas, a atividade nos órgãos viscerais, e mesmo os comportamentos estereotipados. Em *Os Princípios da Psicologia*, James cita – a partir de observações detalhadas de Darwin e de outros que notaram a ligação íntima e a correlação constante entre afeto e resposta corporal –, a palidez do medo, os punhos cerrados de raiva, o nóculo na garganta de tristeza e o largo sorriso de alegria. Essas respostas, assim como a liberação de muco e fibras, são geralmente pensadas como surgindo depois que uma emoção é desencadeada, como quando dizemos que é a tristeza nos faz chorar, e não o contrário. De acordo com James e Lange, a psicologia popular define as coisas em um sentido inverso. As emoções são estados mentais que registram as mudanças corporais. O argumento central para ambos os autores envolve um exercício de subtração mental. Imagine uma emoção intensa e, depois, elimine sistematicamente cada sensação corporal associada a esse estado. Quando a última sensação corporal for removida, não haverá nada para o estado que se poderia chamar de uma emoção. Emoções são percepções sentidas de mudanças corporais.

A teoria James-Lange compartilha algo importante com a teoria de Hume. De acordo com ambos, as emoções podem ocorrer sem juízos ou outras cognições. As emoções não têm uma ligação essencial com a cognição. Mas a teoria James-Lange também se diferencia em vários aspectos da posição de Hume. Mais

obviamente, a teoria James-Lange enfatiza os concomitantes corporais das emoções. Em segundo lugar, eles não dizem que as emoções são individualizadas por suas relações causais com outros estados mentais. Em terceiro lugar, eles implicitamente rejeitam a distinção sensação/reflexão, tratando as emoções como uma classe de sensações. E, finalmente, eles assumem que emoções são representações; elas representam estados do corpo.

As ligações entre as emoções e as mudanças corporais não são, contudo, arbitrárias. As emoções correlacionam-se com as mudanças corporais que servem para fins úteis. As mudanças somáticas associadas ao medo podem ser entendidas como preparações para respostas comportamentais ao perigo; fugimos, congelamos, e lutamos quando sentimos medo. A ereção dos folículos capilares corporais é um vestígio dos longínquos ancestrais dos mamíferos; quando os pelos ficam eriçados em criaturas peludas, o tamanho aparente do corpo é aumentado, e os predadores são desencorajados. A aceleração do coração envia sangue às extremidades do corpo para facilitar a fuga. A respiração tensa nos torna mais silenciosos e difíceis de detectar. Olhos e bocas abertos coletam informações visuais e olfativas. Para James e Lange, portanto, o medo é a percepção dessas mudanças úteis de um ponto de vista comportamental.

Outra possibilidade é que o medo seja um conjunto de comandos corporais. Alguns psicólogos têm sugerido que as emoções são, na verdade, tendências de ação. Elas orquestram as mudanças corporais que facilitam a ação, e elas nos dispõem a agir. A teoria da tendência de ação preserva a suposição da psicologia popular de que as emoções causam respostas corporais, ao invés do contrário. Essa teoria pode realmente ser conciliada com o ponto de vista de James e Lange se afrouxarmos a distinção entre *inputs* e *outputs*. Essa distinção, às vezes, é turva no sistema nervoso. Considere a recente enxurrada de pesquisas sobre "neurônios espelho" (Rizzolatti e Craighero, 2004). Essas são células em áreas pré-motoras do córtex frontal que respondem quando uma criatura (macaco ou homem) executa uma ação manual ou quando essa mesma ação é visualmente percebida. Essa parece ser uma estratégia relativamente comum no cérebro. Células que causam ações também detectam, algumas vezes, ações. Além disso, pode haver redes

neurais com as quais alguns neurônios geram *outputs* e outros registram *inputs*, mas todas estão tão intimamente ligadas que não faz muito sentido descrever qualquer uma delas como parte de um sistema de *input* ou de um sistema de *output*. O máximo que poderíamos dizer dessa rede é que ela é um sistema de *input-output* – um consórcio unificado de células contribuindo para a percepção e para a ação. Se as redes neurais que fundamentam as emoções têm esse caráter, então as emoções são tanto percepções das mudanças corporais, quanto causas das mudanças corporais. Isso colapsaria a distinção entre sensações somáticas e tendências de ações. Nessa abordagem, poderíamos dizer que as emoções são "sinais somáticos", pois o termo "sinal" é neutro entre um efeito e uma causa.

2.1.2 O que está em jogo no Emocionismo?

Vimos que existem duas abordagens principais para explicar as emoções. Os defensores das teorias cognitivas ou bem buscam equiparar emoções com estados cognitivos, tais como juízos, ou bem afirmam que as emoções incluem necessariamente estados cognitivos. Para Spinoza, por exemplo, as emoções são juízos, assim como dor ou prazer, e para muitos psicólogos contemporâneos, as emoções são estados de excitação ou sentimentos causados por juízos de avaliação. Os teóricos não-cognitivistas negam que as emoções requerem cognição. Para Hume, as emoções são uma classe de sensações *sui generis*, individualizadas pelas impressões e ideias com as quais elas interagem causalmente. Para James e Lange, elas são sensações de mudanças corporais. Para outros, elas são tendências de ações, ou comandos corporais. Antes de decidir esse debate, eu quero indicar o que está em jogo nas teorias emocionistas da moralidade e, também, argumentar que os emocionistas deveriam favorecer o não-cognitivismo.

As teorias cognitivistas da emoção podem enfrentar dificuldades para harmonizarem-se com o emocionismo. Se tais teorias estão corretas, as emoções contêm pensamentos ou juízos. Se o emocionismo está correto, então juízos morais contêm emoções morais. Essas duas hipóteses são difíceis de harmonizar.

Se as teorias cognitivas da emoção estão corretas, então emoções morais contêm juízos. Mas que tipo de juízos elas poderiam conter? A resposta natural é que elas contêm juízos morais. Por exemplo, a culpa pode ser definida como contendo o juízo de que eu fiz algo moralmente errado. Se esta proposta estivesse correta, os emocionistas estariam enredados em um argumento circular. Eles teriam o ponto de vista infeliz de que juízos morais contêm emoções morais, e que emoções morais contêm juízos morais.

Para evitar esse círculo, os emocionistas que são atraídos para teorias cognitivas da emoção poderiam tentar argumentar que as emoções morais contêm *juízos não-morais*. Por exemplo, a culpa poderia conter o juízo de que eu prejudiquei alguém, ao invés de conter o juízo de que eu fiz algo moralmente errado. Não existe nada circular na suposição de que juízos morais contenham o julgamento de que eu prejudiquei alguém. Mas essa proposta enfrenta outra dificuldade. Ela é vítima de uma versão do argumento da questão em aberto de Moore. Alguém pode coerentemente perguntar se é errado causar danos. Se o conceito errado contém culpa, e culpa contém o juízo de que eu prejudiquei alguém, então "prejudicar é errado" equivale à tautologia de que prejudicar é prejudicar. Portanto, nessa abordagem, ninguém poderia questionar se prejudicar é errado. O argumento funciona para *qualquer* análise de juízos subjacentes às emoções morais. Cada análise cognitiva das emoções morais, quando combinada com o emocionismo, implica que o conceito ERRADO possa ser decomposto em um juízo de algum tipo. Para qualquer juízo, deveria ser impossível crer consistentemente que algo é errado enquanto se pergunta se o juízo também se aplica. Mas isso é difícil de harmonizar com o fato de que as análises de ERRADO em termos não-morais sempre pareçam soar surpreendentes ou informativas. Se houvesse um juízo não-moral que constituísse o juízo de que algo é errado, então não deveria parecer uma questão em aberto se algo está errado quando esse juízo se aplica. Mas parece que, para qualquer juízo não-moral, poderia existir uma questão em aberto.

Em suma, teorias cognitivas da emoção apresentam um dilema para o emocionista sobre a moralidade. Se as teorias cognitivas são verdadeiras, então

emoções morais ou contêm juízos morais, ou contêm juízos não-morais. Se elas contêm juízos morais, então a consideração emocionista dos conceitos morais é circular. Se elas contêm juízos não-morais, então os emocionistas teriam de enfrentar uma versão do argumento da questão em aberto. Há, sem dúvidas, modos pelos quais um cognitivista a respeito das emoções poderia tentar escapar do dilema. Mas parece que existe uma tensão entre emocionismo e teorias cognitivas das emoções. Ao invés de escapar, o emocionista poderia estar melhor se procurasse por uma teoria não-cognitivista das emoções. Felizmente, penso que as teorias não-cognitivas são independentemente mais plausíveis.

2.1.3 O Processo de Materialização

Vou apresentar agora um caso contra o cognitivismo e a favor da teoria não-cognitivista. Para começar, vamos perguntar se os estados cognitivos, tais como juízos, são necessários ou suficientes para constituir um estado emocional. Os teóricos puramente cognitivistas estão sozinhos ao defender que as cognições são suficientes para a emoção, e creio que suas considerações são singularmente implausíveis. Tome qualquer pensamento que quiser e pergunte “eu poderia ter este pensamento sem ter uma emoção?”. Considere uma pessoa que pensa ter sido insultada. Ela poderia ter esse pensamento sem estar com raiva? Eu diria que a resposta é um óbvio sim. Pode acontecer que a maioria de nós fique irritado quando consideramos esse pensamento, mas não precisamos. Os seguintes indivíduos são suscetíveis de desviar de insultos sem qualquer agitação: uma pessoa com um ego especialmente forte, alguém que recebe insultos frequentemente, uma pessoa com uma disposição estoica, uma pessoa prestes a cair no sono e incapaz de reunir a agitação necessária, alguém ávido por punição, e uma pessoa envolvida em um jogo de trocar desconsiderações verbais inteligentes. Talvez fosse necessário apelar para um juízo diferente se quisermos tornar plausível a teoria puramente cognitiva. Talvez nenhum juízo ordinário realize essa proeza; é preciso ter, ao invés disso, um juízo *avaliativo*. Um teórico puramente cognitivista poderia defender que a raiva envolve o juízo de que eu fui insultado e

que ser insultado é uma coisa ruim, pois representa uma ameaça ao meu bem-estar. Creio que essa proposta fornece pouco progresso, pois um estoico bem treinado poderia fazer esse juízo de forma desapaixonada. É claro, uma pessoa que realmente levou a sério esse juízo, que realmente se sentiu ameaçada por um insulto, sentiria raiva. O ponto é que a raiva não pode ser totalmente constituída pelo juízo. Uma máquina poderia detectar insultos ameaçadores sem se sentir ofendida. Quando pensamos sobre as ameaças ao nosso bem-estar, experimentamos emoções, mas isso é porque reagimos a esses pensamentos de um modo particular.

Se não são suficientes, poderiam os pensamentos ser ao menos necessários para as emoções? A maioria dos filósofos responderia isso de modo afirmativo. A resposta depende do que se concebe por um pensamento. No mínimo, eu concebo os pensamentos como representações mentais que contêm conceitos. O pensamento de que fui insultado contém o conceito INSULTO. Um conceito é uma representação mental que pode ser combinada com outras representações mentais pela pessoa (ou criatura) que o possui. Dizer que as emoções são necessariamente cognitivas é dizer que alguém não pode ter uma emoção sem possuir e simbolizar certos conceitos. O quão plausível é isso? O quão plausível é dizer que não se pode estar com raiva sem ser capaz de pensar sobre insultos (ou ameaças ou alguma outra forma de provocação)?

Eu penso que os requisitos cognitivos são demasiadamente exigentes. Eles colocam as emoções muito acima das escalas filogenéticas e ontogenéticas. Requisitos cognitivos também parecem pouco eficazes diante do imediatismo emocional. Nós frequentemente nos emocionamos sob pressão com pouco tempo para refletir. A raiva pode ser provocada por um olhar, alguém pode pressionar você para fora da estrada quando está dirigindo, ou alguém pode esbarrar em você no corredor do supermercado. A emoção parece ocorrer antes que se tenha tempo para avaliar, ou mesmo classificar, a situação.

Existe um longo debate empírico sobre a existência ou não de pré-requisitos cognitivos nas emoções. Em uma discussão clássica, Zajonc (1984) cita uma série de evidências contra as teorias cognitivistas. Seus exemplos mais impressionantes

envolvem casos nos quais as emoções são induzidas através de processos físicos diretos. Estudos sobre o *feedback* facial mostram que mudar a expressão facial de uma pessoa pode desencadear inadvertidamente o correspondente estado afetivo (Strack et al, 1988). Estudos de processos opostos sugerem que intensas emoções positivas causam um estado emocional negativo após o efeito, como na suspensão de um medicamento, e o inverso (Solomon, 1980). E, é claro, as emoções podem ser alteradas por drogas, sexo e *rock and roll*.

Zajonc também cita evidências anatômicas. Suas citações estão um pouco desatualizadas, mas o trabalho recente confirma sua suspeita. LeDoux (1996) mostrou que um estímulo visual pode desencadear uma resposta emocional antes do envolvimento do neocórtex. Existe uma via subcortical dos olhos até a amígdala, mediada pelo tálamo. A amígdala é uma estrutura localizada no fundo do lobo temporal que associa sinais perceptivos com outras estruturas subcorticais que regulam as respostas corporais. Quando a amígdala desencadeia uma mudança corporal, essa mudança é experimentada como uma emoção. O tálamo é um centro sensorial onde os órgãos sensoriais enviam sinais a caminho do neocórtex, onde o reconhecimento de objetos é obtido. Representações do tálamo de estímulos visuais são relativamente não-processadas, mas elas são suficientes para a identificação de certos objetos emocionalmente salientes. LeDoux dá o exemplo de uma cobra enrolada. Trata-se de um sinal visual distintivo que pode desencadear uma reação de medo antes que o córtex tenha tempo para permitir um juízo perceptual. Morris et al. (1999) apresentaram evidência de que a via amígdala-tálamo também subscreve respostas emocionais para expressões faciais rapidamente apresentadas. Esta via permite até mesmo respostas emocionais nos rostos de pessoas cegas. Indivíduos com cegueira possuem lesões no córtex visual. Eles reagem emocionalmente a rostos apresentados em seu campo cego mesmo que eles não tenham nenhuma experiência visual desses rostos (de Gelder et al., 1999). Seria totalmente insustentável afirmar que o tálamo ou a amígdala abrigam conceitos. Eles não parecem conter qualquer coisa como representações livremente recombinaíveis de insulto, ou de perigo, ou de perda, ou de ego, como requereria uma teoria cognitivista. Esses resultados sugerem fortemente que as

emoções podem surgir sem juízos, pensamentos, ou outros mediadores cognitivos.

O argumento que acabei de apresentar é mais ou menos este: as evidências da neurociência mostram que não precisamos do neocórtex para desencadear uma resposta corporal que, quando experimentada, seria identificada como uma emoção. Os defensores das teorias cognitivistas podem tentar objetar isso do seguinte modo. O trabalho de LeDoux mostra que os juízos não são necessários para *iniciar* um estado que experimentamos como uma emoção, mas talvez os juízos estejam envolvidos em uma fase posterior de processamento. Talvez um estímulo perceptual que é processado de modo subcortical desencadeie uma resposta corporal, e essa resposta corporal é, então, percebida no sistema nervoso central e o significado é atribuído por um processo de avaliação cognitiva: a cobra deixa nossos cabelos em pé, e interpretamos essa resposta como uma reação ao perigo. O teórico cognitivista pode dizer que o medo surge quando, e somente quando, essa avaliação cognitiva é adicionada à nossa experiência da resposta somática.

A melhor maneira de responder a essa objeção é mostrar que sinais somáticos são suficientes para as emoções. A prova demonstrativa dessa afirmação não é, contudo, fácil de ser encontrada, mas a evidência empírica é bastante sugestiva. Considere, novamente, o caso do *feedback* facial. Levenson et al. (1990) têm mostrado que fazer uma expressão facial emocional causa a ocorrência de duas coisas. Isso faz com que um padrão de mudanças corporais (por exemplo, alterações do ritmo cardíaco) seja associado com a emoção expressa, e faz com que alguém relate a experiência dessa emoção. O padrão corporal característico sugere que pelo menos algumas emoções possam ser diferenciadas somaticamente. O auto relato sugere que a pessoa está em um estado emocional induzido. Mas não existe qualquer evidência para a avaliação nesse processo. Convido você a realizar o seguinte experimento. Sorria e verifique se você se sente um pouco mais feliz. Agora, supondo que você o fez, pergunto: você está pensando que seus objetivos foram cumpridos? Em seguida, faça uma careta como se estivesse muito irritado, e verifique se você está sentindo raiva. Pergunte a si mesmo, então, se você está cogitando algo como ter sido insultado.

Eu suspeito que você responderá estas questões negativamente. É claro que as avaliações inconscientes *poderiam* estar ocorrendo, mas supor que elas estão é *ad hoc* e mesmo desnecessário. Podemos reconhecer nossas emoções introspectivamente nesses casos de *feedback*, e parece que estamos a fazê-lo pela introspecção de nossos estados corporais. Se podemos reconhecer uma emoção através de um estado corporal, então não existe razão para pensar que juízos avaliativos não-introspectivos estejam contribuindo para nossos autorrelatos. Concluo que os sinais somáticos são suficientes para certas emoções.

Mas eles são necessários? Defensores das teorias puramente cognitivistas querem nos fazer acreditar que uma emoção pode existir sem qualquer perturbação do corpo, real ou imaginada (por exemplo, Solomon, 1976). Eles querem nos fazer acreditar que algumas emoções são, neste sentido, desmaterializadas. Considere emoções que perduram durante longos períodos de tempo. O amor pode durar por toda a vida, assim como o medo de aranhas. Alguém pode ficar triste por meses, e mesmo incessantemente irritado com o estado atual do mundo. Seria ridículo assumir que o corpo sustenta um estado de perturbação ao longo da duração dessas emoções constantes. Por outro lado, cada uma delas carrega uma disposição para sofrer uma mudança corporal. Questionaríamos a sinceridade de uma aracnofobia autoproclamada por quem nunca estremeceu na presença de sua inimiga de oito pernas. Termos emocionais, assim como termos mentais, possuem um uso disposicional. Em comparação, a alegação de que Jones crê que Deus existe não implica que Jones compulsivamente mantenha esse pensamento em cada minuto de vigília. Podemos distinguir entre estados emocionais e características emocionais, e reservar a consideração somática para os primeiros, desde que reconheçamos que as características são disposições para estados. Essa concessão não é nenhuma vitória para o oponente da visão James-Lange.

Oponentes da visão James-Lange devem defender a afirmação duvidosa de que um estado emocional ocorrente pode ser desmaterializado. Nesse contexto, poderíamos considerar as paixões calmas, discutidas por Hume. A apreciação estética, por exemplo, não está regularmente associada às mudanças corporais. Devemos dizer que ela é uma emoção desmaterializada? Eu penso que não. Em

primeiro lugar, não está claro que a apreciação estética seja uma emoção; ela certamente não é um caso paradigmático. Em segundo lugar, a apreciação estética pode ter um componente corporal. Encontramos obras de arte de tirar o fôlego. Elas fazem nossos olhos se arregalarem e provocam verdadeiros arrepios em nossas espinhas. Obras de arte podem também causar um jardim variado de emoções, tais como a tristeza e a alegria. Essas não são apenas especulações de gabinete. Estudos de neuroimagem têm confirmado que as emoções estão envolvidas quando as pessoas olham para a arte, e as áreas do cérebro envolvidas incluem o cíngulo anterior, que é conhecido por estar envolvido na regulação corporal (Kawabata e Zeki, 2004; Vartanian e Goel, 2004).

O oponente de James e Lange deve olhar para além da apreciação estética para encontrar emoções desmaterializadas. É tentador olhar para essas emoções que parecem filogeneticamente mais avançadas. A apreciação estética se enquadra nessa categoria, mas existem outros exemplos. Uma emoção especialmente pertinente é a culpa. Frequentemente se diz que a culpa não possui um padrão corporal específico, que não há uma expressão da culpa ou qualquer sinal revelador. A culpa destaca-se como uma candidata plausível para uma emoção desmaterializada.

Após a reflexão, contudo, a plausibilidade desse modelo diminui. Em geral, falamos de dores de culpa. Uma pessoa pode ser arruinada pela culpa, assombrada pela culpa, e levada ao desespero pela culpa. A culpa nos faz evitar os olhos daqueles que traímos e nos estimula a buscar o perdão. Nenhuma dessas observações familiares, contudo, prova que a culpa é materializada, mas elas são sugestivas. Imagine o assassino que atesta sentir-se culpado, mas que parece totalmente não-incomodado. Se cada resposta autonômica estivesse baseada nesse patamar, nós certamente suspeitaríamos. A suspeita é derivada do fato de que reconhecemos tacitamente uma perturbação corporal quando nós mesmos sentimos culpa.

Testes mais diretos para a materialização da culpa também estão disponíveis. Em um estudo informal, eu apresentei aos indivíduos a seguinte história:

Após poucos meses na faculdade, você conhece algumas pessoas fantásticas e se envolve em muitas atividades. O ensino médio, assim, parece algo como uma memória muito distante. Um de seus melhores amigos da escola escreve-lhe várias e longas cartas, mas você tem estado tão envolvido com a vida acadêmica que sequer encontra tempo para respondê-las. Meses se passam desde que a primeira carta foi recebida. Pense sobre o sentimento que você tem quando pondera sobre o fato de que você ainda não a respondeu.

Os sujeitos foram convidados a avaliar a adequação de seis expressões faciais. Todos os sujeitos ofereceram avaliações significativamente elevadas para um rosto carrancudo. Os indivíduos também identificaram o sentimento em questão como culpa. Estudos mais cuidadosos precisariam ser feitos para extrair conclusões fortes a partir disso, mas creio que ele levanta dúvidas sobre a hipótese de que a culpa não tenha expressão. Outra evidência para a natureza material da culpa vem de um estudo da neuroimagem. Shin et al, (2000) coletaram a emissão de pósitrons em imagens tomográficas de sujeitos enquanto eles recordavam episódios de culpa a partir de suas próprias vidas. Os resultados mostraram a ativação em áreas que se tornaram muito familiares aos neurocientistas que estudam a emoção: no córtex cingulado anterior, no córtex insular, e nos polos temporais. Como eu tenho observado, o cingulado anterior e a insula também estão implicados na regulação e na percepção do corpo. O fato de que ambas apareçam nesse estudo é uma evidência forte para a hipótese de que a culpa é uma emoção incorporada. Eu proponho que nenhuma emoção *bona fide* é de fato desmaterializada, onde cada candidato aparente demonstra seu núcleo visceral.

Tenho defendido que os sinais somáticos são necessários e suficientes para as emoções. Isso vai contra os teóricos cognitivistas, que insistem na necessidade de juízos, pensamentos ou conceitos. Isso também vai contra Hume, que identifica as emoções com impressões *sui generis*, ao invés de impressões de mudanças corporais. Isso parece ser uma vitória sólida para James e Lange. Contudo, a história não é tão simples assim. A teoria somática deles cai nas garras de um par de graves objeções. Discutirei essas objeções nas próximas duas seções.

2.1.4 Avaliação Racional e Intencionalidade

A primeira objeção à teoria James-Lange decorre de uma simples observação. Costumamos falar de emoções utilizando termos racionais. Dizemos que as emoções são justificadas ou injustificadas, avaliadas ou não-avaliadas, adequadas ou inadequadas. O medo de víboras venenosas circulando livremente é sensato, contudo, o medo de cobras do milho quando estão em cativeiro, não. Como Pitcher (1965) notou, essas categorias de avaliação não se aplicam às sensações corporais normais, tais como cócegas, ou coceiras ou câimbras estomacais. Uma emoção pode ser racional ou irracional, enquanto uma mera pontada ou fisgada é sempre arrracional. Chamo isso de o Problema de Avaliação Racional.

Observe que uma teoria cognitivista não possui qualquer um desses Problemas. Se cada emoção é parcialmente composta de um juízo de pensamento, então cada emoção pode ser distinguida mesmo que seus sintomas somáticos sejam semelhantes. Raiva e indignação implicam diferentes juízos: uma envolve um juízo sobre uma ameaça ou ofensa, e a outra envolve um juízo sobre uma injustiça. A avaliação racional de emoções pode ser fixada sobre estes juízos também. A raiva injustificada [*unwarrant*] é a raiva que surge quando nenhuma ofensa foi realizada, ou quando a ofensa é, em si mesma, o resultado de uma provocação anterior. É a raiva que contém um juízo de que é errado ou injustificado [*unjustified*].

Isso nos deixa certamente em um lamaçal. O Problema de Avaliação Racional nos aponta para uma teoria cognitivista, mas há evidências fortes de que a cognição não é necessária para a emoção. As emoções parecem ser significativas, mas não cognitivas. Há evidência de que elas representam, mas há evidência igualmente boa de que elas não requerem o desdobramento de conceitos.

Felizmente há uma maneira de sair desse lamaçal. A representação não requer cognição, pois podemos ter estados mentais significativos sem lançar mão de quaisquer conceitos. Teorias cognitivistas ignoram essa possibilidade. Elas assumem que cada emoção deve acompanhar um pensamento explícito e descrever alguma relação entre o organismo e o ambiente. A raiva acompanha o

pensamento de que fui insultado. Mas outra possibilidade é que a raiva é, em si mesma, um estado não-conceitual que possui conteúdo representacional.

Para tirarmos esse propósito da lama, precisamos de uma teoria da representação. Seria o caso de recorrer às teorias da representação que não foram desenvolvidas para explicar as emoções. Essas teorias têm a virtude de estarem independentemente motivadas. Devemos descobrir como as emoções representam, ao descobrir como as representações em geral funcionam. Devemos perguntar como, em geral, os estados mentais representam alguma coisa?

Há uma série de respostas a essa questão na literatura. Uma das abordagens mais populares foi desenvolvida por Fred Dretske (1988; ver também Fodor, 1990; e Millikan, 1984). De acordo com Dretske, uma representação mental, *M*, representa aquilo que ela tem a função de detectar de modo confiável. Grosso modo, *M* representa aquilo para o qual ela foi configurada para desencadear. Alarmes de fumaça são um bom exemplo desse tipo de representação que ocorre fora da mente. Um sinal sonoro representa fumaça, porque é confiavelmente causado pela fumaça e foi projetado para ser disparado pela fumaça. Presume-se que as representações mentais funcionam do mesmo modo. Um conceito de água representa água, pois é confiavelmente ativado quando a água é encontrada, e também porque ele foi adquirido para esse propósito. A dor representa doenças físicas porque é confiavelmente causada por elas, e foi desenvolvida para esse propósito. A abordagem de Dretske de representação é amplamente considerada como a mais promissora na literatura. Em outros lugares tenho defendido uma versão dessa abordagem (Prinz, 2000). Tenho argumentado também que essa teoria da representação do propósito geral se aplica igualmente bem às emoções (Prinz, 2004).

Para ver como isso funcionaria, precisamos fazer duas perguntas. O que faz com que emoções ocorram de forma confiável? E, de todas as coisas que causam emoções, quais coisas atribuem a elas a função de detectar? Considere a tristeza, por exemplo. A tristeza pode ser induzida por uma ampla variedade de coisas. A morte de ente querido, a perda de um suéter favorito, uma crítica em um jornal, um olhar cruel de alguém amado. A tristeza também pode ser causada pelo mau

tempo, pelo *Réquiem* de Mozart, pelo álcool, e por franzir a testa ou outras mudanças corporais. Cada uma delas é uma causa confiável. Mas quais dessas múltiplas coisas faz a tristeza ter a função de detectar? Para responder a essa pergunta, tomemos a hipótese plausível de que a tristeza é uma capacidade emocional inata – o produto da seleção natural. Se isso estiver correto, então a questão da função é uma questão biológica: das múltiplas coisas que causam tristeza, quais de suas causas permitiram que a tristeza aumentasse a aptidão de nossos antepassados e os fizeram passar essa capacidade para as gerações futuras? Se a tristeza fosse uma emoção culturalmente construída ou uma resposta aprendida, colocaríamos uma questão paralela: quais causas da tristeza foram responsáveis por sua aquisição e transmissão social? Questões sobre a função biológica são difíceis de responder, pois não temos um registro das condições que levaram à propagação de muitos fenótipos entre nossos ancestrais. Mas podemos, ainda assim, fazer uma estimativa ou aproximação. Suponha que a tristeza detectasse mudanças corporais e nada mais que isso. Teria a tristeza nos conferido alguma vantagem de sobrevivência? Provavelmente não. Não há vantagem em meramente detectar as mudanças corporais associadas com tristeza se essas mudanças não têm, nelas mesmas, qualquer significado biológico. Se testas franzidas não são importantes, a detecção de testas franzidas seria relativamente inútil. Poderia a tristeza ter a função de detectar certos tipos de música? Isso é pouco provável, pois a música pode ter surgido muito mais recentemente em nossa história evolutiva do que a tristeza. É mais provável que a música nos cause tristeza em virtude de sua semelhança com algum indutor fundamental de tristeza, tal como o som do choro (Panksepp, 1995).

Poderia a tristeza ter evoluído como uma resposta ao mau tempo? Aqui começamos a esquentar. O mau tempo sinaliza uma redução potencial na acessibilidade de recursos essenciais, de modo que se torna vantajoso detectar o mau tempo. Mas, se essa é uma função da tristeza, ela não é a única função. É igualmente vantajoso detectar a perda de um ente querido, a falta de sucesso, e o tratamento frio de alguém amado. Todas essas coisas constituem uma ameaça ao bem-estar, e a pessoa que está triste por elas vai buscar estratégias de

enfrentamento. Assim, intuitivamente, a tristeza pode ter sido selecionada por uma série de coisas diferentes, incluindo o mau tempo, mortes, fracassos, divórcios e artigos perdidos. Observe que essa coleção arbitrária de indutores é realmente unificada por um princípio único e coerente. Cada item na lista é um tipo de perda. A perda da vida, a perda de recursos, a perda de alguém amado, a perda de coisas – todas essas perdas induzem tristeza. Assim, podemos dizer que a tristeza evoluiu como uma resposta às perdas em larga escala. Tristeza é um detector de perdas. É também um detector de mudanças corporais e de acordes D-menores, mas ela não tem a função de detectar essas coisas. A tristeza provavelmente evoluiu como uma resposta à perda.

Em suma, a teoria independentemente motivada da representação de Dretske oferece uma resposta bastante satisfatória para a pergunta sobre o que a tristeza representa. Ela simplesmente decorre da teoria de Dretske de que a tristeza representa perdas. O mais importante sobre esse resultado é que a teoria de Dretske não depende de quaisquer suposições sobre a forma ou formato das representações. Não há qualquer condição em sua teoria que exija que os estados mentais aos quais ela se aplica sejam proposicionais, conceituais ou cognitivos. A teoria de Dretske aplica-se às representações altamente conceituais, mas também pode ser aplicada a dores, imagens sensoriais, e mesmo anticorpos no sistema imunológico. Pode ser aplicada às emoções, mesmo que as emoções sejam identificadas com os estados que regulam e registram as mudanças corporais.

Pense nisso dessa maneira. Os sistemas somático-sensoriais do cérebro possuem estados internos que evoluíram para detectar e regularizar os sistemas no corpo. Esses estados evoluíram para detectar o impacto de estímulos em nossos corpos, para identificar posições musculares para o controle comportamental, e para manter a homeostase e as mudanças adaptáveis em nossos órgãos. Mas, para além de cada mudança discreta em cada órgão ou músculo individual, o cérebro pode detectar padrões de mudanças corporais. Alguns desses padrões evoluíram como preparações para respostas comportamentais, tais como a fuga ou a agressão. Os padrões que atendem àqueles fins comportamentais evoluíram também para ocorrer sob condições específicas. Um desses padrões – aquele que

experimentamos como tristeza – é desencadeado por pensamentos sobre perda, esforços fracassados, olhares frios de alguém amado, e pela sobrecarga de céus cinzentos. Os estados corporais que ocorrem provavelmente nos preparam para a retirada e diminuição da atividade, para que possamos armazenar energia, finalizar buscas infrutíferas, e identificar outras formas de enfrentar a situação. A percepção desse padrão é, pelas razões que eu tenho mencionado, uma representação de perda, mesmo que cada parte do padrão seja também uma sensação de alguma mudança corporal específica. O mesmo ocorre com outras emoções. A fuga é uma resposta desencadeada por ruídos repentinos, a aproximação de predadores, e outros riscos; quando percebemos esse padrão, a percepção corporal é uma representação de perigo. A agressão é uma resposta desencadeada por ataques não-específicos, e outras ameaças e provocações; a percepção corporal da agressão é uma representação de uma afronta. Ao primeiro chamamos de estado de medo; ao segundo, de raiva. Ambos são percepções corporais que representam relações específicas entre organismo e ambiente.

James e Lange dão a entender que as emoções são sensações de estados corporais e que, portanto, são representações desses estados. A teoria de Dretske, e de outros como ele, sugere que James e Lange estavam errados. As emoções são causadas confiavelmente pelas mudanças corporais, mas elas representam coisas como a perda e o perigo, que são as relações organismo-ambiente que induzem essas mudanças em nós. James e Lange estavam corretos sobre a forma que essas emoções tomaram, mas estavam errados sobre seus conteúdos. Em uma palavra, as emoções representam preocupações [*concerns*].

Teóricos cognitivistas assumem que os não-cognitivistas não conseguem explicar como as emoções são causadas por preocupações. Se emoções contém juízos, podemos facilmente explicar como o medo é causado pelo perigo. O medo contém juízos do tipo “estou em perigo”, e esse juízo tende a ocorrer quando estou em perigo. Os não-cognitivistas não negam que esses juízos possam desencadear emoções; eles simplesmente negam que tais juízos sejam os únicos desencadeadores. O medo pode ser desencadeado quando ouço um ruído alto, quando sinto a perda de apoio, quando vejo uma cobra, *ou* quando julgo que estou

em perigo. Cada um desses estados internos pode ser considerados como elementos de um arquivo mental. Eles são funcionalmente unificados por suas capacidades de causar o padrão corporal que experimentamos como medo. O medo representa perigo em virtude do fato de que ele tem a função de detectar perigo, e essa função é sustentada por esse arquivo. O arquivo calibra o medo segundo o perigo. Ele pode ser chamado de um arquivo de calibração. Arquivos de calibração possuem uma grande variedade de representações, tanto cognitivas como não-cognitivas, e essas representações podem mudar ao longo do desenvolvimento cognitivo. Não existe razão para pensar que esses arquivos de calibração sejam partes componentes das emoções. Como Hume afirmou, eles são muito contingentes e variáveis. Além disso, quando uma representação a partir de um arquivo de calibração desencadeia uma emoção (por exemplo, quando ouvir um som alto aciona o medo), essa representação ocorre antes do aparecimento do emoção, e pode perdurar por um período mais breve que a emoção. As representações em arquivos de calibração são causas de emoções, e não os seus componentes. Elas são mecanismos que tomam as emoções para servir como detectores significativos de nossa relação com o mundo.

Em suma, podemos explicar como as emoções representam preocupações sem supor o que as emoções de fato são ou contêm, ou ainda, se envolvem essencialmente juízos ou não. Essa conclusão é extraída da teoria da representação de Dretske e de outros como ele. As emoções representam coisas tais como perdas e perigos porque elas foram criadas para serem desencadeadas por essas coisas. Elas representam essas coisas, mesmo que elas não tenham conceitos ou ideias constituintes. Assim como um sinal sonoro de um detector de fumaça, as emoções podem representar sem descrever.

Isso salva a teoria James-Lange da refutação, pois fornece uma resposta ao Problema da Avaliação Racional. Os não-cognitivistas tiveram dificuldades para explicar a intuição de que as emoções podem ser razoáveis ou não-razoáveis. Hume diz que as emoções são arracionais. Hume afirmou que as emoções podem ser causadas por ideias razoáveis ou não-razoáveis, mas elas não podem ser razoáveis ou não-razoáveis em si mesmas. Isso é insatisfatório, pois podemos

avaliar as emoções como razoáveis ou não-razoáveis mesmo quando elas não estão associadas a juízos ou ideias. O medo causado pela visão de uma cobra enjaulada não é razoável, mesmo se nenhuma ideia precede ou se segue da emoção. James e Lange não fazem melhor do que isso. As emoções representam mudanças corporais em suas abordagens, e nada mais. Eu ofereci uma emenda à teoria James-Lange que pode resolver esse problema. Se as emoções representam preocupações, elas podem estar corretas ou incorretas. Se Jones teme uma cobra em cativeiro, ele está literalmente representando a cobra como perigosa. Isso é um erro, mas ele estaria perfeitamente justificado em temer a víbora venenosa quando esta se encontra enrolada em sua perna.

A questão é, na verdade, um pouco mais complicada. Se as emoções são passivas, ou fora de nosso controle direto, então as normas epistêmicas podem ser mais fracas do que elas seriam para juízos. Podemos ser acusados de raciocinar mal, mas é difícil culpar alguém por estar deprimido. Podemos dizer que um pensamento não é razoável e que um pensador não é razoável por ter esse pensamento. Com emoções, fica mais difícil criticar a pessoa, pois exercemos menos controle sobre o que sentimos do que sobre as inferências que traçamos. Dizer que uma emoção é irracional não é sempre uma acusação ao seu portador. O medo que experimentamos em montanhas-russas é irracional, dada a sua relativa segurança, mas nós, enquanto agentes, não somos irracionais por ter esse medo. Emoções irracionais implicam agentes irracionais apenas quando os agentes podem exercer algum controle sobre essas emoções, como quando as emoções são desencadeadas por crenças ao invés de percepções de níveis mais baixos. Medo de voar implica a irracionalidade por parte de um agente de uma forma que o medo experimentado em uma montanha-russa não. Esse último implica a agência irracional somente se ele é usado para embasar certas decisões ou crenças ("Recuso-me a andar em montanhas-russas porque elas são perigosas"). Aqui reside a falha do agente a respeito da inferência a partir da emoção, e não na emoção em si mesma.

Em suma, tenho argumentado que a teoria James-Lange pode ser modificada para resolver o Problema da Avaliação Racional. Como Hume, James e

Lange nunca explicaram como as emoções poderiam representar preocupações. Consequentemente, suas teorias falham em explicar a avaliação racional das emoções. Ao complementar a teoria James-Lange com uma consideração de como as emoções representam, esse problema poderia ser resolvido. Eu chamo isso de *teoria da avaliação materializada*. As emoções são materializadas, pois elas são sinais somáticos, assim como defenderam James e Lange. Mas as emoções são também avaliações, na medida em que representam preocupações, como defendem as teorias cognitivas usuais (Prinz, 2004).

2.1.5 Semelhança Somática, Emoções Básicas e Calibração

Eu afirmei que a teoria James-Lange enfrenta duas graves objeções. Acabo de oferecer uma resposta à primeira delas. As emoções podem ser avaliadas racionalmente, pois representam preocupações. E elas representam preocupações mesmo que não contenham constituintes conceituais. Isso significa que a equação de James das emoções com os sinais somáticos pode ser preservada. O problema remanescente a essa visão é que simplesmente não existem estados corporais suficientes para circundá-los.

Estritamente falando, o corpo tem um certo número de estados distintos. Em cada minuto de nossas vidas, nossos corpos estão provavelmente em um estado um tanto quanto diferente daqueles em que já estiveram antes. Cada arranjo de músculos, cada ritmo de fluxo de sangue, e mesmo sequência de respirações, é único. Mas, todas essas variações podem ser classificadas em padrões repetitivos. O problema para a abordagem James-Lange não é que existem também muitos padrões corporais, mas que há muito poucos. Se nosso grupo de padrões semelhantes são tomados em conjunto, então a matriz de padrões associados às emoções parece bastante limitada. Emoções diferentes estão associadas às mesmas ou a semelhantes mudanças somáticas. Considere a culpa novamente. Anteriormente eu disse que a culpa está associada com a testa franzida. Ela é uma emoção de baixo consumo energético, oprimida. Isso deve soar familiar. O mesmo pode ser dito sobre a tristeza. Se estudamos as expressões corporais de culpa,

podemos descobrir que elas se sobrepõem consideravelmente às expressões de tristeza. Ou, para tomar um exemplo mais óbvio, considere a raiva e a indignação justificada. Essas são emoções diferentes. Alguém pode estar com raiva sem estar indignado (apesar de que, curiosamente, o inverso seja menos plausível). Se houvesse um padrão corporal único para cada emoção, então esperaríamos que a raiva e a indignação fossem somaticamente distinguíveis. É improvável, no entanto, que este seja o caso. Às vezes, emoções diferentes correspondem aos mesmos padrões corporais. Chamo a isso de Problema de Semelhança Somática.

Esse problema representa uma grave ameaça para a teoria de James-Lange, ao menos como foi originalmente formulada. Essa teoria corre o risco de deturpar o número de emoções que possuímos. Isso implica que a raiva e a indignação são uma única emoção, e o mesmo ocorre para vários outros pares aparentemente distintos de emoções: tristeza e culpa, repulsão física e desgosto moral, alegria e orgulho. Felizmente há uma saída. A teoria James-Lange reformulada que estou recomendando tem uma solução simples para o Problema da Semelhança Somática.

Na teoria da avaliação materializada, as emoções representam preocupações, tais como perdas, perigos e crimes. Utilizando os tipos de analogias que Dretske tornou popular, eu comparei as emoções com os alarmes de fumaça. Mais precisamente, elas são como os tons emitidos pelos alarmes de fumaça. Tons, ao que parece, podem desempenhar uma dupla função. Tons acusticamente idênticos podem ser emitidos por diferentes dispositivos. Imagine uma casa que tem tanto um alarme de fumaça quanto um alarme de monóxido de carbono, e imagine que ambos emitam tons muito semelhantes. Quando um deles dispara, os proprietários podem não saber qual é o alarme que está soando, mas existe um fato que diz qual é o tom. Cada tom tem um significado, tal como é determinado pelo mecanismo ao qual ele está ligado.

Por analogia, quero afirmar que um sinal somático do mesmo padrão corporal por ter significados diferentes em ocasiões diferentes, dependendo dos mecanismos mentais que permitiram a formação do padrão. Imagine que um mecanismo detector de fumaça e o mecanismo detector de carbono estejam

ligados ao mesmo sinal sonoro. O que estamos dizendo sobre o significado de um tom em qualquer ocasião dada? Existe uma forte tentação para dizer que ele tem um significado disjuntivo, fumaça-ou-monóxido-de-carbono. Eu acho que isso está errado. Ao ouvir o tom, os proprietários poderiam de fato dizer “esse tom significa que há ou fogo ou gás”, mas eles poderiam também dizer “esse tom está provavelmente indicando a presença de apenas um desses dois perigos, e é melhor descobrir qual”. Creio que o tom tem um significado unívoco quando ele ocorre, exatamente da mesma maneira que uma palavra ambígua tem apenas um significado quando é proferida em um ato de fala. A palavra “morcego” refere-se a um animal se ela se insere em uma entrada léxica no sistema de linguagem da mente, e a uma peça de equipamento esportivo se ela se insere em outra entrada lexical. Colocando em termos da teoria de Dretske, cada símbolo do som “morcego” é na verdade um símbolo de um tipo diferente. Os sons se inserem em arquivos mentais separados, um dos quais tem a função de rastrear roedores voadores e o outro a de acompanhar jogadores de beisebol.

Vejamos o que se segue disso para as emoções. Suponha que a raiva e a indignação tenham os mesmos sintomas corporais. Suponha ainda que esses sintomas corporais possam ser produzidos por meio de dois mecanismos diferentes. Um mecanismo é uma coleção de impressões e ideias relativas aos insultos e ameaças, e o outro mecanismo é uma coleção de impressões e ideias relativas à injustiça. E, finalmente, suponha que os sistemas cognitivos em que residem esses mecanismos possam algumas vezes determinar quais dos dois mecanismos foram recentemente operados. O sistema pode recordar se ele foi recentemente simbolizado por pensamentos de insulto ou de injustiça. Imagine um caso onde a experiência de ler um boletim da Anistia Internacional cause uma perturbação corporal. A representação dessa perturbação é de tipo ambíguo, mas esse símbolo particular é fácil de diagnosticar. Ela foi causada por pensamentos de injustiça; é um caso de indignação. Se alguém com amnésia não conseguisse lembrar quais pensamentos desencadearam a resposta emocional, ainda assim haveria uma questão de fato. A resposta conta como raiva se ela tem um tipo de causa, e como indignação, se ela tem outro tipo. Frequentemente não recordamos

das condições que desencadearam nossas emoções, e, quando isso acontece, podemos ficar inseguros a respeito de qual emoção de fato estamos sentindo. Um psicoterapeuta poderia ajudar um paciente a reconhecer que a emoção não-identificada que ele está experimentando é a culpa.

A proposta da calibração fornece uma solução para o Problema da Semelhança Somática. Dois símbolos do mesmo tipo de sinal somático podem constituir emoções diferentes, pois eles podem estar sob o controle de diferentes arquivos de calibração, e representam preocupações diferentes. Dessa forma, um pequeno grupo de padrões corporais pode subscrever um grande grupo de emoções.

Isso levanta uma questão. Por que toda a duplicação? Por que os mesmos padrões corporais são usados repetidas vezes? Por que há tão poucos padrões estáveis de mudanças corporais e tantas emoções? A resposta a essa questão tem algo a ver com o local de onde vêm as emoções. Creio que estamos naturalmente equipados com um pequeno repertório emocional. Nossas emoções "básicas" são reutilizadas em novas formas para criar novas emoções. Através da experiência, mecanismos são criados para correlacionar uma resposta corporal existente com um novo conjunto de condições evocativas, uma nova preocupação.

A ideia de emoções básicas não é, contudo, nova. Descartes acreditava em emoções básicas, e assim também o fez Spinoza. As emoções básicas também foram um tema comum na pesquisa empírica moderna. McDougall (1908) postulou um conjunto de emoções básicas em seu pioneiro trabalho em psicologia social. Outra lista foi oferecida por Magda Arnold em sua teoria da avaliação seminal (Arnold, 1960). A mais ampla lista de emoções básicas pesquisadas, contudo, surge com Paul Ekman e seus colaboradores, que as usou em seus estudos seminais do reconhecimento de emoções entre culturas (Ekman, 1972). A lista clássica de Ekman contém seis elementos: felicidade, tristeza, raiva, medo, surpresa e aversão. Creio que isso pode precisar de uma pequena revisão, mas é uma boa primeira aproximação, principalmente quando se considera esses rótulos de emoção como se referindo a grandes famílias de emoções, que podem ter manifestações sutilmente diferentes entre as culturas. A ideia é que temos um vocabulário

emocional universal, que é rapidamente cooptado pela experiência e sintonizado às condições desencadeadoras culturalmente específicas.

Existem duas maneiras de uma emoção básica poder ser utilizada para criar novas emoções. Em primeiro lugar, ela pode ser combinada com outra emoção básica para criar uma composição emocional. O desprezo pode ser uma combinação de raiva e aversão. A euforia pode ser a combinação de alegria e medo. Em segundo lugar, uma emoção básica pode ser atribuída a um novo conjunto de condições desencadeadoras. Em casos normais, elas serão um subconjunto das condições desencadeadoras iniciais que tenham sido elaboradas, por meio da experiência, para formar um mecanismo desencadeador independente. Eu chamo isso de recalibração. Anteriormente eu disse que as emoções estão associadas a arquivos mentais que as calibra às preocupações. Na recalibração, um novo arquivo mental é estabelecido, o que estabelece uma emoção existente com um novo conjunto de condições desencadeadoras. Por exemplo, o orgulho pode ser a alegria recalibrada para os sucessos próprios de alguém.

Em muitos casos a combinação e a calibragem conspiram juntas. A nostalgia pode ser uma combinação de alegria e tristeza dirigida ao passado. A pena pode ser uma triste aversão dirigida a outro indivíduo, que é considerado inferior. O ciúme pode ser uma mistura de raiva, tristeza, medo e aversão, sintonizada para situações de suspeita de infidelidade sexual. Esses métodos de geração de novas emoções são cruciais para explicar as emoções morais. Por exemplo, a indignação pode ser a raiva calibrada para a injustiça, e a culpa pode ser a tristeza calibrada para os danos autocausados. Vou desenvolver e defender essas propostas na próxima seção, mas, primeiro, quero salientar que essa abordagem pode ser usada para superar uma dificuldade levantada na seção 2.1.2. Lá eu argumentei que é difícil conciliar o emocionismo com as teorias cognitivas da emoção. Uma razão para isso tinha a ver com o argumento da questão em aberto de Moore. Suponha que a indignação seja definida como o juízo a respeito de que houve uma distribuição desigual de recursos. E suponha que a indignação é parte do conceito de ERRADO. Em outras palavras, suponha que alguns casos simbólicos do conceito ERRADO contenham literalmente um estado de indignação. Se esse fosse o caso, então,

quando se usa tal símbolo, seria incoerente perguntar se a desigualdade estava errada, pois a resposta a essa questão seria uma verdade conceitual. A abordagem de calibração contorna essa dificuldade, pois os arquivos de calibração não são componentes das emoções que eles calibram. A indignação não contém um arquivo de calibração. A indignação é apenas um sinal somático. Sua ocorrência é causada por um arquivo de calibração, e esse arquivo contém presumivelmente juízos sobre a desigualdade e vários outros tipos de injustiça. O arquivo de calibração pode conter também registros perceptuais armazenados de eventos que são injustos. Por exemplo, se você já observou ou imaginou alguém roubar de um morador de rua, uma imagem desse evento poderia ficar armazenada no arquivo de calibração que desencadeia a indignação. Os itens desse arquivo são causas de indignação, mas não são seus componentes. Portanto, os símbolos do conceito ERRADO que contém indignação não contém qualquer uma das imagens ou ideias no arquivo que calibra a indignação. Nenhuma instância do conceito ERRADO contém o conceito DESIGUALDADE, por exemplo. Assim, perguntar se a desigualdade é errada não se trata de uma tautologia.

Antes de iniciar uma discussão mais aprofundada das emoções morais, façamos um balanço. Tenho desenvolvido uma teoria das emoções que se baseia na abordagem somática de James e Lange. A teoria original James-Lange enfrenta duas objeções sérias: o Problema da Avaliação Racional e o Problema da Semelhança Somática. A teoria da avaliação materializada é capaz de resolver esses dois problemas. Essa teoria segue a teoria de James-Lange ao afirmar que as emoções são sinais somáticos. As emoções não são estados cognitivos. Mas a teoria se afasta de James e Lange em um aspecto crucial: as emoções representam preocupações. Isso significa que uma emoção pode ser não-garantida ou garantida, e ela também permite que duas emoções possam ter conteúdos diferentes apesar de serem somaticamente indistinguíveis. Em suma, argumentei que as teorias cognitivistas estão corretas sobre o conteúdo das emoções, mas que as não-cognitivistas estão corretas sobre a forma das emoções.

2.2 AS EMOÇÕES MORAIS

Como uma definição difícil, direi que as emoções morais são emoções que surgem no contexto da conduta moralmente relevante. Mais especificamente, as emoções morais promovem ou detectam as condutas que se conformam ou que violam uma regra moral. A partir da leitura das páginas dos primeiros moralistas britânicos, pode-se ter a impressão de que existem apenas duas emoções morais. Hutcheson fala sobre aprovação e condenação. Ele diz que essas "são provavelmente ideias simples, que não podem ser mais explicadas" (1738:67). Hume sugere uma visão semelhante quando se refere à aprovação e desaprovação. Adam Smith critica a posição de Hutcheson por não perceber que os sentimentos morais são derivados de emoções mais básicas (1759:7, III, iii). Creio que Smith está certo. A aprovação e a desaprovação não são os nomes de sentimentos simples; elas envolvem uma variedade de emoções diferentes. Smith diz que as aprovações envolvem coisas como simpatia e gratidão, enquanto a desaprovação envolve a antipatia e o ressentimento. Minha análise é um pouco diferente. Começarei com a distinção entre emoções reativas e reflexivas. Elas correspondem, aproximadamente, ao que alguns autores chamam de emoções de "culpa-do-outro" e de "autoculpa" (Bem-Ze'ev, 2000), mas elas incluem também as emoções de louvor. Eu recorrerei a diversos tipos ou cuidados morais, incluindo a simpatia.

2.2.1 Emoções Morais Reativas

Eu defino as emoções morais reativas como emoções que surgem quando outra pessoa (ou grupo) é interpretada como estando em conformidade ou como violando uma regra moral. Emoções morais reativas dividem-se em duas classes: culpa e louvor, ou, em termos humeanos, em aprovação e desaprovação. Começarei com a culpa.

Anteriormente eu mencionei um exemplo de uma emoção de culpa direcionada aos outros. Tratava-se da emoção de indignação. A indignação não é uma emoção básica; ela é derivada da raiva. A indignação é a raiva calibrada para a

injustiça. Sentimos indignação quando o governante falha em cuidar dos governados. Em um nível mais local, podemos sentir algo semelhante quando injustiças menores são cometidas pessoalmente contra nós. Se alguém não retribuir ou reconhecer um favor (dizendo "muito obrigado"), podemos nos sentir indignados. Usamos a forma adjetiva da palavra ao invés da forma nominal, talvez porque o nominal signifique injustiças grandiosas ou mais elevadas.

A raiva é também a fonte de outra emoção moral. Quando uma pessoa inflige dano físico em nós, experimentamos algumas vezes um tipo de raiva que não requer qualquer pensamento sobre injustiça. Em algum sentido moral, nos sentimos injustiçados. Uma raiva semelhante é sentida quando alguém toma algo de nós. Em ambos os casos, sentimos que nossos direitos foram violados. Não há nome preciso para essa raiva em inglês. Vou chamá-la de raiva justificada para enfatizar a ligação com direitos. A raiva justificada difere-se da indignação, pois algumas violações de direitos não são violações de justiça. Como pretendo utilizar o termo, a justiça envolve princípios que tem a ver com imparcialidade, equidade e proporção. Quando alguém rouba alguma coisa, não pensamos necessariamente em imparcialidade. Uma pessoa gananciosa, despreocupada com a imparcialidade ou com a justiça, poderia se sentir violada quando alguém rouba algo dela. Ela poderia dizer "Como você se atreve!", mas não "Como você pôde!". Em muitos casos, contudo, direitos e justiça coincidem. O ladrão que leva algo de uma pessoa necessitada faz algo injusto. Nesses casos, as violações evocam *indignação* justificada. No caso de violações menores, nos sentimos justificadamente *indignados*, como quando alguém nos joga para fora da estrada ou esbarra em nós na rua sem pedir desculpas. Vou me referir a essa classe inteira de emoções como raiva moral.

Raiva e indignação justificadas são diferentes porque representam preocupações diferentes. Mas ambas são derivadas da mesma emoção básica. Por que isso? A resposta é que suas preocupações se sobrepõem às preocupações que desencadeiam a raiva em sua forma mais básica. A raiva comum é uma resposta a insultos ou ameaças. O mesmo ocorre com a raiva moral. A raiva moral surge quando ocorre uma injustiça ou quando os direitos de alguém são violados. Em

ambos os casos, há normalmente uma vítima: alguém foi prejudicado ou insultado, por ser tratado de um modo que indica que ele não tem direito ou valor. Quando reagimos emocionalmente à vitimização, a raiva em resposta é natural, pois ela evoluiu para lidar com ameaças e nos predispõe à agressão. Se você tem uma capacidade inata para a raiva, é provável que ela seja recrutada quando você se torna sensível à injustiça e às violações de direitos.

Alguns filósofos supõem que a raiva é sempre uma emoção moral. Eu discordo. A raiva surge quando alguém olha para você, ou desafia você a lutar, ou deprecia você. Todas essas coisas podem ser consideradas como imorais, mas elas não precisam ser. Dois pugilistas em uma luta podem sentir raiva, mesmo que eles não pensem que é errado o outro ser atingido nesse contexto. Ou considere este exemplo de Baier (1967). Um casal guardou dinheiro para a educação de seu filho, mas agora eles decidem pegar esse dinheiro e gastá-lo em férias extravagantes. O filho está irritado sobre isso, mas ele também é moralmente generoso; ele diz a seus pais "Vocês não têm qualquer obrigação moral de pagar pela minha educação, mas eu estou com raiva de vocês por essa escolha". O filho sente que seus pais o privaram egoisticamente de algo, mas ele não crê que o fazer fosse imoral para eles; eles não tinham obrigação moral. Baier pensa que esse exemplo representa uma ameaça para o emotivismo (e por extensão ao emotivismo epistêmico). Se juízos morais eram meramente uma questão de ter emoções tais como raiva, então o filho, nesse cenário, não poderia dizer consistentemente que ele está com raiva sem também afirmar que seus pais fizeram algo imoral. Mas essa dificuldade pode ser resolvida se distinguirmos a raiva ordinária da raiva moral. A raiva do filho, nesse caso, não surge de um arquivo de calibração que contenha ideias referentes à injustiça ou às violações de direitos.

A raiva moral é central para a moralidade, pois muitas de nossas regras são proibições contra violações de direitos e injustiças. Tais regras dizem respeito a vítimas; elas sempre envolvem ameaças ou danos às pessoas. Mas existem também regras morais que têm pouco a ver com justiça e com direitos. Considere, por exemplo, os tabus sexuais. Condenamos moralmente aqueles que se envolvem em certas práticas sexuais: bestialidade, incesto, adultério, e, em tempos anteriores,

sexo antes do casamento, sexo homossexual e masturbação. Em alguns casos esses atos envolvem a violação de direitos (o incesto pai-filha é um exemplo especialmente preocupante), mas em outros casos não há nem mesmo uma leve ameaça. O incesto consensual entre irmãos, mencionado no capítulo 1, é um exemplo. A intuição de que isso é moralmente errado não pode depender de interpretá-lo como uma ameaça leve ou grave. Ninguém é prejudicado por incesto consensual se as partes ofensivas desfrutaram da experiência e não falam a ninguém sobre isso. Consequentemente, a emoção sentida pelas partes ofensivas não pode ser a raiva moral. Ao invés disso, como mencionado no capítulo 1, a emoção desencadeada é uma forma de aversão.

A aversão moral não é uma emoção básica. Ela é uma consequência da aversão comum. A aversão comum diz respeito à contaminação física. Paul Rozin fez uma extensa pesquisa empírica demonstrando que a aversão surge em situações onde há um risco comum de danos devido a germes e outros perigos ocultos (Rozin et al., 1993). Seus desencadeadores primitivos incluem fluidos corporais (todos, exceto lágrimas), criaturas não-humanas, coisas mortas, lixo, ou objetos que entraram em contato com qualquer uma dessas coisas. O princípio de contato é importante. Germes representavam uma ameaça mesmo antes que tivéssemos qualquer concepção sobre eles. Para garantir a nossa segurança, a evolução nos forneceu respostas negativas a coisas que poderiam abrigar germes, incluindo certos materiais biológicos e coisas que foram tocadas por esses materiais. A aversão é uma emoção digestiva. Sua base corporal envolve um padrão associado à rejeição digestiva na boca, nariz e intestino.

Rozin et al. (1993) demonstraram que a aversão moral é derivada da repugnância física, mostrando que ela tem a mesma base corporal e a mesma lógica de contaminação. Não gostamos de ter contato físico com objetos que foram tocados por uma pessoa que consideramos moralmente repugnante. Por exemplo, não gostaríamos de viver na antiga casa de um pedófilo condenado.

Costumes sexuais são candidatos óbvios para a aversão moral, pois o sexo é um ato carnal que notadamente envolve a transferência de fluidos corporais. Uma vez que essas coisas podem desencadear aversão por conta própria, não é

surpreendente que as violações de regras sexuais sejam consideradas como nojentas. A repugnância moral é também dirigida aos assassinos em massa, talvez porque eles estão associados à mutilação e à morte, que são desencadeadores primitivos de aversão. Em um experimento, Rozin et al. (1993) descobriram que os indivíduos não estavam dispostos a experimentar um suéter que, segundo eles, pertencia a Hitler. Sua recusa continuou mesmo após ofertas monetárias terem sido feitas, e eles expressaram aversão. Rozin e Singh (1999) também descobriram que a aversão está associada com tabus morais contra o tabagismo. O fumo é um poluente, logo ele pode facilmente desencadear aversão.

A poluição é, muitas vezes, metaforicamente estendida ao domínio moral. Por exemplo, certas tradições religiosas pensam que a alma se tornará poluída se você agir de maneira que viole as leis divinas. Shweder et al. (1997) demonstraram que os ideais de pureza teleologicamente compreendidos são fundamentais para o sistema moral em Orissa, na Índia, onde eles realizaram um extenso trabalho de campo. Existem regras morais que são relativas ao que você pode comer, com quem pode falar, com o que você pode vestir em templos, onde você pode ir se você é mulher e está menstruada, e quais estados psicológicos você pode ter. Em todos esses casos, os violadores são considerados impuros ou contaminados. Por exemplo, é uma forma de impureza moral ter ódio em seu coração. Em alguns desses casos, existe uma ligação direta com os desencadeadores de aversão física, mas, na maioria, a impureza é metafórica. A pureza física é transformada em pureza espiritual. Shweder et al. (1997) referem-se a sistemas morais que enfatizam a pureza como "ética da divindade", porque, muitas vezes, desempenham um papel central nas sociedades não-seculares, e a pureza é normalmente conceituada de uma forma sobrenatural: atos imorais contaminam a alma. Quando sua alma está contaminada, você não está necessariamente exposto a qualquer perigo físico (ao menos neste mundo), mas você está corrompido aos olhos de Deus ou dos deuses. Nas sociedades seculares, a ênfase tende a ser colocada em outro lugar: na justiça e nos direitos. Shweder et al. chamam isso de "ética de autonomia". A preocupação com a justiça e com direitos decorre da visão de que os indivíduos estão autorizados a fazer certas coisas, e atos imorais são interpretados como crimes

contra as pessoas. Dentro de uma ética da autonomia, o conceito moral central é o de dano, ao invés do de pureza. Comportamentos que não causam danos (como atos sexuais consensuais, dieta ou vestimenta) são menos propensos a ser moralizados em sociedades que possuem uma ética de autonomia.

Shweder et al. (1997) também identificam um terceiro tipo de sistema ético que é bem central em algumas culturas. Eles a chamam de "ética de comunidade", e suas regras governam as relações pessoais com os grupos sociais. Regras que governam a posição social, o respeito pela família e o tratamento de recursos comuns caem sob essa categoria. Sociedades coletivistas tendem a enfatizar regras desse tipo. Elas reservam um valor especial para o lugar de cada indivíduo na sociedade, e a impossibilidade de desempenhar os papéis associados a esse lugar são moralmente proibidos. Quando as pessoas saem da linha, elas perturbam a ordem social. Em tais sociedades, o desrespeito é uma transgressão cardinal, e é frequentemente mais enfatizado do que a impureza e o dano. Shweder e seus colaboradores não creem que essas sociedades operam sob um único sistema moral: a divindade, a autonomia ou a comunidade. Em vez disso, todas as comunidades tendem a ter normas em cada categoria. As diferenças interculturais manifestam-se no grau em que qualquer categoria é forçada, e, em alguns casos, a mesma forma de conduta será interpretada sob diferentes categorias. Por exemplo, em nossa sociedade temos a tendência de pensar o incesto como prejudicial, enquanto que em uma sociedade não-secular é mais provável que o incesto seja visto como impuro, e, em uma sociedade coletivista, ele poderia ser interpretado como uma ameaça à estrutura familiar. É claro que, na maioria das sociedades, todas as três conceituações desempenham o mesmo papel. Quando somos confrontados com casos de incesto consensual, em que ninguém é prejudicado, voltamos a pensar o incesto como um tipo de contaminação.

Em um estudo muito importante, Rozin et al. (1999) mostram que essas três classes de regras estão associadas a diferentes emoções. Eles apresentaram vinhetas a partir de uma série de descrições comportamentais, baseadas em experimentos psicológicos, a japoneses e americanos e solicitaram a eles que identificassem a resposta emocional apropriada. Eles utilizaram exemplos em que

havia ou um caso claro de dano (por exemplo, o homicídio), ou um caso claro de desrespeito (por exemplo, um adolescente jantar antes dos outros serem servidos), ou um caso claro de algo que tende a ser considerado como poluição do corpo (por exemplo, uma adolescente fazer sexo com um homem de setenta anos de idade). Rozin et al. descobriram que as violações de autonomia tendem a estar associadas com a raiva, as violações de pureza com a aversão, e as violações da comunidade com o desprezo (uma emoção que discutirei em breve). Eles chamam isso de modelo CAD, por desprezo [*contempt*], raiva [*anger*] e aversão [*disgust*]. Eles observam que coincidentemente o modelo CAD é também uma sigla formada a partir de comunidade [*community*], autonomia [*autonomy*] e divindade [*divinity*]. O modelo CAD é uma das realizações mais elegantes em psicologia moral recente. Mas ele não é inteiramente novo. A ideia de que emoções diferentes podem ser desencadeadas por diferentes tipos de transgressão já foi sugerida antes. Hume (1739: III. iii. iv) explicitamente argumenta que diferentes tipos de atos suscitam emoções morais diferentes, e, em outra observação presciente, ele escreve:

[Em] deliberações morais devemos estar familiarizados de antemão com todos os objetos, e todas as suas relações cada um com o outro; e a partir de uma comparação do conjunto completo, fixar nossa escolha ou aprovação.... A aprovação ou censura que então se segue, não pode ser a obra de um juízo, mas do coração; e não é uma proposição especulativa ou afirmação, mas um sensação [*feeling*] ativo ou sentimento [*sentiment*].... [A] mente, a partir da contemplação do todo, sente alguma impressão nova de afeto ou aversão, estima ou desprezo, aprovação ou censura. (Hume, 1751: Apêndice I, ii).

Aqui Hume parece antecipar as características específicas do modelo CAD: se substituirmos a palavra "censura" por "raiva", podemos ver que ele assume que a raiva, o desprezo e a aversão são emoções morais primárias, e cada uma surge após a deliberação sobre a natureza da transgressão em consideração.

Eu concordo com o modelo CAD, mas quero oferecer algumas sugestões sobre como ele deveria ser entendido. Minha primeira questão é terminológica. Rozin et al., seguindo Shweder, referem-se às violações de pureza como

pertencentes à ética da divindade, mas, ironicamente, nenhuma de suas vinhetas nessa categoria envolvem violações religiosas. Isso sugere que o rótulo de Shweder para a categoria é muito estreito. A aversão pode surgir em resposta a assassinatos em massa, especialmente quando corpos são mutilados, e às violações dos hábitos sexuais, mesmo entre pessoas com valores seculares. Assim, a aversão moral não pertence exclusivamente ao domínio da pureza espiritual. Ela pertence, ao contrário, ao que poderia ser chamado de "a ordem natural". Dentro da cosmologia religiosa, a natureza é subsumida pelo divino; a ordem natural é a ordem que foi estabelecida pelos deuses. As violações da natureza divina desencadeiam aversão porque elas são violações contra a natureza, e não vice-versa. Tais violações são primeiramente, e mais fundamentalmente, atos antinaturais. Por isso, direi que a raiva moral é dirigida às transgressões contra a autonomia, e a aversão moral é dirigida às transgressões contra a ordem natural percebida.

O próximo ponto para o qual quero chamar a atenção diz respeito ao desprezo. Rozin et al. descobriram que essa é a resposta emocional dominante às violações de normas comunitárias. As pessoas sentem desprezo para com o adolescente que começa a comer antes dos demais serem servidos. As pessoas também sentem desprezo por aqueles que desrespeitam os idosos, ou que destroem propriedade pública, ou que deixam de agir de acordo com o seu lugar prescrito na sociedade. Existe uma estreita relação entre desprezo e classe. Os ricos têm desprezo para com os pobres, e os pobres têm desprezo para com os ricos. Ao discutir a raiva moral e a aversão moral, eu argumentei que ambas são extensões de emoções não-morais. Agora quero questionar sobre as origens do desprezo. O desprezo é uma emoção básica ou ele emerge de emoções que não estão inicialmente ligadas à moralidade?

Minha resposta a essa questão já foi indicada em minha discussão sobre as emoções combinadas. Lá eu sugeri que o desprezo é uma mistura de raiva e de aversão. Sendo assim, isso parece difícil de conciliar com a descoberta de Rozin et al., de que o desprezo é desencadeado por um tipo fundamentalmente diferente de transgressão moral de raiva e aversão. Em uma análise mais detalhada, contudo, o relato sobre a combinação realmente encontra confirmação em seu modelo. A raiva

é dirigida às transgressões cometidas contra pessoas; a aversão é dirigida às transgressões cometidas contra a natureza; e o desprezo é dirigido às transgressões contra a comunidade. A comunidade pode ser pensada como um sistema dinâmico organizado de indivíduos. Dentro de uma comunidade, cada pessoa tem um lugar particular e faz certos tipos de contribuições ao todo orgânico. A este respeito, a comunidade é como a natureza. A ordem comum é a ordem natural de uma coletividade humana. Se essa interpretação estiver correta, uma violação contra a comunidade deveria desencadear raiva, porque é uma violação contra as pessoas, e aversão, porque é uma violação contra a ordem natural. A hipótese de que o desprezo é uma combinação de raiva e de aversão explica o motivo pelo qual o desprezo é desencadeado por violações contra a comunidade.

A proposta da combinação enfrenta três objeções. Em primeiro lugar, pode-se argumentar que o desprezo parece ser mais do que uma combinação de raiva e de aversão. Uma característica central do desprezo é um sentimento de superioridade. Olhamos para pessoas a quem temos desprezo através de uma posição mais elevada. Sentimentos de superioridade não estão geralmente associados com a raiva e aversão. Se o desprezo fosse uma combinação, ele não poderia ter características emergentes.

Essa objeção pode ser respondida ao refletirmos ainda mais sobre a natureza da raiva e da aversão. Consideramos as pessoas repugnantes como defeituosas, ou animais. Não nos sentimos necessariamente superiores, mas apenas nos sentimos repelidos. Aqueles que provocam nossa raiva são considerados como insensíveis a nós enquanto indivíduos. Agora, considere o que acontece quando pessoas repugnantes desrespeitam você. Nesse caso, não parece mais possível simplesmente evitá-las, pois elas invadiram seu espaço. Ser ameaçado por uma pessoa repugnante é diferente de ser ameaçado por um igual. Ela chama nossa atenção para a diferença de *status* social e instiga um sentimento de superioridade. Quando uma pessoa apenas nos repugna, olhamos para longe; mas quando elas causam aversão e raiva em nós, não podemos desviar o olhar, por isso olhamos com desprezo para elas, repelindo-as, como elas são, por sua animalidade. Não

estou sugerindo que pensamos explicitamente desse modo. A arrogância moral é desdenhosa em seu próprio direito. A sugestão é a de que as violações contra a comunidade normalmente envolvem fatores que promovem automaticamente uma espécie de presunção. Qualquer um que nos deixe irritados e repugnados estará mais propenso a nos fazer sentir superiores também. Em vista disso, poderia ser melhor descrever o desprezo como uma combinação de toda a raiva, aversão, e superioridade (assumindo que exista uma emoção de superioridade).

Uma segunda razão para resistir à análise da combinação é que o desprezo parece ter a sua própria expressão facial. Ekman e Friesen (1986) argumentaram que o desprezo está associado à dobra unilateral do lábio, e que essa expressão é reconhecida entre as culturas. Se o desprezo tem uma expressão facial universal, então é provável que ele seja uma emoção básica, pelo menos é isso que Ekman e Friesen querem nos fazer acreditar. Essa objeção, contudo, pode ser combatida. Por um lado, as pesquisas existentes sobre a expressão de desprezo são muito controversas. Wagner (2000) argumentou que os estudos com os quais o desprezo é reconhecido apresentam falhas graves. Nesses estudos, os indivíduos são forçados a escolher entre etiquetas de emoção, incluindo a palavra "desprezo" ou seus sinônimos. Quando a escolha é aberta, ao invés de forçada, os indivíduos não etiquetam a imagem do rosto como "desprezo". Por outro lado, se existe uma expressão de desprezo, ela poderia ser uma expressão mista derivada de expressões de componentes mais básicos. O desprezo está associado com um olhar fixo e com a sobrancelha reta ou abaixada, os quais também são consistentes com a raiva. A dobra unilateral do lábio pode ser derivada da expressão de aversão: a aversão envolve um enrugamento ascendente do nariz, o que normalmente puxa os lábios para cima também. Mas por que a dobra unilateral do lábio? Uma possibilidade é que a aversão e a raiva demandam lugares diferentes da boca. Na raiva, os lábios estão ou normalmente fechados, com o lábio externo superior inchado, ou abertos como em caretas em "forma de caixa" para revelar os dentes. Talvez a presença simultânea da exibição da aversão facial com as exibições de raiva leve a uma ondulação parcial do lábio. A assimetria do lábio por desprezo poderia ser derivada do fato de que as expressões faciais de raiva e aversão estão

sobre o controle de diferentes hemisférios do cérebro. Existe alguma evidência de que as emoções de aproximação estão mais associadas ao hemisfério direito, enquanto que as emoções de afastamento estão mais associadas ao esquerdo (Davidson e Irwin, 1999). Se a raiva é uma emoção de aproximação (a agressão é um tipo de comportamento de aproximação), e a aversão é uma emoção de afastamento, pode-se esperar que a presença simultânea dessas emoções seja manifestada de modo assimétrico no rosto. Outra possibilidade é que a dobra unilateral do lábio envolva a introdução de uma terceira expressão além da raiva e da aversão. Em particular, a dobra do lábio poderia sugerir uma interação com os mesmos grupos musculares utilizados no sorriso. Por que sorrir? Muitas vezes, pensamos que o sorriso é um sinal de felicidade, mas sorrisos e risos são também conhecidos pelo sinal dominante (Provine, 2000). Observe também que há uma estreita relação entre o desprezo e a zombaria, ou escárnio, as quais utilizam o sorriso de uma forma humilhante. A dobra unilateral do lábio é um tipo de zombaria, e os escárnios estão associados com sentimentos de superioridade. É possível que a combinação de raiva e aversão tenha uma propensão a ocorrer de modo concomitante com a dominância, pois quando sentimos aversão por certas pessoas, nós as vemos como animais, e quando sentimos raiva, em geral, queremos agredi-las. Além disso, muitas normas comunitárias envolvem classes sociais, de modo que os sentimentos de superioridade surgirão naturalmente. Se isso estiver correto, então a combinação de desprezo pode normalmente envolver uma emoção de dominância, e essa emoção pode contribuir para a expressão. O sorriso que sinaliza a dominância pode ser assimétrico, pois ele pode ser controlado pelo hemisfério que abriga as emoções negativas. Tudo isso é amplamente especulativo, é claro. O ponto principal é que uma expressão facial universal de desprezo é compatível com a hipótese de que o desprezo é uma emoção composta.

A análise da combinação de desprezo pode ser contestada em uma terceira via. O desprezo está frequentemente associado com violações de posições nas sociedades organizadas de modo hierárquico. A posição social não é, contudo, algo original de nossa espécie. A maioria dos mamíferos sociais vive em grupos

organizados de modo hierárquico. Qualquer emoção associada a violações de posições é provavelmente bastante antiga. A sugestão de que o desprezo é a combinação de emoções mais básicas parece entrar em conflito com a observação de que a posição social tem uma longa história filogenética. Poderia talvez ter surgido como resultado da pressão da seleção natural para o surgimento de uma emoção moral diferenciada capaz de manter as pessoas estratificadas.

Farei três observações em resposta a isso. Em primeiro lugar, se o desprezo é uma combinação de raiva e de aversão, pode-se esperar que ele surja em qualquer espécie que possua essas duas emoções mais básicas. A posição social não exige a evolução de quaisquer novas emoções se as emoções que evoluíram anteriormente podem fazer o trabalho. Em segundo lugar, a sugestão de que o desprezo é uma combinação de raiva e de aversão é consistente com a sugestão de que o desprezo é algo evoluído. Podemos ter mecanismos evoluídos que facilitam e promovem a combinação de certas emoções. Em terceiro lugar, não existem razões para assumir que as violações de posições provocam a mesma emoção em nós do mesmo modo como ocorre em outros mamíferos. Quando há uma disputa animal por posição, o animal dominante reage com raiva (ou, pelo menos, com um sentimento de agressão), ao invés de reagir com desprezo. Se o animal rivaliza sem sucesso, então ele pode experimentar medo e submissão. Eu suspeito que os animais não-humanos não possuem uma concepção suficientemente rica de ordem social para interpretar violações de posições como desprezíveis da maneira que sugeri. Animais não-humanos possuem regras relativas à posição, mas elas são aplicadas através do uso de um repertório emocional mais restrito. Dito de outro modo, para os animais não-humanos, as violações de posições são mais parecidas com violações de autonomia.

Concluo que existem duas classes fundamentais de emoções morais direcionadas aos outros (a raiva moral e a aversão moral) e uma classe derivada (o desprezo moral). Todas essas emoções, a meu ver, são emoções não-básicas. Elas são extensões naturais de emoções básicas que surgem em contextos de transgressão. A emoção provocada por uma transgressão será determinada pela sobreposição entre essa transgressão e situações que provocam emoções em

contextos não-morais. Violações de direitos são ameaças às pessoas; costumes sexuais envolvem ameaças à ordem natural; e violações de posições são ameaças à ordem natural das pessoas.

2.2.2 Emoções Morais Reflexivas

Desprezo, raiva e aversão são emoções de culpa direcionadas aos outros. Quando culpamos a nós mesmos, no entanto, seguem-se diferentes emoções. Os três sistemas éticos de Shweder e o modelo CAD sugerem que podem existir diferentes tipos de autculpa. As emoções que experimentamos quando violamos regras podem depender da natureza de tais regras.

Considere o que acontece quando violamos uma regra referente à justiça ou aos direitos. Você acabou de roubar alguma coisa; você acabou de ferir alguém; você tomou mais do que sua parte justa. Eu suponho que a resposta emocional padrão nesses casos é a culpa, especialmente quando a vítima é um membro do grupo com o qual você está afiliado. Se você tomar algo de um grupo com o qual você não tem qualquer relação ou mesmo prejudicar um inimigo, você pode não sentir qualquer tipo de autculpa. Mas quando você prejudicar um membro de dentro de seu grupo, é provável que você sinta culpa; Baumeister et al. (1994) mostraram que a culpa está mais frequentemente associada com ações que ameaçam os indivíduos a quem se tem alguma relação de apego. A culpa pode representar a preocupação expressa por: "eu violei uma regra de autonomia contra um membro de um grupo com o qual eu sinto uma ligação".

Uma emoção diferente é sentida quando violamos regras que dizem respeito à ordem natural. Os violadores de tabus sexuais e de regras de pureza religiosa estão mais propensos, eu poderia prever, a sentir vergonha. Às vezes é sugerido que a culpa é dirigida ao ato, enquanto a vergonha é dirigida à pessoa. Uma pessoa culpada pode sentir que suas ações eram condenáveis sem se sentir uma pessoa má. Uma pessoa que sente vergonha normalmente se sentirá suja, indigna ou corrupta. Isso faz perfeito sentido dentro do contexto de transgressões contra a natureza. A pessoa que se envolve em atos sexuais ultrajantes ou a pessoa que

mutila outro ser humano pode se sentir como uma pessoa sub-humana. A pessoa envergonhada se sente horrível e impura.

O modelo que tenho proposto prevê que as pessoas que investem contra a comunidade sentirão uma combinação de culpa e de vergonha. Não temos nenhuma palavra para essa emoção, mas eu suspeito que ela, de fato, existe. Pense em como você se sentiria se maltratasse uma pessoa idosa (uma das normas da comunidade que permanecem autônomas em nossa cultura). Nesse caso, tanto a vergonha quanto a culpa parecem emoções apropriadas. Se alguém descobrisse seu crime, você iria querer se esconder como se estivesse envergonhado. Mas você também poderia refletir sobre o mal que causou dirigindo-se à ação, como se ela fosse a culpada.

Para avaliar essa extensão do modelo CAD, eu realizei um estudo no qual os sujeitos liam cenários morais em segunda pessoa. Os cenários incluíam crimes contra pessoas ("Suponha que tu tomaste algo de alguém e nunca devolveste"), crimes contra a natureza ("Suponha que, em um momento de fraqueza, tu permita que uma pessoa que é realmente muito velha te beije romanticamente"), e crimes contra a comunidade ("Suponha que tu, sem pensar, começas a comer teu jantar em família antes que os demais tenham se sentado à mesa"). Em cada caso, os sujeitos são questionados se eles estariam mais propensos a experimentar culpa ou vergonha. Para os crimes contra a natureza, os sujeitos predominantemente escolheram a vergonha, e, para os crimes contra as pessoas, eles escolheram esmagadoramente a culpa. Contudo, não houve diferença estatisticamente significativa entre culpa e vergonha para os crimes contra a comunidade.

Eu disse que emoções morais de culpa atribuídas aos outros são emoções não-básicas. O mesmo é verdade sobre as emoções morais de culpa atribuídas a si mesmo. Como mencionado anteriormente, eu penso que a culpa é uma extensão da tristeza. Ambas estão associadas à testa franzida e à sensação de ser oprimido. É digno de nota que a culpa excessiva é um sintoma de diagnóstico de depressão clínica (*American Psychiatric Association, 1994*). Por outro lado, como observado no capítulo 1, os psicopatas têm uma deficiência de tristeza e uma deficiência de culpa. A hipótese de que a culpa é uma extensão da tristeza prevê a concomitância

de tais comorbidades. Mas agora precisamos de uma explicação do porquê a culpa emerge da tristeza, e não de outra emoção. A resposta emerge de minha extensão do modelo CAD. A culpa está associada com violações de regras morais concernentes aos direitos e à justiça. Ela está especialmente associada com os danos autocausados. Mais especificamente, ela está associada com os danos contra pessoas por quem se sente um sentimento de apego. O apego está intimamente relacionado à tristeza. O mais poderoso desencadeador de tristeza é a perda de um ente querido. Sentimos, geralmente, uma sensação de perda ou de tristeza quando os outros são prejudicados, independentemente do modo como o dano ocorreu. Assim, se eu prejudicar alguém com quem me importo, eu provavelmente me sentirei triste com o sofrimento dessa pessoa. Além disso, se eu prejudicar alguém de meu próprio grupo, os demais membros do grupo ficarão desapontados comigo ou com raiva de mim; eles retirarão o afeto que nutriam por mim. No capítulo 1, eu sugeri que isso poderia desempenhar um importante papel no desenvolvimento moral. Se os cuidadores retiram o amor de crianças que se comportam mal, essas crianças vão se sentir tristes, pois a retirada do amor é um tipo de perda. Esse sentimento de tristeza, originalmente associado com o carinho perdido dos cuidadores, passa a ser associado com os tipos de conduta que provocam uma perda. Se a recusa de compartilhar algo com sua irmã deixa seus pais furiosos, então você ficará triste. A tristeza será agravada se seus pais chamarem a sua atenção para o fato de que seu comportamento fez sua irmã chorar. Inicialmente, a tristeza é uma resposta ao afeto perdido e à angústia sofrida, mas, por fim, ela é transferida para as ações que levam a essas consequências. É o próprio pensamento de ferir, roubar ou se recusar a compartilhar que nos torna tristes, e, não, o dano suposto ou a punição que ele trará. Uma vez que essa transferência ocorra, podemos dizer que temos a emoção de culpa, ao contrário do mero pesar, que é sentir-se mal a respeito de uma ação à luz de suas consequências. A culpa é a tristeza que foi calibrada para atos que prejudicam as pessoas com quem nos importamos.

Creio que a vergonha é derivada do constrangimento. O constrangimento surge quando recebemos a atenção indesejada dos outros. É um sentimento

negativo, mas frequentemente negativo de uma forma despreocupada. Rimos com o constrangimento. A vergonha, contudo, é algo mais grave. Rir é uma maneira de manter as aparências quando houve um pequeno *passo em falso* [*faux pas*] ou uma estratégia profilática contra uma opinião severa em um fórum público. A vergonha surge quando fazemos algo mais sério, algo de que nenhum riso poderia nos salvar. A vergonha moral é uma espécie de constrangimento aversivo que foi calibrado para normas que dizem respeito à ordem natural. As violações de tais normas provocam confiavelmente intenso constrangimento, pois elas nos tornam repugnantes aos olhos dos outros. A reação desencadeada por um deslize menor é ampliada quando nos envolvemos em comportamentos que são considerados como animais ou não-naturais. Quando expelimos gases intestinais através do ânus, nos sentimos embaraçados, mas sentimos vergonha quando fantasiamos sobre atos sexuais desfavorecidos. O constrangimento pode ser uma emoção básica (Miller, 1996). A vergonha é provavelmente uma emoção não-básica, pois é uma forma de constrangimento que poderia surgir somente depois de nossa espécie ter começado a estender metaforicamente a classe de impurezas físicas para o domínio social e comportamental. A vergonha surge inevitavelmente após uma cultura começar a etiquetar certos atos como não-naturais ou desviantes. A vergonha também surge quando as culturas desenvolvem sistemas rígidos de estratificação social. Membros de tais culturas, às vezes, sentem um constrangimento intenso e aversivo na presença de indivíduos hierarquicamente superiores, pois tais encontros chamam a atenção para sua inferioridade.

Na história que tenho contado, as emoções morais reativas e reflexivas se dividem em duas categorias fundamentais e uma categoria mista. Todas essas emoções morais são não-básicas. Há também uma classe de emoções morais que se situa em algum lugar entre as emoções reativas e reflexivas. Considere casos em que a transgressão tenha sido cometida por alguém extremamente próximo a você, como um pai ou um filho. Nesse caso, a resposta moral é dirigida ao outro, mas também dirigida a si mesmo. Você vê a conduta dessa pessoa como refletindo algo sobre você. Quando isso ocorre, creio que, frequentemente, experimentamos emoções que são semelhantes às emoções de autoculpa. Se seu filho comete um

crime contra a justiça ou direitos, você pode se sentir magoado. Mágoa e culpa se sobrepõem, pois elas são ambas derivadas da tristeza. Se seu filho comete um ato não-natural, você pode se sentir envergonhado. Se seu filho comete uma violação contra a comunidade, você pode sentir uma combinação de mágoa e de vergonha – uma sensação melancólica de desgraça. O pai de um membro de uma gangue, cujo filho destruiu uma propriedade pública, poderia sentir uma dor cortante de desapontamento, enquanto ficaria ruborizado diante dos vizinhos descontentes. Todas essas amplificações do modelo CAD estão sumarizadas na Tabela 2.1.

Tabela 2.1. Emoções de culpa em função de transgressão e transgressor

		Transgressor		
		Eu	Um ente querido	
Transgressão	Contra as Pessoas	Raiva	Culpa	Mágoa
	Contra a Comunidade	Desprezo	Culpa/Vergonha	Magoado/Envergonhado
	Contra a Natureza	Aversão	Vergonha	Envergonhado

2.2.3 Emoções Positivas e Moralidade

Não terá escapado à atenção o fato de que tenho me concentrado apenas em emoções negativas. Cada emoção na Tabela 2.1 é negativa. As emoções positivas, contudo, também são importantes para a moralidade. Temos emoções reativas positivas que são direcionadas àqueles que fazem boas ações, e experimentamos emoções positivas autodirigidas quando fazemos boas ações para nós mesmos. Em contraste com emoções de culpa, esses bons sentimentos podem ser descritos como emoções de louvor.

Emoções de louvor não têm sido amplamente estudadas, e elas parecem se comportar de forma um pouco diferente das emoções de culpa. Não parece haver uma emoção positiva diferente para cada um dos três tipos de regras éticas que

tenho discutido. Uma razão para isso é o simples fato de que esses domínios éticos são definidos pelos tipos de transgressão, e não está claro qual o papel que as emoções positivas teriam de exercer. Culpamos alguém por roubar, mas não oferecemos uma medalha quando ele se abstém de roubar. Não esbanjamos louvores aos não-pedófilos pelo bom comportamento. Em outras palavras, temos a tendência de *esperar* que as pessoas se comportem moralmente.

Isso sugere uma assimetria geral entre emoções positivas e negativas na moralidade. O comportamento desejável está mais propenso a ser moldado pelas emoções negativas do que pelas positivas, e, como resultado, somos mais generosos com a culpa do que com o louvor. Isso pode ser demonstrado experimentalmente. Por exemplo, Malle e Bennett (2002) observaram que as pessoas estavam mais propensas a culpar as más intenções do que a louvar as boas intenções, e Knobe (2003) constatou que os sujeitos estavam mais propensos a considerar as ações como intencionais quando elas produziram efeitos colaterais ruins do que quando elas produziram efeitos colaterais bons. A origem dessa assimetria ainda não foi estabelecida. Uma possibilidade é que as sociedades têm um interesse maior em eliminar o mau comportamento do que em promover um comportamento especialmente bom. E quando se trata de eliminar o mau comportamento, a punição pode ser mais eficaz do que o louvor. Não está inteiramente claro, contudo, que a punição deva ser mais eficaz que o louvor. Pode ser um fato bruto sobre nós que os bons sentimentos sejam passageiros (especialmente quando comparados às tentações ruins), enquanto os maus sentimentos estariam mais propensos a deixar uma marca duradoura na memória e na motivação.

Se isso estiver correto, a idoneidade comum ou a conformidade às regras morais depende muito mais de emoções negativas do que de emoções positivas. Mas as emoções positivas podem desempenhar um papel na promoção de comportamentos que vai muito além da idoneidade ordinária. As emoções positivas podem nos levar a fazer coisas que são generosas, nobres ou de mesmo tipo. Tem sido demonstrado que os afetos positivos podem levar a comportamentos pró-sociais. Isen e Levin (1972) descobriram que as pessoas ficam amplamente mais

dispostas a ajudar se um bom humor é induzido antes de terem uma oportunidade de ajudar. Uma possível explicação para isso é que as emoções positivas nos tornam mais extrovertidos, e isso aumenta a nossa consciência sobre as necessidades dos outros. Outra explicação intrigante é que quando uma emoção positiva é induzida, queremos mantê-la, e o comportamento de ajudar pode colaborar para isso. Nessa explicação, ajudar é interpretado como uma fonte de afeto positivo. Nos sentimos bem quando fazemos o bem. Isso pode desempenhar um papel importante na educação moral. Imagine um cenário onde você tem uma oportunidade de salvar alguém em necessidade. Um motivo para fazer isso é que nos sentiremos culpados se não o ajudarmos, mas a imensa literatura sobre os "efeitos do espectador" sugere que as pessoas estão muito mais dispostas a assistir passivamente enquanto as demais clamam por ajuda. Mas aqueles que intervêm podem fazer isso por causa da recompensa antecipada pelo comportamento nobre. Eles podem antecipar sentimentos de felicidade quando eles ajudam. O suporte para isso vem de um estudo realizado por Valdesolo e DeSteno (2006). Eles solicitaram aos sujeitos que considerassem um caso padrão do [problema do] bonde, no qual um bonde está acelerando em direção a cinco pessoas, mas eles podem salvar essas vidas empurrando alguém para o caminho do bonde. Como ocorre com a maior parte dos sujeitos em experimentos do bonde, a esmagadora maioria pensou que era errado empurrar uma pessoa para frente do bonde, matando uma e salvando cinco. Mas alguns sujeitos consideraram esse cenário após assistir um videoclipe engraçado no *Saturday Night Live*. Esses sujeitos mostraram-se três vezes mais propensos a julgar que era apropriado empurrar uma pessoa no caminho do bonde. As emoções positivas podem ter alterado a resposta de duas maneiras: elas podem ter reduzido as emoções negativas associadas a matar alguém, e elas podem ter aumentado as emoções positivas associadas com a recompensa antecipada de salvar cinco vidas. Essa interpretação é especulativa, mas, se estiver correta, ela sugere que emoções positivas podem desempenhar um papel dominante em "ajudar", enquanto as emoções negativas desempenham um papel dominante em simplesmente em não "prejudicar".

Tabela 2.2. Emoções de louvor como uma função de quem é agente e de quem é paciente

Agente		Paciente	Emoção
Outro	Outro	Outro	Admiração
Outro	Eu	Eu	Gratidão
Eu	Outro	Outro	Gratificação
Eu	Eu	Eu	Dignidade

Até agora eu tenho falado muito genericamente sobre as emoções positivas, e disse realmente muito pouco sobre quais emoções específicas estão envolvidas. O campo da psicologia moral precisa de uma taxonomia de paixões positivas, e me esforçarei em oferecer uma primeira tentativa aqui. Eu já observei que as emoções positivas podem não corresponder aos três domínios éticos discutidos pelos proponentes do modelo CAD. Mas uma dimensão que eu discuti nos parágrafos anteriores – eu *versus* outro – pode ser importante. No caso das boas ações, não importa quem está fazendo a boa ação, e também parece importar pouco quem se beneficia com a boa ação. Por essa razão, creio que é útil classificar emoções morais positivas ao longo das dimensões de agente e paciente: de quem está agindo e de quem está sofrendo a ação.

Imagine ler sobre alguém que realizou um ato altruísta, tal como doar generosamente para caridade ou ajudar uma pessoa necessitada. Meu palpite é que a resposta emocional padrão seria a admiração. Agora suponha que essa alma caridosa faça algo de bom para você. Quando alguém lhe faz uma boa ação, você provavelmente sentirá gratidão. Assim, emoções de louvor moral direcionadas aos outros incluem a admiração e a gratidão, e a diferença depende apenas de quem se beneficia com o ato. Considere agora casos nos quais você é o autor de uma boa

ação. Se você faz algo muito gentil, você poderá sentir um tipo de gratificação. É bom fazer o bem. Ou considere um caso no qual suas boas ações são dirigidas não para outra pessoa, mas para você mesmo. Talvez você resista a uma tentação especialmente forte, ou talvez você, por si mesmo, elimine um vício ou altere sua maneira de trabalhar a partir da pobreza, ou adquira uma habilidade que o melhore como pessoa. Em tais casos, você poderia sentir um tipo de autorrespeito ou dignidade. Essas emoções estão sumarizadas na Tabela 2.2.

Não estou sugerindo que essas sejam as únicas emoções positivas que contribuem de algum modo para a moralidade, pois podem haver outras. Haidt (2003) propôs a existência de uma emoção que ele chamou de elevação, que é descrita por ele como uma resposta positiva à beleza moral. Quando uma ação nobre ocorre, nos sentimos elevados por ela. Penso que a elevação é uma emoção intrigante e que merece uma investigação mais aprofundada. Em meu esboço, contudo, a elevação poderia se referir à intensa admiração. Haidt indica que podemos nos sentir elevados por nossas próprias boas ações e pelas boas ações dos outros. Estou inclinado a pensar que o eu e o outro tendem a ser objetos de emoções sutilmente diferentes. Se esse é o caso, a elevação autodirigida pode ser uma forma intensa de gratificação moral, ou talvez uma combinação de gratificação e dignidade (nos sentimos bem fazendo o bem, e isso nos faz sentir melhor sobre nós mesmos).

Além das emoções positivas que eu tenho discutido, a conduta moral pode também beneficiar os sentimentos positivos que recebemos através da filiação social. Essas emoções promovem a bondade e a conformidade. O bom comportamento capta as boas graças dos membros da comunidade, e o desejo de se sentir conectado pode ser uma fonte de motivação. Na próxima seção, discutirei algumas construções afetivas relacionadas à filiação: cuidado, simpatia e preocupação. Como veremos, algumas dessas construções não são positivamente valenciadas.

As sugestões anteriores precisam de mais apoio experimental para obter qualquer tração. Seria razoável que as emoções morais positivas começassem a receber mais atenção (ver, por exemplo, McCullough et al., 2001, sobre gratidão). O

fato de que elas tenham sido negligenciadas pode apoiar a conjectura de que as emoções negativas possuem um papel desproporcional na conduta moral. No que se segue, continuarei a me concentrar em emoções negativas, mas as emoções positivas não devem ser esquecidas.

2.2.4 Cuidado, Simpatia e Preocupação

O mau comportamento é fácil de explicar. As pessoas agem mal por que o mau comportamento muitas vezes carrega uma recompensa imediata. Alcançamos bens através do roubo e prazer através de atos sexuais condenáveis. O mais profundo mistério é o bom comportamento. Por que temos de obedecer às regras morais? Uma resposta tem a ver com o preço do mau comportamento. Culpa e vergonha são desagradáveis, e o bem-estar é ameaçado quando suportamos a cólera dos outros ou quando evocamos sua aversão. Outra resposta, já mencionada antes, é que o bom comportamento forja afiliações valiosas. É benéfico ser confiável. Mas ambas as respostas são limitadas, pois presumem que nos comportamos de acordo com o autointeresse. Às vezes, fazemos coisas boas por que nos preocupamos com outras pessoas, mesmo quando não há recompensas previstas. Esse *insight* é o coração da filosofia moral de Hume, e ele aparece também de modo proeminente no trabalho de Adam Smith (1759).

Hume enfatizou a simpatia. A simpatia nos leva a cuidar das pessoas necessitadas ou em sofrimento. A simpatia pode ser definida como uma resposta emocional negativa ao sofrimento dos outros, pois uma pessoa simpática sente-se mal pelo fato de você se sentir mal. Não está empiricamente claro se a simpatia sempre se refere à mesma emoção subjacente. Se assim for, provavelmente ela é um tipo de tristeza. Nesse sentido, a simpatia é diferente da empatia, a qual se manifesta de modo diferente de contexto para contexto (compare com Darwall, 1998). Na empatia, sentimos a mesma emoção que alguém está sentindo; nos colocamos no lugar da outra pessoa. Se você está com medo, uma pessoa empática vai sentir medo também. Respostas empáticas podem ser medidas, por exemplo, através da comparação da ativação cerebral em indivíduos enquanto eles

observam outras pessoas em perigo. Nas respostas empáticas, as áreas do cérebro que são ativadas no observador são aquelas que correspondem ao tipo de perigo que está sendo observado. Por exemplo, durante as respostas empáticas observadas com relação à dor, os centros de dor do cérebro estavam ativos (Singer et al., 2004). Curiosamente, nestes estudos o observador parece adquirir o componente afetivo da dor, mas não o componente sensorial; ele adquire o sofrimento da dor sem necessariamente imaginar a sensação que produz esse sofrimento. A simpatia não tem sido estudada através da neuroimagem, pois a maioria das pessoas tende a criar empatia, mas poderíamos esperar de uma pessoa simpática que ela mostrasse sinais de afeto negativo ao testemunhar a dor, mas, necessariamente, deveria ser o mesmo tipo de afeto negativo que surge durante a experiência da dor. Uma pessoa que congenitamente carece de dor, devido a uma doença genética, poderia simpatizar com alguém que sente dor, mas não poderia sentir empatia. Ou, para mudar de exemplos, os pais poderiam simpatizar com o medo de seu filho de que existe um monstro no armário, mas eles mesmos poderiam não sentir qualquer tipo de medo.

Simpatia e empatia devem ser diferenciadas da angústia indireta. Os bebês, por vezes, choram quando ouvem o choro de outros bebês, mas isso não significa que eles tenham simpatia ou empatia. Eles estão angustiados *pelo* choro dos outros bebês, mas não estão angustiados *por* aquele chorar ou *sobre* os eventos que levaram o outro bebê a chorar. Quando simpatizamos ou sentimos empatia, nosso descontentamento é bloqueado pelo conhecimento dos motivos pelos quais a outra pessoa está sofrendo. Nos casos de angústia indireta, nosso descontentamento é bloqueado pelos sintomas da outra pessoa que sofre, mas não pelo sofrimento em si mesmo. Um adulto, ao experimentar uma angústia indireta, deixará a sala para evitar o contato com uma pessoa que está sofrendo, e assim a angústia indireta irá diminuir. Um adulto, ao experimentar simpatia ou a empatia, torna-se mais propenso a oferecer ajuda. Deixar a sala não elimina a fonte de infelicidade; isso apenas torna a fonte menos saliente. Psicologicamente, isso significa que as respostas simpáticas e empáticas exigem uma atribuição de sofrimento ao invés da mera observação do sofrimento. Ao reconhecer que a outra

pessoa está sofrendo, as pessoas empáticas ou simpáticas se sentirão compelidas a se engajarem em comportamentos pró-sociais.

Simpatia e empatia interagem de maneiras importantes com o cuidado. O termo "cuidado" é ambíguo. Ele pode nomear uma atitude ou um comportamento ("cuidar"). Esses dois sentidos relacionam-se entre si: se você se importa com uma pessoa, é provável que você cuide dela. Essa ligação pode normalmente ser mediada pela simpatia ou empatia. Se você se preocupa com uma pessoa, você prestará atenção nas necessidades dela. Se ela está necessitada e você simpatizar ou sentir empatia por ela, então esses sentimentos vão impulsioná-lo a ajudá-la. O cuidado nem sempre provoca simpatia ou empatia, pois você também pode cuidar da arte ou cuidar do meio ambiente. Cuidar é uma disposição para prestar atenção em algo de uma forma que o motiva tanto a interagir com essa coisa, como a comportamentos que promovem o florescimento dessa coisa. A simpatia e a empatia são como que veículos para o cuidar quando o cuidado é dirigido a outros seres humanos. Simpatia e empatia podem também promover o cuidado. Sentir-se mal por causa da angústia do outro pode orientar você em direção a essa pessoa, de modo que isso faça você se preocupar com o bem-estar dela.

Assim como a empatia, o cuidado não é qualquer emoção específica. Ele se manifesta de modo diferente em diferentes ocasiões, dependendo do objeto com o qual você se importa e do estado atual desse objeto. A este respeito, o cuidado pode ser diferente da preocupação. Eu creio que a preocupação é uma espécie de medo. Estar preocupado com alguém é preocupar-se com o seu bem-estar. O cuidado está relacionado com a preocupação, pois nos preocupamos com aqueles que amamos. A preocupação surge quando algo ou alguém que amamos está em perigo. Assim como a simpatia e a empatia, a preocupação pode também ajudar a vincular o cuidado-sobre com o cuidado-para. Se você se preocupa com alguém, e essa pessoa está em perigo, você ficará preocupado; e se você ficar preocupado, você poderá tentar cuidar dessa pessoa. Mas, ao contrário da simpatia e da empatia, podemos estar preocupados com alguém que não está experimentando por si mesmo quaisquer emoções negativas. Se seu amigo é um toxicod dependente

feliz, você pode não simpatizar ou sentir empatia por ele, mas você pode ficar preocupado.

Em suma, há uma série de maneiras com as quais podemos reagir negativamente quando coisas ruins acontecem a outras pessoas. Podemos experimentar angústia indireta, caso em que meramente capturamos seus sentimentos do mesmo modo que pegamos um resfriado (nenhuma atribuição é necessária). Podemos experimentar a simpatia, caso em que nos sentimos mal por elas estarem passando por um momento difícil. Podemos experimentar a empatia, caso em que nos sentimos mal por elas, colocando-nos emocionalmente em seus lugares. Ou podemos experimentar a preocupação, caso em que vamos nos preocupar com o bem-estar delas mesmo se elas não conseguem compreender que, de fato, alguma ruim aconteceu. Todas as angústias, exceto as indiretas, são promovidas tanto pelo cuidado quanto pelo comportamento cuidadoso.

2.2.5 Sentimentos Morais

Eu afirmei que o cuidado não é uma emoção, mas, sim, uma disposição para experimentar diferentes emoções. Se você se preocupa com alguém, você se deleitará com suas conquistas e ficará angustiado com o seu sofrimento. Usarei o termo "sentimento" para me referir a uma disposição emocional (ver Prinz, 2004). O cuidado é um sentimento. Outros exemplos incluem gostar, não gostar, amar e odiar. Se você gosta de alguma coisa, você ficará feliz quando ela for alcançada ou quando estiver próxima, e triste quando ela for perdida ou estiver longe. Gostar pode também tornar você invejoso (ela tem o que eu gosto!), assustado (posso perder o que eu gosto!), ou esperançoso (posso conseguir o que gosto!).

Ao dizer que sentimentos são disposições, não estou dizendo que eles não são reais, enquanto estados fisicamente implementados pela mente. De acordo com o modo como usarei o termo, uma disposição psicológica é um estado de repouso de um organismo que pode se manifestar como um estado ocorrente. A distinção repouso/ocorrente é comumente utilizada em filosofia. Em termos psicológicos, as disposições psicológicas podem ser usualmente identificadas com

codificações da memória de longo prazo que podem ser recuperadas pela memória operacional, e serem mantidas lá durante o processamento mental explícito. Em termos neurocomputacionais, as disposições são usualmente identificadas com as conexões ponderadas entre neurônios que podem ativar os conjuntos de neurônios que elas conectam. Todas essas formas de falar parecem capturar a ideia básica de que as disposições são estados internos que nem sempre participam do processamento de informações, mas que podem se tornar contribuidoras ativas sob circunstâncias adequadas. Um sentimento é uma disposição cujas manifestações ocorrentes (ou codificações da memória operacional, ou padrões de ativação neural) são emoções.

Sentimentos, muitas vezes, manifestam-se como diferentes emoções em diferentes ocasiões. Se você ama tiramisu, você ficará feliz, quando o encontrar no menu do jantar, e desolado, quando você descobrir que o restaurante acaba de ficar sem tiramisu. Em qualquer ocasião, a emoção específica que você experimenta é uma manifestação de seu sentimento. Alguns outros sentimentos se manifestam através de emoções iguais ou semelhantes em cada ocasião. Considere o caso das fobias. Se você tem medo de voar, você não experimentará esse medo a cada momento de sua vida. O medo é um estado de repouso que se tornará uma experiência ocorrente quando, por exemplo, você embarca em um avião.

A semântica dos sentimentos é um pouco diferente da semântica de suas emoções constituintes. As emoções, como argumentei, representam preocupações. Preocupações são relações organismo-ambiente que incidem sobre o bem-estar. Essas relações podem ser caracterizadas sem mencionar as emoções que elas nos causam. Em outras palavras, preocupações não são qualidades secundárias (Prinz, 2004). A raiva representa a ofensa, não a propriedade de estar irritado. O medo representa o perigo, não a propriedade de ser assustador. A tristeza representa a perda, não a propriedade de estar depressivo. A felicidade representa a conquista ou a realização, não a propriedade de ser agradável. Algo poderia ser ofensivo, perigoso, uma perda, ou uma conquista, sem nos deixar com raiva, com medo, tristes ou felizes. As emoções são definidas pelas preocupações que elas tem a

função de detectar, e essas preocupações podem ser definidas em termos não-emocionais.

Agora, contraste isso com os sentimentos. Gostar nos fornece um bom exemplo. Gostar é parcialmente constituído por uma disposição para a felicidade, mas gostar não representa a conquista ou a realização. Gostar é também uma disposição para a tristeza, como quando você perde o objeto que gosta. Mas gostar não representa a perda, pois gostar não se manifesta como uma emoção única, não podendo representar qualquer preocupação singular. Ao invés disso, o gostar aprisiona as coisas que acontecem para instigar essas diferentes emoções em você. Em uma palavra, gostar representa a propriedade de ser amável. A amabilidade é, plausivelmente, uma qualidade secundária. Ela pode ser definida por referência ao sentimento de gostar. Os traços que contam como amável dependem do que estamos dispostos a gostar. Assim, os sentimentos podem representar qualidades secundárias mesmo que suas emoções componentes não representem. Isso pode ser verdadeiro até mesmo para sentimentos que se manifestam através da mesma emoção em cada ocasião. Considere o caso das fobias novamente. Se você tem uma fobia de voar, sua fobia sempre se manifestará provavelmente como medo; nós a chamamos de medo de voar. O medo representa perigo, mas, sem dúvida, a fobia em si mesma representa a temeridade. A temeridade é uma propriedade secundária. Algo é assustador se ele provoca medo. Se você disser a alguém que você tem medo de voar, você está na verdade dizendo que voar é assustador para você; voar tem a qualidade secundária de assustar você. Eu menciono esse tópico sobre fobias aqui como algo à parte; quando eu falar sobre sentimentos, eu geralmente tomarei por foco os casos em que as manifestações emocionais variam de acordo com os contextos.

Até agora, eu tenho me concentrado nos sentimentos não-morais. Um sentimento moral é aquele que se manifesta em emoções morais. Gostar e não gostar não são sentimentos morais. Como eu defino o termo, cuidar também não é um sentimento moral, ainda que ele possa promover comportamentos pró-sociais, e que o comportamento pró-social seja moralmente valorizado. Cuidar não

costuma manifestar-se em emoções morais, e podemos cuidar de coisas e de pessoas sem qualquer significado moral.

Há uma variedade de diferentes sentimentos morais. Alguns deles são direcionados às pessoas. Considere o ressentimento. O ressentimento não é uma emoção, mas, sim, um sentimento. É uma disposição para sentir algo como o rancor, a raiva ou o desprezo. O ressentimento, contudo, pode ser algo específico do domínio moral. Geralmente nos ressentimos com aqueles que violam as regras morais. À primeira vista, pode parecer que são instâncias não-morais de ressentimento. Podemos nos ressentir com aqueles que possuem mais do que nós, mas, aqui, o ressentimento pode ser entendido como pertencente à injustiça. Nos ressentimos pelo fato de que há uma distribuição desigual de bens, ou também pelo fato de que algumas pessoas têm o que não merecem, e isso, por si mesmo, é um conceito moral.

Outro sentimento moral dirigido às pessoas é o fenômeno subjacente de guardar rancor. Rancores estão mais ligados à ação do que ao ressentimento. Guardamos rancor do que as pessoas fizeram, mas não do que elas possuem. Rancores são normalmente dirigidos contra aqueles que violaram regras de autonomia. Rancores são uma espécie de raiva moral disposicional.

Alguns sentimentos morais possuem escopos mais abrangentes. Eles se aplicam não apenas às pessoas, mas, também, às formas de conduta. Pode-se ter um sentimento moral negativo em relação à mentira, por exemplo. No próximo capítulo, argumentarei que esses sentimentos são a espinha dorsal da moralidade. Alguns emocionistas chamam a si mesmos de sentimentalistas, mas poucos realizam a distinção entre emoções e sentimentos. Eu creio, no entanto, que os sentimentos são essenciais para a compreensão dos conceitos morais. Meu objetivo neste capítulo foi estabelecer o alicerce para uma teoria sentimentalista, de modo que agora é hora de expor os tijolos e a argamassa.

3. SALVANDO A SENSIBILIDADE

3.1 TEORIAS DA SENSIBILIDADE

No capítulo 1, comecei a apresentar o caso do emotivismo. Eu argumentei que as propriedades e conceitos morais envolvem, essencialmente, emoção. Neste capítulo, apresentarei e defenderei uma teoria emotivista que acomode os dados do capítulo 1 e que conduza a uma aproximação com as emoções tal como estas foram desenvolvidas no capítulo 2. A teoria que defenderei aqui é uma teoria da sensibilidade. Ela é prima de outras teorias na literatura (Dreier 1990; Johnston 1989; D. Lewis, 1989; McDowell 1985; McNaughton 1988; Wiggins 1987; Wright 1992). Argumentarei que as teorias da sensibilidade podem explicar algumas intuições populares sobre a natureza da moralidade. Mas as teorias da sensibilidade enfrentam um número de sérias objeções. Na segunda parte deste capítulo, voltar-me-ei para dez objeções que podem ser encontradas na bibliografia. Argumentarei que uma teoria da sensibilidade apropriadamente formulada tem fundamentos para superar cada uma delas.

3.1.1 Teorias-padrão da sensibilidade

No capítulo 1, defini as teorias da sensibilidade a partir do seguinte esquema:

(S1) *Tese metafísica*: Uma ação tem a propriedade de ser moralmente certa (errada) unicamente no caso de causar sentimentos de aprovação (desaprovação) em observadores normais sob certas condições.

(S2) *Tese epistêmica*: A disposição para sentir as emoções mencionadas em S1 é uma condição de posse acerca do conceito normal de certo (errado).

A teoria da sensibilidade oferece uma maneira de acomodar a evidência de que o juízo moral é emotivo [*affect-laden*]. Esta é uma de suas muitas vantagens.

Mas esta não é a única aproximação que une juízos morais e emoções. O emotivismo, por exemplo, pode se gabar de fazer o mesmo. Para motivarmos a adoção da teoria da sensibilidade, precisamos mostrar que ela tem outras vantagens.

Um atrativo é que a teoria da sensibilidade acomoda a evidência usada para dar suporte ao intuicionismo moral sem tomar seu conteúdo. Os intuicionistas tipicamente asserem três pretensões (ver Ross, 1930; Audi, 2004): há uma pluralidade de valores morais básicos, ao invés de um único princípio moral; estes valores podem ser descobertos pela intuição, sem se esboçar qualquer inferência; e, quando intuímos um valor moral desta forma, ela é autojustificatória (isto é, a forma como apreendemos um juízo moral também conta como justificativa para aquele juízo). Estas pretensões são todas fenomenologicamente plausíveis. Parece que intuímos fatos morais. O problema é que o intuicionismo está usualmente vinculado a uma bizarra metafísica, bem como a uma bizarra epistemologia. Se fatos morais não são conhecidos por inferência ou por observação, então eles devem ser acessados de uma forma misteriosa, e eles devem ser fatos muito incomuns.

A teoria da sensibilidade acomoda as pretensões essenciais do intuicionismo, sem sua bagagem. Há uma pluralidade de valores básicos porque podemos desenvolver cada uma de nossas atitudes emocionais individualmente: podemos adquirir uma repulsa pelo homicídio de uma forma, e o repúdio ao incesto de outra forma. A sensibilidade explica o caráter intuitivo dos valores básicos porque as emoções que constituem aqueles valores podem ser conhecidos de forma não inferencial; sente-se que matar é simplesmente errado. Juízos morais são auto justificativos, pois as emoções que experienciamos quando apreendemos aqueles juízos são também responsáveis por torna-los verdadeiros: fatos morais são consequências de nossas reações emocionais. Logo, a teoria da sensibilidade pode ser vista como uma forma de intuicionismo, mas ela é metafísica e epistemologicamente inócua. A intuição não é, nessa aproximação, uma faculdade misteriosa, mas, ao invés disso, uma instância da capacidade mundana de olhar para dentro, para nossos estados emocionais, nossas reações viscerais. Fatos morais são gerados por tais sensações, e eles não são mais peculiares

metafisicamente do que fatos envolvendo propriedades tais quais o engraçado, o assustador e o embaraçador.

A teoria da sensibilidade também oferece uma resposta atrativa para o desafio colocado por J.L. Mackie (1977). Mackie sustenta que os conceitos morais são desesperançosamente confusos. Por um lado, agimos como se fatos morais fossem parte da estrutura do mundo. Por outro lado, tendemos a pensar que fatos morais são guias da ação. Mas estas duas assunções são difíceis de conciliar. Como pode um fato sobre o mundo ser prescritivo desta forma? Na terminologia de Mackie, os conceitos morais significam a designação de propriedades que são absolutamente estranhas. Não temos qualquer razão para pensar que tais propriedades existam. Portanto, os conceitos morais são vazios. Juízos éticos são falsos, ou sem sentido, porque eles atribuem propriedades não mais reais do que fadas ou o flogisto.

Os teóricos da sensibilidade respondem a Mackie demonstrando que as propriedades morais não são, afinal de contas, tão estranhas. Confira o seguinte enigma sobre cores: cores parecem estar no mundo, localizadas na superfície das coisas, e, ainda assim, elas também parecem experiências, pois elas podem variar mesmo quando as superfícies não sofrem qualquer transformação física. Uma abordagem lockeana da cor pode explicar isso. Locke (1690) define as cores como qualidades secundárias. Isto pode ser capturado pela seguinte definição:

(C) A propriedade de ser azul (vermelho, amarelo, etc.) é aquela que causa sensações azuis (vermelhas, amarelas, etc.) em observadores normais sob certas condições.

Em outras palavras, a abordagem lockeana diz que as cores são forças que certas superfícies e meios têm para causar experiências em nós. Forças são propriedades relacionais de coisas físicas. De certo modo, as cores estão localizadas no mundo, pois as forças em questão são possuídas pelas coisas externas à mente (bananas e narcisos, por exemplo). Mas cores são também constituídas pelas experiências, pois essas forças são individuadas pelos efeitos

em nós. Uma banana poderia não ser amarela se ela não causasse certas experiências em nós. As cores são, portanto, coisas externas, mas elas são individuadas pelos estados subjetivos.

Agora considere as propriedades morais. Ao invés de apelar a experiências visuais, nós podemos definir as propriedades morais como forças que causam emoções em nós. Emoções têm força motivacional, de tal forma que algo que causa emoções é motivante. Se as propriedades morais são forças que causam emoções, então as propriedades morais são motivantes. De certa forma, as propriedades morais são constituídas por estados motivacionais em nós, mas propriedades morais são também aspectos do mundo. Certas situações têm o poder de causar as emoções relevantes. Essas situações existem fora da mente, e elas provocam respostas emocionais em nós. Visto assim, há fatos morais, tal como há cores factuais. Podemos dizer verdadeiramente que uma ação foi certa ou errada. Mas quando dizemos essas coisas, empregamos conceitos morais que possuem componentes emotivos. Os componentes emotivos são guias da ação. Não podemos reconhecer fatos morais como tais se não estamos dispostos a agir, da mesma forma que não podemos reconhecer cores (de uma forma padrão) sem ter experiências visuais de um tipo particular.

As emoções trazem consigo algo mais, que é essencial à moralidade: elas trazem força prescritiva. Se julgo suas ações erradas, eu irei experienciar uma forma de desaprovação em relação a você. Minha desaprovação não apenas descreve o que você fez; ela prescreve que você aja de outra forma. A desaprovação dirigida a outra pessoa impõe uma ameaça àquela pessoa, a qual pode promover comportamentos compensatórios, desculpas, e uma melhor conduta no futuro.

Os teóricos da sensibilidade respondem a Mackie construindo uma ponte conceitual entre metafísica e motivação (ver McDowell, 1985). S1 nos fala das propriedades morais, e S2 nos fala sobre a motivação moral. Eis como segue a história. Um conceito é a representação de uma propriedade. Conceitos representam ao serem confiável ou legitimamente causados por instâncias das propriedades. As propriedades morais são forças que causam emoções em nós. Como representamos mentalmente tais propriedades? A resposta óbvia é que nós

as representamos emotivamente. Representamos as cores por meio das experiências com cores que as cores causam em nós. Por analogia, representamos as propriedades morais mediante as emoções que elas causam. Conceitos morais incorporam as emoções que são causadas pelas propriedades morais, e, assim, servem como detectores confiáveis para aquelas propriedades. Emoções são motivadoras. Portanto, simbolizar um conceito moral nos dispõe a agir. Juízos morais nos colocam a agir porque eles contêm conceitos morais. Passamos da metafísica à motivação via conceitos. Colocando Mackie à prova, uma análise das intuições populares não nos força a concluir que conceitos morais são vazios. Eles são impregnados afetivamente, mas não fatalmente confusos.

3.1.2 Aprovação e desaprovação

A definição esquemática da teoria da sensibilidade necessita de alguns refinamentos. Um problema com (S1) é que ela faz referência às emoções humanas de aprovação e desaprovação. No capítulo 2, eu notei que esses não são os nomes de emoções individuais. "Aprovação" nomeia uma série de emoções positivas, e "desaprovação" nomeia uma série de emoções negativas. Os moralistas britânicos do Séc. XVII desenvolveram teorias plausíveis acerca da aprovação e da desaprovação. Nós estamos, agora, em posição de preencher aquela lacuna com propostas substantivas, empiricamente fundadas.

Começemos com a desaprovação. As emoções negativas que subjazem à desaprovação moral variam de acordo com os contextos morais. Como vimos no capítulo 2, há uma divisão grosseira entre o culpar-se a si mesmo e o ser culpado por outros. Se eu perpetro uma transgressão moral, eu vou sentir emoções de censura autoimposta, notavelmente de vergonha ou de culpa. Se você perpetrar a transgressão, eu sentirei emoções de culpa voltada para você. A emoção de censura dirigida ao outro depende de o outro ser um amigo (desapontamento) ou um estranho (ira moral, desprezo, ou aversão). Diferentes emoções de autocensura e de censura aos outros correspondem a diferentes tipos de transgressões, como

ênfatisado em meu acr scimo ao modelo CAD (ver figura 2.1), de Rozin et al (1999). Dadas tais complexidades, dever amos dizer o seguinte sobre o erro moral:

(S1'-W) Uma a o tem a propriedade de ser moralmente errada apenas no caso de ela causar sentimentos no espectro de ambos os tipos de emo o, na de autocensura e de censura aos demais, em observadores normais sob certas condi es.

Para simplificar, me referirei a isso como o "espectro de desaprova o". Ser  importante manter em mente que, para que algo se qualifique como errado, esse algo deve poder causar tanto emo es de autocensura quanto de censura aos demais. N o se trata de dizer que qualquer coisa que nos irrita ou repugna seja errado. Repugna-nos a ideia de comer larvas de insetos, mas n o consideramos tal refei o uma transgress o moral. Irritamo-nos com crian as chorando em avi es, mas n o condenamos sua conduta como moralmente errada. Ajuizar moralmente requer a disposi o de ter emo es tanto de censura a si mesmo quanto de censura dirigida aos demais no espectro de desaprova o. Algumas das a es que nos repugnam quando observamos outros as realizarem poderiam, tamb m, fazer-nos sentir vergonha se f ssemos realiza-las n s mesmos (considere o canibalismo). E algumas das a es que nos causam ira quando observamos outros as realizarem poderiam nos fazer sentir culpa (imagine um adulto fazendo um imenso escarc u em um avi o). Quando me refiro a a es que causam emo es no espectro de desaprova o, eu quero designar a es que poderiam provocar tanto censura dirigida a si mesmo quanto dirigida aos demais. Essas s o, em minha perspectiva, a es moralmente erradas.

Em adi o ao espectro de desaprova o, h , tamb m, um "espectro de aprova o". Esse provoca as emo es de louvor listadas na Tabela 2.2 do cap tulo 2. A asser o "moralmente certo"   utilizada de diferentes formas.  s vezes, ela funciona como sin nimo de "dever", como quando afirmamos "o que deve ser feito nesta situa o   [...]". Quando usada dessa forma, a palavra "certo" n o expressa necessariamente algo positivo. Raramente enaltecemos as pessoas em virtude de elas terem feito o que a moralidade demanda delas. Uma pessoa que n o rouba em

lojas está fazendo a coisa certa, mas não costumamos nos sentir gratos em relação a ela. Para os presentes propósitos, irei ignorar essa noção de "certo", pois ela pode ser definida nos termos do "errado" (veja, abaixo, as observações sobre o "dever"). Em acréscimo a essa noção, às vezes, nós usamos a frase "moralmente certo" para expressar uma atitude positiva. Quando julgamos um bom feito como moralmente certo, nós estamos enaltecendo aquele feito, não pretendendo que ele seja obrigatório. Nesses casos, a seguinte definição pode ser utilizada:

(S1'-R) Uma ação tem a propriedade de ser moralmente correta apenas no caso em que sua realização cause uma emoção, no espectro de aprovação, em observadores normais sob certas condições.

No que se segue, eu irei focar quase inteiramente em casos de erros morais, ao invés de em acertos morais. Meu palpite, não examinado aqui, é que o conceito de errado desempenha um papel mais importante na moralidade do que o conceito de certo: nós condenamos com mais frequência do que elogiamos.

Ambas, (S1'-R) e (S1'-W) carecem de explicação adicional. Essas definições se referem a "observadores normais" e a "certas condições". Mas quais observadores são normais, e quais condições importam? Começamos com as condições. Como uma primeira aproximação, pode-se cair na tentação de dizer que as condições que importam são aquelas nas quais os observadores têm conhecimento dos fatos relevantes, não estando sob influências emocionais e cognitivas que não sejam relevantes para o caso em questão. Esse não é um mau ponto de partida, mas ele não é totalmente satisfatório. De um lado, ele usa a enganadora noção de relevância, e, por outro, ele pode, em alguns casos, dificultar (como argumentarei posteriormente) a distinção entre conhecimento factual e crenças valorativas. Mas, ao invés de aperfeiçoar a caracterização das condições ideais, pode-se simplesmente alterar a definição de tal forma que se evita qualquer apelo a quaisquer que sejam as condições. O artifício é invocar sentimentos.

No capítulo 2, eu disse que o termo "sentimento" designa uma disposição emocional. Eu penso em disposições psicológicas como estados mentais fisicamente realizados. Se o realismo acerca da mente é verdadeiro, então há um

fato de saber se alguém tem um sentimento; há algo no cérebro que é a realização daquela disposição. Suponhamos que eu tenha um sentimento negativo concernente ao homicídio. Não importa quão frequentemente pensamentos acerca do homicídio me deixem irado, bem como não importa em que condições eu preciso estar para que pensamentos sobre matar provoquem tal reação. O ponto crucial é que, em algum lugar de meu cérebro, há algo que me inclina a ter essa reação negativa diante do homicídio. Podemos definir propriedades morais apelando à existência de tais sentimentos e contornando qualquer necessidade de especificar quando tais sentimentos podem realmente ser fundados sobre a produção das emoções que eles nos incitam a produzir. Assim, ao invés de dizer que as propriedades morais existem em virtude de causarem certas emoções sob certas condições, podemos dizer que elas existem em virtude do fato de que alguns observadores têm sentimentos que os incitam a ter aquelas emoções.

Mas que observadores? O termo "observadores normais" é problemático porque normalidade é um termo valorativo. Tipicamente, alguém que faz um juízo moral considera-se a si mesmo, e aos demais como ele, normal. Isso, como argumentarei adiante, concerne ao que julga, o qual conta como normal, e à maioria de nós, pois um observador normal é simplesmente "eu". Logo, o apelo à normalidade pode se mostrar vazio quando se pensa no domínio moral. Podemos definir propriedades morais em termos de observadores, e deixar de lado a palavra normal. Integrando esses pontos, proponho a seguinte formulação:

(S1') Uma ação tem a propriedade de ser moralmente errada (correta) somente no caso de haver um observador que tenha um sentimento de desaprovação (aprovação) voltada para ela.

Aqui, aprovação e desaprovação são entendidas como a disposição para sentir emoções no plano da aprovação e da desaprovação. (S1') oferece uma boa definição de trabalho acerca da visão que defenderei neste capítulo. Essa formulação é consistente com a visão segundo a qual as propriedades morais são forças para causar estados emocionais. Se um observador tem sentimentos relativos a certas ações, e sentimentos o incitam a ter emoções, então as ações em

questão têm a força para causar emoções. (S1') pode ser considerada como uma variante notacional da visão da qualidade secundária; ela simplesmente se esquivava da necessidade de especificar as condições sob as quais a incitação da emoção pode ser tomada como evidência para a instanciação de uma propriedade moral. Esses aspectos de (S1') obviamente introduzem uma forte forma de relativismo que é, às vezes, subestimado pelos outros teóricos da sensibilidade (contudo, veja Wiggins, 1987: n. 24). Permitam-me notar desde já que, todavia, muitos observadores podem convergir para os mesmos sentimentos. A formulação é estritamente consistente com a existência de uma única moralidade, pois, em princípio, pode acontecer de todos concordarem. E se isso não ocorre na prática, pode ser ainda que um imenso número de pessoas dentro de uma cultura ou subcultura concordem. Eu penso que isso é bem provável, e há boas razões para esperar tal convergência.

3.1.3 Conceitos morais

Os refinamentos de (S1) apontam para um refinamento de (S2). Os teóricos da sensibilidade precisam considerar os conceitos morais. Eu tenho sugerido que conceitos morais incorporam emoções. Deve-se dizer algo mais sobre o que consiste aquela incorporação. Uma consideração plena dos conceitos morais deveria ser desenvolvida contra o pano de fundo de uma teoria dos conceitos independentes da motivação. Eu defendo alhures uma teoria dos conceitos, mas uma explicação e defesa plenas ultrapassariam o objetivo desse projeto (Prinz, 2002). Restringir-me-ei a uns poucos e breves comentários.

Seguindo a tradição da ciência cognitiva, eu considero que conceitos são representações mentais. Representações mentais podem ser individualizadas por aquilo que elas representam e por aquilo que podemos chamar de suas propriedades veiculares. Propriedades veiculares são as propriedades que as representações mentais têm enquanto especificidades mentais. Elas incluem estrutura e formato representacional. Em algumas teorias, conceitos são redigidos em um formato similar ao da linguagem. Eles são símbolos não modais na cabeça.

Dizer que um conceito é não modal é dizer que ele não está em um formato que seja específico para um sistema que envolva *inputs* e *outputs*. Dizer que um conceito é simbólico é dizer que suas propriedades veiculares são arbitrariamente relacionadas aos conceitos que elas representam. Aqueles que privilegiam, no que concerne aos conceitos, a analogia com a linguagem, às vezes, assumem que muitos de nossos conceitos são desestruturados; eles não são constitutivamente relacionados a quaisquer outros conceitos. Fodor (1975) está mais ligado a essa visão. Ele afirma que a maioria dos conceitos que poderíamos expressar utilizando uma única palavra do inglês (conceitos lexicais) são, eles mesmos, similares à palavra: eles não se decompõem em partes significativas. O conceito que representa cães, por exemplo, poderia ser um símbolo primitivo na linguagem do pensamento.

Uma visão distinta dessa poderia dizer que a maioria dos conceitos lexicais são entidades estruturadas redigidas em formatos modais específicos. Essa é a visão que os empiristas britânicos adotam, e eu penso que eles estão certos (Prinz, 2002). Segundo essa visão, o conceito que representa cães poderia ser um conjunto de aspectos perceptuais armazenados a partir de nossos vários encontros com cães. Se poderia perguntar: como pode uma representação perceptual de um cão representar cães e apenas cães? Afinal de contas, as mesmas características perceptuais poderiam ser observadas em lobos e gatos bem disfarçados, bem como pode haver cães tão incomuns em sua aparência que dificilmente seriam reconhecidos perceptualmente. Claro, poder-se-ia colocar a questão paralela acerca da explicação referente à linguagem do pensamento. Como um símbolo arbitrário representa cães? Como se verifica, ambas as teorias podem prover a mesma resposta. Um símbolo mental e um conjunto de aspectos perceptuais podem representar cães da mesma forma. Em particular, eles podem representar cães entrando em confiáveis (mas imperfeitas) relações causais com a propriedade de ser cão [*doghood*]. Mais acuradamente, imagens de cães, tanto quanto símbolos de cães, podem referir por terem a função de identificar cães. Essa é a abordagem dretsiana (1988) da semântica que eu endosso no capítulo 2. Fodor (1990) oferece uma teoria estreitamente relacionada a essa para seus símbolos mentais, e meu

ponto, aqui, é que as teorias empiristas dos conceitos podem se servir das mesmas teorias das referências que seus oponentes endossam. Teorias empiristas e teorias da linguagem do pensamento estão semanticamente em sintonia (por assim dizer).

Esse ponto de similaridade pode também ser usado para trazer uma importante distinção. Para que um símbolo mental arbitrário entre em uma relação causal confiável com uma propriedade do mundo, devemos ser capazes de detectar aquela propriedade. Para detectar uma propriedade que está, pelo menos parcialmente, fora de nós, precisamos percebê-la ou perceber algo que esteja regularmente correlacionado com ela. Em outras palavras, um símbolo mental arbitrário surge sob o controle causal de algo externo por meio de aspectos perceptuais. Assim, conjuntos de aspectos perceptuais são necessários em ambas as abordagens. O empirista identifica os conceitos com aqueles aspectos, e os defensores da linguagem do pensamento os consideram como intermediários entre os símbolos e o mundo. Defensores da linguagem do pensamento poderiam expressar sua perspectiva dizendo que conjuntos de aspectos perceptuais são *concepções*, ao invés de conceitos. Uma concepção é uma forma contingente e variável de pensar sobre algo. Duas pessoas podem ter os mesmos conceitos (por exemplo, símbolos representando as propriedades que caracterizam algo como um cão) e diferentes concepções (por exemplo, imagens de como os cães tipicamente se parecem). O empirista pode dizer que essas concepções são conceitos e negar que tenhamos qualquer necessidade de símbolos em uma linguagem do pensamento.

Eu prefiro o lado empirista do debate, mas o que digo sobre conceitos morais pode ser recolocado como uma teoria das concepções morais. Concepções morais são um aspecto central da cognição moral, e isso independentemente de as concepções serem conceitos ou não. Quando formulamos juízos morais, formamos uma concepção de como as coisas são, e aquela concepção terá um impacto sobre subsequentes juízos e ações.

Essa digressão nos permite colocar a questão que nos interessa. Se conceitos são detectores de propriedades morais fundados na experiência, então que são conceitos morais? Se (S1') está correta, a resposta é simples. S1 diz que as

propriedades morais são aquelas coisas acerca das quais algum observador tem sentimentos de aprovação ou desaprovação. Sentimentos são disposições emocionais, e, se a teoria das emoções defendida no capítulo 2 está correta, então emoções se qualificam como estados perceptuais; na teoria da avaliação corporificada, emoções são percepções de mudanças estampadas no corpo, as quais carregam informações sobre nossa relação com o mundo. Portanto, podemos capturar a ideia de que os conceitos morais são detectores de propriedades morais fundados perceptualmente postulando que conceitos morais são constituídos por sentimentos. Assim, por exemplo, o conceito de errado poderia ser caracterizado da seguinte forma:

(S2'-W) O conceito padrão de errado é um detector para a propriedade do que seja errado que envolve um sentimento que predispõe aquele que o possui a experimentar emoções no plano da desaprovação.

Observe que (S2'-W) usou a expressão "conceito padrão". Essa se refere a conceitos que as pessoas comuns ordinariamente usam, deixando aberta a possibilidade para que seja possível designar a mesma propriedade mediante meios não ortodoxos (ver capítulo 1). Por exemplo, psicopatas podem seguir a propriedade do estar errado respeitosamente, aplicando a palavra "errado" a tudo aquilo que algum não psicopata rotularia com essas palavras. Aquele não seria um conceito padrão. Minha conjectura é que pessoas que não possuem desordens mentais ordinariamente pensam sobre o estar errado mediante sentimentos. Acreditar que algo é errado de uma maneira não deferencial é ter um sentimento de desaprovação com relação a ele.

Outros conceitos morais podem ser analisados de forma similar. O conceito de CERTO será definido com referência a ambas, aprovação e desaprovação. Os conceitos de MORALMENTE MAU e MORALMENTE BOM seriam análogos aos de ERRADO e de CERTO, mas, enquanto os dois últimos são usados para se referir a ações, os primeiros podem ser aplicados a ações, pessoas e feições. Esse é um afastamento do modo como alguns consequencialistas lidam com esses conceitos. Eles dizem que devemos definir o BEM e o MAL antes de definirmos o CERTO e o

ERRADO, apelando a alguma qualidade intrinsecamente boa. O CERTO é, então, usualmente definido em termos de maximização do bem. Eu penso que isso é um afastamento do uso ordinário. Frequentemente louvamos ações como corretas independentemente de elas maximizarem alguma coisa, e, eu suspeito, as pessoas usam esses pares de termos morais intercambiavelmente em muitos contextos. Uma má ação é uma ação errada, por exemplo. Como sugerido, penso que a principal diferença é que BOM e MAU são noções mais amplas, as quais podem ser aplicadas a coisas que não sejam apenas ações. Não obstante, eu não quero ser demasiadamente rígido sobre isso. A análise conceitual é um negócio intrincado, e nós provavelmente usamos termos morais em uma variedade de formas distintas. Eu fico satisfeito, aqui, em capturar mais ou menos um dos meios centrais pelos quais entendemos os termos morais, permitindo que existam outros.

O único outro conceito que mencionarei aqui é o de DEVER. Como ponto de partida, considere a análise feita por Westmarck (1900), outro emotivista epistêmico. Ele sugeriu que devemos fazer o que seria errado não fazer. Essa é uma proposta elegante, pois reduz o conceito de DEVER ao conceito de ERRADO. Ela, contudo, tem a consequência interessante de que juízos de dever não são comandos, pois não é um comando afirmar que é errado não fazer algo. Westmarck argumenta que essa é uma virtude de sua abordagem. Ela dá sentido à ideia de que podemos emitir juízos de dever sobre ações passadas ("eu deveria ter doado mais à Oxfam no ano passado"). Quando fazemos tais juízos não podemos emitir mandamentos, pois é impossível mudar o passado. É claro, nossos juízos sobre o passado têm implicações motivacionais. Quando julgo que deveria ter doado mais à Oxfam, eu sinto culpa, e essa culpa pode influenciar meu comportamento atual.

Eu penso que a análise de Westmarck acerca do dever está quase certa, mas necessita de um pequeno acréscimo. É uma consequência de sua visão que há uma equação simples entre certos juízos de dever e certos juízos acerca do erro. Se digo que não debes bater em teu cão, então nada mais estou dizendo senão que é errado bater em teu cão. Mas essas duas pretensões parecem ligeiramente diferentes. A primeira é um comentário sobre o *status* moral de se bater em um cão. Ela tem impacto motivacional, mas não é uma pretensão clara acerca do que você deveria

fazer. Nossas pretensões são claramente sobre ações. Quando inferimos que você não deve φ a partir da premissa de que φ é errada, parece que estamos fazendo uma inferência substantiva, e não uma variante notacional. Colocado em outros termos, um juízo sobre o erro pode lhe colocar em um estado motivacional, mas um juízo de dever diz que aqueles impulsos motivacionais devem ser observados. Com efeito, um juízo de dever exprime o fato de que uma norma *tem autoridade* sobre o comportamento da pessoa à qual se dirige aquele juízo. No capítulo 5, retornarei a essa ideia de autoridade e a explicarei mais extensivamente. Também retornarei, abaixo, ao tema da motivação.

De acordo com (S2'-W), conceitos morais nos dispõem a sentir emoções de culpa em relação aos outros quando somos vítimas de transgressões, e emoções de culpa relacionada a nós quando somos nós que transgredimos. Se você acredita que φ é errado, nessa abordagem, você tem uma memória de longo prazo representacional, a qual lhe incita a sentir culpa ou vergonha se φ , e ira, desprezo ou repulsa se alguém mais realiza φ . Em qualquer ocasião dada, quando você tem o ocorrente pensamento de que φ é errado, unicamente uma dessas emoções irá se manifestar. Se lhe surpreendo fazendo φ , e digo "Isso é errado!", a palavra "errado" pode expressar meu estado atual de ira, o qual é uma manifestação de minha complexa disposição emocional. Podemos pensar no sentimento da memória de longo prazo como uma crença permanente, e na emoção da memória atuante como uma crença ocorrente. Ou, para introduzir uma peça terminológica útil, podemos chamar o sentimento de *regra moral*, e podemos chamar uma manifestação emocional particular daquele sentimento de *juízo moral*.

Essa proposta é ilustrada na figura 3.1. A imagem mostra uma pessoa observando um batedor de carteiras em ação. Após, há uma série de eventos mentais. Primeiro, o observador interpreta o furto como uma instância do ato de roubar. Esse é um estágio de categorização; ele envolve se basear em um conhecimento prévio para classificar uma ação observada. Roubar é algo em relação ao qual o observador tem um sentimento moral a partir de sua memória de longo prazo. Seu sentimento em relação ao roubo constitui uma regra. Quando a ideia de roubar entra em sua mente, a regra faz com que o sentimento seja ativado.

Esse estágio é de recuperação da regra. Nesse estágio, o sentimento ainda não é experienciado como emoção. Sentimentos são disposicionais. Uma vez que um sentimento é ativado, fatores contextuais são usados para determinar qual emoção será evocada. Uma vez que o observador não é o autor da ação, e uma vez que roubar é uma norma da autonomia, a emoção evocada é a ira. Esse é o estágio da evocação da emoção. A ira que surge no observador não é simplesmente uma ira à toa. A ira foi disparada pela experiência do furto, e está ligada à sua representação. O resultado é um estado composto: ira com relação ao furto. Esse composto constitui o juízo de que realizar o furto é errado, pois a emoção que ele contém foi gerada por um sentimento moral. Esse é o estágio do ajuizamento. Em suma, há uma progressão que segue da categorização à recuperação da regra, à evocação da emoção, ao juízo. Regras morais guiam a transição marcada com 1 na Figura, e juízos morais passam a existir com a transição marcada com 4, na qual uma emoção ativa é aplicada a uma ação que está sob consideração.

O modelo descrito aqui tem vários aspectos interessantes, os quais são dignos de menção. Primeiro, ele ajuda a diagnosticar casos nos quais juízos morais podem ser ditos errôneos. Considere o estudo de Wheatley e Haidt (2005) descrito no capítulo 1. Eles descobriram que algumas pessoas que foram hipnotizadas para sentir repugnância acabaram condenando moralmente um indivíduo perfeitamente inocente. Penso que tais condenações podem ser qualificadas como erros porque elas não foram causadas por sentimentos oriundos da memória de longo prazo, mas, ao invés disso, por fatos estranhos. Elas não se qualificam como expressões legítimas das atitudes morais do indivíduo. Em resumo, uma ação errada é uma ação contra a qual um observador tem uma regra moral. Se uma ação é condenada em virtude do repúdio induzido hipnoticamente, então ela não pode ser qualificada como errada.

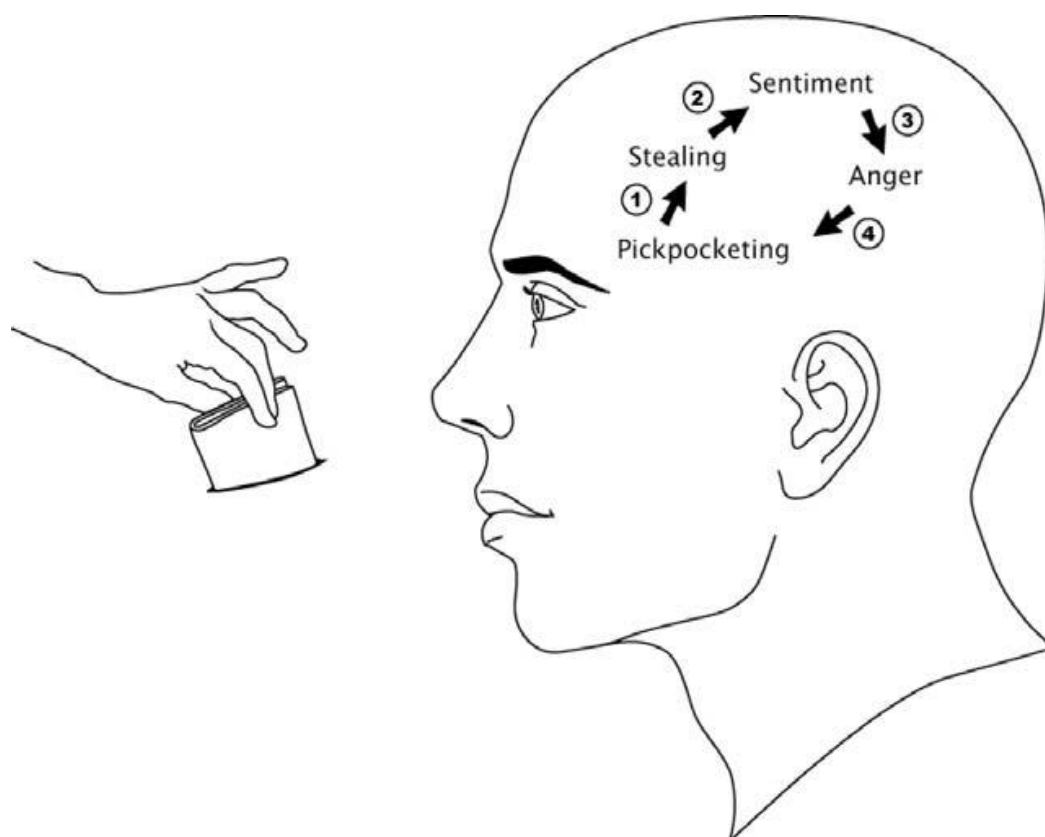


Figura 3.1. Estágios do processamento de informações que conduzem ao juízo moral. (1) Um evento percebido é categorizado; (2) uma regra é recuperada da memória, a qual ativa um sentimento; (3) o sentimento evoca uma emoção sensível ao contexto; (4) a emoção é associada à ação percebida

Uma segunda vantagem do modelo é que ele imediatamente supera uma objeção comum às teorias emotivistas. Críticos gostam de reiterar que nossas emoções oscilam de uma forma que nossos valores morais não oscilam. Por exemplo, Mele (1996) destaca que pessoas indiferentes ou depressivas podem não se sentir muito estimuladas quando pensam sobre assuntos morais, mas seria pouco caridoso sugerir que elas não julgam que o homicídio seja errado. Ele argumenta que isso mostra não haver conexão essencial entre motivação e crença moral. Mas o modelo que estou recomendando é um modelo disposicional, ele não requer que nossas convicções morais sempre se manifestem emocionalmente (compare com Dreier, 1990). Eu posso reconhecer que tenho um sentimento negativo sem sentir aquele sentimento (como quando eu digo, sem repugnar, que

não gosto de fígado fatiado). O modelo que estou recomendando também tem elementos para lidar com uma objeção semelhante colocada por Millgram (1999). Ele nota que reações emocionais diminuem após várias exposições. Uma piada está menos propensa a nos fazer rir na décima vez em que a escutamos, e, quando isso acontece, dizemos que a piada é menos divertida. Similarmente, uma exposição reiterada a, digamos, pessoas sem-teto reduzirá nossos sentimentos de escândalo. Mas, nesse caso, não dizemos que a falta de habitação é menos errada do que quando encontramos um sem-teto pela primeira vez. Como uma primeira linha de resposta, estou tentado a ser firme e dizer que uma superexposição (ou, até o ponto que interessa, subexposição) pode alterar nossos sentimentos e, portanto, alterar nossas visões sobre o que seja certo ou errado. Seria interessante traçar como as atitudes das pessoas em relação à falta de habitação mudam após vários encontros. Mas meu modelo pode também explicar o que Millgram pensa que toda teoria deve explicar: ele pode permitir a continuidade da moral nas flutuações do afeto. Em alguns casos, a exposição repetida deixa nossos sentimentos intactos e simplesmente diminui sua probabilidade de gerar emoções intensas. A exposição repetida diminui o processamento profundo; quando vejo o sem-teto pela milésima vez, não me sinto perplexo ao ponto de refletir sobre o fato de que uma sociedade próspera não deveria jamais permitir tal destituição extrema, mas, se refletisse, emoções seriam evocadas. Devo notar que penso que uma análise similar está disponível para o humor. É uma piada menos divertida depois de reiteradas exposições? Às vezes, dizemos que sim, mas seguimos contando a piada para todos nossos amigos. Seria melhor dizer que a piada segue sendo divertida, mas ela já não me faz rir. Por que ela permanece sendo divertida? Porque ela segue tendo propriedades (um inteligente jogo de palavras, por exemplo) que me incitam a rir.

Minha abordagem do juízo moral é similar a outras propostas sustentadas empiricamente dentro da literatura. Ela está relacionada a ideias que têm sido desenvolvidas por Haidt (2001) e Nichols (2004a), mas há importantes diferenças. Haidt defende um modelo de "intuicionismo social" acerca do juízo moral. Uma intuição é um estado mental afetivo de atração ou de repulsa, o qual aparece

subitamente sem qualquer consciência de se estar em um processo deliberativo. Certas coisas simplesmente parecem erradas ou corretas. Haidt não equipara juízos morais com essas intuições. Ele diz, ao invés disso, que juízos morais são ordinariamente causados por intuições (Haidt, 2001: 814). Nichols equipara valores morais a “regras sentimentais”. Ele afirma que regras sentimentais envolvem dois mecanismos: uma “Teoria Normativa” e um sistema emocional de resposta. Teorias Normativas são listas de prescrições sobre o que se deveria, ou não, fazer. Essas prescrições não possuem conteúdo emotivo. Elas se tornam “respaldadas pelo afeto” quando unidas ao sistema emocional de resposta, e Nichols coloca especial ênfase nas emoções que são ativadas pela percepção do sofrimento de outra pessoa – emoções de preocupação. Assim, em termos gerais, uma regra sentimental é uma prescrição livre de afecção, combinada com uma disposição para se sentir preocupado quando aquela prescrição é violada. Pode-se pensar nisso como um modelo bidirecional causal: quando uma norma proibindo causar dano é violada, sentimos-nos preocupados, e quando vemos alguém sentindo dor, a preocupação resultante pode nos fazer recuperar uma norma relevante da memória. Esses dois componentes são dissociáveis; psicopatas têm Teorias Normativas sem preocupação emocional, e crianças muito jovens tem preocupação sem Teorias Normativas (Nichols, 2004a: 18 segg.).

Minha aproximação difere dos modelos de Haidt e de Nichols. Diferentemente de Haidt, eu não quero insistir que atitudes morais surjam tipicamente na ausência de um raciocínio deliberativo; pode haver muita inferência antes que possamos ver uma ação de tal forma que ela dispare uma reação emocional (ver abaixo). Diferentemente de Nichols, eu não penso que preocupação seja a emoção moral mais fundamental; as emoções que figuram nos juízos morais expressam censura àquele que age errado, e não simpatia pela vítima. Hume (1739) pensa que censura e simpatia estão intimamente ligadas. Ele diz que nós julgamos que alguma ação é errada por primeiramente sentirmos simpatia por uma pessoa que tenha sido agredida pela ação. Nossa empatia para com o sofrimento, a qual pode ser equiparada com o sentimento de preocupação de Nichols, causa, então, um sentimento de desaprovação em nós. Eu penso que Nichols e Hume colocam

demasiada ênfase nas respostas empáticas. Podemos condenar moralmente uma ação sem pensar na vítima (considere o incesto consensual), bem como frequentemente pensamos na vítima apenas após uma ação ter sido condenada (considere sonegação de impostos).

Essas são diferenças menores. Haidt, às vezes, admite que precisamos raciocinar antes que experienciemos uma intuição moral, e Nichols, às vezes, discute emoções de censura (especialmente de repugnância) e minimiza o papel da preocupação. Há, todavia, uma diferença mais importante entre esses modelos e a aproximação da qual sou a favor. Ambos, Haidt e Nichols, pensam que há uma relação *causal* entre juízos morais e emoções. Juízos morais não contêm emoções como partes; eles estão meramente fundados causalmente em emoções. Logo, esses modelos permitem que se forme um juízo moral sem emoções. Segundo minha aproximação, emoções são *integrantes* dos juízos morais, pois emoções constituem instâncias simbólicas de conceitos tais quais os de certo e de errado. Aqui minha visão ecoa a de Hume (1739: III, i, ii): "Não inferimos que um caráter seja virtuoso porque ele agrada: mas após sentirmos que ele agrada de uma maneira peculiar, nós, de fato, sentimos que ele é virtuoso."

Minhas razões para preferir um modelo integrador em relação a um modelo causal já foi esboçado no capítulo 1. Um modelo causal implica que emoções e juízos morais são duas coisas separadas que podem ser colocadas separadamente. Isso parece implausível, dado que emoções ocorrem concomitantemente com os juízos morais, e parecem ser necessárias para o desenvolvimento de juízos morais. Deixe-me acrescentar, aqui, uma evidência adicional. Em um estudo não publicado, perguntei a sujeitos que considerassem dois cenários nos quais emoções são confrontadas com um comportamento verbal. Em um cenário, um estudante em uma fraternidade insiste que não é moralmente errado fumar maconha, mas ele se sente enojado quando vê seus irmãos de fraternidade fumando, e envergonhado quando ele mesmo fuma. No outro cenário, um estudante em uma fraternidade conta a seus irmãos que fumar maconha é moralmente errado, mas ele nunca se sente enojado quando os vê fumar, bem como não se sentiria envergonhado se fumasse. Em cada caso, os sujeitos são

questionados sobre se os valores morais do estudante são refletidos pelo que ele diz ou pelo que ele sente. Os sujeitos, em ambas as condições, foram muito mais propícios a dizer que o valor correspondia às emoções: o estudante que se sentia envergonhado por fumar maconha é moralmente oposto a isso, e o estudante que não sentia vergonha, não. Isso sugere que a moralidade popular estabelece um elo conceitual entre valores e emoções. Se um modelo causal estivesse correto, então um valor moral poderia ser constituído por uma atitude proposicional livre de afecção. Se essa possibilidade fosse consistente com a moralidade popular, então as respostas para esses dois cenários seriam menos claras.

Eu não nego que nossas atitudes emocionais são geralmente acompanhadas por representações proposicionais livres de afecção. Quando julgamos que algo é errado, tendemos a verbalmente expressar esse juízo em voz alta ou em um discurso interior. Ao vermos um furto, podemos pensar "aquilo é errado!". Tais pensamentos verbalizados podem ocorrer na ausência de emoções. É assim que as pessoas provavelmente pensam sobre a moralidade, quando elas são indiferentes ou superexpostas. As pessoas também podem empregar pensamentos verbalizados em uma conversação causal quando não estamos experienciando diretamente atos imorais. Se você pensou sobre o furto anteriormente, você armazenou uma representação verbal do tipo "furar é errado", a qual você pode recuperar sem se conturbar. Mas a evidência empírica sugere que emoções são ordinariamente evocadas mesmo em uma reflexão casual (ex., Moll et al., 2002). Não obstante, eu quero insistir que essas verbalizações ganham seu sentido dos estados emocionais subjacentes. Geralmente falamos como se verbalizações de juízos morais fossem juízos morais por direito próprio. Vou me referir a sentenças do tipo "furar é errado" como um juízo moral verbalizado. Mas esse rótulo é realmente redutor. "Furar é errado" não é um juízo; é uma cadeia de palavras. Ao chama-lo de juízo moral verbalizado, eu quero dizer que ele é uma forma verbal que pode ser usada para expressar um juízo moral.

Juízos morais verbalizados são muito úteis, pois eles nos permitem raciocinar sobre valores morais aos quais frequentemente não damos suporte. Por exemplo, podemos refletir sobre condicionais tais quais, "se matar é bom, então

Stalin foi um santo". Quando levamos em consideração esse condicional, usamos o juízo verbalizado, ao invés de formar um genuíno juízo moral, pois é psicologicamente difícil para nós manter uma atitude positiva em relação ao homicídio. Mas a palavra "bom" poderia não ter significado algum em contextos condicionais se não atribuíssemos significação emocional a ela em contextos incondicionais. Compreendemos o significado de "certo" e de "errado", ou de "bom" e de "mau" em contextos incondicionais por meio de emoções morais subjacentes. Em analogia, eu posso compreender a sentença "se matar é divertido, então Stalin foi um comediante" sem me sentir entretido. Mas, para entender o que significa "divertido", deve haver sentenças em que eu profira a palavra fora de um contexto condicional, e, quando o faço, a palavra expressa meu entretenimento. Da mesma forma, eu posso compreender a fala moral condicional apenas se os termos morais expressam emoções para mim quando usadas em contextos incondicionados.

Eu acabo de dizer que juízos morais verbais incondicionais são *expressões* da emoção. Eles exprimem como nos sentimos sobre as coisas. Nesse sentido, minha abordagem está intimamente relacionada com as teorias expressivistas, tais como o emotivismo, as quais enfatizam o caráter expressivo ou o discurso moral. Mas os expressivistas argumentam que juízos morais verbalizados são *meras* expressões de sentimento; eles não têm conteúdo representacional. Segundo vejo, juízos morais verbalizados simultaneamente expressam como nos sentimos e representamos coisas. Quando dizemos "Isto é errado!", expressamos nossos sentimentos e também almejamos asserir um fato. "Furtar é errado" representa o fato de que o furto tem a propriedade de estar errado. O estar errado moralmente, eu tenho sugerido, é a propriedade de ser objeto de desaprovação. Isso corresponde ao que eu disse sobre o erro ser uma qualidade secundária.

Expressivistas dizem que conceitos morais não denotam, e eu digo que eles denotam qualidades secundárias. É importante notar que, quando usamos conceitos morais, nós não temos necessariamente quaisquer crenças explícitas sobre o tipo de propriedades que elas selecionam. O conceito de errado não é uma descrição de uma qualidade secundária enquanto tal. Ele representa uma qualidade

secundária em virtude de sua relação causal com o mundo. Conceitos que denotam causalmente não precisam descrever seus referentes.

Eu penso que conceitos morais têm uma camada adicional de conteúdo representacional que não foi percebida pelos teóricos da sensibilidade. Como acabo de observar, sentimentos representam qualidades secundárias, mas sentimentos também nos inclinam a experienciar emoções, e emoções não representam qualidades secundárias; elas representam preocupações (ver capítulo 2). Logo, quando um estado emocional é gerado a partir de um sentimento, ele representa tanto a qualidade secundária acompanhada pelo sentimento *quanto* uma preocupação. Considere as fobias. Fobias são sentimentos que representam a qualidade secundária da temeridade, mas fobias se manifestam na emoção do medo. Quando você tem uma reação fóbica a algo, você está atribuindo a ele, simultaneamente, a propriedade de ser assustador e a propriedade de ser perigoso, que é o conteúdo representacional do medo. Sentimentos morais envolvem emoções morais, e, seguindo Rozin et al. (1999), argumento que emoções morais diferentes correspondem a diferentes domínios: ordem pessoal, ordem social e ordem natural. Portanto, o juízo segundo o qual alguma ação é errada expressa duas coisas sobre ela. O juízo expressa a visão conforme a qual a ação causa certas respostas emocionais em mim, e que ela é uma transgressão de algum tipo. Conceitos morais avaliam e classificam.

Se estou certo acerca da semântica dos juízos morais, então Bernard Williams (1985) está enganado quando diz que conceitos como os de CERTO e de ERRADO são "tênués". De acordo com Williams, um conceito moral "tênué" faz duas coisas: ele expressa uma atitude valorativa e classifica. Exemplos incluem atitudes generosas, heroicas e cruéis. Quando dizemos que uma ação é "generosa" nós simultaneamente classificamos a ação como uma instância do dar coisas aos outros e a avaliamos como positiva. Williams argumenta que conceitos como os de CERTO e de ERRADO, ou de BOM e de MAU, são comparativamente tênués: eles expressam atitudes sem classificar. Eu penso que Williams está equivocado. Na maioria das ocasiões, quando julgamos que algo é mau, por exemplo, estamos experienciando uma emoção específica, e aquela emoção serve para classificar a

ação sob apreciação. Se eu sinto desprezo quando afirmo que um político corrupto é "mau", então meu desprezo serve para classificar sua conduta como um mal para a comunidade. O conceito MAU, nessa ocasião, representa tanto uma propriedade valorativa dependente de resposta, quanto uma categoria de transgressão moral. Da mesma forma, quando julgo que o incesto é errado, eu não apenas o condeno; eu experiencio repulsa, a qual serve para classificar o incesto como uma violação de uma regra que pertence à ordem natural. Eu não pretendo sugerir que nós pensamos explicitamente na comunidade quando sentimos desprezo, ou na ordem natural quando sentimos repulsa. Emoções representam preocupações em virtude de suas relações causais com o mundo. Quando julgamos que algo é errado, a única coisa diretamente disponível para nós, na consciência, é que ele faz com que nos sintamos de certa forma. Esses sentimentos têm duas camadas de conteúdo semântico, mas, tipicamente, nós não estamos explicitamente cientes delas.

Essa discussão sobre conceitos morais pode ser sumariada da seguinte maneira. Se você internalizou uma regra moral contra, digamos, o incesto, você tem um sentimento moral em relação ao incesto em sua memória de longo prazo. Aquele sentimento representa a qualidade secundária de causar desaprovação em você (e outros como você) sob boas condições epistêmicas. Aquela regra é uma atitude permanente em relação ao incesto. Ela se torna um juízo moral ocorrente quando você pensa sobre alguém cometendo um ato incestuoso. Quando isso ocorre, você experiencia repulsa e esse sentimento, aplicado ao seu pensamento atual sobre um ato incestuoso, constitui seu juízo segundo o qual o ato em questão é errado. A repulsa é causada pelo sentimento na memória duradoura, e, como tal, ela representa a propriedade do estar errado (i.e., a qualidade secundária de causar desaprovação). Mas repulsa também representa violações da ordem natural, a qual é um subgrupo dos erros. Quando você julga que o incesto é errado, você tem a experiência da repulsa, mas você pode não se aperceber que esse sentimento representa uma qualidade secundária, ou mesmo que tenha qualquer coisa a ver com a ordem natural. Tudo o que você sabe a partir do sentimento é que você pensa que o incesto é errado. Você não sabe em que consiste o ser errado.

3.1.4 Motivação Moral

Eu tenho argumentado que os conceitos morais são constituídos por sentimentos, e esses sentimentos se manifestam em distintas emoções em diferentes ocasiões, dependendo, por exemplo, do domínio da regra que foi violada. O apelo a sentimentos tem importantes implicações para o elo entre juízo moral e motivação. Eu disse que uma grande atração da teoria das emotividades é que ela pode explicar a observação segundo a qual juízos morais tendem a motivar. Se juízos morais contêm conceitos morais, e conceitos morais têm uma composição emocional, então juízos morais motivam a ação, pois emoções são estados emocionais. A teoria da sensibilidade implica internalismo, o qual, como argumentei no capítulo 1, é um resultado desejável. Aqui eu quero dizer algo sobre como o tipo particular de teoria da sensibilidade que sustento contribui para a motivação.

A primeira coisa a notar é que minha teoria faz a predição internalista genérica que, excetuando-se a indiferença e outros impedimentos, as pessoas serão motivadas a se comportar de acordo com seus valores morais. Se você contempla uma ação que você considera imoral, você irá antecipar as consequências emocionais negativas e evitar aquele comportamento. Durante a deliberação sobre a ação, o conceito de errado funciona como aquilo que Damasio (1994) chama de marcador somático: ele nos faculta experimentar os sentimentos emocionais e, portanto, os corporais, que surgiriam se perpetrássemos ações imorais. Se você contempla uma ação moralmente boa, você pode antecipar uma resposta afetiva positiva, e você pode agir de acordo.

Meu tipo de teoria da sensibilidade também vai além. A maior parte das discussões filosóficas do internalismo discute apenas dois tipos de comportamento: fazer o bem e evitar o mal. Mas, uma vez que tenhamos identificado as emoções específicas que constituem nossos conceitos morais, podemos começar a fazer outras predições comportamentais altamente específicas. Emoções diferentes têm consequências motivacionais diferentes, e um signo do conceito de ERRADO nos motivará diferentemente, dependendo de qual emoção se manifesta em uma ocasião em particular.

Para ilustrar, deixe-me focar em quatro das emoções que são cruciais para o modelo CAD: ira, repulsa, culpa e vergonha. Essas emoções são associadas, na literatura da psicologia, a diferentes estados motivacionais. No modelo que estou defendendo aqui, o conceito de ERRADO irá tipicamente se manifestar na emoção de ira, quando surpreendermos outra pessoa roubando, causando dano ou agindo injustamente. Esses são crimes contra pessoas, e a ira é a resposta dominante. Ira é associada à agressão, então o modelo atual prediz que moralizadores poderiam ser motivados a agredir aqueles que cometem crimes contra pessoas. Por exemplo, o modelo poderia prever que pessoas terão atitudes retributivas para com aqueles que causam danos a outros. Isso está de acordo com descobertas empíricas. Por exemplo, Carlsmith et al. (2002) relata que as pessoas tendem a endossar uma "retribuição justa" ao invés de dissuasão ao escolher punições para criminosos violentos.

Agora, suponha que você mesmo tenha perpetrado um crime contra outra pessoa. O modelo prevê que você sentirá culpa. Do ponto de vista comportamental, a culpa é associada a um comportamento conciliador. Culpa nos impele a reparar apegos ameaçados (Baumeister et al. 1994). Há duas maneiras de se obter conciliação se você causou dano: uma é confessar seu delito, outra é fazer reparos realizando algo bom. Estritamente falando, essas estratégias de enfrentamento não são incompatíveis, mas há alguma evidência de que, se você confessar, você se sentirá menos obrigado a fazer reparos. Para demonstrar isso, Harris et al. (1975) organizaram um estande de doações para a Fundação *March of Dimes* próximo a um confessionário de uma Igreja Católica. Eles descobriram que as pessoas estavam duas vezes mais inclinadas a doar se elas estivessem indo confessar do que se elas já tivessem confessado.

Ira e culpa são respostas a crimes contra pessoas. Repulsa e vergonha são respostas a violações a regras que têm a ver com a ordem natural, e elas têm diferentes efeitos sobre a motivação. Repulsa é, do ponto de vista comportamental, associada à rejeição e ao afastamento. Não queremos ter contato com qualquer um que nos cause repulsa. Isso sugere que as pessoas terão fortes tendências de afastamento quando julgarem, por exemplo, que outra pessoa se engajou em um

comportamento sexual imoral. Se alguém é acusado de bestialidade, não o queremos perto de nós, e podemos instar o sistema penal a manter aquela pessoa trancafiada e longe de nossos vizinhos. Ironicamente, esse modelo prevê que as pessoas podem preferir viver perto de um ladrão do que de uma pessoa que mantém relações sexuais com animais, mesmo que o ladrão seja uma ameaça maior. A emoção "vergonha" surge quando você comete um crime contra a natureza. Do ponto de vista comportamental, vergonha é associada ao desvio do olhar, ao manter a cabeça baixa, e outros esforços para se esconder do olhar dos outros. Pode-se esperar ver tais mudanças corporais em uma pessoa que considera seu próprio comportamento sexual como imoral. Tal pessoa seria mais propensa a evitar os outros do que a buscar conciliação pela confissão ou pela realização de compensações.

Essa discussão ilumina a milhagem explanatória que não podemos deixar de lado ao explorar as emoções específicas que contribuem para o juízo moral. O aspecto central de minha abordagem é que eu distingo emoções de sentimentos, e eu argumento que nossos valores morais estão fundados em sentimentos em nossa memória de longo prazo, os quais se manifestam mediante diferentes emoções de uma maneira sensível ao contexto. Minha abordagem é uma teoria da sensibilidade, mas outros teóricos da sensibilidade não distinguem emoções de sentimentos, e eles são geralmente vagos sobre quais emoções nossos juízos morais incluem (expressivistas são, em geral, mais específicos; ver especialmente Gibbard, 1990). Wiggins (1987) segue Hume ao falar sobre sentimentos de aprovação, sem falar em que tais sentimentos consistem. McDowell (1985;1987) compara a relação entre juízos morais e fatos morais com a relação entre entretenimento e a propriedade de ser entretido, bem como entre medo e a propriedade de ser amedrontado, mas, tal como Wiggins, ele não especifica quais emoções dão suporte à moralidade. McDowell e Wiggins talvez suponham que as emoções morais são *sui generis*, enquanto eu penso que as emoções morais têm um fundamento não moral. Como veremos abaixo, McDowell e Wiggins também endossam uma abordagem metacognitiva, de acordo com a qual valorar é uma questão de *acreditar que as emoções são apropriadas*. Minha abordagem é de

primeira ordem. Há outras teorias disposicionais de valor na literatura, mas recentes defensores do disposicionalismo não tendem a definir valores em termos de respostas emocionais. Firth (1952) evoca a aprovação (a qual pode ser ou não uma emoção), David Lewis (1989) evoca desejos de segunda ordem, Dreier (1990) se refere a "estados motivacionais", e Smith (1994) equipara o bom com o racionalmente recomendável.

Minha visão compartilha muito em comum com as teorias desenvolvidas pelos moralistas britânicos, ou, mais acuradamente, com os moralistas escoceses, tais como Hutcheson, Hume e Smith. É claro, há muitas diferenças nos pormenores. Uma importante diferença é que eu coloco menos ênfase na simpatia e na benevolência. Para Hutcheson, benevolência é a fonte de toda a virtude; sentimentos de benevolência são o que nos motiva a tratar bem os outros, e tais comportamentos benevolentes são aqueles que descobrimos como moralmente bons. Hume parte de Hutcheson ao colocar a simpatia como mais fundamental. Para Hume, simpatia é a causa da benevolência (somos motivados a ajudar aqueles em necessidade porque simpatizamos com seu sofrimento), e a simpatia também engendra virtudes que não envolvem benevolência. Por exemplo, Hume considera a sagacidade uma virtude porque nós simpatizamos com a satisfação que ela causa. Hume também argumenta que a simpatia, mas não a benevolência, fundamenta a virtude da justiça (a qual Hume geralmente tem em mente como direito à propriedade); temos um prazer simpatético nos sentimentos daqueles cujos direitos à propriedade são protegidos, e um desprazer na dor daqueles cujos direitos são violados, e aqueles sentimentos simpatéticos nos motivam a nos comportar com justiça e a condenar a injustiça.

Eu penso que Hume está certo ao acusar Hutcheson de inflar o papel da benevolência como um motivador moral, mas ele erra igualmente ao inflar o papel da simpatia. A evidência empírica sugere que a simpatia é apenas um motivador moral fraco. Em uma meta-análise, Neuberg et al. (1997) descobriram que a simpatia está apenas modestamente correlacionada com o oferecer ajuda, e é improvável que ela nos motive quando os custos de ajudar são altos. Em contraste, considere a culpa. Em um estudo, Carlsmith e Gross (1969) mostraram que sujeitos

que administraram choques elétricos em outra pessoa foram três vezes mais solícitos em uma subsequente campanha ambiental para arrecadação de fundos do que sujeitos que não administraram os choques, sugerindo que a culpa estimula significativamente o comportamento altruístico. Em outro estudo, Zhong e Liljenquist (2006) pediram a sujeitos que esses se lembrassem de um evento em que eles tivessem se comportado de forma não ética, como forma de induzir à culpa e à vergonha. Em seguida, eles fizeram com que alguns sujeitos limpassem as mãos com lenços umedecidos, enquanto que aos outros não foi dada tal oportunidade. Foi pedido a todos os sujeitos que participassem voluntariamente de outro estudo, sendo que 73.9 por cento dos sujeitos que limpavam as mãos concordaram, ao passo que 40.9 por cento dos que não se limpavam concordaram. Isso sugere que a culpa é um motivador moral forte, e que a eliminação da culpa (nesse caso, literalmente a limpando) pode reduzir a solícitude dramaticamente.

A ira é também um grande motivador moral. No capítulo 1, notei que Lerner et al. (1998) lograram aumentar os desejos punitivos mostrando aos sujeitos vídeos induzindo à ira. Isso tem sido corroborado por outros estudos. Usando um "jogo público dos bens", Fehr e Gächter (2002) mostraram que jogadores mesquinhos causam ira em jogadores cooperadores, e, quando irados, jogadores cooperadores desejam arriscar significativos custos monetários para punir a mesquinhez, mesmo que não haja a possibilidade de que eles sejam compensados por aqueles custos. Tais descobertas sugerem que a simpatia é um motivador menos importante do que outras emoções.

Contudo, emoções como as de culpa e de ira parecem ser suficientes para a motivação moral na ausência de simpatia. Em verdade, há muitos tipos de juízos morais que não parecem depender de forma alguma da simpatia. Considere casos em que a pessoa que julga tem sido vítima de um crime (você não pode simpatizar com você mesmo). Ou considere casos em que a vítima não é importante, mas o crime sim (p. ex., ações que causarão dano ambiental em um futuro distante). Ou considere casos em que um ato moralmente questionável não causa sofrimento à vítima (p. ex., incesto consensual e canibalismo de corpos de pessoas que morreram de causas naturais). Eu também questiono a sugestão de Hume segundo

a qual a simpatia é o fundamento de nossa preocupação com a justiça. Simpatia pode ser um impedimento para a justiça, pois, como Hume observou, nós frequentemente simpatizamos mais com os que são próximos e nos são caros. Para superar essa propensão, Hume nos encoraja a tomar um ponto de vista geral, mas isso pode ser algo muito difícil. Pode ser comparativamente fácil condicionar as pessoas a sentirem diretamente ultraje diante da desigualdade mesma, sem que elas tenham que sentir qualquer sensação de eco emocional com relação às vítimas da injustiça. Todavia, Hume talvez possa ter levado a causação para a direção errada: em muitos casos, pensar sobre a injustiça pode causar simpatia, ao invés do contrário.

Em suma, eu penso que uma adequada compreensão da motivação moral requer uma investigação cuidadosa, empiricamente informada, das emoções específicas que contribuem para o juízo moral. Eu penso que os moralistas britânicos podem ter colocado muita ênfase na simpatia e na benevolência, não tendo colocado ênfase suficiente na ira ou na culpa. Em capítulos posteriores, partirei dos moralistas britânicos adotando outros caminhos, rejeitando uma ética das virtudes (no capítulo 4), defendendo o relativismo (capítulos 5), e negando a pretensão de Hutcheson de que há um senso moral inato (no capítulo 7). Por ora, eu quero classificar as diferenças e enfatizar as similaridades. Tal como os moralistas britânicos, eu penso que juízos morais têm uma base emocional, e a motivação moral deriva daquelas emoções subjacentes. Eu tomo isso como a mais importante contribuição da filosofia moral britânica, e os teóricos da sensibilidade são herdeiros dessa tradição.

O restante desse capítulo é dedicado à defesa. As teorias da sensibilidade têm sido amplamente criticadas. Eu quero mostrar que minha versão pode enfrentar a crítica. Eu considero, aqui, dez objeções. Algumas outras objeções serão abordadas nos capítulos seguintes. Ao longo do caminho, oferecerei alguns refinamentos da teoria que delineei nessa seção.

3.2 DEFENDENDO A SENSIBILIDADE

3.2.1 A objeção expressivista

As teorias da sensibilidade são similares, em alguns aspectos, às teorias expressivistas, de tal forma que, ao argumentarmos em defesa de uma teoria da sensibilidade, faz-se importante ver se o expressivismo é uma alternativa melhor. A principal diferença entre as abordagens é semântica: expressivismo é a visão segundo a qual os conceitos morais não denotam coisa alguma; eles meramente expressam as atitudes daqueles que os usam. O emotivismo é uma versão dessa abordagem (Stevenson, 1937; Ayer, 1952). Os emotivistas dizem que sentenças do tipo "matar é errado" deveriam ser interpretadas como "vaias para o matar!" Esse expletivo não pode ser nem verdadeiro nem falso; ele meramente expressa um sentimento. Nas últimas décadas, formas mais sofisticadas de expressivismo têm surgido em cena (Blackburn, 1984, Gibbard, 1990). Gibbard, por exemplo, analisaria "matar é errado" como algo do tipo, "eu aceito uma norma que condena ou permite culpa ou ultraje se eu ou alguém mais for capaz de matar sem motivação suficiente". As teorias expressivistas e as teorias da sensibilidade são primas próximas, pois ambas implicam as emoções em conceitos morais; ambas são formas de emocionismo epistêmico. A teoria da sensibilidade que endosso está também comprometida com a visão de que o discurso moral ordinário é expressivo. Símbolos de conceitos morais são normalmente constituídos por emoções, assim, quando eu digo, "matar é errado", eu estou tendo uma resposta tipicamente emocional, e essa emoção é expressa, de alguma maneira, em minha asserção. A diferença mais importante entre a teoria da sensibilidade e o expressivismo é que os expressivistas tradicionalmente negam a existência de propriedades e de fatos morais. Eles não pensam que juízos morais sejam passíveis de serem verdadeiros (Gibbard, 2003, talvez seja uma exceção, mas sua visão, a esse respeito, afasta-se do expressivismo). Como vimos no capítulo 1, Hume fomentou tais dúvidas. Em contraste, os teóricos da sensibilidade são realistas. Mais acuradamente, eles são realistas internos: propriedade morais existem, mas elas dependem de nós. Os

expressivistas desafiam os teóricos da sensibilidade a defender essa pretensão. Se juízos morais expressam como nos sentimos sobre as coisas, o que acrescenta dizer que há fatos morais? Isso não complica nossa ontologia adicionando um nível de realidade que é fundamentalmente diferente de rios, rinocerontes e ferrovias? Por que multiplicar propriedades além do necessário?

A resposta é que propriedades morais vêm de graça. Nossas melhores teorias da intencionalidade dizem que qualquer representação mental representa suas autênticas causas. Se sentimentos morais estão seguramente ligados a assassinatos, a assaltos e a molestamentos, então aquelas coisas caem dentro do âmbito dos conceitos morais. Não se faz necessário introduzir qualquer maquinaria semântica extra para se chegar a esse resultado. Simplesmente se sai de uma teoria semântica que tenha sido motivada independentemente para explicar a representação mental em geral. Não há, aqui, qualquer ofensa a Ockham. Algumas coisas causam o engajamento de nossos sentimentos morais. E esse é um fato que expressivistas não podem negar. Logo, há uma propriedade de causar o engajamento de nossos sentimentos morais. Certas formas de conduta têm aquela propriedade. Elas têm o poder de fazer com que nos sintamos culpados, irados ou envergonhados. Dado que os expressivistas estão igualmente comprometidos com a existência dessa propriedade, a diferença entre a visão deles e a teoria da sensibilidade não é uma diferença envolvendo parcimônia. É uma diferença envolvendo a semântica. Teóricos da sensibilidade defendem que conceitos morais representem essas propriedades reais, e os expressivistas discordam. Teóricos da sensibilidade vencem essa disputa, pois as teorias semânticas predominantes endossam sua posição.

Expressivistas poderiam replicar argumentando que as teorias semânticas predominantes são elaboradas para explicar como conceitos denotam tipos naturais, e não para explicar como conceitos denotam propriedades dependentes de resposta. Eles podem pretender que os teóricos da sensibilidade não podem se aproveitar das teorias semânticas predominantes, e devem, ao invés disso, admitir teorias semânticas *ad hoc* para explicar como os conceitos morais denotam.

Eu não acho essa resposta convincente. Primeiro, deve haver uma teoria viável sobre como conceitos denotam propriedades dependentes de resposta, pois tais conceitos obviamente existem. Exemplos incluem a ADORÁVEL, o ASSUSTADOR e, talvez, o DIVERTIDO, bem como o VERMELHO. Teóricos da sensibilidade podem dizer que conceitos morais denotam da mesma forma que esses conceitos denotam, seja como for. Segundo, a teoria dreetskiana da denotação que endosso é suficientemente geral para abranger conceitos que denotam propriedades dependentes de resposta. Segundo essa teoria, conceitos denotam aquilo que eles têm a função de credivelmente identificar. Possivelmente, a única propriedade que assassinatos, assaltos e molestamentos têm em comum é sua capacidade de causar certas emoções em nós. Logo, é plausível que conceitos morais denotem essa propriedade dependente de resposta, pois essa é a causa mais confiável de nossos conceitos morais (para uma discussão mais circunstanciada, ver capítulo 4). Não há uma barreira evidente para aplicarmos a teoria semântica dreetskiana aos conceitos morais.

Além disso, é importante lembrar que a alternativa expressivista não é atraente. Se conceitos morais nada denotassem, então seria difícil dar sentido ao discurso moral. Afinal de contas, falamos como se houvessem fatos morais, e usamos proposições morais assertóricas em conjunção com proposições não morais. Considere o seguinte argumento: se matar é um mal, então Jack, o estripador, era mau; se Jack, o estripador, era mau, então ele está no inferno; se ele está no inferno, ele está, nesse momento, sofrendo; logo, se matar é algo mal, Jack, o estripador, está, nesse momento, sofrendo. Para que esse argumento funcione, todas essas sentenças precisam ser tratadas como padrões indicativos condicionais. Geach (1965) celebrenemente argumentou que expressivistas não podem resgatar essa exigência. Se o expressivismo está certo, as sentenças morais asseridas no indicativo são secretamente expletivas: elas realmente expressam as convicções do falante. Se traduzimos sentenças morais em sentenças expletivas, argumentos como aquele recém apresentado ruirão. Sentenças que parecem estar ordinariamente no indicativo condicional se transformam em monstruosidades gramaticais. Os expressivistas têm tentado enfrentar esse desafio criando sistemas

lógicos que tratem das sentenças morais como se elas estivessem no indicativo, ainda que elas não estejam (Blackburn, 1984, 1988; Gibbard, 1990). Alguns desses lógicos podem ficar às voltas com o problema técnico, mas podem querer se perguntar que direito os expressivistas têm de se servirem de lógicas indicativas que tratam das sentenças morais como pseudoindicativas. E, mais pontualmente, se podemos encontrar uma teoria segundo a qual as sentenças morais são realmente indicativas, ao invés de pseudoindicativas, não deveria isso ser preferível? Teorias da sensibilidade oferecem exatamente isso. Elas estão muito próximas do expressivismo ao enfatizar a natureza emocional dos juízos morais, mas elas dizem que juízos morais denotam fatos reais. Quando expressivistas questionam as pretensões semânticas subjacentes às teorias da sensibilidade, eles eles atiram pedras tendo tetos de vidro. Teóricos da sensibilidade podem contar com teorias semânticas motivadas independentemente, enquanto expressivistas são forçados a construir novos sistemas lógicos misteriosos para explicar o caráter aparentemente denotante do discurso moral.

3.2.2 As objeções das cores

Os expressivistas não vão simplesmente aceitar essas queixas. Eles pensam que as teorias da sensibilidade têm falhas fatais. Se eles estão certos, então talvez devamos, afinal, tomar a rota expressivista, junto com o drástico programa de uma revisão lógica. Mas, quais são essas falhas fatais? Uma família de preocupações envolve a tese segundo a qual as propriedades morais são qualidades secundárias (McDowell, 1985; Wiggins, 1987). Essa é a intuição fulcral da teoria da sensibilidade, e ela jaz na pretensão de que as propriedades morais parecem se comportar de maneiras sugestivamente similares às propriedades tomadas amplamente como qualidades secundárias, especialmente as cores. A analogia com a cor é instrutiva, como temos visto, mas é, também, arriscada. Críticos têm identificado várias formas nas quais a moral difere muito das cores, e aqueles contrastes podem ser usados para levantar dúvidas sobre a totalidade do programa. Eu concordo que a analogia com a cor é deficiente. Mas analogias são, geralmente, defeituosas, e os

teóricos da sensibilidade estão cientes disso. Eu penso que diferenças entre moral e cores poderiam, em verdade, ser previstas pela teoria da sensibilidade que venho defendendo. Para apresentar esse caso, quero considerar seis desanalogias que são enfatizadas por Blackburn (1985) em uma crítica a Wiggins e McDowell. Focarei em alegadas desanalogias entre moral e cores.

Primeiro, Blackburn pretende que propriedades morais sobrevivem às qualidades naturais ou primárias de uma maneira muito diferente da forma como as cores sobrevivem. A relação entre cores e suas bases em objetos físicos é uma verdade científica, e não conceitual. Suponha que o amarelo seja a força para causar experiências amarelas em nós e que aquela força é possuída por certos comprimentos de onda no espectro eletromagnético. Alguém poderia ter o conceito de AMARELO sem saber que o amarelo sobrevém a esses comprimentos de onda (junto com fatos sobre nossos sistemas visuais); alguém poderia conceber um mundo no qual esses mesmos comprimentos de onda (bem como nossos sistemas visuais) existissem, mas no qual eles não fossem amarelos. Mas ninguém que possua o conceito de ERRADO poderia conceber um mundo no qual, digamos, o matar recreativo não seja errado.

Essa é uma intrigante assimetria, mas ela não solapa o ponto de vista da qualidade secundária. Pelo contrário, ela sugere que cores talvez não possam ser uma analogia perfeita para a moral. Mas cores não são os únicos exemplos plausíveis de qualidades secundárias. Considere a repugnância. Possivelmente, a repugnância é uma qualidade secundária. Ela é a força para causar um certo sentimento em nós. Na verdade, ela é, precisamente, o tipo de qualidade secundária ao qual os teóricos da sensibilidade deveriam apelar ao defenderem sua teoria. Se propriedades morais são dependentes de respostas emocionais, elas devem se comportar da forma que as outras propriedades dependentes de emoção se comportam. Agora, considere como a repugnância sobrevém ao mundo. Não procuramos a ciência para que ela nos diga o que é a repugnância. E assumimos que as coisas que nos causam repugnância neste mundo seriam repugnantes em qualquer mundo. Você pode imaginar que uma sopa de vermes seja algo além de repugnante? Nesses aspectos, a repugnância parece se comportar como o erro.

Eu também quero enfatizar que Blackburn exagera na extensão da crença das pessoas acerca de que a moralidade sobrevém ao mundo. Ele diz não podemos imaginar que, digamos, o homicídio recreativo não seja errado em alguns mundos. Isso é parcialmente verdadeiro e parcialmente falso. O caso da repugnância é, novamente, ilustrativo. Alguém pode, após ser exposto à multiplicidade cultural, vir a reconhecer que a mesmíssima coisa (mesmo uma sopa de vermes) pode ser repugnante para algumas pessoas e não para outras. Uma pessoa que tenha esse discernimento viria a reconhecer que a repugnância não sobrevém ao mundo externo apenas, mas também a nós. Tal pessoa poderia, em algum sentido, conceber um mundo em que uma sopa de vermes seja saborosa. O mesmo pode ser dito sobre a moralidade. Após fazer o levantamento dos registros antropológicos, pode-se aperceber que alguns grupos culturais (talvez os caçadores de cabeça Llongot, da ilha de Luzon) não pensam que o assassinato recreativo seja errado. Isso pode conduzir alguém a adotar a visão de que a moralidade não sobrevém aos meros fatos externos, mas também depende de nós. Mas tanto no caso dos vermes quanto do assassinato, há um limite para o que podemos maginar. Podemos reconhecer, de uma forma tranquila e intelectual, que há mundos nos quais ninguém é repelido por sopa de vermes ou por assassinato, mas não podemos conceber essa possibilidade imaginativamente, ainda que em uma simulação em primeira pessoa. Se sopa de vermes causa repugnância em você, você não será capaz de se imaginar pensando que a sopa de vermes é deliciosa (embora você possa se imaginar dizendo essas palavras). Da mesma forma, você não seria capaz de se imaginar achando aceitável seguir em uma matança. Em ambos os casos, a estabilidade das próprias respostas emocionais limita a extensão até aquela em que você possa realmente imaginar mundos possíveis nos quais as mesmíssimas coisas externas mudam seus status valorativos. Esse limite sobre a imaginação promove a impressão de que fatos valorativos sobrevém ao mundo externo. Blackburn concordaria, eu penso, mas ele falha na apreciação de que esse ponto, sobre a aparente superveniência das propriedades morais, é consistente com uma visão sobre a qualidade secundária.

Agora, voltemos-nos para o segundo contraste de Blackburn. Ele diz que cores requerem mecanismos receptores especiais que possam ser seletivamente dificultados. Cores são vistas pelos olhos. Não há órgão da sensação moral, e assim se desfaz a analogia.

Assim como com a superveniência, a analogia com as cores se desfaz nesse ponto, mas não a comparação com as emoções. Há um mecanismo receptor para as emoções? Sim e não. Por um lado, se a teoria da avaliação corporificada está certa, emoções são sensações de mudanças no padrão corporal, e elas envolvem os receptores que monitoram eventos somáticos. Por outro lado, as preocupações que disparam nossas respostas emocionais podem ser representados de várias formas distintas. Podemos descobrir um perigo mediante a visão, a audição, o olfato ou pela luz da razão. Neste aspecto, não há órgão sensorial das paixões. Exatamente o mesmo pode ser dito sobre a moral. Sentido algum é dedicado a perceber questões de relevância moral, mas uma vez que tais questões tenham sido identificadas, elas nos perturbam de formas que podem ser percebidas mediante nossos sentidos corporais.

O terceiro contraste de Blackburn envolve as condições sob as quais as cores poderiam desaparecer. Ele diz que o azul poderia deixar de existir se nossas mentes mudassem de tal forma que todas as coisas que uma vez nos atingiram como azuis aparecessem, agora, como vermelhas. Em contraste, ele diz que o mau trato a animais poderia seguir sendo errado, mesmo que começássemos a considera-lo permissível. Segundo entendo isso, esse contraste está intimamente relacionado com o primeiro ponto de Blackburn sobre a superveniência. Minha réplica aqui ecoa minha réplica àquela objeção. Segundo as intuições *desinformadas*, nós realmente pensamos que o azul não sobreviveria a qualquer mutação em nossos sistemas visuais. Similarmente, intuições *desinformadas* nos levam a pensar que a crueldade com animais seria errada independentemente de nossas inclinações. Essas últimas intuições são recalcitrantes porque acreditamos ser difícil imaginar a crueldade contra animais sem sentir uma pontada de horror.

Em seguida, Blackburn contrasta as formas pelas quais as cores e a moral poderiam estar culturalmente ligadas. Pessoas formadas em diferentes culturas

poderiam vir a classificar as cores ligeiramente diferente, mas ninguém argumentaria que os termos de cores usados em outras culturas estão errados. Há muitas formas compatíveis de organizar o espaço da cor. Em contraste, quando a moral varia pelas culturas, os valores divergentes são vistos como incompatíveis, e os debates morais são impossíveis de se resolver. Para usar uma famosa frase da filosofia, os debates morais são "essencialmente contestáveis" (ver Wiggins, 1987: n. 19).

As pressuposições de Blackburn sobre o relativismo das cores e da moral são, ambas, altamente litigiosas, mas eu estou preparado para fazer concessões a ambas. Essa é uma genuína desanalogia entre cores e moral, mas note que uma analogia forte pode ser esboçada entre os debates morais e os debates que surgem de sentimentos divergentes. Se uma pessoa fóbica insiste que insetos são assustadores, há pouco que alguém possa fazer para dissuadi-la. Da mesma forma, se você gosta de música *country*, e eu não, você terá muita dificuldade para me fazer concordar que música *country* é agradável. Ser assustador ou ser agradável são qualidades secundárias. O fato de os debates morais serem essencialmente contestáveis não ameaça a pretensão de que os conceitos morais representem qualidades secundárias. Ele tão somente mostra que os conceitos morais se comportam como os conceitos de ASSUSTADOR e de AGRADÁVEL, e isso é exatamente o que a teoria da sensibilidade prevê.

Em um quinto ponto de contraste, Blackburn argumenta que podemos ser indiferentes às cores, mas que não podemos ser indiferentes à moral. Ele está bem certo ao enfatizar esse ponto, mas isso tem um papel nas mãos do teórico da sensibilidade. Se a moral é constituída pelas nossas respostas emocionais, então deveria ser muito difícil ser indiferente a ela.

O último contraste de Blackburn é que os predicados valorativos, diferentemente dos predicados sobre cores, são tipicamente adjetivos. Em outras palavras, "x é bom" implicitamente significa "x é bom para esta e para aquela categoria". Portanto, alguém pode ser um bom presidente e um mau pai. Em contraste, Blackburn o reivindica, coisa alguma pode ser um tomate vermelho sem ser um objeto vermelho. Eu não concordo com essa análise. Eu penso que

predicados de cor são adjetivos. Cabelo vermelho não é uma coisa vermelha, e, da mesma forma, isso vale para marés vermelhas, granito vermelho, ouro vermelho, barro vermelho e areia vermelha. Blackburn seleciona cuidadosamente seu exemplo (por assim dizer) escolhendo tomates vermelhos. Tomates podem exibir cores tão saturadas que eles poderiam ser classificados como vermelhos sob qualquer categoria de objetos à qual eles pertençam.

Eu concluo que as alegadas desanalogias entre cores e moral não colocam qualquer ameaça às teorias da sensibilidade. Alguns dos aparentes contrastes desaparecem em um exame atento, e outros podem ser explicados. Quando a analogia da cor se desfaz, há ainda uma forte analogia entre conceitos morais e conceitos que denotam qualidades secundárias tais como as de assustador e agradável. Teóricos da sensibilidade não deveriam se ater tanto à comparação entre propriedades morais e cores. Ao invés, eles deveriam comparar propriedades morais com propriedades dependentes da emoção. Afinal, essas devem ser as propriedades morais em uma teoria da sensibilidade.

3.2.3 A objeção do mérito

Há outra aparente desanalogia entre cores e moral, a qual merece atenção especial. Geralmente, assume-se que algo tem uma certa cor se ele causa um certo tipo de experiência de cor. Mas um modelo causal de propriedades morais parece problemático. Parece que causar desaprovação não é nem necessário, nem suficiente para ser errado. Imagine um atirador que dispara contra multidões, o qual está passando por um processo de regeneração. Ele talvez não experiencie qualquer sentimento de desaprovação em relação ao assassinato, mas ele, não obstante, julga que é errado assassinar. Isso indica que sentimentos não são necessários. Ou considere uma pessoa que foi criada para pensar que a homossexualidade é errado, mas que agora rejeita aquela visão. Ela pode ainda abrigar sentimentos negativos incontroláveis quando pensa sobre a homossexualidade, ainda que ela insista que a homossexualidade é moralmente aceitável. Isso sugere que sentimentos não são suficientes para juízos morais.

É tentador acomodar esses contraexemplos partindo de um modelo meramente causal das relações entre propriedades morais e os sentimentos morais. McDowell (1985) escreve: “A desanalogia, agora, é que uma virtude (digamos) é concebida não para ser meramente algo que incite a uma atitude apropriada (como a cor é meramente algo que causa as experiências apropriadas), mas, ao invés disso, algo que nos torne merecedores dela”. Um ponto paralelo é feito por Wiggins (1987), e algum tipo de apelo ao mérito é agora considerado como um aspecto das teorias contemporâneas da sensibilidade (Darwall, et al. 1992; D’Arms e Jacobson, 2000). Mas que faremos sobre o contraste entre incitar e merecer? A sugestão mais popular é que a moralidade não requer respostas emocionais atuais, mas, ao invés disso, normas sobre quais respostas morais deveríamos ter. Mais especificamente, teóricos contemporâneos da sensibilidade têm se mostrado propensos a trocar (S1) por esta alternativa metacognitiva:

(S1-M) Uma ação tem a propriedade de ser moralmente correta (errada) apenas no caso de os observadores *a estimarem como apropriada* para envolver sentimentos de aprovação (desaprovação).

Correspondentemente, um juízo moral, sob essa perspectiva, não é, em si mesmo, uma resposta emocional, mas é, ao invés, um juízo sobre o efeito que seria apropriado a uma resposta emocional. Em outras palavras, símbolos de conceitos morais *aludem* a emoções, ao invés de as *usarem*. McDowell e Wiggins parecem endossar perspectivas dessa espécie, e Gibard (1990) tem defendido uma proposta próxima a essa. A alternativa metacognitiva acomoda o antigo atirador e o antigo homofóbico. O antigo atirador não sente desaprovação em relação ao assassinato, mas ele estima a desaprovação como apropriada. O antigo homofóbico sente desaprovação em relação à homossexualidade, mas ele não estima a desaprovação como apropriada.

(S1-M) poderia ser uma revisão tolerável da teoria se não fosse um irritante problema com a questão da apropriabilidade. Podemos estimar algo como apropriado de diversas formas. Por exemplo, algo pode ser apropriado porque ele garante a realização da tarefa, ou porque ele permitirá evitar constrangimento, ou

ele é apropriado porque foi o que outros fizeram no passado, e assim por diante. Que tipo de apropriabilidade é requerida para (S1-M)? A resposta natural é uma apropriabilidade moral: roubar torna a desaprovação *moralmente* apropriada. Mas isso introduz um círculo vicioso. Conceitos morais não podem ser definidos como crenças sobre quais emoções são moralmente apropriadas.

D'Arms e Jacobson (2000) têm desenvolvido uma versão dessa objeção. Eles vêm com casos nos quais diferentes tipos de condutas apropriadas se desfazem. Considere uma piada ofensiva. A piada pode ser boa enquanto piada. Ela pode ser inteligente e divertida. Nesse sentido, ela pode merecer o regozijo. Mas a piada pode ser tão ofensiva que a estimamos como inapropriada para que ela divirta em outro sentido. Qual a diferença entre esses dois tipos de apropriabilidade? Bem, podemos dizer que o divertimento é socialmente inapropriado e comicamente apropriado. O apelo à apropriabilidade cômica parece suficientemente inócua, mas ela nos coloca um problema se definimos "divertida" como aquilo que torna a diversão comicamente apropriada. Essa definição nos leva a lugar algum, pois "cômico" e "divertido" são sinônimos próximos. Parece que nos encontramos em uma situação similar quando definimos a moral como aquilo que merece certas respostas emocionais. Suponha que eu queira me manter em uma disputa em que eu sei que estou do lado errado. Culpa poderia ser considerada moralmente apropriada nessa situação, mas taticamente inapropriada. O sentido relevante de apropriabilidade é difícil de ser capturado sem circularidade.

Penso que essa dificuldade deveria ser atacada de início. Não há razão alguma para introduzir um componente metacognitivo na definição de propriedades morais. (S1-M) deveria ser rejeitada. Segundo a visão que endosso, (S1'), algo é errado se há um sentimento de desaprovação com relação a ele. Esse é um aperfeiçoamento em relação a um modelo causal muito simples, pois implica que algo pode estar errado, em alguma ocasião, sem causar efetivamente emoções negativas. Se estou deprimido, eu posso não me sentir indignado com as atrocidades relatadas no noticiário da noite, ainda que eu tenha um sentimento que, em estados menos depressivos, me disporia a sentir indignação. Pode haver momentos, como no caso da indução hipnótica à repulsa, em que eu sou levado à

indignação na ausência de um sentimento. Assim, a atual incitação de uma emoção não é nem necessária, nem suficiente para a erroneidade moral segundo minha visão. Mas as coisas ficam um pouco mais complicadas com o antigo atirador e com o antigo homofóbico. Eu penso que esses casos requerem um pouco mais de escrutínio.

Com respeito ao atirador, sinto-me tentado a dizer que, se ele realmente não tem um sentimento negativo relativo ao assassinato, então ele não pensa realmente que assassinar seja errado. Ele está apenas oferecendo conversa fiada ao pensar assim. Em comparação, se penso que eu deveria gostar de ópera, mas ópera não incita realmente uma resposta em mim, então eu realmente não gosto disso. Mas o antigo atirador talvez tenha algo que se aproxime do juízo moral conforme o qual assassinar é errado. Ele talvez tenha um juízo moral para o efeito de que é errado aprovar o assassinato. Ele talvez se sinta culpado quando não se sente culpado sobre o assassinato. Portanto, ele ainda não acredita realmente que assassinar seja errado, mas ele acredita que ele está errado por não ter essa crença, e esse juízo moral metacognitivo irá derradeiramente leva-lo a adquirir sentimentos necessários a um juízo em primeira pessoa. Minha versão de (S1) me permite esboçar a distinção entre moralização de primeira e de segunda ordem. Essa distinção está perdida na versão metacognitiva de (S1).

O antigo homofóbico convida a uma análise similar. Talvez a pessoa que esteja rejeitando sua formação homofóbica ainda seja, durante esse período de transição, homofóbica; se ela sente repulsa pelo comportamento homossexual, estaríamos autorizados a dizer que ela abriga uma atitude moral negativa em relação à homossexualidade. Mas, ela talvez acredite que essa atitude é injustificada, e ela talvez esteja trabalhando duro para muda-la. Nesse sentido, podemos dizer que ela é uma homofóbica em recuperação (tal como ocorre com um alcoólatra em recuperação). Ela condena o fato de que ela condena a homossexualidade, e aquela atitude metacognitiva irá, podemos ter esperança, derradeiramente erodir suas disposições de primeira ordem. Isso repercute em mim como a maneira correta de descrever o caso, e a versão metacognitiva de (S1) é menos equipada para descrevê-lo acuradamente.

Eu concluo que não há razão para adotar a visão metacognitiva, e muitas razões contra ela. Ela se faz presa das preocupações erguidas por D'Arms e Jacobson, bem como oblitera a distinção entre moralização de primeira e de segunda ordem.

Essa linha de resposta deixa uma questão sem resposta. Ainda não expliquei a intuição de McDowell, segundo a qual o errado *merece* desaprovação. Note que a visão de uma banana não merece uma experiência amarela. Respostas morais, diferentemente das respostas a cores, têm mais do que uma relação meramente causal com as coisas que elas representam.

Como posso acomodar essa diferença entre cores e sentimentos morais? De onde vem essa intuição sobre o mérito? Talvez ela tenha algo a ver com o fato de que o elo entre emoções e suas causas é menos direto do que o elo entre cores e suas causas. No caso da cor, a conexão causal é fixada por nossa biologia. Experienciamos cores da forma em que as experienciamos porque somos construídos de certa forma, e seria difícil, ou mesmo impossível, mudar isso mediante treinamento. No caso das emoções, exercemos mais controle. Em alguns casos, emoções requerem bastante deliberação antes de surgirem. Podemos pensar em uma decisão em política externa por horas antes que ela provoque condenação ou louvor. Em casos nos quais as emoções ocorrem mais espontaneamente, elas ainda são abertas à alteração mediante treinamento. Podemos nos treinar para sermos destemidos quando pulamos de paraquedas, ou para sermos imperturbáveis quando somos insultados. Quando dizemos que algo merece uma emoção, estamos implicando que uma pessoa que falha em ter a emoção poderia ser considerada responsável. Nós não podemos considerar uma pessoa cega para as cores responsável por falhar ao distinguir entre o vermelho e o verde, mas podemos manter que uma pessoa emocionalmente saudável é responsável por temer estrangeiros ou por falhar no temor aos efeitos da fumaça do cigarro.

Se essa análise está certa, então uma resposta emocional pode ser merecida, em oposição a ser *meramente* causada, sem ser estimada como apropriada. Uma resposta psicológica é merecida por sua causa se (a) ela se aplica à sua causa, e (b)

se o agente pode ser considerado responsável, em algum nível, por ter, ou falhar em ter, aquela resposta. Cores não merecem experiências de cor porque elas satisfazem apenas a primeira condição. Elicitores de emoção merecem emoções porque eles caracteristicamente satisfazem também a segunda condição. Respostas emocionais conectadas podem ser exceções a isso (um ruído alto e súbito garante causar espanto?), mas muitas respostas emocionais têm um grau de plasticidade. Isso é verdade, *a fortiori*, para os sentimentos. Objetos de agrado e de desagrado podem mudar ao longo do tempo, mediante treino e experiência. Emoções morais são meritórias por suas causas na medida em que elas se aplicam às suas causas e na medida em que temos alguma responsabilidade por nossas respostas morais. Podemos deliberar mais, adquirir o conhecimento de mais fatos, nos expormos a mais experiências, e nos submetermos a mais treinos. Eu penso que podemos explicar nossas intuições de que o mérito está envolvido na resposta moral sem adotarmos uma versão metacognitiva de (S1). O fato de que ações merecem emoções não garante que tenhamos quaisquer pensamentos metacognitivos sobre quais emoções são apropriadas; ele é derivado meramente do fato de que respostas emocionais podem ser controladas até certo ponto. Pensar que algo é errado não significa estimar emoções negativas como apropriadas; é estar disposto a ter tais emoções. Apenas ocorre de parecer que tais disposições são relativamente plásticas.

A objeção de D'Arms e Jacobson ainda nos deixa com um quebra-cabeça. Mesmo que estimar emoções apropriadas não seja colocada na definição do correto e do errado, ela ainda é um fenômeno que carece de explicação. As pessoas claramente distinguem entre dois tipos diferentes de apropriabilidade. Elas distinguem entre casos nos quais culpa é demandada pela moralidade e quando ela é taticamente inconveniente. Como esboçamos essa distinção?

Penso que a teoria da sensibilidade que endosso pode oferecer uma resposta natural a essa questão sem qualquer circularidade viciosa. Estimar a culpa como moralmente apropriada é ter um sentimento moral em relação a ela. É ter um metassentimento. Às vezes, nos sentimos culpados por não nos sentirmos culpados. Se suprimirmos a culpa com vistas a manter a reputação em um

argumento, podemos sentir culpa sobre a supressão. Emoções morais são frequentemente implementadas por um segundo nível de emoções morais. Isso se encaixa na minha análise do antigo atirador e do homófobo em recuperação. Metassentimentos podem ser usados para mudar valores morais com o passar do tempo. Metassentimentos podem também desempenhar um importante papel na sustentação de normas. Se nossas normas de primeira ordem estão fundadas em metanormas, elas serão mais difíceis de perder.

Essa digressão revela um genuíno contraste entre o caso das cores e o caso da moral. Emoções morais não são meramente causadas; elas são se fazem merecer por suas causas e elas são reguladas por metaemoções. Há um sentido segundo o qual estimamos nossas emoções morais como apropriadas, mas isso não é essencial para aquelas emoções, ou para nosso conceito de moralidade. Ela é, ao invés disso, um mecanismo poderoso para dar sustentação aos nossos sentimentos. Nesse sentido, pensar sobre o mérito revela limitações na analogia com as cores. Mas isso em nada solapa o tipo de teoria da sensibilidade que venho defendendo.

3.2.4 A objeção do desenvolvimento

A abordagem apresentada na seção 3.1 apela aos sentimentos morais. Um sentimento moral é uma disposição para experienciar emoções morais. Nichols (2004a: Cap. 4) desenvolveu recentemente uma crítica empírica às abordagens que apelam a algumas dessas emoções. Ele foca na culpa. O maior golpe de sua crítica é que a culpa se desenvolve depois da competência moral, sendo a competência moral mensurada pela habilidade para distinguir normas morais de normas convencionais. Crianças podem determinar quais normas são morais antes de aprender a fundo a culpa, e desordens que prejudicam o desenvolvimento da culpa não prejudicam necessariamente o desenvolvimento da moralidade.

O alvo principal de Nichols é Gibbard (1990). De acordo com a teoria de Gibbard, considerar uma ação como errada é aceitar uma norma que nos faculta ou proíbe sentir culpa quando performamos aquela ação (e ira ou ressentimento

quando alguém outro a performa). A teoria de Gibbard é metacognitiva. Moralizadores precisam ter atitudes sobre a culpa e outras emoções morais. Se eles não compreendem essas emoções, então eles não podem ter aquelas atitudes. Nichols tenta mostrar que algumas pessoas, incluindo crianças pequenas, se moralizam antes de compreenderem aquelas emoções.

Nichols usa a sensibilidade para a distinção moral/convencional como medida da competência moral. Smetana e Braeges (1990) têm mostrado que crianças de dois anos e de dez meses já distinguem entre erros morais e convencionais. Eles consideram erros morais como comparativamente independentes de autoridades. Crianças não parecem compreender a culpa até que estejam mais velhos. Nichols cita um estudo feito por Nunner-Winkler e Sodian (1988) que parece dar suporte a essa conclusão. Em um de seus experimentos, eles perguntaram a crianças como uma criança se sentiria se ela empurrasse outra criança de um balanço. As crianças menores de seis anos disseram que a criança que agrediu deveria se sentir feliz. Nichols e os autores concluem que crianças menores de seis anos não compreendem a culpa. Se eles estão certos, compreensão da culpa vem vários anos depois da emergência da competência moral básica.

Nichols também argumenta em prol de uma dissociação entre compreensão da culpa e competência moral em adultos autistas. Pessoas com autismo são sensíveis à distinção moral/convencional (Blair, 1995), mas eles sabidamente têm deficiências em sua compreensão de certas emoções. Estudos têm mostrado que pessoas com autismo não compreendem totalmente orgulho e constrangimento, e Nichols infere disso que eles provavelmente também têm dificuldades para compreender a culpa. Essa é uma inferência razoável. Culpa é uma emoção fundamentalmente interpessoal; ela surge em contextos nos quais acreditamos ter causado dano a outra pessoa, e isso promove esforços para reparar a relação com a pessoa lesada. Pessoas com autismo têm dificuldades para entender o domínio interpessoal, e isso pode limitar sua habilidade para reconhecer e raciocinar sobre a culpa.

Nichols coloca um sério desafio a Gibbard e a outros que pensam que a posse de conceitos morais requer a habilidade de formar crenças sobre emoções morais. Gibbard poderia replicar dizendo que pessoas com autismo e crianças pequenas não possuem conceitos morais. Ele poderia dizer que a distinção moral/convencional não é suficiente para testar a mestria no uso dos conceitos morais. Mas essa réplica é *ad hoc*. Pessoas com autismo e crianças pequenas são relativamente boas na identificação de transgressões morais, e elas tendem a se comportar de acordo com as regras morais.

Felizmente, os argumentos de Nichols não ameaçam o tipo de teoria da sensibilidade que eu endosso. Como insisti na seção 3.2.3, minha teoria não é metacognitiva. Algo está errado se há um sentimento de primeira ordem contra ela. Uma pessoa pode ter atitudes morais sem apreender conceitos que denotam emoções morais. O crucial não são os conceitos de emoções, mas as emoções elas mesmas. Moralizadores devem estar dispostos a sentir culpa quando cometem alguma transgressão. Não se requer que eles tenham a crença de que a culpa é apropriada.

A evidência experimental que Nichols discute não demonstra uma falta de emoções morais em crianças ou pessoas com autismo. Há comparativamente poucos dados disponíveis sobre emoções morais no autismo, mas há muitos estudos de emoções morais em crianças. Esses estudos confirmam, reiteradamente, que crianças têm emoções que se assemelham às emoções morais de adultos (ver, para um exame, Eisenberg, 2000). Eles sentem-se mal quando fazem coisas ruins contra outros, por exemplo. Michael Lewis (1998) argumentou que a culpa emerge em torno do terceiro aniversário, mas ela pode aparecer, efetivamente, mais cedo. Comportamentos reparadores já estão presentes no segundo ano, especialmente quando a criança causa sofrimento em que a está cuidando (Hoffman, 1998). No capítulo 1, eu descrevi experimentos que inquiram a culpa precoce usando o paradigma da má sorte (Barrett et al., 1993; Kochanska et al., 2002). Crianças recebem um objeto (uma boneca ou uma camiseta) que se diz ter valor especial para seu dono. As crianças são advertidas a terem cuidado com o objeto, mas uma situação é armada para que ocorra a má sorte (a boneca quebra

ou respinga tinta na camiseta). Kochanska et al. (2002) descobriram que crianças jovens com vinte e dois meses mostram sinais de culpa quando a má sorte ocorre; elas se retorcem, esquivam o olhar, cobrem seus rostos, e assim por diante. Barrett et al. (1993) distingue sinais de culpa dos de vergonha, e percebe evidência de ambas em crianças com dois anos. Algumas crianças exibem um comportamento evasivo, indicativo de vergonha, e outras tentam reparar e se desculpar, o que indica culpa. Isso sugere que culpa e vergonha (ou precursores dessas emoções) estão disponíveis bem cedo na vida.

Alguém pode jogar com as palavras e questionar sobre se as respostas emocionais recém descritas se qualificam como instâncias *bona fide* de culpa. Eu não penso que isso importe. As emoções no espectro da desaprovação deveriam ser, na verdade, descritas mais amplamente. Essas emoções podem variar um pouco de cultura para cultura, e nas diversas faixas etárias. Elas estão unidas por uma similaridade familiar. Em todas as culturas parece haver emoções de autocensura e outras emoções de vergonha. O caráter e escopo delas podem diferir de formas sutis. Por exemplo, nas chamadas culturas da vergonha, reações semelhantes à vergonha podem ser mais predominantes, e o conteúdo semântico daquela resposta pode ser mais inclusivo do que as espécies de vergonha que temos em nossa cultura (ver Prinz, 2004). Porém, ainda podemos identificar essas emoções como pertencendo à família das emoções semelhantes à vergonha, caracterizadas por uma resposta aversiva à atenção dos outros. Pode haver mais uma família de emoções similares à culpa sutilmente diversificada, cada uma delas podendo ser caracterizada como sofrimento causado por dano a outra pessoa. Crianças com autismo não exibem os característicos sinais de culpa, mas elas exibem, sim, emoções aversivas quando violam regras (Peter Hobson, comunicação pessoal). É como se elas tivessem um análogo da culpa que tivesse sido despido de seu caráter interpessoal. Onde uma criança em desenvolvimento normal talvez tente se engajar emocionalmente com a pessoa à qual ela causou dano, uma criança autista poderá se afastar ou ficar aborrecida sem tentar reestabelecer a conexão social ameaçada. Poderíamos relutar em chamar essa resposta aversiva de culpa, mas ela tem uma similaridade familiar. Não há evidência alguma, pelo que

eu saiba, de competência moral em um indivíduo que não tenha análogo algum com a culpa ou com a vergonha.

3.2.5 A objeção de Eutífron

Teorias da sensibilidade são teorias subjetivistas. O valor moral de uma ação depende de respostas subjetivas em nós. É, às vezes, argumentado que o subjetivismo caminha de trás para frente. Lembre do Eutífron, de Platão. Sócrates pergunta a Eutífron se as ações são piedosas porque os deuses as aprovam ou se os deuses as aprovam porque elas são piedosas. Sócrates argumenta em prol da última opção. Considerações similares podem ser trazidas como suporte contra o subjetivismo. Considere as questões: é uma ação errada porque a desaprovamos, ou a desaprovamos porque ela é errada? A segunda opção parece mais convincente intuitivamente.

Há um sentido em que os teóricos da sensibilidade podem acomodar a intuição de Eutífron, pelo menos em parte. Subjetivistas dizem que dizem que propriedades morais são engendradas por nossas atitudes, mas eles podem também dizer que nossas atitudes são engendradas pelas propriedades morais. Eis como. O errado é aquilo que provoca emoções de um certo tipo. Que tipo de emoções? Bem, emoções morais, é claro. Uma emoção moral é uma emoção que ocorre em resposta a algo que se conforma, ou falha em se conformar, a uma regra moral. E, logo, temos um círculo. Não se trata de um círculo vicioso, mas, ao invés disso, um círculo de cocriação. Emoções morais criam normas morais, as quais, simultaneamente, criam emoções morais. Wiggins (1987:195) e McDowell (1985:219) defendem uma ideia que se relaciona a isso, mas eles vão além. Eles dizem que não há maneira alguma de individuar as emoções que constituem nossos sentimentos morais sem denotar o que essas emoções designam. Essa “tese da não prioridade” introduz um grau de circularidade que eu não quero aceitar. Temos muitos tipos de sentimentos (amor, ódio, carinho, desaprovação moral, e assim por diante). Se cada um desses pudesse ser individuado apenas pelas propriedades que eles denotam, e essas propriedades pudessem ser individuadas

unicamente pelos sentimentos que elas evocam, então ficaria difícil ver como poderíamos sequer distinguir uma classe de qualidades emocionais secundária da outra. Não poderíamos distinguir a qualidade de ser amável da qualidade de ser moralmente repreensível, se não pudéssemos distinguir as emoções envolvidas no amor e na vergonha moral. Para evitar isso, precisamos de alguma forma de denotar emoções morais sem denotar o que estamos moralizando. Ao mesmo tempo, precisamos de uma forma de capturar a codependência entre emoções e fatos morais.

Pode-se pensar sobre esse processo como um efeito catraca. Aqui está a origem de um mito. Imagine que certos comportamentos causam emoções que ainda não são específicas do domínio moral. Um ato de crueldade pode causar ira por parte da vítima, e simpatia dentre os demais. O responsável pelo ato poderia ser isolado, criticado e punido. Isso poderia fazer com que o responsável pelo ato sentisse tristeza. Se essas respostas são consistentes, então a crueldade é regida por algum tipo de regra. A regra consiste no fato de que crueldade é desencorajada como um resultado dessas respostas emocionais. As emoções garantem um padrão previsível de comportamento. É menos provável que a crueldade ocorra, e quando ela ocorrer, certas emoções e o comportamento correspondente se seguirão. Após esse padrão ser estabelecido, as emoções que uma vez não tiveram qualquer significância moral, passam a ter um novo sentido. Tristeza não é apenas uma genérica resposta à perda, mas um sentimento associado à violação de uma regra. Ira não é uma resposta genérica à ameaça, mas um sentimento dirigido àqueles que violam as regras. Culpa e ira justificada nascem. No exato momento em que essas emoções nascem, a regra adquire um novo sentido. Ela é, agora, uma regra fortalecida pelas emoções morais. Ela é uma regra moral.

Esse círculo de codependência evita o problema enfrentado pela tese da não prioridade, de Wiggins e McDowell, mas ele me faculta estabelecer uma mais benigna pretensão de não prioridade em resposta à intuição de Eutífron. Se você pergunta, "por que a crueldade te enfurece?", eu posso responder que é porque ela é imoral. Se você pergunta, "Por que a crueldade é imoral?", eu posso responder corretamente dizendo que ela viola uma regra moral. Ela é errada porque há um tipo

especial de regra que a proíbe. Podemos dizer que a crueldade nos enfurece porque ela é errada. Podemos, inclusive, especificar que tipo de erro ela é, capitalizando o fato de que a moralidade divide-se em dois diferentes tipos de transgressão. Erros que nos fazem muito bravos envolvem violações à autonomia. Quando você pergunta por que a crueldade me enfurece, eu posso responder que ela viola a autonomia da vítima. É claro, há uma outra questão que você pode colocar: "De onde vem aquela regra?". Aqui, o subjetivismo vem à tona. A regra vem de nós; é a nossa regra; a regra existe porque há certas coisas que provocam nossa indignação. A questão de Eutífron pergunta: algo é errado porque o julgamos moralmente indignante, ou ele é moralmente indignante porque é errado? Ambas as respostas estão corretas.

O objetor pode se queixar de que essa solução incompleta é insatisfatória. Ela não vai suficientemente longe. A intuição de Eutífron demanda que não tenhamos algo a ver com a elaboração da moral. Diante disso o teórico da sensibilidade tem duas opções: negar a intuição ou admitir que as intuições populares sobre a moralidade sejam confusas. Eu prefiro a primeira opção. É intuitivamente óbvio que a moralidade não depende de nós? Penso que não. É fácil reunir a intuição segundo a qual a moralidade é uma construção humana. Algumas pessoas lamentam que tal visão poderia reduzir a força das pretensões morais. Eu discutirei essa preocupação extensamente em um capítulo posterior, mas me deixe oferecer uma breve resposta aqui. Não rebaixaria a moralidade dizer que as regras morais emanam das preferências humanas. Em verdade, se a moralidade vem de nós, então a moralidade é fundamentalmente importante. Podemos ser neutros sobre coisas que não dependem de nossas respostas, mas não podemos ser neutros sobre a moralidade. A moralidade é, em virtude de sua natureza mesma, algo com o qual nos importamos.

Em resumo, os teóricos da sensibilidade podem concebê-la das duas maneiras. Eles podem dizer que condenamos as coisas más porque elas são más, e eles podem dizer que as coisas são más porque as depreciamos. Ambas as pretensões têm suporte intuitivo. Teorias que apresentam apenas a primeira pretensão desligam a moralidade dos interesses humanos. Elas dão à moralidade

um tipo de prioridade que não temos razão para abraçar. Em verdade, se a moralidade tem tal prioridade sobre nossos interesses, haveria ainda uma questão sobre por que nos preocupamos com isso tudo.

3.2.6 Debate moral

A teoria da sensibilidade que estou defendendo é visivelmente relativista em sua formulação. Propriedades morais são definidas relativamente a observadores, e parece plausível que dois observadores possam ter sentimentos conflitantes. Se for assim, então não haverá um único fato sobre o qual se possa decidir quem está certo, mas, ao invés, dois fatos correspondendo ao sistema valorativo de cada observador. Terei muito mais a dizer sobre o relativismo no capítulo 5, mas quero dirigir uma objeção óbvia, a qual coloca uma ameaça às teorias da sensibilidade. Se moralidade é definida relativamente aos observadores individuais, então não é de todo claro por que os sujeitos deveriam se engajar em debates morais. Suponha que um abolicionista diga a um proprietário de escravos que "escravidão é errada". Se a moralidade é relativa aos observadores, o dono de escravos poderia responder que "talvez, para você, seja errado, mas, para mim, a escravidão está bem, de tal forma que não há desacordo entre nós". Mas não é isso o que acontece. Engajamo-nos incessantemente em debates morais. Se a sua perspectiva moral difere da minha, eu vou querer lhe trazer para o meu lado. Em contraposição, se eu descobrir que você e eu gostamos de diferentes tipos de música, eu talvez não tente lhe persuadir a gostar do que eu gosto. A teoria da sensibilidade precisa explicar essa diferença. Se valores morais são sentimentos, por que debatemos a moralidade quando não debatemos outras questões que dependem de sentimentos, como o gosto pela música?

Eu penso que o teórico da sensibilidade tem muitos recursos para lidar com essa questão. Primeiro, algumas pessoas talvez tenham uma falsa crença de que o absolutismo moral é verdadeiro. Essas pessoas irão se engajar no debate porque elas pensam que os juízos morais de seus oponentes não têm qualquer pretensão à verdade. Um debate moral entre dois absolutistas com valores conflitantes é, eu

creio, confuso, mas a confusão pode ser perfeitamente inocente. As pessoas não têm necessariamente acesso consciente direto aos seus conceitos, e, portanto, mesmo que seja uma verdade conceitual que a moral esteja relacionada aos observadores, os usuários comuns dos conceitos morais falharão no reconhecimento desse fato.

Segundo, nós, muito frequentemente, engajamo-nos em debates morais com indivíduos que compartilham de nossos valores morais básicos. É possível que as pessoas se equivoquem sobre aquilo que seus valores implicam. Má-fé, equívocos sobre os fatos, interesses conflitantes e vários outros fatos podem nos prevenir de experimentar a indignação moral em relação a algo moralmente indignante *de acordo com os nossos próprios padrões*. Quando isso acontece, um reformador moral pode nos mostrar o erro de nossos caminhos. Quando William Lloyd Garrison denunciou a escravidão norte-americana, ele não estava sugerindo que donos de escravos deveriam mudar seus valores morais. Ao invés disso, ele estava dizendo que os donos de escravos eram hipócritas porque estavam engajados em um comportamento que eles mesmos consideravam errado em outros contextos.

Terceiro, valores morais são implementados por emoções dirigidas a outras coisas. Se eu penso que você está moralmente errado, eu sentirei ira (ou outra emoção de censura) em relação a você, e essa emoção talvez me conduza a um conflito com você. Minha ira talvez me faça iniciar uma discussão com você, ou algum outro tipo de disputa, mesmo que a disputa não possa ser decidida racionalmente. Considere como as emoções conduzem às brigas entre pessoas que se amam. Em alguns casos, uma luta emotivamente instigada pode ter uma resolução racional, mas nem sempre esse é o caso. Não obstante, nossas emoções nos levam a conflitos.

Quarto, sentimentos morais estão fundados em metaemoções, de tal forma que não apenas valoramos as ações; nós também valoramos o valorar daquelas ações. Se seus valores discordam dos meus, eu terei uma resposta negativa tanto com respeito ao seu comportamento, quanto aos seus valores. O primeiro talvez me leve a tentar alterar seu comportamento diretamente, mas meu dissabor pelos seus

valores requererá que eu tente alterar como você se sente. O debate moral é uma ferramenta para afetar os estados psicológicos dos outros.

Quinto, mesmo que reconheçamos que os valores morais são relativos aos observadores, talvez haja algo em jogo quando conduzimos debates morais. Em analogia, considere o relativismo sobre o gosto. Se você ama música *country* e eu não, então eu normalmente não tentaria mudar seu gosto. Mas suponha que eu saiba que você está selecionando a música para uma festa vindoura. Nesse caso, eu tentaria persuadir você a mudar o gosto. No caso da moralidade, estamos frequentemente nesse tipo de posição. Valores morais são usados para guiar políticas ou práticas sociais. Se seus valores diferem dos meus, você vai tentar dar apoio a políticas que conflitam com meus valores, e essas políticas terão um impacto sobre mim. Conseqüentemente, eu serei motivado a tentar mudar suas perspectivas. Em geral, não toleramos os valores de outras pessoas (morais, estéticos ou outros) quando esses valores nos afetam. Questões morais são frequentemente questões de vida ou morte, de tal forma que temos um forte interesse especialmente em mudar as mentes daqueles com valores conflitantes com os nossos.

Sexto, há um reverso do ponto cinco. Da mesma forma que tenho um interesse em mudar suas visões, você tem um interesse em mudar as minhas. Isso significa que você tentará alterar meus valores a partir de quaisquer meios disponíveis, incluindo o debate moral.

Sétimo, consideramos a moralidade como refletindo profundamente quem somos. Uma pessoa com valores diferentes é estranha e antipática. Valores alternativos nos ameaçam, pois eles podem nos conduzir a reconsiderar as coisas com as quais mais nos importamos, e tal reconsideração poderia nos levar a uma perda de identidade. Em geral, resistimos à conversão e nos sentimos mais confortáveis quando as pessoas ao nosso redor têm uma perspectiva de vida similar. Esses fatos sobre a psicologia da identidade nos dá uma razão adicional para nos engajarmos no debate moral.

3.2.7 Os outros distantes

Há aqui, ainda, um enigma. Eu acabo de argumentar que temos ampla razão para nos importarmos com os valores morais daqueles que estão próximos. As pessoas em nossa comunidade têm um impacto em nossas vidas. Temos motivação para persuadir qualquer um que esteja dando forma a políticas que podem nos afetar. Mas e quanto aos lugares distantes? Os valores morais das pessoas em terras estrangeiras deveriam ser de menor importância se o que dizemos está certo. Essa predição acaba por ser verdadeira. A maioria das pessoas é, de alguma forma, indiferente aos valores morais de pessoas em terras distantes. Mas nós não somos totalmente indiferentes. Quando ouvimos sobre práticas que estimamos como repreensíveis (por exemplo, “mutilação genital feminina” ou “tortura”), tendemos a sentir indignação moral.

Há, ao menos, formas de explicar esse fato. Primeiro, podemos estar erroneamente estendendo excessivamente nossos valores. Se valores morais são disposições para reagir emocionalmente a certas formas de conduta, então qualquer exemplo de tal conduta poderia evocar tais reações. Essa generalidade da resposta poderia ser, em parte, uma simples solução para um problema de construção. Como construímos um sistema que puna moralmente ações imorais dentro da comunidade? Uma solução é construir um sistema que primeiro estabeleça a composição da comunidade, e, então, as punições. Um sistema mais simples pularia o primeiro passo, e puniria independentemente de qualquer coisa.

Segundo, às vezes, punimos os outros distantes porque, a despeito da distância, eles podem representar uma ameaça contra nós. Quando temos a ocasião de considerar as atividades em outras comunidades, isso frequentemente ocorre porque estamos interagindo com aquelas comunidades. Tais interações podem ter uma influência sobre coisas aqui em nosso lar. Assim, importa, para nós, assegurarmo-nos de que as outras comunidades estão agindo como nós agimos. Em algumas circunstâncias, sabemos de outros distantes que dificilmente algum dia encontraremos. Suponha que estejamos considerando a etnografia de um antropólogo sobre pessoas remotas, cujos costumes poderiam ser fortemente

proibidos em nossa comunidade: canibalismo amazônico, por exemplo. Nesses casos, a teoria da sensibilidade acertadamente prevê que talvez não fiquemos tão perturbados. Podemos considerar suas atividades com certo tipo de desapego moral, pois os vemos como absolutamente separados de nós.

Ainda assim, deve ser admitido que algumas pessoas ficam moralmente indignadas mesmo quando o mal ocorre em pessoas de culturas muito remotas. Esse pode ser um caso de errônea extensão excessiva de nossos valores, mas há também uma terceira história a ser contada sobre o porquê de nos importarmos com os outros distantes. Como Hume demonstrou, as pessoas são naturalmente propensas à simpatia (ou, talvez, à empatia). Hume pensou que a simpatia era muito paroquial: naturalmente simpatizamos apenas com pessoas semelhantes a nós, ou que estejam muito próximas a nós. Mas talvez tenhamos estado errados sobre isso. Talvez estejamos propensos a simpatizar com todos (ou com quase todos os seres humanos), como a pesquisa em mecanismos neurais da simpatia têm sugerido. Quando vemos outra pessoa sofrendo, isso nos faz sofrer (Singer et al., 2004). Condenamos ações que conduzam ao sofrimento dos outros, pois simpatizamos com as vítimas. Isso é verdade mesmo quando as vítimas estão distantes. Uma pessoa que se sente indignada com o canibalismo amazônico pode estar sendo conduzida pela simpatia. É claro, essa simpatia pode ser indevida; talvez as vítimas sintam-se honradas em serem devoradas. O ponto não é que estamos certos ao simpatizar, mas que talvez o façamos automaticamente, e isso pode nos levar a nos preocuparmos moralmente com pessoas que, muito improvavelmente, terão qualquer impacto direto sobre nossas vidas.

3.2.8 A objeção do raciocínio

As teorias da sensibilidade emergiram da tradição humeana. Hume é famoso (infame) [(in)famous] por dizer que o raciocínio moral está submetido às nossas respostas emocionais – a razão é escrava das paixões. Alguns críticos pensam que esse quadro é incompatível com a observação comum de que frequentemente chegamos a conclusões morais mediante um processo de deliberação racional.

Tentamos mudar os juízos morais uns dos outros apresentando argumentos. Na reflexão ética privada, tomamos cuidado especial ao pensarmos sobre todos os fatos relevantes e evidências de peso. Se emoções estivessem conduzindo a moralidade, os meios racionais da deliberação moral seriam pobremente apropriados para levar a juízos morais.

Para me reportar a essa objeção, quero considerar o papel do raciocínio em sentimentos não morais. Considere o sentimento de gostar. Suponha que eu queira lhe persuadir a gostar de alguém. Para fazê-lo, eu devo lhe convencer que a pessoa é passível de ser objeto do gostar. Em princípio, você não está convencido. Você foi dissuadido pelas maneiras arrogantes e rudes dela. Eu talvez tente amenizar esses aspectos. Talvez sua arrogância seja apenas uma defesa em virtude de sua insegurança. Eu lhe digo que ela é uma pessoa inteligente e cativante na conversação, bem como que tem um senso de humor sarcástico. Eu explico que ela é passional acerca das causas sociais e uma leitora ávida com amplos interesses. Nesse momento, você talvez conceda que fez um juízo muito apressado. Você talvez conclua que, se eu estou certo, e nenhum outro traço negativo é encontrado, então é possível gostar dela. Isso não significa que você instantaneamente começará a gostar dela, mas munido com uma nova interpretação acerca de sua arrogância e com novas informações sobre outras características, você está disposto a começar a gostar dela.

Esse é um exemplo de debate racional. Eu apresentei razões para uma conclusão, e aquelas razões fundamentam a conclusão. Se você aceita minhas razões, você está preparado para aceitar minha conclusão. Não obstante, há dois sentidos nos quais esse debate parte de exemplos paradigmáticos do discurso racional. Primeiro, você talvez conclua que a pessoa em questão é objeto do gostar sem, todavia, vir a gostar dela. Nesse sentido, o debate se aproxima mais do raciocínio prático do que do teórico. No raciocínio prático você talvez conclua sobre um curso de ação que seja razoável sem, contudo, agir.

Segundo, o debate sobre gostar pressupõe que certos traços são objetos do gostar. Um senso de humor é um traço passível do gostar. Inteligência e consciência social são passíveis do gostar. Arrogância é desagradável. Suponha

que voltássemos nossa atenção para esses traços. Suponha que meu interlocutor negue que ter um senso de humor sarcástico seja objeto do gostar. Sagacidade sarcástica, ele pode defender, mostra um tipo de cinismo que indica tendências elitistas e pessimistas. Poderíamos debater sobre isso. Mas agora suponha que meu interlocutor diga que ele não gosta de pessoas com senso de humor em geral. Ele pode dar razões para isso ("a vida é muito séria para piadas"). Podemos ter um debate sobre se um senso de humor revela uma ausência de cuidado ou de seriedade sobre os problemas do mundo. Mas se meu interlocutor simplesmente declara um desagrado pessoal pelo humor, eu concluiria que ele e eu gostamos de coisas diferentes. Eu concluiria que o que é objeto do gostar para mim não o é para ele. Eu não poderia persuadi-lo racionalmente. Seria mais ou menos como tentar persuadir alguém que odeia amendoins a gostar de manteiga de amendoim.

Debates sobre o gosto funcionam quando se tenta estabelecer que algo ou alguém tem traços que são objetos do gostar. O que conta como um traço passível do gostar depende do que se gosta efetivamente. Pode-se racionalmente demonstrar que alguém tem traços passíveis do gostar, mas é muito mais difícil mostrar racionalmente que um traço é passível do gostar. Em alguns casos, pode-se fazê-lo mostrando que um traço T1 é uma instância de outro traço T2, do qual se gosta. Contar piadas é um sinal de crueldade ou de *joie de vivre*? Eventualmente, cairemos por terra. Em um certo ponto, os traços dos quais gostamos não são objetos do gostar por várias razões. Eles têm a ver com como somos constituídos. Talvez eu pudesse vir com uma razão para gostar de pessoas com senso de humor. Essa razão poderia ser *post hoc*. Não se trata da minha razão para gostar de tais pessoas; eu simplesmente gosto delas. O humor é passível do gostar para mim. Se você não gosta do humor, somos constitutivamente diferentes.

Eu penso que os desacordos morais são assim. Há muito espaço para o debate racional. Frequentemente precisamos usar a razão para demonstrar que uma ação cai sob uma categoria moral (lembre-se do estado de categorização na Figura 3.1). Suponha que eu diga que a oração nas escolas seja errada. Eu posso acrescentar uma razão: é errado porque discrimina ateus e membros de religiões minoritárias. Você pode replicar que não discrimina. Ou você pode replicar que

proibições contra orações na escola discriminam os membros da religião majoritária, e, portanto, discrimina contra mais pessoas do que se a proibição fosse abandonada. Esse é um debate racional. Poderíamos assentar sobre uma definição inicial de discriminação, e prover evidência para nossas respectivas perspectivas. Se eu lhe persuado sobre minha perspectiva, sua atitude emocional em relação à oração na escola poderia não mudar instantaneamente, mas ela poderia estar propensa a mudar. Mas agora suponha que você não está persuadido. Suponha, ao invés disso, que você diga que discriminação não é moralmente errada. O melhor que posso fazer é encará-lo incredulamente. Se você pensa que discriminação não é errada, então somos constituídos diferentemente.

Alguns autores que enfatizam o papel das emoções subestimam o papel do raciocínio no juízo moral. Por exemplo, Haidt (2001) diz que o raciocínio tipicamente entra no processo apenas como uma justificação *post hoc* dessas intuições. Eu concordo com Haidt que o raciocínio frequentemente desempenha esse papel, mas o raciocínio pode também desempenhar um papel importante na determinação de quando certos eventos caem sob uma categoria acerca da qual temos um sentimento moral. Geralmente temos que raciocinar, por exemplo, para determinar se é o caso que algo seja uma discriminação. O modelo de Haidt permite isso, mas ele não o enfatiza. Esse é um descuido digno de nota, pois vários debates morais podem envolver desacordos morais sobre como categorizar as coisas. Mas, uma vez que tenhamos categorizado algo como sendo o caso de discriminação, o raciocínio pára. Há uma reação emocional imediata. Isso é o pouco que Haidt enfatiza, e é o essencial para minha versão da teoria da sensibilidade também.

No capítulo 1, eu disse que podemos pensar sobre debates morais como sendo arranjados em uma série de questões sobre os porquês (compare com Anscombe, 1957, sobre as intenções). Por que φ é errado? Trata-se de uma instância de ψ . Por que ψ é errado? Bem, ele simplesmente o é. Se ψ é o mais profundo ao qual eu posso chegar, eu ficarei perplexo se alguém perguntar por que ele é errado. Por que sonegação fiscal é errada? Porque nós colhemos benefícios dos impostos, e é errado colher benefícios da contribuição de outras pessoas se você não contribuiu com coisa alguma; isso seria caronismo. Por que é errado ser

um caronista? Essa questão é difícil de responder. Ela é fundamental para nossa perspectiva moral. É como responder por que é errado matar pessoas inocentes. Bem, simplesmente é. Alguém que tenha dúvidas sobre isso tem uma concepção diferente do que seja errado. Eu chamo esses valores mais profundos de “normas fundacionais”.

Podemos agora colocar o papel do raciocínio no discurso moral da seguinte forma. Quando duas pessoas discordam sobre se algo é errado, elas podem ter um debate racional sobre se ele cai sob uma categoria acerca da qual elas tenham uma norma fundacional. Deixar alguém morrer é o mesmo que matar? São os abrigos fiscais uma forma legalmente sancionada de caronismo? Mas quando se chega a normas fundacionais, o debate racional é impossível. Se duas pessoas têm diferentes normas fundacionais, elas devem recorrer a outros meios de persuasão.

Se estou certo sobre normas fundacionais, então uma grande parte do trabalho filosófico em ética está baseado em uma pressuposição falsa. Filósofos estão preocupados com a ideia de que os juízos morais podem servir como *razões* práticas. A invocação de um juízo moral supostamente nos provê um razão para a ação, e razões são supostamente justificadoras. Eu penso que há uma tênue razão consoante a essa abordagem que está certa. Se você tem uma norma fundacional que exige certo comportamento, então você tem uma razão para agir dessa forma. Mas isso é análogo a dizer que, se você gosta de chocolate, então você tem uma razão para comer chocolate. Normas fundacionais não podem servir como razões em um sentido mais robusto. Minhas normas fundacionais não seriam vistas como razões para alguém que não as compartilhasse comigo. Filósofos que quisessem mostrar que normas morais fornecem razões robustas para o agir estariam embarcando em uma missão de tolos. Por um lado, se há demandas independentes da preferência acerca do agir, não há absolutamente qualquer razão para pensar que nosso vocabulário moral se refere àquelas demandas. Por outro lado, o discurso moral não parece pressupor que normas fundacionais tenham esse *status*. Se normas fundacionais fossem construídas como independentes de preferências, os interlocutores morais deveriam sentir alguma compulsão para justificá-las em termos puramente racionais. Mas, eu acharia que tais tentativas de justificativa não

têm papel significativo no discurso moral ordinário. Não é o caso que eu valorize a vida humana em virtude de um bem trabalhado argumento racional, e eu não sinto qualquer obrigação para gerar um tal argumento (exatamente como eu não tenho que argumentar em prol do deleite do chocolate). Se encontro alguém que cruamente assere que a vida humana não tem valor, eu assumo que essa pessoa é perversa, e não estúpida. E eu não respondo com a razão, mas com o punho cerrado.

3.2.9 A objeção do propagador de modos

Algumas pessoas realmente se importam com a etiqueta. Essas pessoas se sentem incomodadas quando as pessoas colocam os cotovelos sobre a mesa, ou quando mastigam goma de mascar com a boca aberta. Elas talvez se sintam embaraçadas quando falham ao tentar passar o sal na direção certa, ou quando tentam dizer as palavras apropriadas de desculpas após um espirro. Mas certamente essas não são regras morais. Certamente esses não são valores morais. Teorias da sensibilidade parecem muito liberais. Elas implicam que regras intuitivamente não morais são morais, e isso nos deixa sem uma boa forma de distinguir entre o moral e o convencional. Teóricos da sensibilidade podem esboçar uma distinção moral/convencional apenas naqueles casos em que as regras convencionais não são colocadas em vigor pelas emoções morais, mas isso deixa em aberto casos nos quais as pessoas ficam indignadas acerca da etiqueta.

Há muitas coisas que podem ser ditas em resposta a casos como esse. A primeira é que pessoas que expressam emoções morais fortes sobre etiqueta podem estar moralizando. Elas podem estar tratando as maneiras como se essas fossem morais. As pessoas geralmente moralizam sobre coisas que são meramente convencionais. Exemplos podem incluir aquele em que um dos pais faz com que a criança se sinta culpada, ou envergonhada, por não praticar o piano. Às vezes, as pessoas meramente fazem de conta que estão moralizando com o propósito de alcançar um efeito desejado. Por exemplo, fingir indignação pode ser uma forma de obter cumplicidade ou controle. Considere o cliente insatisfeito que

reclama ao gerente da loja que foi enganado (Greenspan, 2000). Esses casos não são uma preocupação para o teórico da sensibilidade.

Em toda cultura, há provavelmente alguns modos que são frequentemente moralizados. Na cultura norte-americana contemporânea, tendemos a ter uma atitude moral em relação àqueles que deixam a mesa enquanto o outro ainda está comendo, bem como em relação à linguagem obscena. Esses são comportamentos que as pessoas de fora dessa cultura poderiam imediatamente reconhecer como convencionais, mas há uma tendência para termos sentimentos morais plenos em relação a eles. Na Inglaterra vitoriana, a lista de modos moralizados poderia ser muito maior, e outras culturas, em outros tempos, poderiam ter suas próprias listas.

Rupturas nos modos podem ser tomadas bem seriamente, mas, mesmo quando nos indignamos, podemos reconhecer, às vezes, uma diferença entre maus modos e erros morais. Dizer "saúde" é realmente uma exigência da moralidade? Há uma tensão em nossas intuições que precisa ser resolvida. Por um lado, nós punimos moralmente aqueles com maus modos, por outro, muitas de nossos modos são convenções arbitrárias.

Eu penso que a teoria da sensibilidade pode lidar com a tensão. O truque é empregar a noção de normas fundacionais, as quais introduzi acima. Normas fundacionais são normas que tendemos a considerar como livres da necessidade de qualquer explicação. Em nosso sistema de valores, exemplos incluem proibições contra o assassinato, contra o roubo e contra o incesto. Quando alguém pergunta "por que matar é errado?" respondemos com um olhar incrédulo, ou nos atrapalhamos com uma explicação *post hoc*. Apenas lunáticos e filósofos colocam tais questões. Mas outras normas não são tão fundamentais. Se declaramos que sonegação fiscal é errada, precisamos tornar claro o porquê, e a resposta envolverá demonstrar que a sonegação fiscal é uma instância de uma norma fundacional. Talvez ela viole uma norma de equidade. Talvez ela seja uma forma de roubo. Talvez ela seja um tipo de desrespeito pelos outros. Ou todas essas coisas.

Agora considere a etiqueta. Talvez ter bons modos seja uma norma fundacional. Ou talvez modos derivem de uma norma mais básica, a qual tem a ver com o respeito. Se você não obedece às regras de etiqueta, você desrespeita

aqueles que se importam com tais regras. Logo, pode ser imoral ter maus modos, mas unicamente de forma derivativa. Violações específicas de etiqueta são imorais, na medida em que elas são violações de alguma norma mais fundamental.

Isso explica por que modos se mostram, em algum sentido, como morais para nós, bem como por que eles se mostram como não morais – como meras convenções. Violações de etiqueta são explicitamente reconhecidas como dependendo de costumes locais. Talvez haja uma norma fundacional proibindo maus modos, mas essa norma pouco nos guia na determinação de quais comportamentos deveriam ser vistos com maus olhos. A inferência a partir de uma proibição geral contra os maus modos para algumas normas específicas de etiqueta requer uma premissa auxiliar sobre quais comportamentos são localmente considerados maus. Por exemplo, você precisa saber que, na cultura local, as pessoas dizem “saúde” quando alguém espirra, bem como que elas tiram os chapéus quando estão comendo. Pessoas que se pautam por esses modos usualmente reconhecem que eles são condicionais aos costumes locais. Quando eles falham em reconhecer isso (considere atitudes relativas a cuspir em público), eles tomam o comportamento em questão como moralmente repreensível, não como estupidez. Isso oferece uma clara análise da distinção moral/convencional. Uma norma é moral se ela é uma norma fundacional ou se ela é derivada de uma norma fundacional sem apelo necessário a uma premissa acerca de costumes. Uma norma é convencional se ela depende de um apelo a costumes. Nós, às vezes, talvez nos enervemos com maus modos, mas nossa resposta depende da crença de que certos comportamentos são costumeiramente proibidos. Se mudássemos nossas crenças sobre os costumes locais, a ira feneceria. Teóricos da sensibilidade não precisam se sentir constrangidos quanto ao fato de que nós, às vezes, ficamos emotivos acerca de erros convencionais.

Antes de deixar de lado esse tópico, deve-se notar que a distinção moral/convencional não está isenta de críticas. Kelly et al. (2007) tem recentemente levado a cabo uma poderosa crítica empírica. De acordo com Turiel, Nucci, Smetana e outros psicólogos que estudam essa distinção, regras morais são, supõe-se, independentes de autoridade. Violações morais deveriam permanecer erradas

mesmo quando os costumes locais as autorizam. Kelly et al. descobriram que não é o caso. Seus pacientes estavam dispostos a tolerar erros morais fundamentais (danos físicos) em diferentes contextos culturais. Por exemplo, foi pedido a pacientes que considerassem o capitão de um navio que espancasse um marinheiro bêbado. Em uma versão, a vinheta ocorre no presente e, em outra, ela ocorre há 300 anos, quando a punição corporal era amplamente tolerada. Os pacientes estavam inclinados a dizer que estaria tudo bem quanto ao espancamento para um capitão há 300 anos, mas não agora. Isso sugere que normas morais não são independentes da autoridade. Isso também sugere uma preocupação quanto à minha análise da distinção, pois implica que crenças sobre costumes locais incluem juízos sobre casos morais paradigmáticos. Kelly et al. pensam que não haverá maneira segura de esboçar uma distinção entre normas morais e convencionais.

Eu penso que esses resultados são intrigantes, e eles levantam preocupações sobre a forma na qual a distinção moral/convencional tem sido operacionalizada na literatura. Mas eu não penso que os resultados solapem a distinção moral/convencional como eu a tenho esboçado. Primeiro, proibição contra a punição corporal pode ser considerada como uma norma convencional para muitas pessoas. Segundo, quando sujeitos julgam que está tudo bem um capitão espancar um marinheiro poucos séculos atrás, eles talvez não estejam sugerindo que esse é um comportamento moralmente aceitável; pelo contrário, eles talvez estejam sugerindo que devemos perdoar o capitão por sua conduta, porque, à época, as pessoas eram menos esclarecidas. Terceiro, os sujeitos talvez creiam que havia, à época, fatores que tornavam moralmente aceitável o espancamento; por exemplo, eles podem imaginar marinheiros de eras passadas como sendo mais desregrados e propensos a se amotinar do que os altamente profissionalizados marinheiros do comércio marítimo contemporâneo.

O teste crucial para a forma que tenho esboçado a distinção seria perguntar aos sujeitos por que é errado espancar alguém. Se estou certo (e a literatura sobre a distinção moral/convencional mostra-o bem), as pessoas ordinariamente não apelam aos costumes locais quando justificam normas que permitem a agressão.

Elas não dizem que é errado espancar porque as pessoas ao redor não costumam fazê-lo. Em contraste, se você pergunta às pessoas por que é errado usar um chapéu quando se está comendo, elas provavelmente mencionarão os costumes locais. Se crenças sobre costumes adentram as deliberações sobre os dois capitães marítimos, tais crenças não desempenham um papel na derivação do juízo segundo o qual certo comportamento é errado. Pelo contrário, elas servem como uma desculpa quando se está decidindo que uma forma dúbia de comportamento pode ser tolerada sob certas condições. O fato de que considerações sobre costumes podem prevalecer sobre juízos morais é consistente com a pretensão de que normas morais são aquelas que as pessoas não justificam apelando aos costumes.

3.2.10 O imperativo categórico

A moralidade impõe demandas sobre nós. Isso é algo difícil de negar. A questão interessante concerne à natureza dessas demandas. Os kantianos insistem que elas são categóricas. Diz-se que isso apresenta um sério desafio a qualquer um que fundamente a moralidade sobre emoções. Imperativos categóricos são pensados como tendo três aspectos dignos de nota. Primeiro, diz-se que eles nos governam independentemente de nossas inclinações, paixões e desejos. Isso é o que se quer dizer quando se afirma que eles são categóricos. Segundo, diz-se que eles são comandos da razão. Essa é a melhor explicação sobre como eles nos governam independentemente de nossas paixões. Terceiro, diz-se que eles são universais. Um comando da razão deve ser aplicado a todas as pessoas racionais. A universalidade está no centro da concepção de Kant. Ele pensa que há apenas um imperativo categórico, a partir do qual exigências mais específicas são derivadas. Deveríamos fazer aquilo que poderíamos querer como uma lei universal. Se imperativos morais são categóricos, as teorias da sensibilidade não vão alçar voo. Esse é o principal ponto de contenda entre kantianos e descendentes de Hume.

Seria impossível desenvolver um argumento completo contra a ética kantiana aqui. Eu posso apenas erguer algumas preocupações familiares. Os

kantianos não se deixarão convencer. Meu objetivo não é refutar Kant, mas lembrar aos leitores que o programa kantiano permanece altamente controverso, e eu enfatizarei que algumas das estratégias para preencher os buracos dos argumentos kantianos trazem as paixões de volta para o quadro. Eu também tenho um objetivo mais importante. Eu quero explicar como uma visão humeana pode dar conta de aspectos centrais da intuição de que há algo categórico sobre os imperativos morais. Em verdade, eu argumentarei que humeanos podem fazer um trabalho melhor do que os kantianos na apreensão do *status* normativo distintivo das pretensões morais. Portanto, mesmo que haja uma forma de derivar imperativos da razão, isso não evitaria a necessidade de uma história humeana acerca dos conceitos morais.

A teoria da sensibilidade pode ir razoavelmente longe no sentido de dar à moral um sabor categórico, mas não pode ir até o fim. Começemos com a independência à paixão. O teórico da sensibilidade tem ao menos duas estratégias diferentes para mostrar que uma ação é errada mesmo quando as pessoas não reagem negativamente a ela. Primeiro, quanto ao tipo de teoria da sensibilidade que endosso, os valores morais são constituídos por sentimentos, e sentimentos são disposições. Em algumas ocasiões, a disposição para ter uma emoção negativa não é atualizada. Se estou distraído, cansado, deprimido, extático ou bêbado, posso testemunhar um crime com total indiferença. Isso não quer dizer que o crime é permissível naquele caso. Ele ser errado depende do que estou disposto a sentir, não dos sentimentos que ocorrem de eu ter. Segundo, as paixões não serão evocadas quando eu falho em me aperceber que uma ação particular é incompatível com minhas normas fundacionais. Considere, novamente, a escravidão. Imagine que todos, em uma comunidade de donos de escravos, tenham uma norma fundacional que proíba a escravidão, mas eles não empregam aquela norma apropriadamente. Membros daquela sociedade caem sob a norma que proíbe a escravidão mesmo que eles falhem em se aperceber disso. Dado que eles não se apercebem disso, a escravidão não provoca um sentido de indignação. A norma vale mesmo quando eles não consideram, como uma questão de fato, que a

escravidão seja abominável. Logo, há um sentido em que a norma é independente da paixão.

A independência da paixão tem seus limites. Os teóricos da sensibilidade não podem dizer que algo poderia ser errado mesmo que as pessoas estivessem dispostas a desaprovar isso. É esse um problema? Penso que não. Intuições sobre a independência em relação à paixão derivam de casos nos quais as pessoas ignoram a moralidade em favor de outras preocupações. Uma pessoa rica que não dá coisa alguma aos necessitados pode ser moralmente obrigada a dar, mesmo que ela tenha um forte desejo de manter seu dinheiro. Demandas morais não desaparecem toda vez que algo mais desejável surge. Teóricos da sensibilidade podem facilmente acomodar esse tipo de independência em relação à paixão. Paradoxalmente, mesmo propriedades dependentes do afeto são independentes em relação à paixão nesse sentido. Algo pode ser assustador mesmo se, em uma ocasião em particular, nós não nos assustamos com ele. Algo pode ser deprimente, mesmo se que isso não nos entristecer. O mesmo é verdade para as coisas das quais se podem gostar. Eu gosto de sushi, e, logo, sushi é objeto do gostar para mim. Mas há ocasiões em que eu não quero sushi. Eu não gosto de comer sushi enquanto estou comendo sorvete de chocolate, e eu não iria querer sushi após uma refeição abundante. Mas sushi é objeto do gostar para mim em todos os momentos. A capacidade de gostar é uma propriedade disposicional. Ela independe de desejos transitórios, mas não é independente de paixões *tout court*.

Uma forma mais forte de independência em relação à paixão implicaria que uma ação pudesse ser errada mesmo se ninguém tivesse uma psicologia que promovesse sua desaprovação sob condições de total informação e de nenhuma demanda emocional concorrente. Imagine uma comunidade isolada de psicopatas, vivendo em um mundo com nenhum outro ser inteligente. Eles podem mentir, enganar e matar com regularidade. Eles reconhecem as regras convencionais, e eles estabelecem muitas leis para desencorajar esses tipos de comportamentos. As leis conferem uma espécie de estabilidade, como no estado hobbesiano, mas ninguém considera imoral sua violação. Punições são avaliadas como prudenciais (e, às vezes, como entretenimento), mas não como justificadas moralmente. A

categoria de justificação moral não existe. Que diremos do assassinato nessa comunidade? É ele moralmente errado? Penso que não. Não é mais imoral para um psicopata assassinar, nesse mundo horrível, do que é para uma leoa matar um gnu nas savana africana.

Considerarei tais casos em maior detalhe no capítulo 4. Aqui eu apenas quero registrar uma intuição. Eu penso que nossas regras morais não abarcam aqueles que não estão abertos aos sentimentos morais. O kantiano talvez concordasse argumentando que a comunidade de psicopatas não é uma comunidade de pessoas. Eles não estão constituídos de uma forma que lhes permita ver uns aos outros como fins. Eles não são, em sentido relevante, humanos. É possível que assim seja. Mas essa concessão poderia ser uma vitória para o teórico da sensibilidade. A coisa mais importante que distingue psicopatas de nós (casos paradigmáticos de pessoas) é sua indiferença emocional. Colocá-los à parte da moralidade é conceder que a independência em relação à paixão pode ser tomada somente até esse ponto.

Voltemo-nos, agora, para o segundo aspecto dos imperativos categóricos. De acordo com Kant, eles são demandas da razão. A razão exige que evitemos a conduta imoral. A conduta imoral é irracional. Humeanos podem acomodar essa intuição até certo nível. Eles podem dizer que o comportamento imoral é incompatível com a razão *prática*, onde razão prática é definida em referência aos nossos desejos. A conduta imoral pode ser como a fraqueza da vontade. Ela pode envolver agir a partir de desejos que conflitam com outros desejos que teríamos se tivéssemos toda informação relevante. Mas humeanos negam que a moralidade seja oriunda da razão *pura*. A ideia de uma "razão suficiente" é, para Hume, um oxímoro. Assim como conceitos são cegos sem perceptos, a razão prática é sem propósito sem paixão. Kant diz que a razão pode prover fins, assim como meios, sem qualquer contribuição da paixão. Humeanos discordam.

Como pode a razão prover fundamentos suficientes para a ação? O argumento mais famoso de Kant jaz sobre a pressuposição muito instável segundo a qual alguns cursos de ação são literalmente incoerentes ou contraditórios quando seguidos por todos. Seu clássico exemplo, mencionado no capítulo 1, é o da falsa

promessa. Se todos fizessem promessas sem a intenção de mantê-las, então a prática da promessa ruiria. As promessas não podem existir se elas não são mantidas durante alguma parte do tempo. Se todos fizessem falsas promessas, não haveria promessas, e, portanto, falsas promessas. Não podemos querer que todos os outros façam falsas promessas sem contradição, e, Kant diz, não podemos querer racionalmente fazer aquilo que não podemos consistentemente querer que todos os outros façam. Portanto, não podemos fazer uma falsa promessa. Talvez todos os erros morais sejam erros em virtude de serem contradições nesse sentido.

Mas isso não pode estar certo. Primeiro, há muitos erros aparentes que não geram contradição alguma quando universalizados. Se todos chutassem crianças sempre que as vissem na rua, a prática de chutar crianças não seria considerada incoerente. Segundo, algumas ações paradigmaticamente boas se tornam contraditórias quando universalizadas. Suponha que há recursos suficientes para alimentar todas as pessoas necessitadas no mundo. Agora considere o mandamento moral: dê comida àqueles em necessidade. Se todos fizerem isso, na medida em que são capazes, a necessidade desapareceria, e a prática de dar aos necessitados seria considerada vazia. Isso significa que dar aos necessitados é errado? Ou, para usar outro exemplo, considere o mandamento: trabalhe para abolir a escravidão. Se todos fizessem isso, a escravidão desapareceria, e a abolição seria impossível. Ou considere o comando: "boicote as leis de Jim Crow". Um bem sucedido boicote solaparia o futuro boicote. Um kantiano poderia protestar que os mandamentos de dar comida aos necessitados, de combater a escravidão e de protestar contra leis discriminatórias permanecem válidas mesmo que não existam pessoas necessitadas, escravos e discriminação. Mas isso viola a restrição segundo a qual dever implica poder. Reconhecidamente, poder-se-ia universalizar uma máxima que dissesse, "combata a escravidão se há alguma", mas poder-se-ia provavelmente universalizar também uma máxima que dissesse, "minta, se isso te serve". Uma vez que condicionalizamos nossas máximas, o teste de universalização falha na filtração do mau comportamento.

Poder-se-ia também discordar da derivação do imperativo categórico. Eis uma interpretação. De acordo com Kant, a única coisa que é intrinsecamente boa é

uma boa vontade. Uma boa vontade é uma vontade que age por dever. Nós, humanos, sozinhos, podemos agir por dever, pois nós, sozinhos, podemos nos representar como se estivéssemos agindo tendo, como base, razões; animais não humanos simplesmente seguem suas inclinações. Para que o dever controle a vontade, e nos libere das inclinações, ele deve suprir a vontade com razões que sejam suficientes para a ação. Mas que tipo de razão pode ser suficiente para a ação? A resposta de Kant é que podemos alcançar a liberdade unicamente agindo por razões que poderíamos querer como leis universais. Para descobrir mandamentos da razão, precisamos perguntar se um dado mandamento poderia governar todas as criaturas racionais. Essa é a base do bom.

Há muito espaço para dúvidas aqui, mas me limitarei a dois pontos. Primeiro, Kant esboça um altamente questionável elo entre nossos conceitos de liberdade e de ação por regras universais da razão. Tais regras não parecem nem conceitualmente necessárias, nem suficientes para a liberdade. Fenomenologicamente, podemos sentir um senso de liberdade quando podemos agir de acordo com nossos desejos, de tal forma que agir pela razão não é necessário. Dúvidas sobre a suficiência surgem quando consideramos o fato que Kant toma a razão como exigindo um único curso de ação. Mas, se a razão nos dá apenas uma opção, então se pode dizer que agir a partir da razão nos priva da liberdade, ao invés de nos outorga-la. O fato de que podemos construir nossas vontades como livres não nos impele a construir nossas vontades como limitadas pela razão. Logo, se nós necessariamente procuramos liberdade, disso não se segue que nós necessariamente queremos atentar para as demandas da razão.

Minha segunda reclamação acerca de Kant concerne à primeira premissa: a pretensão de que a boa vontade (concebida como o agir por dever) é realmente boa intrinsecamente. A ideia toda de benevolência intrínseca é um pouco enigmática. Eu acho bizarro supor que algo pode ser bom sem ser bom para um agente, ou estimado como bom. Deixe-me colocar de lado essa preocupação. Quanto mais urgente é essa preocupação, se há coisas boas intrinsecamente, então a boa vontade não parece um exemplo especialmente convincente, muito menos o único exemplo. Poder-se-ia argumentar que a boa vontade é valiosa unicamente em

virtude do fato de que ela pode trazer bons efeitos. Exemplos intuitivamente mais plausíveis de coisas intrinsecamente boas incluem felicidade, amizade, coragem, saúde e conhecimento. Esses parecem fins que poderíamos procurar por eles mesmos. Kantianos podem protestar que esses bens carecem de valor intrínseco porque os desprezamos quando eles são possuídos pelos perversos. Mas, se vamos levar a sério a ideia de valor intrínseco, devemos focar nesses bens à parte de quem os têm. Considerados isoladamente, saúde pode parecer uma coisa boa. Em verdade, é ruim que o perverso seja saudável, precisamente porque a saúde é boa, e nós não queremos que o perverso tenha coisas boas. Nesse ponto, kantianos podem objetar que a saúde e a felicidade não são bens *morais*. Mas como eles vão demonstrar que a boa vontade, concebida como agir a partir da razão, é um bem moral? Deveríamos valorar moralmente a devoção aos deveres morais? O sr. Spock é o nosso modelo de perfeição? A resposta está distante do óbvio. Se a boa vontade tem valor moral, ela pode ser derivativa. Ela pode se originar do fato de que boa vontade traz outros bens que valoramos moralmente. Se for assim, o dever racional não pode ser o fundamento da moral.

Muitos devotos contemporâneos de Kant admitiriam que o argumento da *Fundamentação* é repleto de falhas. Mas ao invés de rejeitar o seu programa, alguns têm oferecido argumentos melhorados que não dependem de pretensões dúbias sobre a natureza da liberdade ou sobre a benevolência do dever. Um dos melhores desses argumentos kantianos melhorados é colocado por Gewirth (1994). Eis uma reconstrução:

P1. Eu sou um agente.

P2. Se eu sou um agente, eu aceito que a liberdade e o bem-estar são necessários para mim.

C1. Portanto, eu aceito que eu devo ter liberdade e bem-estar. (De P1 e P2.)

C2. Portanto, eu aceito que é inadmissível não ter liberdade e bem-estar. (De C1, pois sua negação é a contradição daquela premissa.)

C3. Portanto, eu aceito que é inadmissível que os outros removam ou interfiram em minha liberdade e bem-estar. (De C2.)

C4. Portanto, eu aceito que os outros devem conter-se de remover ou interferir em minha liberdade e bem-estar. (De C3.)

C5. Portanto, eu aceito que eu tenho um direito à liberdade e ao bem-estar. (De C4, em virtude da correlação entre deveres e direitos.)

C6. O mesmo argumento é aplicado a todos os agentes. (Generalização universal de P1.)

C7. Portanto, todos os agentes aceitam que eles têm o direito à liberdade e ao bem-estar. (De P3.)

C8. Portanto, eu aceito que todos os agentes têm o direito à liberdade e ao bem-estar. (De P1 e C7.)

Se esse argumento é robusto, então temos uma razão suficiente para respeitar a liberdade e o bem estar dos outros.

Mas o argumento se articula a partir de um equívoco. A necessidade evocada em P2 é conceitual. Para ver a mim mesmo como um agente, eu devo ver a mim mesmo como tendo liberdade e bem-estar. Isso já é uma pretensão altamente problemática, a qual considero implausível. Mas vamos conceder, a título meramente argumentativo, que ela seja verdadeira e que, portanto, C1 seja verdadeiro. Mas agora considere C2. Segue-se, do fato de que eu deva ver a mim mesmo como livre, que eu deva também acreditar que aos outros *não é permitido* causar dano a mim? Absolutamente não. Segue-se apenas que, se ocorre de eles me causarem dano, eu não poderei mais me ver como um agente. Gewirth ilicitamente passa da necessidade conceitual para a necessidade moral. Para escapar desse problema, a necessidade em P2 tem que ser concebida normativamente. Tem que ser o caso que eu pense que eu tenho minha liberdade e bem-estar em virtude da necessidade moral. Do ponto de vista dessa interpretação, P2 é uma petição de princípio, bem como, provavelmente, falsa. Em princípio, eu poderia ver a mim mesmo, de uma maneira hobbesiana, como um agente em um mundo onde impera a lei do mais forte. Os outros estão autorizados a me esmagar, e eu estou autorizado a esmagá-los. Que o melhor esmagador vença.

Poder-se-ia remendar o argumento ressaltando que, enquanto *uma questão de fato*, nós sentimos que estamos autorizados a ter liberdade e bem-estar. Ocorre

de P2 ser verdadeira. Nagel (1986:63) argumenta apelando ao fato de que nós nos *ressentimos* com relação aos que passam dos limites contra nós. A emoção de ressentimento revela que nós pensamos que os outros não têm o direito de nos causar dano. Isso sugere que C1 é verdadeira, e, em verdade, que a premissa C5 de Gewirth é verdadeira. Consideramos que temos direitos. Se me considero como tendo direitos, eu não posso, sob risco de incorrer em contradição, negar que os outros têm direitos. Ou assim Nagel argumenta. Isso prova que as exigências morais podem ser derivadas da razão pura? Eu penso que a resposta é negativa, e o é obviamente.

O argumento de Nagel está baseado em emoções, e não na razão. Não se trata de uma verdade conceitual sobre a ação que os agentes têm direitos. Pelo contrário, é uma verdade emocional sobre mim e sobre os outros agentes que assumimos como tendo direitos. Poderíamos (como no cenário hobbesiano) desistir dessa pressuposição e continuar a nos conceber como agentes. A questão de Nagel sobre o ressentimento mostra que nós internalizamos uma norma que nos autoriza a considerar os outros como responsáveis por nos causar danos. Uma vez que essa norma é internalizada, segue-se que nós somos também comprometidos a respeitar o bem estar dos outros, mas a norma, ela mesma, está fundada em nossas emoções. O argumento de Nagel sobre o ressentimento solapa o argumento de Gewirth ao assentar nosso senso de titularidade sobre normas emocionais. Se nosso senso de titularidade está fundado emocionalmente, e não é derivado de premissas puramente racionais, então Gewirth não tem elementos para mostrar que a razão pura é suficiente para a preocupação moral. O argumento de Gewirth é representativo dentre as tentativas de atualizar o programa kantiano, e suas deficiências sugerem que talvez jamais possamos derivar as exigências morais apenas da razão.

Eu penso que há um problema ainda mais sério com o programa kantiano. O racionalismo frígido de Kant não faz justiça às principais intuições subjacentes à hipótese de que os mandamentos morais são categóricos. Essa hipótese deriva, eu suspeito, da intuição de que as exigências morais respondem a uma questão normativa fundamental. Quando nos contam que uma ação é exigida pela

moralidade, não sentimos que precisamos de qualquer fato adicional para que essa ação tenha *status* normativo. Sentimos que *devemos* realizar a ação. A questão normativa pergunta por que é esse o caso. Afinal de contas, quando apresentados a regras externas à moralidade, nós, às vezes, sentimo-nos compelidos a dizer, "eu entendo que essa é a regra, mas por que devo eu segui-la?". Regras não morais deixam um abismo entre a regra e a ação. Regras morais não parecem deixar um abismo. Korsgaard (1996) pensa que Kant tem uma explicação. Se regras morais são regras da razão, então toda vez que haja uma regra moral teremos razão suficiente para agir. O abismo entre regra e ação se fecha. Com regras não morais, a regra é insuficiente por ela mesma. Eu penso que essa solução kantiana falha em seus próprios termos. Kant não pode responder a questão normativa. Eis por que.

O projeto todo de Kant é desenhado para libertar as regras morais das paixões. Eu penso que tal libertação viria com um preço terrível. Se a moralidade fosse puramente racional, nós correríamos efetivamente o risco de nos colocar de volta à mercê do capricho. Uma regra puramente racional não pode ter efeito motivacional em nós. Ela nos deixaria frios. E se pudéssemos ser emocionalmente indiferentes à moralidade, poderíamos falhar em agir moralmente. "Por que obedecer as regras da moralidade?", pergunta o sensível valete. A resposta de Kant é que seria irracional não o fazer. Para isso, o valete tem uma simples réplica: "por que ser racional?". Ou, mais exatamente, o valete poderia dizer: "eu reconheço que é inconsistente valorar a mim mesmo enquanto trato mal os outros, mas eu simplesmente não me importo com ser inconsistente acerca disso. Afinal de contas, ao favorecer a mim mesmo, eu tenho tanto a ganhar!" Aqui, Kant deve ficar em silêncio.

Hume tem uma análise mais satisfatória do caso do valete. Ele diz que pessoa alguma, *constituída como nós*, pode desobedecer, sem custo emocional, a moralidade. Pessoas comuns, felizmente, não levantam questões engenhosas. Essa resposta é paralela à resposta de Hume ao ceticismo. Pessoas comuns não levantam questões céticas. Não podemos evitar acreditar em um mundo externo. Similarmente, alguns de nós não podem evitar ser influenciados pela simpatia e pela indignação moral. A resposta de Hume explica por que regras morais parecer

ter poder sobre nós. Ele fecha o abismo entre regra moral e ação argumentando que, sempre que vemos uma regra como moral, nós somos, dessa forma, movidos a agir a partir dela. Eis porque regras parecem categóricas. A abordagem de Kant sugere que nós podemos ser indiferentes às demandas da moralidade. Portanto, ele falha em responder a questão normativa.

Em resposta, o kantiano irá argumentar que mandamentos morais têm maior peso se eles são fornecidos pela razão. Para Hume, mandamentos morais são mandamentos da paixão, e, como tais, eles podem ser superados por outros estados afetivos. Para Kant, mandamentos morais são razões, não importa o que desejemos. Isso faz com que os mandamentos morais pareçam mais poderosos. Eles se sustentam, não importa o que aconteça. Mas esse movimento sai pela culatra. Primeiro, a visão de Kant tem a mesma limitação prática da de Hume. A moralidade pode ser superada pelo desejo. Alguém pode ignorar um imperativo moral porque deseja outra coisa. Um mandamento moral pode falhar em ter algum impacto. Segundo, a visão de Hume tem a mesma vantagem da de Kant. Mandamentos morais permanecem no lugar quando falhamos em lhes dar atenção. Quando um mandamento moral falha em sua batalha pelo controle, ele não desaparece no esquecimento. Ele espera avidamente em seu canto pelo próximo *round* no ringue. Terceiro, Hume tem uma vantagem sobre Kant. Para Hume, mandamentos morais concorrem diretamente pela atenção da vontade. Para Kant, eles podem existir sem qualquer inclinação para que os sigamos. A visão humeana dá aos mandamentos morais mais peso, e não menos.

Eu tenho tentado levantar dúvidas sobre o caráter racional dos imperativos morais, mas há uma última dimensão da visão kantiana à qual eu não me dirigi diretamente. Kantianos defendem que a moralidade é universal. Verdade morais valem para todos os seres racionais. Da parte das teorias da sensibilidade, a moralidade não precisa ser universal. Mas há um sentido de universalidade que seria uma lástima perder. Esse tem a ver com a natureza prescritiva da moral. No que diz respeito a um módulo de sociedade de psicopatas, tomamos nossas regras morais como aplicáveis aos outros. Se soubermos que alguém tem sido torturado

em terras distantes, condenamos a ação, não importa que juízos morais a pessoa que realiza a tortura possa fazer. Podem as teorias da sensibilidade permitir isso?

É claro que podem. De acordo com as teorias da sensibilidade, o errado é constituído por nossas reações emocionais, e essas reações poder ser direcionadas para eventos em lugares distantes. As teorias da sensibilidade, inclusive, prevêm que esse é o caso. Como Hume observou, alguns de nossos sentimentos mais fortes de desaprovação são derivados de nossa profunda capacidade de simpatia. Sentimos pela vítima de tortura, e esses sentimentos alimentam a indignação. Hume notou que nossas simpatias diminuem com a distância, mas esse é parcialmente um efeito da proximidade. Quando o sofrimento dos outros distantes se torna próximo a nós, nossas simpatias despertam. Juízos morais são abrangentes, em virtude de sua natureza passional. Eles são prescritivos, e nós os prescrevemos a pessoas fora de nossas comunidades. Mas juízos morais ficam aquém de serem universais, ou assim argumentarei nos capítulos 4 e 5. O fato de que aplicamos nossas regras morais aos outros não implica que eles são sempre aplicáveis.

Em suma, eu tenho argumentado que uma regra moral pode valer para mim, mesmo que eu não tenha qualquer inclinação para segui-la em algum momento particular no tempo, pois estou distraído, cansado ou insciente de que a regra é embasada por uma de minhas normas fundacionais. Isso dá às regras morais um caráter categórico. Mas regras morais são categóricas somente de modo fraco; eu não teria obrigação alguma de seguir uma regra que não fosse fundada em minhas disposições sentimentais. Isso significa que regras morais são hipotéticas? Regras morais são, certamente, amplamente condicionais; "roubar é errado" não significa " não roube se se abster do roubo satisfaz algum desejo seu". Imperativos hipotéticos têm duas partes: o imperativo (não roube!), que não tem força motivadora intrínseca, e a cláusula condicional que lhe dá força. Em minha perspectiva, normas morais são, intrinsecamente, guias para a ação. Eu não evito roubar em virtude de reconhecer um fim o qual, abster-me de roubar, alcançarei; pelo contrário, abster-me de roubar é um de meus fins. Mas regras morais não são guias para a ação ou autorizadas de qualquer outra forma para aqueles que não

as tenham internalizado. Como imperativos hipotéticos, regras morais dependem de fins, e esses fins são totalmente contingentes. Poder-se-ia dizer que regras morais não são nem fortemente categóricas, nem estritamente hipotéticas. Elas são psicologicamente categóricas (nós as vemos como fins intrínsecos), e metafisicamente hipotéticas (se não tivéssemos esses fins, eles não teriam autoridade). Kantianos que insistem em um *status* categórico forte para a moralidade quebram o elo entre demandas morais e quereres humanos. Na tentativa de livrar as regras morais das inclinações, os kantianos tiram das regras morais a força motivacional que as faz funcionar como imperativos no sentido psicológico. Regras morais se tornam categóricas, mas pouco imperativas.

Concluirei com uma objeção parcial ao programa kantiano. De acordo com Kant, deveríamos evitar ações que não possamos querer como leis universais. Talvez ele esteja certo. Minha queixa final é que há pouca razão para pensar que essa exigência, se ela é uma, é uma exigência moral. O procedimento da universalização de Kant impõe demandas sobre nós que não se alinham às intuições pré-teóricas sobre as exigências da moralidade. Seu tipo de universalidade não é nem necessário, nem suficiente. Não é necessário para a moral, pois alguns erros morais não colocam qualquer ameaça de contradição lógica ou prática quando universalizados. No capítulo 1 mencionei o tabu do incesto consensual entre irmãos como um possível exemplo. A lei de Kant não é suficiente para a moral, pois ela coloca demandas sobre nós que caem fora do domínio moral. Alguns autores notaram que o procedimento da universalização de Kant poderia implicar uma lei categórica pertencente aos imperativos hipotéticos (Hill, 1973). Enquanto nenhum imperativo hipotético *particular* é categórico, talvez haja um imperativo categórico pertencente aos imperativos hipotéticos em geral: se você quer totalmente um fim, você também quer os meios, necessários para aquele fim, que estão ao seu alcance. Se isso é um imperativo categórico, e imperativos categóricos definem a moralidade, então toda vez que queremos os meios tendo em vista os fins que ocorre de termos, estamos fazendo o que a moralidade demanda. Querer uma ida ao bar quando você deseja uma cerveja se torna um ato moral. Alguns autores argumentam que Kant não tem como distinguir esses imperativos

de imperativos morais *bona fide* (Frei, 2005). Se for assim, o método de Kant para obter universalidade estende o domínio moral para além do reconhecimento. Se elaborarmos Kant como oferecendo uma análise de conceitos morais (ao invés de uma análise de algumas outras classes de obrigações), então ele pode ser acusado de perder lamentavelmente seu alvo. A teoria de Kant torna o domínio moral mais largo e mais estreito do que ele deveria ser. Paixões fornecem ao Procrusto o cavalete e a serra.

3.3 Conclusão

Comecei esse capítulo com uma visão geral de uma teoria da sensibilidade, e então verifiquei dez objeções. Tentei mostrar que essas objeções podem ser respondidas. Teorias da sensibilidade são mais resistentes do que presumem os críticos. Mas não quero, ainda, deixar isso de lado. As teorias da sensibilidade enfrentam duas sérias objeções que requerem uma discussão mais detalhada do que a que ofereci nesse capítulo. Primeiro, há uma intuição persistente de que as morais são objetivas. Teorias da sensibilidade tendem a tropeçar nessa intuição, pois elas estão comprometidas com uma forma, sobretudo forte, de subjetivismo. No próximo capítulo, vou me dirigir a essa preocupação. Segundo, teorias da sensibilidade parecem estar comprometidas com o relativismo. Eu tenho mencionado isso ao longo desse capítulo. O relativismo é, às vezes, considerado uma doutrina vil. Diz-se que ela é repreensível, equivocada e mesmo incoerente. Argumentarei o oposto no capítulo 5.

4. CONTRA A OBJETIVIDADE

A teoria da sensibilidade moral é uma teoria subjetivista. Para alguns, essa é uma razão suficiente para rejeitá-la. A moralidade é objetiva, eles insistem. É um fato objetivo que matar pessoas inocentes é errado. Subjetivismo perde isso de vista. Ele torna a moralidade dependente demais de nós.

Subjetivistas podem responder a esta objeção de uma dessas duas maneiras. Uma estratégia é negar que a objetividade e a subjetividade sejam incompatíveis. Outra é a negar que a moralidade seja objetiva. A primeira estratégia tem sido mais popular, mas a segunda é mais viável. Objetividade e subjetividade são compatíveis, mas a moralidade não é objetiva. É isso que pretendo argumentar.

Começarei expondo algumas definições de objetividade. Depois, eu farei um levantamento das teorias éticas que, de uma forma ou de outra, defenderam o objetivismo. Algumas dessas teorias objetivistas são compatíveis com o subjetivismo, outras não. Todas elas, eu defendo, têm falhas insuperáveis. Se eu estiver certo, talvez seja hora de abandonarmos a esperança no objetivismo ético. Por fim, eu concluo expondo que um afastamento do objetivismo não precisa ser um afastamento do realismo. Há fatos morais; eles são apenas puramente subjetivos.

4.1 ESPÉCIES DE OBJETIVIDADE

No capítulo 3, mostrei como os teóricos da sensibilidade podem satisfazer a intuição pré-teórica de que os juízos morais podem ser verdadeiros ou falsos. Uma forma de conduta é verdadeiramente errada para alguém se essa pessoa tem um sentimento de desaprovação em relação a ela. Eu disse que isto captura uma intuição de que a moralidade tem algo a ver com a forma como as coisas realmente são no mundo, sem perder a intuição de que juízos morais derivam de nós. Essas duas coisas parecem incompatíveis, e a alegação de incompatibilidade reside no coração do argumento de Mackie de que os conceitos morais são incoerentes. Ao provar que os juízos morais podem ser tanto verdadeiros quanto intrinsecamente

motivantes, os teóricos da sensibilidade minam o argumento de Mackie (McDowell, 1985). Mesmo assim, um defensor de Mackie pode queixar-se de que isto é insuficiente. Para Mackie (1977: 94), as nossas intuições não exigem apenas que os juízos morais possam ser verdadeiros, mas que possam ser objetivamente verdadeiros: “[J]ulgamentos morais comuns incluem uma alegação de ‘objetividade’ [...]. E eu não penso que seja ir longe demais dizer que esta suposição tem sido incorporada ao significado básico, convencional, de termos morais”. McDowell (1985) pensa que Mackie está correto e tenta mostrar que as teorias de sensibilidade podem permitir a objetividade moral. Eu creio que Mackie esteja errado.

Antes de construir esse argumento, é necessário deixar claro o que é objetividade, e por que o discurso moral pode parecer comprometido com a hipótese de que as verdades morais são objetivas. O termo “objetividade” é usado de diferentes formas. E eu vou considerar algumas.

De acordo com a primeira definição, a objetividade está ligada à imparcialidade. Um fato objetivo é aquele certificado por um juízo objetivo, aquele que uma pessoa faria se tivesse todas as evidências, nenhum preconceito e boa capacidade de raciocínio. Leiter (2001) chama a liberdade de preconceitos de “objetividade epistêmica”. Tem algo a ver com o processo pelo qual obtemos conhecimento. Nagel (1986) invoca, nesse sentido, a visão de lugar nenhum.

A imparcialidade é claramente compatível com a subjetividade. Subjetividade pode ser definida em termos de dependência da mente, e um juízo imparcial pode ser um juízo sobre algo que depende da mente. De fato, a imparcialidade sempre faz referência à mente, na medida em que é uma denominação epistêmica. Juízes, juízos ou opiniões são imparciais; fatos não. Assim, quando dizemos que um fato é objetivo, dentro da presente definição, estamos dizendo que ele é certificado por um método imparcial para fixar crença. Isto significa que fatos objetivos podem depender de nós. Por exemplo, pode ser um fato objetivo que a Missa de Bach em B-menor é evocativa, porque seres como nós reagiríamos emocionalmente a ela se nos livrássemos de quaisquer preconceitos que nos impedem de apreciar a música (por exemplo, uma aversão desdenhosa à composição barroca ou aos temas

cristãos na arte). Ser evocativa é, bastante plausivelmente, uma propriedade subjetiva. Assim, esta característica da Missa é objetivamente subjetiva. Chamarei isso de imparcialidade subjetiva.

Algumas verdades imparciais não são subjetivas. O termo “imparcialidade” é, por vezes, usado para designar juízos livres de respostas subjetivas. Ou, mais especificamente, ele é utilizado para designar juízos desapaixonados, juízos que não derivam de sentimentos subjetivos, mas de regras universais da razão. Chamemos esses juízos de racionalmente imparciais. Tais julgamentos são objetivos, na medida em que eles são imparciais e racionais, na medida em que são desafiados como referência a certos tipos de procedimentos que certifiquem o conhecimento. No entanto, juízos racionalmente imparciais não são subjetivos no sentido que considereí. Eles não designam fatos que se sucedem em virtude de nossas disposições para ter algumas respostas. Juízos racionais representam fatos que seriam verdadeiros, mesmo que não os apreciássemos. Eles são demandados pelas regras da razão, restrições que regem inferências preservadoras de verdade – restrições que nós frequentemente falhamos em obedecer.

Imparcialidade é um construto epistêmico. Leiter contrasta a objetividade epistêmica com uma variedade não-epistêmica que ele chama de objetividade metafísica. Ele caracteriza a objetividade metafísica em termos de independência da mente. Todavia, a noção de independência da mente é, em si, ambígua. Por um lado, a independência da mente pode significar existir independentemente de algum estado mental. Um fato objetivo, nesse sentido, é aquele que poderia acontecer em um mundo sem mente alguma. Ele aconteceria mesmo que mentes nunca viessem a ser e nunca pudessem vir a ser. Chamemos isso de objetividade transcendental. O fato de que olmos perdem suas folhas no outono é transcendentalmente objetivo. Realistas internalistas ou construtivistas radicais podem discordar argumentando que a categoria ‘olmos’ depende de sistemas humanos de classificação, mas não precisamos ser distraídos aqui por esses pontos de vista. A afirmação de que os olmos perdem suas folhas é feita verdadeira, pelo menos em parte, por algo que seria o caso em um mundo sem mentes.

Por outro lado, independência pode significar algo como não-representado. Alguma coisa tem uma propriedade objetiva, nesse sentido, se ter essa propriedade não depende de ser representado enquanto tendo essa propriedade. Chamemos isso de independência de representação. Eu uso o termo 'independência de representação' ao invés de 'independência de resposta', porque a palavra 'resposta' é muitas vezes utilizada, mais estreitamente, para se referir aos estados afetivos ou perceptuais. Eu penso que respostas sejam representações, mas algumas representações não são respostas, por isso ficarei com o termo mais amplo. Uma propriedade pode ser independente de representação no sentido forte ou no sentido fraco. Uma propriedade independente de representação no sentido forte é aquela que pode ser instanciada, mesmo que ninguém jamais tivesse sido disposto a representá-la. Propriedades transcendentemente objetivas são independentes de representação no sentido forte, mas independência forte de representação é uma noção mais fraca. Fatos sobre a psicologia humana são frequentemente independentes de representação, mas não transcendentais. Trivialmente, eles não poderiam existir em um mundo sem mentes. Para dar um exemplo arbitrário, considere o fato de que os seres humanos podem reconhecer objetos dentro de 100 milésimos de segundo de apresentação visual. Este não é um fato transcendental. É um fato psicológico. Mas ele é independente de representação, porque a nossa velocidade de reconhecimento seria a mesma, ainda que nunca tivéssemos sido capazes de medir ou pensar sobre isso. Chamemos os fatos que são independentes de representação no sentido forte, mas não independente da mente, de "fatos psicológicos". Estes são fatos sobre a mente, não fatos criados pela mente.

Fatos psicológicos podem ser contrastados com fatos independentes de representação no sentido fraco. Estes são os fatos criados pela mente. Dennett (1991) dá o exemplo de ser um suspeito de um crime. Alguém pode ser um suspeito apenas como resultado de ter sido suspeitado. Quando Hercule Poirot considera a possibilidade de que o mordomo tenha cometido o assassinato, seu estatuto metafísico se altera. Ele se torna um suspeito. E ser suspeito exige que ele seja representado como tal. Contudo, há um sentido no qual ser um suspeito é perfeitamente objetivo. Uma vez que o mordomo ficou sob suspeita, ele continua a

ser um suspeito até que a suspeita seja oficialmente retirada (por exemplo, quando o verdadeiro culpado é identificado). De fato, ele pode continuar a ser um suspeito, mesmo que Hércule deixe de suspeitar dele ativamente. Suponhamos que Hercule morra repentinamente e ninguém assuma seu caso. O mordomo permaneceria um suspeito. Ser um suspeito não depende da existência de alguém nutrindo suspeitas no presente momento, mas de uma única ação intencional de suspeita feita por uma pessoa com competência oficial de investigação. A propriedade de ser um suspeito pode ser chamada de um 'produto intencional'.

Há muitos tipos diferentes de produtos intencionais. Suponhamos que um explorador viking apelide seu navio de *Skålbåd* e, quando encontrássemos este navio, séculos mais tarde, não houvesse registros de seu nome. É um fato objetivo que este navio é o mesmo *Skålbåd*, ainda que esse fato seja o resultado de uma ação intencional. Isto é verdadeiro no caso de artefatos em geral. Se alguém produz um objeto imaginando-o como um peso de papel e, posteriormente, esquece-se de sua função, ele continua a ser um peso de papel. E a classe de produtos intencionais inclui também alguns fatos institucionais. Se Fred e Wilma se casassem, mas esquecessem-se de seus votos devido à senilidade, eles continuariam casados. O casamento não é estabelecido apenas por intenções. Outros fatores institucionais são necessários, mas esses fatores dependem, em última instância, das intenções e práticas passadas de uma população humana. Eles são, pelo menos parcialmente, produtos intencionais.

Eu identifiquei vários tipos de fatos objetivos. Eles são baseados em tipos de coisas diferentes. Fatos subjetivamente imparciais são baseados em respostas subjetivas; fatos racionalmente imparciais são baseados em regras da razão; fatos transcendentais são baseados em como as coisas são no mundo externo à mente; fatos psicológicos são baseados na psicologia humana; e produtos intencionais são baseados em intenções e atos intencionais humanos. Alguns desses são compatíveis. Fatos verificados com imparcialidade podem ser produtos intencionais ou verdades transcendentais. Alguns são incompatíveis. Fatos transcendentais não podem ser fatos psicológicos, fatos subjetivamente imparciais ou fatos produzidos intencionalmente. E algo pode ser um fato objetivo de qualquer

uma destas espécies sem que seja um fato objetivo de qualquer uma das outras espécies. Então, os teóricos da ética que afirmam que os fatos morais são objetivos devem ser muito específicos sobre o tipo de objetividade que têm em mente. As diferenças entre as principais teorias éticas objetivistas podem ser caracterizadas em termos da taxonomia que ofereço aqui. Por exemplo, Kant está comprometido com a imparcialidade do juízo moral racional, mas não com a ideia de que a moralidade seja independente da psicologia humana. A autoridade que as regras morais têm sobre nós tem algo a ver com a liberdade humana e com a estrutura da vontade. Se não agíssemos a partir de leis, então não necessitaríamos fazer universalizações na busca por razões, e, se nós não o fizéssemos, a razão não nos levaria a buscar o bem e evitar o mal. Tendo examinado o programa kantiano no capítulo 3, não vou discutir novamente imparcialidade racional. Mas alguns teóricos da ética tendem a outros tipos de objetividade. Eu coloco alguns exemplos principais na próxima seção, mas eu não penso que alguma dessas teorias tenha sucesso em encontrar um fundamento objetivo para a moralidade.

4.2. AS VERDADES MORAIS SÃO OBJETIVAS?

4.2.1 Imparcialidade: observadores ideais

A ideia de que a moralidade possa ser tanto objetiva quanto subjetiva tem atraído uma variedade de autores. As intuições populares sobre a moralidade parecem tender a ambas as direções. Ao defender teorias da sensibilidade, evoquei a intuição de que os juízos morais são inextricavelmente vinculados às reações emocionais. A ênfase a esta ligação promove uma espécie de subjetivismo. Por outro lado, tendemos a pensar que as pessoas são capazes de estar erradas em suas reações morais. Verdades morais não são como reflexos patelares. Se nossos juízos mudam depois de considerarmos uma situação com maior profundidade, dizemos que os nossos juízos anteriores eram falsos. Isso dá a impressão de que os juízos morais são objetivos.

Para acomodar essas duas intuições, alguns filósofos têm endossado teorias do observador ideal. Nesta abordagem, algo é certo (ou errado) se provoca aprovação (ou desaprovação) em condições ideais de observação. Hume (1757) é interpretado como defensor de uma teoria deste tipo de juízo estético. Ele diz que uma obra de arte é boa se é analisada favoravelmente por um bom crítico. Um bom crítico é alguém livre de preconceitos e com a delicadeza de gosto, o bom senso e a prática em fazer juízos estéticos. Hume implica que todos os bons críticos concordarão em suas avaliações estéticas. Baille (2000) sugere que a teoria moral de Hume possa ser interpretada em moldes similares. Talvez haja críticos morais ideais (ver Sayre-McCord, 1994, para uma interpretação diferente de Hume). Francis Hutcheson (1738) pode ter sustentado uma visão como essa. Em tempos mais recentes, o mesmo tem sido defendido por Firth (1952). Ele diz que o observador ideal moral é onisciente, onipercipiente, desinteressado e desapaixonado. Naturalmente, observador real algum pode conseguir essas coisas. O melhor que podemos fazer é tangenciar esse ideal. Ao fazê-lo, Firth sugere, temos uma chance maior de emitir juízos morais verdadeiros.

Se as teorias do observador ideal estão certas, a moralidade, então, é objetiva em um dos sentidos discutidos acima. A moralidade é subjetivamente imparcial. Mas estariam certas as teorias do observador ideal? Antes de responder esta questão, quero ressaltar que as teorias do observador ideal não garantem universalidade (ver Brandt, 1954). É bem possível que duas pessoas com acesso a todos os fatos e sem investimentos pessoais possam formar diferentes avaliações morais da mesma situação. Nossas reações emocionais são influenciadas pela biografia, personalidade, valores básicos e por outras diferenças individuais não-cognitivas. Tal variabilidade é óbvia quando mudamos de questões de valor para questões de gosto. Dois observadores plenamente informados podem discordar se um bom Bordeaux é melhor que um bom Burgundy. Enquanto sentimentalista, o defensor da teoria do observador ideal deve admitir que valores não sejam logicamente implicados pelo conhecimento factual, de modo que o acesso à informação idealizada subdetermine juízos avaliativos do observador ideal. Se outros fatores sustentam-se sobre esses juízos e esses fatores são variáveis,

observadores ideais não chegarão a um consenso. Observadores ideais são imparciais, portanto, objetivos em um sentido, mas divergentes em outro.

Em face disso, esta observação não enfraquece a teoria do observador ideal. Ela apenas ilustra a surpreendente compatibilidade entre objetivismo e relativismo. Em uma análise mais minuciosa, no entanto, ela expõe um problema sério. Ao apresentar a teoria do observador ideal, assumi simplesmente que caracterizações populares de observadores ideais qualificam-se como caracterizações de observadores imparciais. Esta suposição pode ser falsa. Pois, se um observador tem um conjunto de valores básicos e esses valores refletem sua educação moral, enculturação ou personalidade, então ele não pode ser considerado livre de preconceitos. Na verdade, a ideia de um avaliador sem preconceitos pode ser incoerente. Avaliar é expressar os próprios preconceitos. Preconceitos são preferências pessoais. O melhor que alguém pode fazer é tentar libertar-se de suas preferências pessoais a partir das características idiossincráticas da sua própria biografia. Mas quão impessoal pode ser um observador? Pode um observador descartar as preferências transmitidas pela cultura? Talvez não. Na prática, ignorar as preferências culturalmente impostas é muito difícil, porque as preferências são construídas em nossas respostas afetivas e estas não são passíveis de controle direto ou deliberativo.

A teoria do observador ideal poderia responder que tais limitações práticas são irrelevantes. Condições de observação ideal são ideais, e elas apenas podem ser abordadas assintoticamente. O que importa é o princípio, não a prática. Mas também pode haver aqui preocupações baseadas em princípios. Por que supor que uma pessoa que poderia, *per impossible*, descartar preferências culturalmente transmitidas teria alguma preferência moral restante? Respostas morais exigem emoções morais, e emoções morais podem surgir através da enculturação. Uma pessoa desprendida da cultura poderia ser incomodada por certas coisas e entristecida por outras, mas não há garantia de que ela teria plenos sentimentos morais, já que estes, como sugeri, são adquiridos através de um regime de punição, retirada de afeto e indução durante o desenvolvimento moral (ver capítulo 1). Uma vez que uma pessoa tenha desenvolvido uma capacidade amadurecida em relação

aos sentimentos morais, não há garantia de que ela tenha alguma opinião sobre como aplicar esses sentimentos sem assumir traços culturais. A visão de lugar nenhum pode ser confusa e moralmente indiferente. Se nós nos livrássemos da cultura, nos livraríamos também da moralidade.

Outra preocupação é que cortar os laços com os preconceitos culturais pode deixar-nos com respostas que são repugnantes aos padrões culturais. Para recuperar as preferências que não são distorcidas pela cultura, temos de voltar a um estado pré-cultural. Pensar em tal retorno desperta curiosas fantasias românticas; contudo, o sonho de Rousseau pode ser um pesadelo. Esquemas de preferência que não são culturalmente instruídos podem autorizar impropriedades, brutalidades e estupro. Se o comportamento dos grandes símios é alguma indicação de que os valores humanos seriam semelhantes sem o edifício da história cultural, os valores da mente desenculturada dificilmente são dignos de deferência. Além disso, a mente sem cultura não é uma mente sem preconceitos. É apenas uma mente cujos preconceitos refletem mais a natureza do que a cultura.

O que tenho sugerido é que é imprudente buscar uma verdadeira imparcialidade. Mas a questão mais importante na avaliação da teoria do observador ideal é descritiva. As pessoas, de fato, consideram autoritários os juízos morais imparciais? Esta questão teria que ser investigada empiricamente, mas estou disposto a apostar que formas fortes de imparcialidade não desempenham um papel proeminente no pensamento moral. Reconhecidamente, nós valorizamos a reflexão cuidadosa, o conhecimento factual completo e o desinteresse ao decidir o que é certo. Os sistemas legais são uma prova disso. Mas também temos uma parte nos valores dos grupos aos quais pertencemos. Consideremos as regras morais que regulam as relações de parentesco. Quando perguntados se as relações monogâmicas são corretas, eu diria que a maioria das pessoas tem pouca tendência a considerar a questão a partir de uma postura transcendental. Elas não concordariam com alguém que dissesse que, em outros contextos culturais, a monogamia nem sempre é valorizada. Elas diriam, ao invés disso, que "nós valorizamos a monogamia" e que é isso o que realmente importa. Colocando a questão de modo mais geral, ao fazer juízos morais, nós não tentamos acomodar o

que qualquer um valorizaria; nós tentamos acomodar o que *nós* valorizamos, onde “nós” refere-se ao avaliador e a seu grupo cultural. Se nós valorizamos a democracia, e as pessoas em outro contexto cultural, não, nós temos pouco interesse em fazer juízos morais a partir de seus pontos de vista, bem como pouca esperança de encontrar uma base comum útil.

Em suma, eu suspeito que as teorias do observador ideal, quando tomadas como se exigissem imparcialidade total, não são descrições precisas de nossos objetivos no raciocínio moral. Imparcialidade completa não é nosso ideal e a imparcialidade parcial não é completamente objetiva. Se eu estiver certo, procuramos boas condições de observação porque queremos discernir o que está de melhor acordo com nossos preconceitos, não porque queremos liberdade dos preconceitos.

4.2.2. Transcendentalismo: espécies naturais

A ideia de que verdades morais são transcendentais no sentido pleno parece peculiar à primeira vista. A moralidade governa as vidas das criaturas inteligentes. Ela não tem lugar em um universo sem mentes. A ideia de que as regras morais possam existir independentemente da mente parece bastante implausível. Em uma reflexão mais aprofundada, no entanto, essa ideia, na verdade, tem uma longa e distinta história. Muitas pessoas acreditam que a moralidade nos tenha sido entregue por um ser sobrenatural. Que a moralidade seja um mandamento divino. Nesse ponto de vista, a moralidade teria conexões importantes com a mente: é destinada às criaturas com uma faculdade para o entender, destina-se a regular os assuntos humanos e é imposta por um ser de inteligência suprema. Em um sentido mais restrito, contudo, uma moralidade do mandamento divino é transcendental. Para constatar isto, deixe-me distinguir entre dois tipos de dependência da mente. Primeiramente, algo pode ser dependente da mente na medida em que sobrevém aos fatos psicológicos; se fosse diferente, a psicologia teria de ser diferente também. A superveniência psicológica subsume casos do que, acima, chamei de dependência de representação. Fatos dependentes de representação não podem

mudar sem que mudemos o que nós representamos (ou o que estamos dispostos a representar). Em segundo lugar, algo pode ser dependente da mente na medida em que é *aplicável* apenas às agências psicológicas sem que seja superveniente sobre elas; por exemplo, se algo é uma "lição", deve ser uma lição para uma mente. Mandamentos divinos são dependentes da mente neste segundo sentido. Eles são aplicáveis apenas a mentes. Mas a teoria do mandamento divino não é totalmente dependente da mente no primeiro sentido. Um ser sobrenatural pode mudar as regras morais sem mudar nossa psicologia. Certamente, os mandamentos divinos sobrevivem aos fatos psicológicos acerca dos legisladores divinos, mas eles não necessariamente sobrevivem nas mentes daqueles que governam. Assim, os mandamentos divinos transcendem as mentes humanas neste sentido. Se restringirmos o transcendentalismo a esta forma (a moralidade não é geralmente superveniente sobre a psicologia humana), então as teorias transcendentalistas da moralidade começam a parecer possíveis.

Teorias do mandamento divino não gozam da popularidade que tiveram outrora. Muitos teóricos da moral já desistiram dos deuses, e muitos outros foram convencidos pelo tipo de raciocínio apresentado no Eutífron de Platão de que, mesmo que os deuses existam, a moralidade deve ter uma fundamentação não-teológica. Na teoria do mandamento divino, a afirmação de que Deus é bom apenas significa que Deus age de acordo com seus próprios desejos. Isso é um falso louvor. Assim, a moralidade deve ter outra fonte. Acontece que o *insight* de Eutífron é ampla, embora implicitamente, reconhecido por pessoas de forte fé religiosa. Nucci (2001) realizou uma série de experimentos (mencionados no capítulo 1) para investigar os valores morais de jovens de diferentes grupos religiosos: católicos, judeus ortodoxos, judeus conservadores e menonitas. Ele queria saber se estes considerariam errado um ato imoral, tal como ferir ou roubar, mesmo na ausência de uma prescrição divina contra eles. Nucci lhes perguntou se estas coisas ainda seriam erradas se Deus não as tivesse proibido. A resposta foi avassaladora. Entre 80 e 100 por cento das crianças religiosas em seus estudos disseram "sim": roubar e ferir é errado, independentemente de mandamento divino. Isso sugere que as teorias do mandamento divino não conseguem captar plenamente os conceitos

morais dos religiosos. A falha principal nas teorias do mandamento divino não são as premissas teológicas sobre as quais se baseiam. O problema é que a maioria das pessoas realmente não acredita que a moralidade dependa de um mandamento divino, mesmo que sejam religiosas ou que concordem com tal afirmação. A teoria do mandamento divino não é uma boa análise dos conceitos morais comuns.

Isso prova que as teorias transcendentais da moralidade estejam erradas? Talvez não. O transcendentalismo não precisa ser teologicamente fundado. A moralidade baseada na razão de Kant chega perto de ser transcendental, mas não percorre todo o caminho necessário, posto que a normatividade depende, para Kant, de aspectos da natureza humana (especialmente nossa capacidade de agir a partir do dever). Para que se veja com que se parece uma teoria transcendental não-teológica, imaginemos um programa que tenha tentado defender verdades morais por meio do apelo às características intrínsecas às situações pelas quais temos consideração moral. Alguém pode tentar argumentar que a crueldade é errada em virtude de suas propriedades intrínsecas, não em virtude de como nós somos constituídos. Nesse viés, erro e correção seriam propriedades naturais ao invés de propriedades sobrenaturais. Realmente, elas seriam como espécies naturais na medida em que suas existências não dependem de fatos a nosso respeito. Certo e errado seriam como urânio e elefantes.

Meu objetivo aqui não é detalhar uma teoria nesse sentido. Abordarei uma questão anterior. Tomamos nossos conceitos morais para referirmo-nos às espécies naturais? Os conceitos morais visam indicar características do mundo que sejam independentes da mente? Uma breve discussão da teoria do mandamento divino já nos aponta uma resposta. Se o garoto menonita representa o resto de nós, então os conceitos morais não pressupõem uma fundamentação transcendental.

A ideia de que conceitos morais são conceitos de espécie natural enfrenta algumas dificuldades. Nós geralmente assumimos que as essências da espécie natural são desconhecidas para nós e que as descobrimos por meio de instrumentos científicos. Mesmo quando não temos à nossa disposição tais instrumentos, temos muitas vezes a sensação de quais são os tipos de instrumentos necessários e para onde devemos olhar. Isto contrasta com o caso

moral. Nós não presumimos que as propriedades morais tenham essências escondidas nesse mesmo sentido. Que ferramentas usamos para descobrir o bem? Por onde devemos começar a olhar? Que tipo de essência tem o erro?

Conceitos morais e conceitos de espécie natural também apresentam diferentes padrões de deferência. Ao lidar com conceitos de espécie natural, nós assumimos que há especialistas com acesso especial às essências das coisas. Nós as deixamos para químicos, biólogos, geólogos e outros que estão no ramo de descobrir essências escondidas. Quanto aos conceitos morais, porém, somos menos propensos a deferir-los. Podemos ouvir atentamente a membros da família, líderes comunitários ou autoridades religiosas, mas também presumimos que, pessoalmente, temos acesso às verdades morais. Autoridades morais podem ajudar a nos orientarmos para o bem, mas, quando o fazem, é para nos ajudar a ver as coisas cuja capacidade para discernir nós já temos por conta própria. Nós nos consideramos autoridades morais. Quando nossas próprias intuições morais divergem da moral sustentada pelos especialistas, tendemos a pensar que os especialistas estejam errados. Nós não afirmamos ter tal autoridade no que diz respeito a tigres ou titânio.

Além disso, existe um contraste semântico importante entre conceitos de espécie natural e conceitos morais. Com conceitos de espécie natural, podemos construir casos do tipo terra-gêmea (Putnam, 1975). Duas pessoas, cópias exatas uma da outra em todas suas propriedades intrínsecas, poderiam ter conceitos referentes a coisas diferentes se vivessem em mundos diferentes. Em um universo paralelo onde as coisas que se parecem com o ouro tem uma constituição química diferente, a palavra "ouro" e seu conceito correspondente referir-se-iam a algo diferente do ouro. Porém, este não parece ser o caso com os conceitos morais (ver Horgan e Timmons, 1990-1). Se considerarmos duas pessoas exatamente iguais e as situarmos em diferentes mundos, parece que seus conceitos morais serão os *mesmos* e seus conceitos de certo e errado referir-se-ão às mesmas propriedades. Não se trata de negar o fato de que alguns juízos morais mudariam valores de verdade se o mundo fosse diferente. Por exemplo, não haveria problema algum na criação confinada se as vacas não fossem sencientes. Mas nenhum fato sobre o

mundo parece sustentar-se em valores morais centrais. Se, para você, é um valor básico que os seres sencientes não devem ser torturados, então o mesmo valor é sustentado por sua duplicata na Terra Gêmea e, intuitivamente, os conceitos que entram neste valor têm as mesmas referências, e nenhum fato sobre o mundo poderia tornar seu valor verdadeiro enquanto o dele é falso. Posto que não é totalmente claro como resolveríamos uma disputa moral entre contrapartes da Terra e da Terra Gêmea, que fatos sobre o mundo poderiam resolver qual lado do debate está certo? Intuições sobre esses casos sugerem que nós não interpretamos conceitos morais da mesma forma que interpretamos conceitos de espécie natural.

Juntas, essas observações sugerem que nós não consideramos que os conceitos morais refiram-se a propriedades com essências profundamente escondidas, mas isso não esgota as opções disponíveis ao objetivista transcendental. Talvez conceitos morais se refiram a aspectos do mundo independente da mente que não estejam ocultos de alguma maneira profunda. Consideremos conceitos que se refiram a qualidades primárias lockeanas, tais como conceitos de forma. As formas são facilmente discernidas. Todos nós podemos ser peritos em formas. Talvez conceitos morais pretendam referir-se a algo assim.

Para que se teste a independência da mente, o que realmente importa não são essências escondidas, mas propriedades morais. Nós assumimos que entidades mentalmente independentes possam existir sem nós. Urânio e ouro existiriam se não existíssemos. O mesmo vale para as formas. Nós assumimos que as formas existiriam sem nós. De fato, elas já têm existido sem nós. O mundo tinha formas antes de existirem as mentes. Coisas eram redondas, angulares ou oblongas antes que houvesse seres sencientes para apreciá-las.

Mas e as propriedades morais? Presumimos que poderiam existir sem nós? Em certo sentido, obviamente não. Um mundo sem seres altamente inteligentes não tem lugar para a moralidade. Ao contrário de formas, as propriedades morais, tais como correção ou erro, parecem ser instanciadas apenas quando há criaturas que podem apreciá-las. Quando uma leoa mata um gnu, não dizemos que ela fez algo imoral. Imoralidade requer a capacidade de compreender as regras morais. Isso

torna a imoralidade um tanto diferente da propriedade de ser redondo ou quadrado.

O defensor da independência da mente pode admitir que as propriedades morais são diferentes das formas e ainda sustentar a esperança em alguma espécie de objetivismo. Lembre-se do meu exemplo anterior sobre uma aula. Aulas não poderiam existir a menos que houvesse alunos, mas ser uma aula não sobrevém aos fatos psicológicos. Talvez propriedades morais sejam assim. São propriedades que se relacionam com mentes sem sobrevir às mentes.

Entretanto, esta resposta não ajudará, posto que aulas diferem da moralidade em um aspecto crucial. Para que seja uma aula, algo só precisa ser dirigido a um aluno. O aluno não precisa ter a capacidade de compreender as aulas; aulas podem ser mal planejadas. Sua conexão com a mente tem a ver com a sua finalidade, não com seu conteúdo, ainda que, em casos de sucesso, o conteúdo seja bem projetado para seu propósito. Regras morais parecem ser diferentes. Elas devem ser compreensíveis para que sejam aplicáveis. Quando uma leoa não tem restrições contra matar criaturas inocentes, não dizemos que ela inadvertidamente cometera uma iniquidade. Se as regras morais fossem produtos de um mandamento divino, poderíamos dizer que são como aulas. Podemos dizer que elss poderiam existir mesmo sem um público que as compreendesse. Mas nós não parecemos considerar as regras morais dessa forma. Nós não consideramos que as regras morais possam existir em um mundo sem criaturas inteligentes. Essa suposição é, na melhor das hipóteses, não-obrigatória e, na pior, incoerente. Se nossos conceitos morais fossem comprometidos com uma forma forte de independência da mente, nós poderíamos abraçar essa suposição sem hesitação.

Defensores da independência da mente podem tentar manter sua posição invocando outras intuições modais. Ao invés de imaginar um mundo sem criaturas inteligentes, imaginemos um mundo onde as criaturas inteligentes simplesmente deixem de apreciar as regras morais. Intuitivamente, matar pessoas inocentes seria errado, mesmo que ninguém reconhecesse que isto é errado. De fato, matar pessoas inocentes seria errado, mesmo se a raça humana tivesse se tornado tão corrompida que não estivessemos sequer dispostos a reconhecer o erro em tais atos. Isso dá a impressão de que o erro goza de um tipo de independência da

psicologia. A intuição invocada é especialmente embaraçosa para as teorias da sensibilidade. O erro não está conceitualmente ligado a disposições responsivas.

Não estou certo se realmente temos as intuições modais evocadas. Algumas pessoas podem dizer tranquilamente que matar pessoas inocentes seria correto se as pessoas não estivessem dispostas a pensar o contrário. Em um mundo pós-apocalíptico, um mundo onde imperasse a lei do mais forte, a moralidade atual pode não ser aplicável. A moralidade talvez não possua lugar no Estado de Natureza hobbesiano. A inteligibilidade desta suposição pode ser empiricamente investigada. Podemos perguntar a indivíduos se as regras morais se aplicam em uma comunidade largamente corrompida. Até onde eu sei, ninguém realizou um experimento que perguntasse essa questão em particular, mas há alguns resultados empíricos que lançam dúvidas sobre a questão dos usuários de conceitos morais estarem comprometidos com o objetivismo. Por exemplo, Kelly et al. (2007) descobriram que os indivíduos estão dispostos a dizer que certas ações (como um capitão batendo em um marinheiro bêbado) não são corretas para nós, mas são aceitáveis quando feitas por membros de outras culturas. Eles utilizam esse resultado para criticar a distinção moral/convencional e, no capítulo 3, eu sugiro que essa conclusão possa ser prematura. Mas seus resultados levantam a possibilidade de que as pessoas possam estar abertas a certo grau de relativismo sobre a moralidade. (Estudos futuros terão de determinar se os indivíduos pensam que os comportamentos realizados em outras culturas são moralmente corretos ou, por algum motivo, apenas perdoáveis).

Shaun Nichols (2004b) realizou um teste mais direto para analisar se as pessoas são objetivistas morais. Em uma série de estudos, os participantes foram instruídos a imaginar primeiramente uma discussão entre uma pessoa de nossa cultura e uma pessoa cuja cultura tolera bater ou agredir pessoas inocentes "apenas para a diversão". Eles foram, então, perguntados se algum participante deste debate estava objetivamente certo. Em cada estudo, uma parcela considerável (até 66 por cento) deu a resposta não-objetivista. Este resultado é consistente com outras pesquisas que sugerem diferenças individuais na ideologia moral. Forsyth (1980) desenvolveu uma escala para medir os tipos de

personalidade moral que tem sido utilizada em um número considerável de estudos que classificam as pessoas entre uma visão relativista ou absolutista. Ele encontrou uma divisão entre aqueles que postularam valores morais universais e aqueles que negam tais valores. Isto sugere que intuições sobre a objetividade são divididas. Se for assim, há duas possibilidades. Ou as pessoas destes grupos têm conceitos diferentes de moralidade, objetivistas para alguns e não-objetivista para outros. Ou, alternativamente, o objetivismo não é uma verdade conceitual, mas, ao invés disso, uma crença independente sobre a moralidade. Esta segunda alternativa é um pouco mais atrativa porque, nela, os objetivistas morais podem coerentemente perguntar se a moralidade é objetiva. Se o objetivismo fosse uma verdade conceitual, isso seria impossível. Talvez conceitos morais sejam normalmente silenciosos sobre a questão do objetivismo e crenças sobre o objetivismo sejam opções adicionais, que, se surgirem, o fazem mais tarde no desenvolvimento moral.

Deixemos este ponto de lado e assumamos que os objetivistas estejam certos sobre as intuições. Assumamos que as pessoas normalmente considerem que regras morais se apliquem mesmo quando ninguém está disposto a reconhecê-las. Pode o teórico da sensibilidade acomodar intuições sobre erros não reconhecidos? Penso que sim. Uma possibilidade é enrijecer. Deixar que o “errado” se refira ao que quer que seja que as pessoas estejam *atualmente* dispostas a condenar na reflexão. As pessoas *estão* atualmente dispostas a condenar o assassinato. Assim, quando dizemos que o assassinato seria errado mesmo que nos corrompêssemos, estamos fundamentando a palavra em seu uso atual. A sentença é perfeitamente verdadeira. Se no futuro nos tornássemos corrompidos, o assassinato ainda seria aquele que (agora) estamos dispostos a condenar. Enrijecimento dá às verdades morais uma espécie de tenacidade sob condições contrafactuais. As intuições modais podem ser apreendidas dessa forma. Eles não apresentam nenhuma ameaça às teorias da sensibilidade.

Este último ponto mostra que as teorias de sensibilidade podem explicar intuições modais tão bem quanto as teorias objetivistas transcendentais. Assim, temos um impasse. Teorias da sensibilidade assumem na dianteira do

transcendentalismo, quando consideramos os argumentos independentes para o subjetivismo no capítulo 1 e as dúvidas sobre o transcendentalismo levantadas anteriormente nesta seção. Por exemplo, eu sugeri que nós tendemos a pensar que a moralidade apenas exista em um mundo de seres que possam apreciá-la. Que a intuição suporta teorias da sensibilidade. De fato, que a intuição se choca com a intuição de que as regras morais existiriam se todos nós nos corrompéssemos. A proposta de enrijecimento reconcilia essas duas intuições, e as teorias da sensibilidade superam o transcendentalismo, porque podem facilmente lidar com ambas as intuições.

Quero concluir esta seção levantando um ponto que repete uma lição do capítulo 1. Até agora, tenho argumentado contra a hipótese de que os conceitos morais visem a indicar propriedades independentes da mente. Contrário ao transcendentalismo, nós normalmente não assumimos que conceitos morais designem propriedades objetivas nesse sentido. Minha última observação é que os argumentos contra o transcendentalismo ainda seriam possíveis se os conceitos morais, *de fato*, objetivassem indicar propriedades independentes da mente. Um conceito pode tentar indicar tais propriedades e falhar.

Consideremos novamente a cor. É bastante plausível que os conceitos de cor visem a objetividade transcendental. Isso quer dizer que nós consideramos que nossos conceitos de cor designem características reais do mundo externo à mente. Nós localizamos as cores na superfície das coisas e assumimos que o mundo era colorido antes mesmo que entrássemos em cena. Mas suponhamos que o mundo não fosse assim. Suponha que não haja características nas superfícies lá fora que sistematicamente correspondam aos nossos conceitos de cor. Suponha que não há maneira alguma de agrupar as propriedades que nos levam a ter experiências de cores particulares em classes organizadas sem mencionar o fato de que elas causam essas experiências. Concluimos que as cores não existem? Alguns filósofos seguem esse caminho (Hardin, 1993), mas essa não é a única opção. Podemos concluir, ao invés disso, que as cores são apenas diferentes do que inicialmente acreditamos. Podemos concluir que as cores são propriedades dependentes de resposta. Para cada conceito de cor, há um conjunto bastante

heterogêneo de propriedades de superfície unificados pela sua capacidade de provocar certas respostas em nós.

Eu penso que a segunda opção seja melhor do que a primeira. É extremamente útil acreditar que as cores existam, porque que elas desempenham papéis sistemáticos em nossas vidas, e poucas de nossas práticas referentes às cores dependem da suposição de que as cores sejam independentes da mente. A este respeito, cor é diferente de "flogisto", postulada em contextos teóricos e que não desempenha papel algum na vida diária, ou de "bruxa", que não poderia desempenhar seu papel pretendido na vida diária se abandonada a suposição de que bruxas têm poderes sobrenaturais. Considere alguns casos comuns nos quais fazemos referência à cor azul. O gerente de uma boutique de roupas poderia colocar todos os itens azuis juntos em uma única prateleira. Uma mulher poderia saber que está grávida quando uma fina linha azul aparece em seu teste de gravidez. Depois que uma testemunha ocular aponta, a polícia poderia parar todas as pessoas dirigindo sedans azuis. Picasso criou uma série de pinturas em tons de azul. "Azul" é um predicado projetável, pelo menos em um nível macro. Se soubermos que Otto gosta de blazers, camisas e calças azuis, podemos seguramente comprar-lhe uma gravata azul. Todas essas práticas efetivamente reificam o azul, elas são condição suficiente para se dizer que o azul existe. Mas notemos que nenhuma dessas práticas depende da afirmação de que o azul é independente da mente. A categoria das coisas azuis pode não ter unidade intrínseca alguma. Seus limites podem ser inteiramente determinados por nós. Nós chamamos certas coisas de "azuis" porque elas exercem efeitos similares sobre os nossos sistemas visuais tricromáticos. Mas mesmo aqueles efeitos subdeterminam os limites de cor. Em russo, não há palavra para o azul. Há palavras para o azul claro (*goluboy*) e para o azul escuro (*siniy*), mas nenhuma palavra para o azul em geral. Eles traçam um limite categórico entre tons que falantes do inglês agrupam (comparemos nosso vermelho e nosso rosa). Se uma falante do inglês gosta de azul escuro, provavelmente gosta de azul claro, mas, se um orador russo gosta *goluboy*, ele pode não gostar de *siniy* (mais uma vez, comparemos o vermelho e o rosa). Exemplos como este indicam a dependência que o espectro de cores mantém em relação à mente. As categorias dependem de

nossas respostas. Esta é uma descoberta interessante sobre as cores, mas que não nos força a ser antirrealistas quanto às cores. Dada a robustez das práticas humanas em relação às cores, é mais plausível dizer que as cores são reais, mas dependentes da mente. Quando os russos classificam as coisas como *goluboy* ou *siniy*, engendram duas propriedades que podem ser usadas para fazer generalizações, orientar o comportamento, bem como para organizar os itens no mundo.

O conceito ordinário de azul visa indicar uma propriedade independente da mente e falha. Mas essa falha não implica numa referência falha. Pense em AZUL como um conceito com um plano de reserva. Ele faria referência a uma propriedade que é independente da mente se tal propriedade fosse confiavelmente relacionada às nossas experiências do azul. Se não há tal propriedade, ele se refere à propriedade de causar azul em nós, independentemente dessa propriedade ter alguma integridade maior. Eu suspeito que muitos conceitos tenham planos de reserva. O conceito CÃO visa indicar uma espécie natural, mas, se nenhuma espécie natural é confiavelmente ligada ao nosso uso deste conceito, ele refere-se, ao invés disso, a uma classe de criaturas que compartilham de propriedades morfológicas e comportamentais similares. Se oceanos e lagos tivessem sido preenchidos com uma grande variedade de produtos químicos, ÁGUA não indicaria uma espécie natural, mas ainda faria referência [a algo]. Ou, para tomarmos exemplos reais, o conceito ERVA pode referir-se às coisas que tomamos como ervas e PNEUMONIA pode referir-se a algumas condições que são similares nos sintomas, apesar das diferenças de etiologia (bactérias, fungos, vírus, micoplasmas, etc.).

Eu afirmo que mesmo que os conceitos morais visassem a indicar propriedades independentes da mente, eles podem falhar, e, se falham, caem em propriedades que são dependentes da mente. Esta observação complementa a resposta a Mackie iniciada no capítulo 3. Lá, comento a idéia de McDowell de que o teórico da sensibilidade faz juízos morais motivados intrinsecamente e capazes de ser verdade. Nós podemos dizer corretamente que a crueldade é errada, e, assim dizendo, estarmos dispostos a agir. Mas, Mackie poderia responder que as espécies de verdade disponíveis ao teórico da sensibilidade não capturam intuições pré-

teóricas. O pensamento popular sobre a moralidade exige que a verdade moral seja independente da mente. Uma vez que alguma coisa não pode ser, ao mesmo tempo, independente e dependente da mente (ou seja, essencialmente motivadoras), conceitos morais devem ser incoerentes e, portanto, vazios. Em resposta, eu digo duas coisas. Conceitos morais não visam a indicar propriedades independentes da mente, como afirma Mackie, e, mesmo que fizessem, eles não seriam vazios.

Mackie supõe que, se os conceitos morais falham em entender as propriedades objetivas, eles são vazios. Esta suposição baseia-se em uma teoria semântica demasiadamente estrita. Mackie não percebe que conceitos podem conter erros. Putnam (1975) dá o exemplo do conceito de GORILA, que, para muitas pessoas, contém a característica FERROZ. (Putnam usou este exemplo no início de 1970. Talvez seja menos verdadeiro agora que a pesquisa de Dianne Fossey já se transformou em sabedoria popular. Nós já não derivamos mais todo o conhecimento sobre gorilas a partir do filme *King Kong*). Se a ferocidade é uma característica de nossos conceitos de GORILA, nossos conceitos, então, contêm um erro. Gorilas são dóceis. No entanto, nossos conceitos sobre gorilas conseguem denotar gorilas porque contêm uma série de outras características que permitem que o conceito se torne correlacionado aos gorilas. Se os conceitos morais contêm erros, isso não significa que eles sejam vazios. Se eles pretendem indicar propriedades objetivas (o que duvido) e essas propriedades objetivas não existem, então eles podem voltar a representar propriedades que estejam de acordo com outras características. Eles se referem a propriedades que são essencialmente motivadoras, mesmo que não sejam objetivos. A utilidade e a projetabilidade das categorias morais, tais como cores e gorilas, atesta sua não-vacuidade.

4.2.3 Fatos psicológicos 1: teoria da virtude

Objetivistas transcendentais afirmam que propriedades morais são fortemente independentes de nós. Eles afirmam que as propriedades morais são independentes de representações e que a moralidade não depende de nenhum fato psicológico. Alguns teóricos da ética abandonam a segunda alegação, enquanto

mantêm a primeira. Eles dizem que as propriedades morais sobrevivem aos fatos sobre a psicologia a um grau significativo, mas que as verdades morais aconteceriam mesmo que nós não estivéssemos dispostos a reconhecê-las. Já argumentei que nós não estamos conceitualmente empenhados com a independência de resposta. Há, no entanto, uma classe de teorias éticas que parece tornar plausível a independência de resposta. Essas são teorias da virtude. Elas dizem que as verdades morais derivam das verdades sobre as virtudes e que as virtudes derivam de fatos sobre a natureza humana, incluindo nossa natureza psicológica. Teorias de virtudes assumem que há uma ligação, portanto, entre moralidade e psicologia. As virtudes são, tipicamente, tomadas como traços de caráter que permitem ou constituem o desenvolvimento humano (cf. Aristóteles, 350 a.C.). Então, não se pode mudar a moralidade, de acordo com uma teoria da virtude, sem mudar o caráter ou algo sobre o desenvolvimento – algo psicológico. Isso significa que a moralidade é, para o teórico da virtude, dependente da mente, mas não dependente da resposta. Teorias de virtude são teorias objetivistas na medida em que fundamentam a moralidade em fatos psicológicos que poderiam ocorrer sem que tomássemos conhecimento dos mesmos.

Teorias da virtude tornaram-se cada vez mais populares nos últimos anos, mas também estão sob ataque empírico. Os críticos argumentam que traços de caráter são empiricamente indefensáveis. Eu acredito que este ataque é, em última análise, malsucedido, mas a teoria virtude sofre de outras falhas fatais.

Teorias de virtude são teorias normativas, mas, ao contrário de outras teorias normativas, elas não estão, primariamente, no campo de ações prescritivas. Teorias da virtude nos dizem que tipo de pessoas deveríamos ser, não como deveríamos agir. A ação, é claro, entra no debate. Deveríamos fazer como pessoas virtuosas fariam, mas devemos fazê-lo, cultivando as virtudes em nós mesmos. Mas o que, exatamente, são virtudes? Segundo as teorias vigentes, elas são traços louváveis do caráter. Um traço de caráter é um aspecto contínuo da personalidade e que induz alguém a agir ou reagir de maneiras particulares. Aristóteles, que desenvolveu uma teoria da virtude especialmente seminal, enfatizou características tais como coragem, magnanimidade, veracidade, temperança, liberalidade com as finanças,

paciência, perspicácia e amizade. Ele afirmou que cada uma dessas situa-se entre extremos indesejáveis. Coragem situa-se entre a covardia e a temeridade; liberalidade, entre extravagância e frugalidade; amizade, entre obsequiosidade e impertinência. Cada uma destas tríades pertence a um domínio diferente. Coragem pertence à confiança face ao perigo; amizade, às interações sociais, e liberalidade pertence ao gasto.

Geralmente, considera-se que a ética da virtude é fundamentada na natureza humana. Os traços teriam a função de capturar algo sobre como somos, ou melhor, sobre como, idealmente, somos. Os seres humanos desenvolvem-se ao máximo, quando expressam sua humanidade mais plena e frutuosamente, quando eles possuem traços de caráter virtuosos. Estritamente falando, alguém pode ser um especialista em ética da virtude sem invocar esses traços de caráter (Thomson, 1997; Harman, 2001). Pode-se falar, ao invés disso, das ações corajosas, amigáveis, ou magnânimas. Mas as teorias que falam sobre ações, no lugar de traços, renunciam ao vínculo entre a moralidade e a natureza humana, tão importante em teorias da virtude mais ortodoxas. Os imperativos de que devemos agir honesta, corajosa, generosamente, e assim por diante, parecem mais com imperativos das principais teorias normativas éticas. É um retorno à visão de que a ética deve nos dizer como agir, não o tipo de pessoas que devemos ser. Eu focalizarei a ética das virtudes baseada em caráter.

Recentemente, éticas da virtude baseadas em caráter têm estado sob ataque em bases empíricas. As virtudes são tradicionalmente concebidas como traços de caráter, mas muitos psicólogos sociais consideram traços de caráter uma construção defeituosa. O comportamento não é determinado pelo caráter, afirmam esses psicólogos, mas por características externas das situações em que nos encontramos (por exemplo, Mischel, 1968; Ross e Nisbett, 1991). Doris (1998, 2002) e Harman (1999) argumentam que as descobertas empíricas que embasam esta abordagem situacionista são fatais para a ética da virtude baseada em caráter (ver também Flanagan, 1991; Merritt, 2000; e Vranas, 2005). Eticistas da virtude definem traços como amplos, permanentes e eficazes: eles influenciam o comportamento de uma forma consistente ao longo de um vasto leque de situações. Evidências de

psicologia social levantam dúvidas sobre a existência de tais traços. Na realidade, o comportamento é fortemente influenciado pelas circunstâncias. Características menores de uma situação podem influenciar o comportamento e as pessoas que afirmam ter traços de caráter diferentes serão igualmente influenciadas. Se os traços não existem, uma teoria moral que nos pede para cultivar traços, então, está pedindo o impossível.

Vejamos alguns dos exemplos usados por Harman e Doris. Ambos citam um dos estudos clássicos de Stanley Milgram sobre obediência (descrito em Milgram, 1974). Os indivíduos deste experimento foram instruídos a fazer a outro voluntário, localizado em uma sala adjacente, uma série de perguntas. Cada vez que o segundo voluntário falhasse em responder corretamente às perguntas, ao indivíduo que fazia as perguntas era solicitado que administrasse um choque elétrico utilizando um botão com voltagens crescentes. Sem o conhecimento do indivíduo, o segundo voluntário era, na verdade, um ator que estava trabalhando com o pesquisador, e o botão de voltagem, um adereço inofensivo. Os atores foram instruídos a cometer erros de tal modo que os indivíduos tivessem que administrar os choques elétricos. Nas fases pré-planejadas, os colaboradores expressariam dor, manifestariam preocupações sobre a segurança do experimento, fariam sons de agonia, bateriam na parede, ou, finalmente, parariam de fazer qualquer barulho. Se um indivíduo transmitisse relutância em continuar a aumentar a voltagem, o experimentador responderia que isso era essencial para que a experiência continuasse. A experiência terminava se, ou quando, um indivíduo recusava-se persistentemente a continuar.

Surpreendentemente, todos os indivíduos de Milgram giraram o botão para 300 volts – rotulado “choque intenso” – ou mais. Ainda mais preocupante, 65 por cento dos indivíduos de Milgram giraram o botão até o nível máximo de 450 volts, o qual era simplesmente rotulado “XXX”. A essa altura, o colaborador estava totalmente silencioso, indicando inconsciência ou mesmo pior. Milgram não encontrou diferenças individuais significativas em seus indivíduos e todos pareciam ser muito suscetíveis à obediência cega quando persuadidos por uma figura de autoridade.

Harman e Doris também discutem um estudo realizado por Darley e Batson (1973). Eles convidaram estudantes de um seminário a fazer um sermão sobre a parábola do Bom Samaritano. Todos os participantes preencheram questionários sobre suas personalidades e sobre suas convicções religiosas. Os pesquisadores asseguraram-se de que alguns indivíduos se atrasariam para o sermão, alguns teriam apenas o tempo suficiente, e alguns teriam tempo mais que suficiente. Na estrada que leva ao local do sermão, eles encontraram um comparsa deitado, caído, tossindo e gemendo de dor. 63 por cento dos estudantes que tinham tempo ajudaram, enquanto apenas 10 por cento dos que estavam atrasados ofereceram ajuda. Variáveis de personalidade e religiosidade não tiveram impacto.

Outro estudo discutido por Doris mostra que os fatores situacionais que influenciam o comportamento podem ser extremamente triviais. Para ver se os transeuntes o ajudariam, Isen e Levin (1972) fizeram com que um colaborador deixasse cair uma pilha de papéis. De antemão, eles organizaram as coisas de tal modo que, pouco antes, alguns transeuntes encontrassem uma moeda em um telefone público. Oitenta e oito por cento das pessoas que encontraram a moeda ajudaram, ao passo que apenas 4 por cento daqueles que não a encontraram fizeram o mesmo. Diferenças dramáticas no comportamento podem ser impulsionadas por uma pequena recompensa inesperada. Doris conclui que, se tais fatores mínimos circunstanciais podem conduzir à diferença entre a decência humana e a indiferença egoísta, o caráter não pode ter tanto impacto sobre o comportamento.

Essas descobertas são apenas alguns exemplos bem publicizados de uma vasta literatura que exhibe resultados semelhantes. Eles parecem mostrar que o melhor preditor de comportamento não é o caráter, mas a situação. Ética da virtude baseada em caráter sustenta-se em pressupostos fracos da psicologia popular.

Os defensores da ética da virtude responderam a esta crítica de várias formas (Kupperman, 2001; Sreenivasan, 2002; Anás, 2003; Kamtekar, 2004). Eu não vou revisar esta crescente literatura, mas quero mencionar três temas recorrentes. Uma resposta é que as virtudes seriam difíceis de serem alcançadas. Claro, esta resposta é uma expressão da fé na ausência de evidência empírica. Seria mais fácil

acreditar que as características virtuosas são possíveis se houvesse uma boa evidência para traços viciosos. Mas, se os situacionistas têm razão, não há uma boa evidência para traços de caráter amplos e eficazes. Se for assim, alcançar a virtude exigiria o cultivo de mecanismos psicológicos que se afastam radicalmente da maneira como as mentes humanas geralmente operam. Isto parece tornar a virtude não apenas difícil, mas quase impossível.

Outra objeção comum feita a Doris e Harman é a de que eles deturpam a natureza dos traços como esses são entendidos dentro da tradição da teoria da virtude. Teóricos da virtude definem traços como capacidade de raciocínio prático, não como reações inflexíveis e automáticas; assim, o fato de que as pessoas respondem de forma diferente a situações não ameaçaria a concepção de traços mencionada. O problema com esta objeção é que as experiências situacionistas levantam dúvidas sobre habilidades robustas de raciocínio prático. Por exemplo, os indivíduos de Milgrim (1974) se envolveram em raciocínio prático extensivo durante o experimento da obediência, mas os fatores situacionais determinam como eles raciocinam e os levam a fazer escolhas erradas.

Uma terceira objeção a Doris e Harman é que os experimentos podem introduzir fatores que podem ser racionalmente considerados boas razões para se comportar de um modo considerado, por outro lado, sem escrúpulos. Por exemplo, os indivíduos de Milgrim podem formar a crença racional de que um cientista de jaleco não os levaria para o mal e os indivíduos de Darley e Bateson podem formar a crença racional de que seria errado manter o seu professor e o público esperando quando eles foram convidados para dar uma palestra. Nesta explicação, os indivíduos desses experimentos não são cegamente inundados por fatores situacionais estranhos, mas, ao invés disso, deparam-se com fatos relevantes que são usados para tomar decisões racionais, talvez até virtuosas. Considero esta resposta especialmente insidiosa, porque implica que devemos isentar de culpa os indivíduos destes experimentos. Certamente, o temor de atraso não é um bom motivo para se abandonar alguém que esteja gemendo de dor; o público e professor perdoariam – e até aplaudiriam – o aluno que chegasse tarde sob estas circunstâncias. E, certamente, a fé em cientistas não é um bom motivo para ignorar

uma pessoa que esteja implorando por misericórdia; indivíduos no experimento de Milgrim relataram mais tarde terem pensado que seus choques poderiam, sim, ser fatais. Além disso, a sugestão de que as experiências situacionistas introduzem variáveis que podem ser usadas para apoiar racionalmente o comportamento observado não explica os resultados de Isen e Levin. Encontrar uma moeda não é uma *razão* para que alguém seja útil aos outros.

Para minar a crítica a Doris e Harman, seria preciso encontrar evidências experimentais que contradigam a avaliação sombria dos traços de caráter encontrada na literatura situacionista. Tais investigações não são difíceis de ser encontradas. Situacionismo não é a visão dominante na psicologia da personalidade. Muitos psicólogos acreditam nos traços, e baseiam sua fé em resultados de estudos que mostram que certos aspectos da personalidade são altamente estáveis e predicativamente válidos. Certos traços parecem manter-se constantes ao longo dos anos e parecem ter um impacto mensurável e confiável sobre o comportamento. Por exemplo, os psicólogos de personalidade identificaram cinco dimensões de traços altamente fortes (Costa e McCrae, 1992). Uma dimensão, o neuroticismo, inclui traços que abrangem, de um lado, a ansiedade elevada e, de outro, à alta estabilidade. Outra, extroversão, vai da sociabilidade elevada à introversão total. A terceira dimensão é a abertura e inclui traços tais como imaginação, em um pólo, e falta de imaginação rígida, em outro. Agradabilidade, a quarta dimensão, vai de confiabilidade e altruísmo a tipos de personalidade desagradavelmente egoístas. A dimensão final, consciência, varia de ordem, disciplina e senso de dever a seu oposto irreconciliável. Numa meta-análise, Roberts e DelVecchio (2000) descobriram que esses traços são muito consistentes ao longo da vida. Na infância, correlações variam de 0,35 a 0,51 e, na idade adulta, as correlações variam de 0,57 a 0,75. Estes traços são também conhecidos por correlacionar uma variedade de comportamentos variáveis, incluindo o desempenho no trabalho (Ones et al., 1993), relações conjugais (Caughlin et al., 2000), assunção de riscos (Krueger et al., 2000) e comportamento antissocial (Lynam et al., 2005). Traços não são o único fator, é claro, mas eles dão uma contribuição mensurável.

Estes resultados contradizem as formas fortes de situacionismo, mas podem ser conciliados com os resultados situacionistas. É evidente que fatores situacionais podem substituir as disposições de comportamento do caráter. Uma teoria adequada do comportamento permitiria interações entre situação e traços e especificaria as condições nas quais cada fator teria as maiores probabilidades de prevalecer. Buss (1989) oferece um panorama útil das condições que fazem a diferença. Fatores situacionais tendem a influenciar o comportamento nas tarefas quando o contexto é novo, as instruções são detalhadas, a escolha é diminuída, a duração é curta ou a resposta é rigidamente definida. Traços tendem a influenciar o comportamento em tarefas onde o contexto é familiar, instruções não são dadas, uma gama considerável de escolhas está disponível, a duração é longa e as respostas são amplamente definidas. Nos estudos de obediência de Milgram, poucas escolhas são dadas aos indivíduos e suas instruções são detalhadas; nos estudos Darley e Bateson sobre o Bom Samaritano, as instruções também são detalhadas, pois os indivíduos estão supostamente dando uma palestra. No estudo de Isen e Levin, a situação é nova e breve. Assim, os conflitos aparentes entre situacionistas e seus oponentes podem ser resolvidos prestando-se atenção às demandas da tarefa. Uma vez que traços parecem desempenhar um papel importante em algumas tarefas, é seguro concluir que as teorias baseadas em traços são potencialmente relevantes para aquelas tarefas. Cultivar o bom caráter em situações onde o caráter pode ter um impacto pode valer a pena. Resultados situacionistas expõem limites ao âmbito e à eficácia do caráter, mas não os tornam nulos ou sem efeito.

Doris e Harman podem se queixar de que os traços de caráter enfatizados na psicologia da personalidade têm pouca semelhança com as características enfatizadas na tradição de ética da virtude. Traços como neuroticismo e extroversão são difíceis de encaixar na lista de Aristóteles e podem ter relevância moral limitada. Se as cinco dimensões de traços que discuto fossem os únicos elementos fortes de caráter, a teoria da virtude ainda estaria em apuros. A réplica pode ser respondida olhando-se para pesquisas que vão além das – assim chamadas – “cinco grandes” dimensões da personalidade. Por exemplo, Paunonen

e Jackson (2000) encontraram evidências de nove outras dimensões. Alterando seus rótulos um pouco, eles encontraram evidências para: religiosidade, astúcia, honestidade, sensualidade, frugalidade, conservadorismo, gênero (masculinidade ou feminilidade), arrogância e perspicácia. Alguns deles coincidem com as virtudes de Aristóteles. Frugalidade, arrogância, honestidade e perspicácia são todas discutidas em seu trabalho.

Como uma moral provisória, então, concluo que traços de caráter possam estar em boas condições e podem incluir traços que são de interesse para os eticistas da virtude. A crítica de Doris e Harman falha porque se baseia na negação dessas afirmações. Dada a existência dos traços moralmente relevantes, eles seriam obrigados a fornecer novos argumentos para a ineficácia desses traços e, neste ponto, a evidência simplesmente não está disponível.

Apesar de tudo isso, eu não quero concluir que as teorias de virtude sejam verdadeiras. Eu penso que elas sofram de outra falha, relacionada, em parte, aos resultados que venho discutindo. A verdadeira questão não é a realidade de traços de caráter, mas seu lugar nos fundamentos da moralidade. Os teóricos da virtude nos querem fazer crer que a bondade e a maldade moral derivem de fatos sobre as virtudes e que essas tenham a sua postura moral de alguma forma não-derivativa. As virtudes deveriam ser um nível moral mais baixo. O problema é que essa conclusão não pode ser sustentada.

A palavra virtude se refere a um bom traço de caráter. Mas o que torna bom um traço? A resposta padrão na teoria da virtude é a de que bons traços são mais naturais que os traços maus. É parte do projeto humano, parte da nossa teleologia natural, ter determinados traços e não outros. Alguém a quem falta virtude não está cumprindo a função enquanto ser humano.

Eu posso identificar duas estratégias para se chegar à ideia de que as virtudes são fundamentadas na natureza humana. Em uma estratégia indireta, pode-se argumentar que determinados traços são favoráveis ou constitutivos daquilo que seria naturalmente considerado bem-estar. Aqueles que carecem de virtudes não se desenvolverão. Na abordagem direta, pode-se argumentar que determinados traços são mais naturais. Deixe-me, pois, considerar estas opções.

O problema com a primeira abordagem é que há pouca evidência apoiando a afirmação de que o bem-estar é uma construção natural. Como observado no capítulo 4, há uma variação transcultural significativa entre o que se considera uma vida boa (para revisões, ver Diener et al, 2003; Tiberius, 2003). Por exemplo, em populações ocidentais, o bem-estar está altamente correlacionado à coisas tais como autossatisfação, liberdade, emoções positivas e realização dos objetivos. Esses fatores tendem a ser menos importante para o bem-estar nas culturas do Extremo Oriente. Lá, as pessoas são mais propensas a enfatizar aceitação pelos outros, fazer os outros felizes e perseguir objetivos. Nenhuma dessas concepções de bem-estar tem a pretensão de ser a mais natural; pois eles simplesmente refletem diferenças culturais em valores. Se as concepções de bem-estar são moldadas pela cultura, não se pode dizer que algum traço de caráter constitua uma forma mais natural de bem-estar.

Voltemo-nos agora à idéia de que as virtudes são apenas mais naturais do que outros traços. Aristóteles apresenta suas virtudes como o meio termo entre os extremos desagradáveis. Este elegante quadro faz seus traços parecerem como que em um equilíbrio natural. Porque estão no meio, as virtudes parecem saudáveis e naturais, assim como a saciedade fica entre fome e excesso de indulgência. Porém, tudo isto é uma ilusão. Enquanto alguns traços de Aristóteles provavelmente existem, suas dimensões talvez não existam. Os traços que ele colocou no meio de dois extremos podem, na verdade, estar nos pólos. Por exemplo, Aristóteles coloca veracidade no meio de uma gama que se estende da ostentação autoengrandecedora ao pudor da autoanulação. Mas a veracidade pode ser, na verdade, um extremo em uma dimensão que se estende desde a honestidade a desonestidade. Os traços que valorizamos podem estar deselegantemente distribuídos ao longo das dimensões da personalidade. Alguns podem estar no meio; outros, perto das extremidades. Quando percebemos que as virtudes têm uma distribuição confusa ao longo das dimensões do caráter, esta história sedutora começa a soar como um canto das sereias, acenando-nos sobre as pedras.

Poder-se-ia, então, abandonar a taxonomia elegante de Aristóteles e argumentar que virtudes são mais naturais em outro sentido: elas são traços que

nos foram dados pela evolução. Esta alternativa darwiniana não é, no entanto, muito promissora. Personalidades humanas variam e essa variação é perfeitamente natural. Algumas das variações nos traços básicos parecem ter uma base genética (Jang et al. 1996) e todos os traços dentro da faixa normal são compatíveis com uma vida produtiva e com sucesso reprodutivo. Pode até haver benefício para a nossa espécie nas diferenças individuais. Uma variedade de traços pode permitir que as pessoas desempenhem papéis diferentes (suponha que todo mundo quisesse ser macho ou fêmea alfa). Mesmo traços que nós desvalorizamos intensamente podem ser altamente adaptativos. Personalidades antissociais reproduzem-se, em média, mais do que as outras pessoas e modelos da teoria dos jogos sugerem que ter uma pequena porcentagem dela pode nos ajudar a manter boas habilidades para detectar trapaceiros e malfeitores (ver Mealey, 1995). O cruel e o astuto podem prosperar, e nós podemos prosperar ainda mais se tivermos alguns deles conosco. Então, quando nos voltamos a questões sobre a natureza humana, encontramos apenas naturezas humanas, no plural (Ehrlich, 2000). Nós descobrimos que uma variedade de traços muito diferentes, com um status moral diferente, são adaptáveis. Temos muitos bioprogramas diferentes. Fundamentar a virtude na natureza humana é uma biologia duvidosa.

Concluo com uma objeção final. Suponhamos, contra todas as evidências, que a natureza humana, de fato, favorece um conjunto de traços de caráter sobre os outros. Suponhamos, contra todas as evidências, que aqueles traços coincidam com os traços que nós, pré-teoricamente, consideramos virtuosos. Nós podemos perguntar ainda por que esses traços são moralmente bons. Certamente ser natural não é suficiente para ser moralmente bom. Defecação não é um ato nobre. Alguma outra coisa deve dar às virtudes seu *status* moral. Teorias da sensibilidade podem fornecer uma resposta. A diferença entre um traço virtuoso e um traço que carece de significado moral é determinada pelos nossos sentimentos. Nós *valorizamos* certos traços. Valorizá-los consiste em disposições emocionais. Até agora, eu foquei em valores morais atinentes à forma como as pessoas devem agir. Mas também podemos ter sentimentos voltados a traços de caráter. Detestamos a ganância e sentimos vergonha da covardia. São sentimentos que fazem esses

traços bons, não a natureza, embora a natureza possa exercer alguma influência sobre os nossos sentimentos. Se eu estiver certo a teoria da virtude não é, então, um caminho para o objetivismo moral. As virtudes atingem alta posição por causa das reações que provocam em nós.

4.2.4 Fatos psicológicos 2: consequencialismo.

Eticistas da virtude tentam assegurar a independência responsiva sem divorciar a moral da psicologia. Nisso, eles não estão sozinhos. Há outra tradição celebrada dentro da teoria ética que tem objetivos semelhantes. Esta, naturalmente, é o consequencialismo. Consequencialistas dizem que o mérito moral depende do resultado. Mais exatamente, eles oferecem uma teoria do que é bom e uma teoria do que é certo. Ao dar uma teoria do bom, consequencialistas tentam identificar as coisas que têm valor intrínseco. Utilitaristas clássicos, como Bentham e Mill, são consequencialistas que identificam o prazer ou a felicidade como o bem mais fundamental. Consequencialistas definem o certo como aquilo que maximiza o que eles definem como bom. O imperativo dos utilitaristas é a maximização da felicidade. Felicidade, claro, é um estado psicológico. Todavia, os utilitaristas não afirmam que a felicidade se torna moralmente boa em virtude de sua representação como tal. A felicidade seria boa independentemente das nossas convicções morais. Assim, maximizar a felicidade não é moralmente correto por conta de tomarmos isso como correto. A maximização herda seu mérito moral do bem daquilo que é maximizado. Se a felicidade é boa, mais felicidade é melhor. Se tivermos de escolher entre dois cursos de ação e um conduz a algo melhor, esse é, então, o curso de ação que devemos perseguir. O utilitarismo é, assim, uma teoria psicológica baseada na moralidade, mas não uma teoria dependente de resposta. Se o utilitarismo é correto, as coisas têm, objetivamente, mérito moral (são boas ou corretas) em um sentido do termo. O mesmo vale para outras formas de consequencialismo; as coisas podem ser boas ou corretas sem que sejam representadas como tal.

Eu penso que as teorias consequencialistas sejam insustentáveis, mas este não é o lugar para uma avaliação completa. A literatura sobre este tópico é vasta. Ao invés de desenvolver uma crítica detalhada, quero sugerir que o consequencialismo não esteja realmente em concorrência com o tipo de teoria que estou propondo. Minha tese neste capítulo é de que a moralidade não é objetiva. Por isso, quero dizer que os conceitos morais, *tal como existem atualmente*, não se referem a propriedades objetivas. A este respeito, o meu projeto é descritivo. Em contraste, o consequencialismo é uma teoria normativa e, nesse sentido, revisionista. Ao invés de capturar os valores morais que as pessoas comuns, na verdade, endossam, consequencialistas oferecem um novo sistema de valores. Assim, se o consequencialismo pudesse identificar uma fonte objetiva de valores, não se seguiria disso que nossos valores sejam objetivos – não se seguiria que os nossos conceitos morais se refeririam às características que consequencialistas definem como os fundamentos da ação correta. Colocando mais incisivamente, se o vocabulário moral é fixado pela forma como usamos termos morais, então, consequencialismo é mais bem entendido, não como uma abordagem de moralidade, mas como uma alternativa a ela. (Eu penso que a ética kantiana deva ser interpretada da mesma forma).

Há muitas razões para se pensar que o consequencialismo seja uma teoria revisionista ao invés de um relato de com o que já estamos moralmente comprometidos. Pode-se ver isso ao perceber que as recomendações das teorias consequencialistas muitas vezes partem de intuições sem instruções sobre o bom comportamento. Consideremos alguns exemplos. Primeiramente, o consequencialismo é difícil de ser conciliado com a intuição de que deveríamos não prejudicar uma pessoa, a fim de salvar os outros. Por exemplo, apenas 8,3 por cento dos indivíduos pensam que é permitido empurrarem uma pessoa no caminho de um bonde que está acelerando em direção a cinco pessoas (Mikhail, 2002). Em segundo lugar, as teorias consequencialistas tipicamente implicam que seria bom distribuir recursos materiais com base em necessidades, mas muitas pessoas se sentem desconfortáveis com essa idéia. Por exemplo, Kluegel e Smith (1986) descobriram que 81 por cento dos americanos preferem um esquema segundo o

qual o rendimento é determinado pela habilidade e treinamento, enquanto apenas 13 por cento são a favor da distribuição com base na necessidade. Há pouco consenso sobre tais questões. Em geral, os liberais apóiam mais a justiça distributiva do que os conservadores, mesmo quando lhes pediram para fazer essa avaliação sob um véu de ignorância (Mitchell et al., 2003). Há também enormes diferenças culturais nas atitudes para programas de redistribuição. Por exemplo, 69,5 por cento dos austríacos pensam que o governo deve trabalhar para alcançar a igualdade econômica, ao passo que 38,3 por cento dos americanos concordam [com isso] (Svallfors, 1997). Em terceiro lugar, as pessoas frequentemente condenam moralmente as más intenções, mesmo quando elas não levam a consequências ruins, o que é inconsistente com a maioria das teorias consequencialistas. Curiosamente, há diferenças entre os grupos na tendência de condenar as más intenções, por exemplo, os protestantes são mais propensos a fazer isso do que os judeus (Cohen e Rozin, 2001). Em quarto lugar, algumas versões do consequencialismo (especialmente o utilitarismo) parecem sugerir que as pessoas que têm níveis mais elevados de bem-estar subjetivo têm maior valor moral. Por exemplo, as pessoas extrovertidas são mais felizes do que as introvertidas, depressivas, e pessoas que pontuam alto em escalas de neuroticismo (por exemplo, Diener et al, 2004); escandinavos e sul-americanos são mais felizes do que os asiáticos (Diener e Oishi, 2003), e as pessoas religiosas são um pouco mais felizes do que os não-religiosos (Witter et al., 1985). Isso significa que, digamos, introvertidos ateus devam ser colocados na linha de frente durante a guerra porque suas vidas têm menos valor do que a vida das pessoas extrovertidas e religiosas?

Teorias consequencialistas parecem colocar sobre nós demandas que vão contra nossos valores atuais. Claro, a literatura consequencialista possui diversas tentativas de acomodação com as intuições morais ordinárias. Mas esses esforços envolvem, por vezes, afastamentos significativos dos princípios consequencialistas. Por exemplo, a maldade das más intenções é acomodada pelo utilitarismo das motivações de Adams (1976), mas, estritamente falando, esta é uma mudança na visão de que as consequências têm prioridades normativas. Mais geralmente, o

padrão generalizado de discrepâncias entre as demandas do consequencialismo e valores ordinários sugere que o consequencialismo nunca pode ser perfeitamente combinado com nossas intuições. De fato, a diversidade das intuições morais sugere que, mesmo que alguma teoria consequencialista complexa e impura possa ser construída para coincidir com as intuições de algumas pessoas sobre o que conta como bom e correto, nenhuma teoria pode acomodar as intuições de todos. Se um consequencialista torna as motivações moralmente importantes, a teoria pode fazer um bom trabalho em capturar as intuições morais de um protestante, mas ele pode fazer um trabalho ruim ao tentar acomodar as intuições de um judeu. Se a teoria coloca limites à redistribuição, os conservadores ficarão felizes, mas os liberais podem recusá-la.

Eu não penso que a maioria dos consequencialistas discordaria com minha avaliação. Consequencialistas não tentam capturar os valores morais populares, pelo menos não com profundidade. A maioria dos consequencialistas provavelmente diria que os valores morais ordinários devam ser abandonados e substituídos por valores consequencialistas. Talvez eles estejam certos. Mas isso mina qualquer argumento do consequencialismo de que a moralidade seja objetiva. Consequencialismo não é uma boa abordagem daquilo que queremos dizer com os nossos termos morais, e, portanto, não acrescenta suporte à afirmação de que os valores morais têm um fundamento objetivo. Na melhor das hipóteses, os consequencialistas podem argumentar que a moralidade deva ter um fundamento objetivo.

4.2.5 Produtos intencionais: convencionalismo.

Quero considerar mais uma forma em que se poderia tentar argumentar que a moralidade é objetiva. A estratégia que eu tenho em mente começa com a suposição de que a moralidade possa estar intimamente relacionada à lei. As leis são regras de conduta, têm força normativa e são objetivas num sentido importante da palavra. É um fato objetivo que os motoristas nos Estados Unidos devam dirigir no lado direito da rua e que atravessar a rua descuidadamente é ilegal. Esses fatos

não dependem das nossas opiniões e atitudes. Eles nem sequer dependem de serem impostos. Fatos legais contam em virtude de sua etiologia. Uma regra é uma lei se ela foi produzida em conformidade com certos procedimentos. A este respeito, as leis são como artefatos, seu status como tal depende da forma como acontecem. Eles são, na terminologia introduzida acima, produtos intencionais. Eles são criados através de um processo intencional, mas, uma vez criados, desfrutam de um grau de autonomia.

Alguém poderia pensar que as regras morais são como normas jurídicas. Philippa Foot (1972) por exemplo, defende uma analogia entre as regras morais e as regras de etiqueta. Regras de etiqueta mantêm-se mesmo que nós não nos preocupemos com elas. Elas são, neste respeito, categóricas. Mas, ao contrário dos imperativos categóricos de Kant, eles são criados por nós e carregam uma força racional pequena para a pessoa que não tem consideração por convenções operatórias. Esta abordagem da moralidade está em harmonia igualmente com certos tipos de contratualismo. O tipo de contratualismo que tenho em mente difere das teorias contratualistas inspiradas por Kant. Contratualistas kantianos argumentam que as regras morais são aquelas que elaboraríamos se as políticas sociais fossem geradas a partir de uma posição epistêmica ideal (como equilíbrio reflexivo). Isso se torna uma espécie de teoria do observador ideal, sendo vulnerável às preocupações sobre os observadores ideais que levantei acima. O tipo de contratualismo que eu quero considerar aqui se centra nas políticas sociais que nós temos *realmente* criado, não aquelas que gostaríamos de criar. A ideia é que temos a obrigação de seguir a lei porque a lei é um tipo especial de produto social. Ela surgiu através de procedimentos que a tornam autoritária, mesmo se nos tornarmos preguiçosos, cansados, críticos ou criminosos.

O contratualismo convencionalista abre a porta para uma forte forma de relativismo, deixando, portanto, de atrair muitos filósofos com tendências objetivistas. No entanto, parece, sim, ser compatível com uma forma de objetivismo que eu ainda devo analisar. Contratos são como artefatos. Eles devem ser criados com intenções certas, mas não são, eles próprios, subjetivos. A regra de direito, por exemplo, não conta como uma regra porque estamos dispostos a representá-la ou

a responder a ela de certa maneira. Ela conta como uma regra devido à forma como foi criada, ainda que, uma vez criada, goze de certo grau de autonomia. Leis são produtos intencionais, e o mesmo pode ser dito sobre contratos sociais em geral. Se as leis têm postura moral em virtude de como foram criadas, então, a moralidade tem um tipo de objetividade.

Teorias convencionalistas, tais como o contratualismo convencionalista, sustentam que há uma conexão entre as regras morais e as regras convencionais que foram criadas pelos membros de grupos sociais. Dentro do escopo dos nossos propósitos, a característica chave dessas teorias é considerar as regras morais como produtos intencionais; uma vez que as regras foram criadas de maneira correta, elas têm força normativa, independentemente de nossas atitudes e disposições. Se uma teoria deste tipo pode ser defendida, então a moralidade será objetiva na mesma forma em que artefatos têm, objetivamente, sua função.

Eu penso que o convencionalismo tenha um núcleo de verdade apenas se abandonarmos o objetivismo. O problema com o convencionalismo é que ele assume incorretamente que as regras morais dependem unicamente de eventos anteriores, não de práticas atuais. Eu penso que as regras morais tornam-se obrigatórias apenas em relação às disposições de respostas atuais. Convenções sociais moralmente significativas não são *meros* produtos intencionais. Ao contrário de martelos e pregos, elas devem ser sustentadas por respostas humanas após sua criação inicial. Uma vez que essas respostas desapareçam, as convenções sociais deixam de ter autoridade moral.

Para demonstrar este ponto, quero distinguir entre dois tipos de “fatos sociais”. Um fato social é, grosso modo, um fato que não pode ser caracterizado sem referência a uma prática social. Práticas sociais são constituídas por comportamentos ou estados mentais coordenados dos membros de um grupo social. A primeira classe de fatos sociais que eu quero considerar pode ser chamada de fatos sociais etiológicos, fatos que dependem etiológicamente de processos sociais. Um fato social etiológico é aquele que não teria ocorrido, [se] não tivesse havido, no passado, algum processo social. Uma vez que um fato social desse tipo aconteceu, os processos sociais relevantes não precisam continuar. O

fato pode se manter sem quaisquer estados mentais ou práticas sociais. O casamento é um exemplo. Duas pessoas não seriam consideradas casadas a menos que se casassem, e uma vez casadas, continuam casadas, não obstante suas intenções e práticas. O casal teria que passar por um processo socialmente prescrito para acabar com seu casamento. Se o casal tornou-se senil junto, esqueceu-se de que se casou e ninguém mais se lembrasse, os dois permaneceriam, não obstante, casados. Fatos sobre quem é casado são fatos sociais etiológicos. Fatos envolvendo nomes são também fatos sociais etiológicos. Considere o fato de que esta ilha é Fiji. Ela é Fiji em virtude de ter sido denominada "Fiji" no passado. O fato de que certas fronteiras delimitam o Canadá pode ser outro fato etiológico, bem como o fato de que uma determinada pessoa seja o presidente.

Contrastemos este tipo de fato social com aqueles que devem ser continuamente sustentados pelas práticas mentais de um grupo social. Fatos sobre valor monetário são exemplos clássicos. O valor de uma moeda depende de práticas sociais. Algumas dessas práticas devem ser etiológicas. Um pedaço de papel só pode constar como um dólar se foi criado sob certas condições. Mas o valor monetário de um dólar também depende de práticas *atuais*. Se todos coletivamente decidissem considerar sem valor as notas de dólar, elas não teriam mais valor. Dólares só têm valor se forem considerados e tratados como valiosos. Chamo esse de fato social imanente. Fatos sociais imaneses devem ser sustentados por práticas sociais. O fato de um babuíno particular ser o macho alfa num bando de babuínos é outro exemplo de um fato social imanente. O mesmo vale para o fato de o Partido Democrata ser mais liberal que o Partido Republicano.

Munidos desta distinção, podemos agora perguntar se as regras morais obtêm sua força normativa de modo imanente ou etiológico. As convenções sociais parecem ter normatividade etiológica. Leis e regras de etiqueta são obrigatórias mesmo que nós não nos preocupemos com elas. É errado, em algum sentido, comer com a boca aberta ou atravessar uma rua tranquila fora da faixa. Mas isso é o erro moral? Eu penso que não. O convencionalista sobre moralidade admite que as regras morais sejam produtos humanos, ao invés de, digamos, mandamentos divinos ou exigências da razão. Então, a questão pode ser colocada na forma

condicional é: se as regras morais são fatos sociais, elas, portanto, são etiológicas? Responder "sim" seria tomar literalmente a analogia entre a lei e moralidade. A idéia seria que nós criamos regras morais e que agora estamos moralmente obrigados a obedecê-las. Se for esse o caso, deve haver algo na maneira como elas são criadas que lhes dá força normativa. Contudo, o convencionalista agora enfrenta um problema. O simples fato de que uma regra fora criada no passado não parece suficiente para dar força àquela regra. Assim, o convencionalista deve admitir que a força normativa da regra moral deriva de algo relacionado às condições sob as quais ela foi concebida.

Para atender essa exigência, convencionalistas podem estipular que uma regra é boa porque foi criada por um processo moralmente autoritativo. Aqui, duas possibilidades vêm à mente. Regras podem ser criadas através de um processo de consenso ou por meio de uma posição de equilíbrio reflexivo. Eu já descartei a última opção como uma versão da teoria do observador ideal, então, deixe-me focar no consenso. Se o consenso pode conferir força moral, então uma ausência de consenso deve miná-la. Em uma sociedade onde as pessoas parassem de se preocupar com regras morais anteriores, essas deixariam de ter força. Assim, se as convenções obtêm sua força moral pelo consenso, sua força moral é contingente sob as disposições atuais da população; é imanente, não etiológica. Regras existem imanentemente apenas enquanto são consideradas regras. Por este motivo, concluo que o convencionalismo só tem êxito como uma teoria subjetivista.

Existe, além disso, um sentido maior e mais profundo no qual o convencionalismo depende do subjetivismo. Ao discutir outras teorias éticas, eu disse repetidas vezes que as regras obtêm seu status moral apenas se elas são objetos de sentimentos morais. Este ponto também se aplica aqui. Suponha que encontremos uma sociedade que seguiu um conjunto de leis e, assim, teve convenções imanentes. Parece um enorme salto chamá-las de convenções morais. Uma sociedade de psicopatas também pode viver de acordo com convenções. O que dá status moral a uma convenção não é simplesmente as pessoas concordarem em segui-la. Consenso não é suficiente. As pessoas podem insistir que uma regra tem autoridade moral, porque os outros concordaram com isso, mas

este apelo ao consenso não é uma verdade analítica; é uma afirmação de valor. O consenso assume autoridade moral apenas para a pessoa que respeita o consenso como um bem moral. Aqueles que tomam uma atitude prática em direção a um consenso (como nossa comunidade de psicopatas) não podem receber o crédito de seguir os ditames da moralidade. Em outras palavras, o consenso não é um bem intrínseco. É bom porque o valorizamos. Ou, mais tipicamente, o consenso é derivado de outras coisas que valorizamos, como a equidade. Se não tivéssemos atitudes morais em relação a essas coisas, poderíamos ainda aceitar as regras convencionais, mas isso não teria qualquer significado moral. Convencionalistas que rejeitam essa intuição devem mostrar porque as convenções têm mérito moral sem apelar a uma norma de fundamento sustentada afetivamente, e eu penso que a dedução será impossível.

4.3 REALISMO CONSTRUTIVO

4.3.1 Realismo sem objetividade

A pesquisa mencionada acima demonstra que muitos filósofos querem encontrar um fundamento objetivo para a moralidade. Eu não penso que tal fundamento exista. Cada uma das principais teorias que eu considere têm falhas graves. Outras teorias objetivistas podem ser concebidas em seu rastro, mas não estou otimista quanto ao seu sucesso. Eu não mostrei que objetivismo na ética seja incoerente ou impossível, mas espero ter mostrado que a fé no objetivismo pode ser inapropriada. Em primeiro lugar, intuições objetivistas, na medida em que nós as temos, podem ser explicadas numa teoria subjetivista. Em segundo lugar, as tentativas de se fundamentar a moralidade em algo objetivo estão repletas de dificuldades. Sou levado à conclusão de que não há nenhuma razão forte para se pensar que a moralidade seja objetiva em qualquer sentido. De fato, eu penso que objetivismo seja uma piedosa esperança.

Na condução desta pesquisa, um tema surgiu reiteradamente. Os fatos objetivos que se propõem a fundamentar a moralidade não parecem ter significado

moral por conta própria. Uma ação pode ser imprudente se não falha ao tentar maximizar utilidade. Um ato pode ser escatologicamente arriscado se for contra o mandamento divino. Uma ação pode ser prejudicial se for contra a natureza humana. Uma ação pode ser irracional se for contra os ditames da razão universal. Uma ação pode ser impopular se for contra o consenso. Para que uma ação seja *moralmente errada*, outra coisa é necessária. Uma ação pode ser errada em um desses outros sentidos e não ser imoral. Da mesma forma, uma ação pode ser racional, prudente, satisfatória e popular sem que seja moralmente boa. Eu tentei mostrar que as afirmações morais são tornadas verdadeiras – moralmente verdadeiras – por meio de reações em nós. Nossas reações são verificadoras [*truthmakers*] morais; elas o são em virtude de quais as reivindicações morais alcançam seu *status* como moral. Isso significa que a moralidade é subjetiva. Tentei mostrar também que as reações relevantes não podem ser idealizadas a uma posição livre de preconceitos e que elas devem permanecer imanentes para que as verdades morais persistam. Isso sugere que a moralidade seja subjetiva de uma forma que impede a objetividade. Os argumentos para o subjetivismo nos capítulos 1 e 3 ajudam a sustentar estas afirmações, bem como os argumentos para o relativismo, que eu apresentarei no próximo capítulo. Concluo, então, que objetivismo moral está enganado.

Que implicações isso tem para debates sobre o realismo moral? Realismo é um conceito complicado. O termo tem sido usado de diferentes maneiras por diferentes ramos da filosofia. Para alguns, o realismo é conceitualmente vinculado à objetividade; ser real é ser, de alguma forma, objetivo. Por exemplo, é dito, às vezes, que o realismo implica algum tipo de independência da mente (por exemplo, Brink, 1989). Quando os filósofos falam sobre o realismo semântico ou realismo científico, querem, por vezes, afirmar a possibilidade de fatos ocorrerem sem quaisquer observadores humanos. Esta definição não abrange todos os usos técnicos ou vernáculos do termo. Berkeley não estava falando um total absurdo quando disse que um mundo constituído por idéias ainda se qualifica como *realidade*. Uma pessoa com uma dor de cabeça não encontraria conforto na insistência de um médico de que suas dores não são *reais*. Rosas são *realmente*

vermelhas mesmo que o vermelho seja uma qualidade secundária. Há, evidentemente, usos do termo "real" e de seus cognatos que não são sinônimos de "objetivo". Há definições alternativas.

De acordo com uma definição, realismo tem algo a ver com a verdade. Um domínio pode ser considerado real se as afirmações sobre esse domínio puderem ser verdadeiras. Chamemos isso de veracidade. A ideia está estreitamente relacionada com outra definição de realismo, concentrada em fatos. Um domínio é real, nesta definição, se ele contém fatos. Isto é o factualismo. Veracidade e factualismo estão atrelados porque, segundo muitos, uma afirmação é verdadeira apenas no caso em que corresponde a um fato. O realismo deste tipo é mais naturalmente contrastado com as teorias não-factualistas ou teorias de erro. Uma afirmação é não-factual quando não tenta afirmar um fato. Teorias expressivistas são um exemplo disso. Expressivismo é o ponto de vista de que as afirmações dentro de um domínio de interesse expressam sentimentos ou atitudes, não fatos. Teorias de erro asseveram que as afirmações dentro de um domínio visam escolher fatos e falham. Elas são aptas à verdade [*truth-apt*], mas sempre falsas.

O termo "realismo" também tem sido associado à noção de eficácia causal (Harman, 1977; Sturgeon, 1985). Podemos imaginar que as coisas reais contribuem para o fluxo e refluxo causal do mundo. Se algo é real, ele deve ser capaz de fazer um trabalho causal. Esta pode ser uma condição mais forte do que a aptidão à verdade [*truth-aptness*] e que o factualismo porque há, sem dúvida, fatos sobre coisas que não são causalmente eficazes. É um fato, talvez, que o Super-Homem é enfraquecido pela kryptonita. A sentença "unicórnios têm um chifre" é, sem dúvida, verdadeira. Mas o Super-Homem e unicórnios não fazem parte do fluxo e refluxo causal do mundo. Eles não são reais. Além disso, pode-se definir o realismo em termos de existência. O Super-Homem e os unicórnios não existem. Se a "existência" é um predicado, esta é, então, uma afirmação substantiva. Se a "existência" é um quantificador, a questão é que não têm condições de quantificar objetivamente Super-Homens e unicórnios. Isto é, não podemos especificar um objeto que é o Super-homem ou um objeto que é um unicórnio. Isto levanta a questão de quando temos condições de quantificação objetiva. Uma resposta

plausível é a de que podemos quantificar objetivamente sobre todas as coisas e tão somente sobre aquelas que são causalmente eficazes – as coisas que precisamos postular a fim de explicar plenamente todos os eventos que ocorrem no mundo. Tais objetos são, em um sentido intuitivo, reais. Realismo deste tipo pode ser contrastado com ficcionalismo e eliminativismo, como estes termos são mais tipicamente usados. Teorias ficcionalistas afirmam que um domínio de discurso pressupõe a não-existência das entidades que descrevem. Essas entidades “existem” apenas num sentido fictional. Teorias eliminativistas afirmam que um domínio do discurso visa a escolher as entidades que literalmente existem, mas que essas entidades não existem realmente. Eliminativismo e teorias do erro andam de mãos dadas, mas elas podem se separar se adotarmos uma abordagem da verdade de acordo com a qual podemos fazer afirmações verdadeiras sobre entidades não existentes (não limitadas a afirmações existenciais negativas).

Agora podemos perguntar: se a moralidade não é objetiva, ela pode, no entanto, ser real em algum sentido? Consideremos, primeiramente, o sentido de realismo associado à verdade e aos fatos. O tipo de subjetivismo que eu delinieei no capítulo 3 foi projetado como uma alternativa ao expressivismo. Ele foi projetado para acomodar a intuição de que afirmações morais são aptas à verdade. Quando dizemos que roubar é errado, não estamos apenas expressando nossos sentimentos, estamos dizendo implicitamente que o roubo tem a propriedade de causar certos tipos de reações negativas em nós. É *verdade* que o roubo nos faz sentir raiva, culpa ou desapontamento. Portanto, a afirmação de que roubar é errado é verdadeira. Da mesma forma, podemos dizer que é um fato que *roubar* cause estas reações em nós, e, portanto, a afirmação de que roubar é errado corresponde a um fato. Subjetivismo deste tipo é uma teoria realista, em certa interpretação do termo.

O subjetivismo que defendo é também realista na interpretação que enfatiza a eficácia causal e a existência. Fatos morais impactam nosso comportamento e fazem isso porque são fatos morais. Consideremos, por analogia, a eficácia causal de outras propriedades indutoras de emoções, tais quais o susto, a repugnância ou a amabilidade. Estas propriedades são trivialmente causais na medida em que são

definidas pela sua tendência de provocar reações emocionais em nós. Elas também fazem contribuições causais ao nosso comportamento. Podemos evitar assistir filmes de suspense porque eles são assustadores; podemos descartar comida estragada porque é nojenta; poderíamos convidar alguém para jantar porque esta pessoa é amável. Em cada caso, a propriedade de indução de emoção está causalmente implicada nas nossas ações.

Alguém pode objetar que esta é realmente uma ilusão. O que afeta nossas ações não são as propriedades indutoras de emoção, mas as próprias emoções. Evitamos filmes de suspenses porque o medo nos leva a evitá-los; nós descartamos comida podre porque somos impelidos pela repugnância; nós convidamos um amigo porque antecipamos o prazer de sua companhia. Susto, repugnância e amabilidade são epifenômenos. Eles são totalmente desnecessários para explicar o comportamento.

Em resposta, eu gostaria de salientar que as propriedades indutoras de emoção estão desempenhando um papel explicativo diferente daquele desempenhado pelas próprias emoções. Há situações nas quais se pode ter uma emoção sem a propriedade indutora de emoção. Suponhamos, por exemplo, que eu forme a falsa crença de que um pote de comida na minha geladeira estragou. Eu poderia descartá-lo porque eu estou repugnado ou porque eu acredito que seja repugnante, mas não porque seja realmente repugnante. Ser repugnante é causar de modo fiável, sob determinadas circunstâncias, repugnância (por exemplo, ao saber a informação relevante). Este tipo de caso faz com que a repugnância pareça o determinante real da ação. Mas suponhamos, agora, que perguntemos de onde vêm esses estados psicológicos. No exemplo atual, eu posso ficar repugnado porque interpretei mal uma data de vencimento ou confundi uma sombra com bolor. Mas agora considere o caso em que a comida, de fato, esteja estragada. Neste caso, minha repugnância é causada pelo fato de que o alimento é repugnante. A propriedade indutora de emoção é causalmente responsável pelo meu estado mental, e, portanto, está implicada no meu comportamento subsequente. Alguém poderia objetar dizendo que a minha repugnância foi, realmente, causada pelo fato de que a comida está estragada, não porque a comida estragada é repugnante. Este

movimento é implausível. Se comida estragada não fosse repugnante para nós, não induziria a uma emoção. Vejamos outro exemplo. Rozin et al. (1986) perguntaram aos indivíduos em um experimento se eles beberiam um copo de suco de laranja depois que uma barata esterilizada tivesse mergulhado nele. A maioria dos indivíduos disse que não. A barata é inofensiva. O fato de que uma barata teve contato com o suco é insuficiente por si só para explicar a nossa resposta. Indivíduos recusam-se a beber porque as baratas têm a propriedade de causar repugnância em nós. Repugnância é causalmente eficaz.

O mesmo é verdadeiro para as propriedades morais na abordagem subjetivista. Estar errado é a propriedade que nos dispõe a ter certas reações negativas. O erro de roubar desempenha um papel causal em nos impedir de roubar e em punir aqueles que o fazem. O erro induz emoções em nós, e essas têm impacto sobre nosso comportamento. Se roubar não fosse errado, nós não evitaríamos fazê-lo.

Concluo que as propriedades morais são perfeitamente reais neste sentido. Elas fazem parte do tecido causal do mundo. Podemos usá-las para explicar o comportamento e podemos quantificá-las em nossas explicações.

Em resumo, eu penso que devemos desistir do objetivismo moral. Propriedades morais não são objetivas em nenhum sentido relevante. Caso ser real implique ser objetivo, então o realismo moral é falso. Nós devemos ser antirrealistas. Contudo o realismo pode ser compreendido de outras formas. Enunciados morais podem ser verdadeiros; existem fatos morais; propriedades morais são causalmente eficazes; e, portanto, podemos quantificá-los. Por essas razões, faz todo o sentido chamar o tipo de subjetivismo que eu defendo uma forma de realismo moral.

Há uma última pergunta que vale a pena abordar. Se as propriedades morais são reais, mas não são espécies naturais, então, que tipo de espécie são elas? A resposta a esta questão foi prenunciada acima. Fatos morais são como dinheiro. Eles são fatos sociais que ocorreriam em virtude de nossas disposições e práticas atuais. Eles são tão reais quanto valores monetários, e ainda mais importantes, talvez, em guiar nossas vidas.

4.3.2 Sentimentalismo construtivo

Valor monetário é, em um sentido perfeitamente simples, uma construção. É perfeitamente real, mas isso depende de práticas humanas e de respostas humanas. Eu penso que a moralidade também seja uma construção. No último capítulo, eu defendi uma versão da teoria da sensibilidade, segundo a qual os fatos morais dependem de sentimentos morais. Agora proponho chamar esse tipo de teoria da sensibilidade de *sentimentalismo construtivo*.

O termo "sentimentalismo" é frequentemente usado como um guarda-chuva para todas as teorias que relacionam conceitos morais a emoções (incluindo, igualmente, expressivismo e teorias da sensibilidade). Mas, dada minha ênfase em sentimentos e a distinção entre emoções e sentimentos, esse termo é especialmente apropriado para a minha opinião. O modificador "construtivo" refere-se à ideia de que as propriedades morais passam a existir sem os sentimentos. Nós, de fato, construímos a moralidade; nós não a descobrimos.

Isso levanta uma questão terminológica. No capítulo 1, mencionei a distinção elaborada por D'Armas e Jacobson (2006) entre as teorias perceptivistas e as teorias projetivistas. Teóricos da sensibilidade devem ser perceptivistas; eles comparam as propriedades morais a qualidades secundárias e implicam que usamos nossas emoções para experimentar o mundo de uma forma avaliativa (ver especialmente McDowell, 1987; McNaughton, 1988). Projetivistas, em contraste, costumam dizer que a moral não pode ser percebida porque ela existe unicamente em nós. Expressivistas se enquadram nesta categoria (ver Blackburn, 1980). Neste ponto de vista, quando conferimos atributos morais a uma ação, estamos simplesmente projetando nossos sentimentos sobre o mundo.

Em face disso, percepção e projeção parecem ser diametralmente opostas. Percepção implica pegar em algo que está fora de nós, e projeção implica que tomemos alguma coisa dentro de nós e que irradiemo-la para fora. Entretanto, numa análise mais aprofundada, pode ser difícil manter estas duas ideias separadas. D'Arms e Jacobson argumentam que ambas são metáforas enganosas: a metáfora da percepção implica que os valores são independentes de valoração, e

a metáfora da projeção ameaça minar a moralidade, o que implica que os valores são apenas “fábulas” de sentimentos. Eu quero desafiar também o contraste entre a percepção e projeção, mas não porque eu penso que ambas as metáforas estejam erradas. Pelo contrário, eu penso que ambas possam ser importantes. Há um sentido no qual nós percebemos as propriedades morais. De fato, pode-se dar sentido literal a essa ideia ao observar que as emoções são percepções de mudanças no corpo, e é por meio de nossas emoções que frequentemente chegamos a descobrir que algo está certo ou errado. Mas há também algo adequado a respeito da ideia de projeção moral. A moralidade não é uma característica objetiva do mundo da mesma forma que podem sê-la, digamos, leões, tigres e ursos. Ela vem de nós.

A ideia de construção resolve o conflito entre as ideias de percepção e de projeção. Coisas que edificamos ou construímos vêm de nós, mas, uma vez lá, são entidades reais que podemos perceber. Isto é obviamente verdadeiro no caso das construções de materiais, tais como edifícios ou artefatos. Mas também é verdadeiro para construções sociais. Valor monetário é criado por nós, mas também é uma característica real do mundo que tem um impacto sobre nós.

Eu penso que os fatos morais sejam um tipo especial de construção: eles são feitos por nossos sentimentos, e, uma vez feitos, podem ser percebidos. Você pode perceber que é errado chutar um cachorro. Quando alguém chuta um cão, provoca uma resposta emocional cuja manifestação é um sentimento que constitui a convicção de que a crueldade animal é errada. O sentimento a faz errada; e também torna o erro palpável para nós quando o encontramos. Construções sentimentais reconciliam projeção e percepção.

Na primeira metade deste livro, eu foquei principalmente no lado perceptivo da construção. Por exemplo, falei sobre as emoções que surgem em resposta às transgressões morais e enfatizei minha lealdade às teorias de sensibilidade, que, como o próprio nome diz, implicam uma comparação com a percepção. Eu não falei muito sobre a fabricação de moral. Em particular, eu não descrevi as formas nas quais fatores culturais podem contribuir para verdades morais. Esta é a questão central que abordo nos próximos capítulos.

PARTE II
CONSTRUINDO A MORAL

5. JANTANDO COM CANIBAIS

“Sentei com Michii e com Darinimbiak, nós três sozinhos perto do fogo e os outros dançando, cantando ao nosso redor. Peguei um pedaço da carne que Michii me estendia e comi e engoli e comi um pouco mais, e entrei no círculo novamente para dançar” (Schneebaum, 1969: 106). A “carne” mencionada aqui tão sem cerimônia é carne humana. O autor é Tobias Schneebaum, um antropólogo americano, que participou de um festim canibal com o povo Akamara, no Peru. Os canibais tendem a nos encher de uma sensação de horror. Chamar o canibalismo de imoral é uma atenuação. É moralmente monstruoso comer a carne de nossas próprias espécies. Assassinos que se dedicam a essa prática são considerados maus e depravados. Ainda assim, o canibalismo tem sido praticado por culturas no mundo todo. Ele é, em alguns casos, a prática cultural padrão, em vez de ser a exceção. O canibalismo é relativamente predominante em algumas sociedades pré-estatais (bandos, tribos ou lideranças supremas), e desaparece com o surgimento da agricultura, da domesticação de animais, e, mais importante, de certas instituições associadas com a estatização (Harris, 1985). Há um abismo moral dividindo culturas que praticam o canibalismo e aquelas que o repudiam. Membros de grupos diversos parecem ter ideias radicalmente diferentes sobre o que é certo e o que é errado.

Neste capítulo, eu exploro a variação moral. Alguns filósofos morais acreditam que a variação é exagerada. Outros pensam que a variação nos valores morais não tem implicações para a variação em fatos morais. Pode haver uma moralidade verdadeira apesar das diferenças significativas quanto a convicções morais. Eu rejeitarei essas duas afirmações. Valores divergem, e essa divergência importa. Eu argumentarei que o relativismo moral é uma consequência direta da teoria da sensibilidade da qual eu sou a favor, que eu chamei de sentimentalismo construtivo. A ligação entre teorias sentimentalistas e o relativismo certamente tem sido notada por outros (ver especialmente Westermarck, 1906), mas teóricos da sensibilidade não despendem muito tempo explorando essa implicação. Eu tento corrigir esse descuido aqui e nos capítulos que se seguem.

É claro, um *modus ponens* de um filósofo é o *modus tollens* de outro. Se a teoria da sensibilidade com a qual eu concordo acarreta relativismo, então muitos a tomarão como fundamento para rejeição. O relativismo enfrenta muitas objeções poderosas. Abordarei quatro dessas na seção 5.2. No capítulo 7, avaliarei criticamente a sugestão de que somos biologicamente dotados de universais morais, e no capítulo 8 eu examinarei as perspectivas para o progresso moral dentro de um quadro relativista.

5.1 O RELATIVISMO MORAL DEFINIDO E DEFENDIDO

5.1.1 Relativismo descritivo e relativismo metaético

O termo "relativismo moral" é notoriamente ambíguo. Por um lado, ele é usado para rotular a tese de que algumas pessoas têm valores morais fundamentalmente diferentes. Seguindo Brandt (1967), eu o chamarei de relativismo moral descritivo, porque isso significa que se trata de um simples fato empírico. Relativismo moral descritivo não é inquestionável (como veremos), mas é amplamente aceito com base no registro antropológico. A controvérsia aumenta quando vamos de afirmações antropológicas para afirmações sobre verdade moral. Defenderei a seguinte tese:

Relativismo metaético: as condições de verdade de um julgamento moral dependem do contexto no qual o julgamento é formado.

Se essa tese estiver correta, então um julgamento que é verdadeiro em um contexto pode ser falso em outro. Muitas pessoas acham essa ideia implausível, incoerente e moralmente repreensível. Eu abordarei tais objeções no tempo devido. Primeiro, eu quero discutir se há alguma razão para pensar que essa tese seja verdadeira.

As pessoas, em geral, tentam argumentar a favor do relativismo metaético apelando para o relativismo descritivo. O argumento chave implicado é assim:

1. O relativismo descritivo é verdadeiro.
2. Se o relativismo descritivo é verdadeiro, então o relativismo metaético é verdadeiro.
3. Logo, o relativismo metaético é verdadeiro.

A segunda premissa está longe de ser trivial. As diferenças nos valores morais não acarretam diferenças em verdades morais em qualquer sentido. Esse ponto é frequentemente levantado ao se traçar analogias à ciência. Diferenças em crenças científicas não acarretam que julgamentos científicos conflitantes sejam igualmente verdadeiros. O fato de as culturas terem visões variadas sobre o formato da Terra, por exemplo, não acarreta que a Terra possua vários formatos. De fato, as disputas científicas não possuem implicações relativistas mesmo quando essas disputas são insolúveis. Suponha que não há como provar decisivamente que o universo tem onze dimensões em vez de, digamos, quatorze dimensões. Isso não acarretaria que ambas as dimensões sejam verdadeiras. Há uma contradição, mesmo se os debates não tivessem resolução. Precisamos ser cuidadosos, então, ao delinear conclusões metaéticas a partir da existência e da persistência de debates morais. O relativista metaético deve oferecer uma razão para pensar que a premissa 2 é plausível quando a premissa paralela seria extremamente improvável em outros domínios.

A premissa 2 é defensável se a teoria da sensibilidade à qual eu sou a favor for verdadeira. O relativismo descritivo é a visão de que as pessoas têm valores morais fundamentalmente diferentes. É a visão de que os valores poderiam diferir mesmo se todos os fatos não-morais estivessem nela. No capítulo 3, eu disse que a retidão ou a incorreção moral poderiam ser definidas em termos de sentimentos. Uma ação é certa ou errada se há um sentimento moral em relação a ela. Um sentimento moral é uma disposição a ter emoções no escopo da aprovação ou da desaprovação. Se o relativismo moral descritivo é verdadeiro, então as pessoas têm diferentes sentimentos morais em relação às mesmas coisas. Se a retidão e a incorreção dependem, metafisicamente, dos sentimentos que as pessoas têm, então a existência de diferenças nos sentimentos das pessoas acarreta uma

diferença nos fatos morais. Assim, o relativismo metaético pode ser derivado do relativismo descritivo.

5.1.2 Avaliador e Agentes

A definição de relativismo metaético com a qual eu tenho trabalhado não é detalhadamente descrito. Antes de avaliar a doutrina, quero ser um pouco mais preciso a respeito daquilo com que estarei me comprometendo. Em particular, quero abordar a seguinte questão: os valores *de quem* são relevantes para determinar as condições de verdade de um julgamento moral?

Quando um julgamento é formado, há uma pessoa formando o julgamento, e uma pessoa (ou pessoas) sobre a qual (as quais) o julgamento é formado. Se uma pessoa julga que ϕ -ar é moralmente errado, ela implica que é errado *para alguém*. Às vezes, esse alguém é a pessoa que faz o julgamento, mas de forma igualmente frequente, nós fazemos julgamentos morais a respeito de outros. Isso significa que há dois contextos potenciais a considerar. Há o contexto do julgador e o contexto do julgado. Em alguns casos, o julgador e o julgado aderem a diferentes valores. Quando isso acontece, podemos perguntar de quem são os valores que importam.

Lyons (1976) define o Relativismo Avaliador como a visão de que a verdade de julgamentos morais depende do contexto de quem julga, e ele define Relativismo Agente como a visão de que o contexto do que é julgado importa. Quando um missionário diz aos membros de uma aldeia Akamara que canibalismo é moralmente errado, as condições de verdade deveriam ser fixadas pelos valores do missionário ou pelos valores dos aldeões?

Lembremos uma observação que eu fiz no preâmbulo deste livro. Se o sentimentalismo está correto, alguém não pode fazer um julgamento da forma "alguém deve ϕ " ou "alguém não deve ϕ ", a menos que esteja disposto a ter sentimentos prescritivos. Sentimentos prescritivos podem ser identificados com as formas de aprovação e de desaprovação descritas no capítulo 3. Uma pessoa que não tem esses sentimentos pode vir a acreditar que algo é obrigado por alguém, mas formar um pensamento prescritivo no sentido de carga afetiva seria

impossível. Chamarei tais pensamentos prescritivos de "deveritades proposicionais" [*propositional oughtitudes*]. Em inglês, esses pensamentos são canonicamente expressos pelo uso da palavra "ought" ("dever").

Suponha que um missionário forme uma deveridade proposicional em relação ao canibalismo Akamara. Por definição, a deveridade contém conceitos morais de carga afetiva. As emoções implícitas nesses conceitos são, é claro, as emoções do missionário. É o missionário que está disposto a sentir raiva ou nojo ao pensar sobre o canibalismo. Consequentemente, o significado da deveridade proposicional que o missionário forma deve ser fixado por suas próprias emoções. São os valores do missionário que estão sendo expressos. Quando deveritades proposicionais são formadas, condições de verdade são relativizadas para o avaliador.

Contraste as deveritades proposicionais com outro tipo de julgamento moral, que pode ser chamado de uma crença a respeito de obrigações. Em vez de formar o julgamento de que os Akamaras deveriam abster-se do canibalismo, um missionário pode formular o pensamento de que os Akamaras estão sob uma obrigação de se abster do canibalismo. Esse pensamento neutro se refere implicitamente a sentimentos, porque obrigações são determinadas por deveritades, mas não expressam um sentimento. Se um julgamento neutro é formado, o missionário pode ajustar o contexto para seus próprios valores ou para os dos Akamara. O ajuste de contexto pode até mesmo ser explícito. O missionário pode julgar que os Akamaras estão sob uma obrigação *sob o seu sistema de valor* para se comprometerem com o canibalismo, mas que eles não têm tal obrigação sob os valores do missionário. Com julgamentos neutros, alguém pode relativizar para agente e para avaliador. Nossa habilidade em fazer isso provém explicitamente certo suporte para o relativismo metaético.

Pragmaticamente, um missionário implicaria seu próprio sistema de valor se ele fizesse uma asserção moral sem especificar um contexto. Se fosse para o missionário dizer "Akamaras são obrigados a abster-se do canibalismo", os ouvintes assumiriam que ele estava expressando uma deveridade ou relatando seus próprios valores. Frequentemente, não afirmamos os valores dos outros, especialmente quando os consideramos repugnantes. Pontos semelhantes podem

ser levantados acerca de julgamentos que seriam expressos usando as palavras "certo" e "errado". Esses julgamentos podem ser fundamentados nos sentimentos que a pessoa faz deles, ou eles podem ser julgamentos sobre obrigações. Quando usamos a palavras "certo" e "errado" sem especificar o contexto, os ouvintes assumem que estamos expressando nossos próprios valores.

A maioria dos relativistas é Relativista Avaliador, mas Gilbert Harman (1975) tem argumentado a favor de um limite ainda mais estrito sobre a aplicabilidade de julgamentos e de termos morais. Ele sugere que não podemos ter deveridades proposicionais em relação a membros de culturas que aderem a diferentes valores. Nessa terminologia, os usos morais da palavra "deveria" expressam "julgamentos internos". Assumimos que a palavra "deveria", quando usada em um sentido moral, aplica-se apenas a indivíduos que são motivados pelos tipos de considerações morais que nos motivam. Se Harman estiver certo, então nenhuma lacuna pode ser aberta entre o avaliador e o agente quando formamos deveridades proposicionais. A evidência de Harman para essa tese vem dessa intuição que soa estranha quando tentamos expandir julgamentos de dever para aqueles com valores diferentes. Teria soado estranho dizer a Hitler que ele não deveria realizar o genocídio. Se estivéssemos sendo invadidos por alienígenas inteligentes de outro planeta, soaria estranho dizer a eles que eles não deveriam nos matar. Harman argumenta que sua peculiaridade mostra que os julgamentos de dever são apenas aplicáveis àqueles que compartilham os valores daquele que faz esses julgamentos.

A tese semântica de Harman vai além da tese que eu sancionei acima. Sugeri que deveridades expressam os valores dos julgadores, mas não disse que esses valores não são aplicáveis a pessoas com diferentes valores. Na forma de relativismo que eu tenho assumido, um falante diz "você deveria \emptyset " expressa o fato que o falante assume valores que requerem \emptyset -ar. Essa é uma forma de Relativismo Avaliador. A tese semântica de Harman acarreta que "você deveria \emptyset " também expressa o fato de que o agente que está sendo abordado assume valores que requerem \emptyset -ar. Esse é um Relativismo Agente. Para ver se o Relativismo Agente é verdadeiro, precisamos examinar como as pessoas usam o termo "deveria". Harman baseia seu caso na intuição de que nós não aplicamos o termo "deveria" a

peessoas como Hitler, mas isso pode ser uma evidência inadequada para a tese que queremos defender.

O argumento de Harman enfrenta duas objeções. Primeiro, ao contrário de Harman, as pessoas frequentemente parecem abrigar deveritudes proposicionais em relação àqueles que têm valores diferentes, pelo menos sob certas circunstâncias. Por exemplo, missionários podem dizer aos Akamaras: "Você deveria abster-se do canibalismo". Ou um pai pode dizer a um filho: "Você não deveria bater na sua irmã", mesmo que a criança claramente tenha uma gama de valores diferentes. Segundo, não precisamos da tese do Relativismo Agente para explicar por que estamos relutantes em usar a palavra "deveria" em relação a alguém como Hitler ou a invasores do espaço sideral. Como Lyons (1976) argumenta, dizer a Hitler que ele deveria evitar o genocídio não teria sentido. Punir os alienígenas seria uma perda de tempo. Não nos incomodaríamos com a falação em relação ao que "deveria" quando estivéssemos lidando com monstros morais, porque não temos a intenção de despertar seu interesse. Se nossa relutância em formar deveritudes nesses casos deriva de nosso reconhecimento de que a negociação não teria sentido, então nossa relutância não pode ser tomada como evidência para o Relativismo Agente.

Essas objeções parecem significar problemas para a visão de Harman, mas penso que elas podem ser respondidas. Começemos com a primeira objeção, que pretende mostrar que nós frequentemente direcionamos julgamentos de dever em relação a pessoas com diferentes valores. Isso poderia significar um problema para o Relativismo Agente se não houvesse outra explicação para os exemplos em consideração. Eu salientei que os missionários podem usar a palavra "deveria" quando falassem com membros dos Akamara. Mas isso não prova que essa palavra, de fato, aplica-se a tais casos. Missionários podem estar cometendo um erro. Ordinariamente, é uma má ideia atribuir erros conceptuais a falantes comuns, mas eu penso que esse é um caso excepcional. Primeiro de tudo, missionários são suscetíveis a serem absolutistas morais. Eles são suscetíveis a terem a visão de que há uma gama de valores que provém do comando divino. Podem também abrigar a crença de que todas as pessoas têm acesso epistêmico a esses valores.

Assim, eles podem considerar os Akamara como membros da mesma comunidade moral e, por isso, estariam sob a jurisdição das mesmas leis morais. Segundo, missionários estão no negócio da conversão, e adotar uma deveritude pode ser uma ferramenta poderosa para influenciar os valores e comportamento de outros. Deveritudes são constituídas por sentimentos morais. Aqueles que violam nossas deveritudes incorrem em nossa ira. A ira pode ser muito persuasiva. Então, por adotar uma deveritude, podemos levar os outros a adotar os nossos valores. Isso pode também explicar por que os pais dirigem julgamentos de dever a seus filhos. Os pais de crianças pequenas sabem que seus filhos não internalizaram valores morais adultos, mas eles dizem a eles o que eles deveriam fazer, em um esforço de ensiná-los como se comportar. Se o Relativismo Agente é verdadeiro, então tais diretivas são semanticamente anômalas, mas taticamente apropriadas. Não estou sugerindo que missionários e pais percebem que suas normas não se aplicam. Eles podem ou não perceber isso. Meu ponto é que, independentemente de as normas se aplicarem, essas pessoas têm boas razões estratégicas para fazer julgamentos de dever. Mesmo se os pais percebessem que foi um erro semântico dizer aos seus filhos o que eles deveriam fazer, eles continuariam emitindo tais diretivas por causa do seu impacto na formação do comportamento. Se você quiser alterar a moralidade de alguém, é uma boa estratégia tratá-lo como se ele fosse obrigado a seguir os seus valores.

Aquele que se opõe ao Relativismo Agente pode admitir esses pontos, mas note que a concessão simplesmente nos deixa em um beco sem saída. Uma observação casual sugere que nós, às vezes, aplicamos a palavra "deveria" a membros de uma outra cultura. Eu recém argumentei que o Relativismo Agente pode explicar esse padrão de comportamento verbal, mas também o podem os oponentes do Relativismo Agente. Precisamos de um argumento para mostrar que essa interpretação está certa. Esse beco sem saída parece ainda pior quando consideramos a segunda objeção para o argumento de Harman que mencionei acima. Às vezes, aplicamos julgamentos de dever a pessoas com diferentes valores (como com missionários e pais), e às vezes não (como com Hitler e os alienígenas). Harman tomou a última evidência para o Relativismo Agente, mas Lyons ofereceu

outra explicação para por que nos abstermos de dirigir julgamentos de dever a alguns indivíduos (Hitler está além da persuasão). Então, há duas explicações para o comportamento verbal, e nós ainda precisamos de uma forma de decidir qual explicação está correta.

Para nos desvencilharmos desse beco sem saída, ajudará refletirmos sobre o conceito *DEVERIA*. No capítulo 3, eu chamei a atenção para uma diferença entre julgamento de dever e julgamentos de incorreção. Se eu disser que você não deveria bater no seu cachorro, eu estou dizendo algo mais do que *bater no seu cachorro é errado*. Apreendi essa sutil diferença ao dizer que um julgamento de dever transmite o fato de que uma norma tem autoridade sobre o comportamento da pessoa a quem aquele julgamento se dirige. Se eu digo que você não deveria bater no seu cachorro, estou dizendo que a regra contra os maus-tratos a cães tem autoridade sobre você, e assim seu comportamento deve sujeitar-se a ela. Não desfiz essa ideia de autoridade no capítulo 3, mas ela pode ser desfeita agora. Missionários presumivelmente pensam que o canibalismo está errado. Esse pensamento pode consistir em seus sentimentos negativos em relação ao canibalismo. Mas algo mais é necessário se forem julgar que esse sentimento de incorreção tem autoridade sobre o dos Akamara. Por que os Akamara atenderiam à norma que condena o canibalismo? Por que essa norma regeria suas ações? A resposta a essa pergunta depende da fonte da normatividade. Se as normas têm autoridade em virtude de alguns fatos universais, independentes da mente, então os missionários poderiam estar certos ao inferirem que a norma contra o canibalismo aplica-se aos Akamara. Mas suponha que as normas sejam locais e dependam dos valores que uma pessoa apoia. Suponha que a qualidade de guiar as ações de uma norma dependa de ter internalizado essa norma. Se for o caso, então a norma contra o canibalismo não é de fato normativa, ou guia de ações para os Akamara, e seria um erro dizer que eles estão sob essa jurisdição. Se eu estou certo em pensar que o conceito *DEVERIA* está ligado à ideia de autoridade normativa, então parece se seguir disso que os missionários estão enganados quando dizem que os Akamara deveriam privar-se da prática do canibalismo, mesmo que eles estejam certos em dizer que o canibalismo está errado.

É fácil motivar essa intuição se mudarmos as coisas e imaginarmos como responderíamos a julgamentos morais dos Akamara. Se o Relativismo Agente fosse falso e o sentimentalismo construtivo fosse verdadeiro, então, quando um membro de outra cultura julga que deveríamos abster-nos de fazer algo, então estaríamos sob a obrigação de evitá-lo. Suponha que os Akamara desaprovem a ideia de manter animais como bichos de estimação. Esse fato significa que nós temos uma obrigação de libertar nossos animais de estimação? O fato de os racistas se oporem à integração significa que devemos ressegregar as nossas escolas? Essa seria uma descoberta muito problemática. Para evitar essa conclusão, devemos insistir que somos obrigados a seguir apenas os valores que realmente aprovamos. A noção de obrigação, que está implicada na palavra "deveria", parece aplicar-se apenas quando uma norma tem autoridade, e, se o sentimentalismo construtivo for verdadeiro, normas têm autoridade apenas sobre aqueles que a seguem. Essas observações me levaram a pensar que o Relativismo Agente é verdadeiro.

Concluo que nossos julgamentos são relativos à cultura de julgadores morais e também aos agentes que são julgados. Quando um falante diz "X deveria \emptyset ", esse enunciado pode ser verdadeiro se e somente se o falante tem valores morais que estabelecem \emptyset -ar, e X tem esses valores. Como sugerido acima, as coisas podem ser um pouco mais flexíveis com julgamentos que expressamos usando palavras como "certo" ou "errado". Como padrão, essas palavras podem referir aos valores do falante, mas não dependem dos valores dos agentes. Podemos dizer que os Akamara estão errados em se engajarem no canibalismo. Podemos dizer isso porque esse julgamento não faz nenhuma afirmação direta sobre o comportamento dos Akamaras. Isso não implica que os Akamara estão sob a jurisdição de nossa norma.

Também pode haver contextos nos quais podemos usar os termos "certo" e "errado" de maneira não prescritiva, como quando dizemos "Para os puritanos, sexo antes do casamento é errado". Mas quando usamos "certo" e "errado" de modo prescritivo, relativizamos para o nosso próprio sistema de valores. Não dizemos "Sally estava errada ao fazer sexo antes do casamento; afinal, ela era uma puritana".

Essas lições levam a uma extensão da definição de relativismo metaético que eu ofereci acima:

Relativismo Metaético (versão estendida): as condições de verdade de um julgamento moral dependem do contexto no qual o julgamento é formado, tal que:

Um julgamento de que *X deveria \emptyset* é verdadeiro se e somente se não- \emptyset é errado nos sistemas de valor de ambos falante e X.

Um julgamento de que \emptyset -ar está errado é verdadeiro se e somente se \emptyset -ar é o objeto de um sentimento de desaprovação entre o(s) indivíduo(s) contextualmente saliente(s) (normalmente, o falante).

Maiores aperfeiçoamentos poderiam ser acrescentados para outros conceitos morais, mas esse será suficiente para a discussão no restante deste capítulo.

5.1.3 Relativismo de Conteúdo e Relativismo de Verdade

Na definição que estou usando, a verdade de um julgamento moral depende do contexto. Como afirmado, essa formulação é subespecificada. Há diferentes formas de sensibilidade ao contexto. Em recente teoria semântica, duas formas de relativismo receberam considerável atenção: o relativismo do conteúdo e o relativismo da verdade. Para a maior parte deste capítulo, a distinção entre essas espécies de relativismo não importa, mas uma teoria relativista completa especificaria qual, se não as duas, é a correta para termos morais. Isso facilitaria também a exposição. Na seção 5.2, revisarei as objeções ao relativismo metaético, e as duas espécies de relativismo oferecem estratégias sutilmente diferentes para abordar algumas dessas objeções. Nesta seção, então, oferecerei algumas razões preliminares para preferir o relativismo do conteúdo ao relativismo da verdade. Primeiro, deixem-me delinear a distinção.

Como o nome sugere, um relativista do conteúdo diz que o conteúdo de um enunciado (ou julgamento) é determinado, em parte, pelo contexto. O termo

"conteúdo" é uma palavra que tem mais de um significado em filosofia, mas a ideia básica é que o contexto tem um papel em determinar o que um enunciado significa; dependendo de sua teoria do significado, esse pode ser mais bem entendido em termos de qual proposição o enunciado expressa. Um enunciado é relacionado a um conteúdo se ele expressa diferentes proposições sob diferentes circunstâncias, e um termo é relacionado a um conteúdo se faz diferentes contribuições à proposição expressa por um enunciado que o contém. Se proposições são complexos estruturados que contêm propriedades e indivíduos, então se pode dizer que um termo é relacionado a um conteúdo se a propriedade ou o indivíduo ao qual contribui para uma proposição varia entre contextos.

Para ilustrar, considere as expressões indexicais. A palavra "eu" em "eu sou filósofo" toma seu conteúdo, em parte, do contexto de enunciação. Se Gertrude dissesse essa sentença, "eu" designaria Gertrude. Um relativista do conteúdo a respeito de termos morais diria que as propriedades expressas por esses termos dependem do contexto. "Matar é errado" atribui uma propriedade diferente de matar quando pessoas diferentes a enunciam.

Há diferentes teorias sobre como termos e conceitos relacionados a conteúdos funcionam. Uma possibilidade é que os usos desses termos e conceitos não representam internamente o fato dos quais eles são relativos. Nessa visão, a palavra "errado" significa algo diferente quando dita por diferentes observadores, mas as pessoas tipicamente não reconhecem esse fato porque o conceito correspondente, ou a entrada lexical, não contém nenhum componente correspondente ao fato de que o significado é sensível ao contexto. Outra possibilidade é que essa sensibilidade ao contexto é representada internamente. Em particular, poderia haver um componente na entrada lexical para termos morais que referem indexicalmente a, digamos, falante e destinatário, ou a algum outro indivíduo saliente. Dreier (1990) foi um campeão nessa abordagem. Sua avaliação de termos morais provém diretamente da celebrada teoria dos indexicais de Kaplan (1989). A palavra "eu", por exemplo, seleciona indivíduos diferentes, mas, em todas as instâncias, há um componente de significado comum, que pode ser explicado pela expressão "o falante deste enunciado". Kaplan chama isso de "característica"

da expressão. Ele usa o termo "conteúdo" para referir o indivíduo determinado por essa característica quando um enunciado é dito em um contexto. Por analogia, pode-se dizer que "errado" tem característica e conteúdo. A característica do termo "errado" pode ser grosseiramente capturada pela expressão: "a propriedade de ser reprovado pelo falante" (ou algum outro avaliador saliente). O conteúdo do termo é, então, determinado contextualmente. Se Gertrude expressa o enunciado "matar é errado", então aquele enunciado atribui a *matar* a propriedade de ser reprovado por Gertrude. O termo "errado" implicitamente apresenta um elemento indexical, que determina conteúdos diferentes em contextos diferentes. A característica do termo "deveria" poderia ser um pouco mais complexo; poderia referir a valores do avaliador e do agente avaliado. Mais especificamente, sentenças na forma "X deveria \emptyset " podem significar que é errado não- \emptyset no sistema de valor do falante, e X cai sob a autoridade dessa norma. O sistema de valor se torna determinado contextualmente por meio de um indexical oculto que aponta para o falante, e há também um elemento (não-indexical) oculto que implica que o agente X deve cair sob aquele sistema de valor.

Vamos agora tratar do relativismo da verdade. Relativistas do conteúdo dizem que julgamentos têm diferentes conteúdos em diferentes contextos, e, por contraste, relativistas da verdade dizem que julgamentos sempre têm o mesmo conteúdo, mas sua verdade varia entre contextos de avaliação. Para ilustrar, considere a sugestão plausível de que muitos enunciados comuns expressam proposições que têm valores de verdade apenas relativos a pontos no tempo. O enunciado "Gertrude está em pé" tem o mesmo conteúdo quando quer que seja dito, mas é verdade apenas quando avaliado nas vezes em que Gertrude está em pé. Um relativista da verdade sobre termos morais diria que, além de parâmetros temporais, proposições são avaliadas em relação a sistemas de valor. Nessa visão, "matar é errado" tem um conteúdo invariável, mas tem diferentes valores dependendo do sistema de valor relativo a qual verdade é avaliada. Na semântica filosófica recente, o relativismo do conteúdo é chamado de "contextualismo" e o relativismo da verdade é chamado apenas de "relativismo" (MacFarlane, 2005), mas essa é uma

terminologia infeliz, já que ambas as posições postulam um tipo de relativismo e um tipo de dependência de contexto.

Há uma literatura bastante considerável e crescente sobre essas duas espécies de relativismo, e pouco do que eu tenho a dizer depende muito fortemente de decidir qual, se não as duas, oferece a análise correta de termos morais. Ainda, para registro e para facilitar a exposição no restante deste capítulo, quero transmitir meu motivo para pensar que o relativismo do conteúdo é preferível ao relativismo da verdade.

Minha queixa contra o relativismo da verdade é que eu não vejo quais propriedades os termos morais poderiam possivelmente expressar na ausência de um contexto. Suponha que Gertrude diga que o canibalismo é errado. O que é expresso por "errado" aqui? Um relativista do conteúdo diria (grosseiramente) que a incorreção refere-se à propriedade de ser o objeto da reprovação de Gertrude. Mas o relativista da verdade não pode dizer isso. Essa propriedade é determinada pelo contexto, e o relativista da verdade diz que a propriedade expressa por um termo relacionado à verdade é o mesmo entre contextos. O relativista da verdade poderia propor que "errado" expressa a propriedade de ser o objeto da reprovação de *alguém*. Mas essa não pode ser a análise correta. Se "errado" se referisse à propriedade de ser reprovado por alguém (qualquer pessoa!), então eu poderia realmente dizer que, digamos, sexo antes do casamento é errado – algo que eu não acredito. Alguém poderia tentar contornar a situação ao sugerir que termos morais referem a "propriedades não específicas". Esse termo é apresentado por Kompa (2002) para descrever propriedades que são invariáveis entre enunciados de um termo, mas o contexto é requerido para determinar quem tem essas propriedades. Mas não entendo quais propriedades são, de modo que elas podem não ser específicas dessa forma. Primeiro, a proposta sugere que propriedades podem ser individualizadas independentemente de suas instâncias, e essa é uma afirmação controversa. Segundo, a visão parece implicar que a mesma propriedade, a incorreção por exemplo, viria e iria dependendo de quem estaria fazendo a afirmação moral; se "errado" sempre referisse à mesma propriedade, por que aquela

propriedade deveria aumentar e diminuir como uma função de contexto linguístico? Eu acho a ideia de propriedades inespecíficas difícil de engolir.

Aqueles que rejeitam o relativismo do conteúdo em virtude do realismo da verdade o fazem porque pensam que o segundo faz um trabalho melhor em explicar as discordâncias morais (Kölbel, 2003; Richard, 2004; Lasersohn, 2005; MacFarlane, 2005; Brogaard, no prelo). Mas eu não acho que essa é uma razão adequada para preferir o relativismo da verdade. Os relativistas do conteúdo possuem estratégias perfeitamente boas para explicar discordâncias morais. Todas as estratégias discutidas no capítulo 3 estão disponíveis para o relativista do conteúdo. Por exemplo, os participantes de um debate, às vezes, compartilham valores morais básicos ou falsamente acreditam que compartilham, e mesmo quando eles não possuem essa crença, eles podem conduzir um debate com a esperança de persuadir seus parceiros de diálogo a adotar um novo conjunto de valores. Além disso, rejeito a afirmação de que relativistas da verdade saem-se melhor em explicar debates. Relativistas da verdade gostam de dizer que, quando as pessoas expressam uma divergência usando termos relativos, elas conseguem se referir à mesma proposição, e assim há um objeto legítimo de disputa. Se Tobias diz “canibalismo é errado” e Michii diz “não, canibalismo não é errado”, então há uma proposição à qual eles dois estão se referindo, e o valor de verdade daquela proposição está sendo disputado. Até agora, tudo bem. Mas note que o relativista da verdade está também comprometido em dizer que os participantes desse debate estão assumindo diferentes contextos para avaliar aquela proposição. Michii o está avaliando para um sistema de valor, e Tobias o está avaliando para outro. Desse modo, estritamente falando, há um sentido no qual eles estão falando de coisas diferentes. Tobias e Michii têm o direito de avaliar suas afirmações morais relativas a seu próprio sistema de valor, então ambos têm o direito de atribuir valores de verdade diferentes para o enunciado “canibalismo é errado”. Consequentemente, o debate parece se situar em algum tipo de confusão. Acrescenta pouco dizer que eles expressam a mesma proposição com o enunciado, uma vez que se admite que aquelas proposições podem ter valores de verdade sensíveis ao contexto. Em uma abordagem relativista da verdade, ter uma disputa moral é um pouco como ter um

debate sobre se Gertrude está em pé, no qual uma parte insiste que ela está em pé a cada vez que ela passa a ficar em pé, e a outra parte insiste que Gertrude não está em pé a cada vez em que ela passa a sentar. Não poderíamos explicar esse debate ao assinalar que “Gertrude está em pé” expressa a mesma proposição quando quer que seja dita. Outros fatores precisariam ser apresentados para dar sentido a por que os participantes do debate falharam em perceber que suas afirmações são perfeitamente compatíveis. Penso que o relativismo do conteúdo e o relativismo da verdade estão em uma posição igual quando se trata de discordância moral. Ambos precisam considerar o fato de que, em alguns debates morais, as pessoas que fazem afirmações consistentes tomam a si próprias como sendo divergentes.

Em resumo, penso que o relativismo do contexto faz mais sentido que o relativismo da verdade, porque ele tem uma história melhor sobre quais propriedades nossos termos morais expressam, e o relativismo da verdade não oferece uma abordagem superior de discordância moral. No que segue, eu assumirei que o relativismo do conteúdo está correto, embora os argumentos que eu darei possam ser aceitos em defesa do relativismo da verdade.

5.1.4 O papel da cultura.

A forma de relativismo que estou defendendo relativiza os julgamentos morais para indivíduos (a pessoa que faz um julgamento e, às vezes, a pessoa ou pessoas a quem se dirige). Mas eu já tive oportunidade de referir aos valores de um *grupo* de pessoas. Falei sobre os valores dos Akamara. O relativismo, às vezes, é apresentado como a visão de que a moralidade varia de cultura para cultura. Estritamente falando, penso que culturas são unidades amplas demais para fundamentar a moralidade, mas penso que a moralidade é profundamente influenciada pela cultura, e que, às vezes, tem sentido falar sobre os valores morais de um grupo cultural. Além disso, muito da evidência para o relativismo descritivo vem das observações de diferenças culturais. Antes de olhar para objeções ao relativismo, quero dizer algo sobre o que as culturas são e como elas contribuem para a moralidade.

Começamos considerando uma definição geral de cultura. A definição primeira de cultura mais influente foi proposta por Edward Burnett Tylor. Ele define cultura como "o todo complexo que inclui conhecimento, crença, arte, lei, morais, costumes, e quaisquer outras capacidades e hábitos adquiridos pelo homem como um membro da sociedade" (Tylor, 1871: 1). Nessa definição, os componentes da cultura são bastante abrangentes. A cultura inclui componentes materiais, comportamentais e psicológicos, e componentes psicológicos são amplos. De fato, tratam-se de muitos componentes. Tylor menciona também os meios pelos quais as culturas são transmitidas. Ele diz que as culturas são "adquiridas pelo homem como um membro da sociedade". Se "sociedade" e "cultura" são sinônimos próximos, isso pode introduzir uma circularidade na definição: cultura é aquilo que é culturalmente adquirido.

Alguém pode melhorar a definição de Tylor ao enfatizar e elaborar sua observação sobre transmissão. Culturas consistem naquelas coisas que são socialmente aprendidas. O aprendizado social é aprendido pela ajuda de outros indivíduos ou por meio de coisas que foram criadas por outros indivíduos. Mais especificamente, as culturas consistem em coisas que são aprendidas por um processo que alguém pode chamar de conformidade: membros de uma cultura fazem coisas em virtude do fato de que outros indivíduos fazem aquelas coisas. Conformidade pode ser consciente ou inconsciente. As culturas consistem em muitos traços psicológicos e comportamentais diferentes (hábitos, habilidades, ideias, valores etc.) porque essas coisas podem ser todas transmitidas socialmente. Objetos materiais, tais como trabalhos de arte, não são partes componentes de uma cultura; em vez disso, eles são produtos culturais. Culturas podem incluir conhecimento e atitudes pertencentes a tais objetos, contudo. Culturas podem incluir coisas que têm uma origem humana e biológica. As habilidades de linguagem podem ser um exemplo. Todas as coisas que aprendemos por meio de outros dependem de capacidades biológicas para o aprendizado. Mas coisas que são adquiridas por meio de aprendizado social genuíno, como eu usarei esse termo, não podem ser inteiramente controladas pela

biologia. Coisas que devem ser socialmente aprendidas são aquelas que não podem ser completamente determinadas por um bioprograma.

Ao colocar o aprendizado social no centro de uma definição de cultura, podemos fazer algum progresso nas questões de individuação social e de pertencimento de grupo. Ser um membro de um grupo cultural é (a) ter traços psicológicos e comportamentais que foram adquiridos por meio de conformidade com membros daquele grupo e (b) estar disposto a continuar a concordar com o grupo. Para simplificar a exposição, na maior parte das vezes eu deixarei de lado a segunda condição. Um grupo cultural *ideal* é um grupo de indivíduos com traços uniformes socialmente aprendidos, cada qual toma todos os seus traços socialmente aprendidos a partir de outros membros do grupo. Grupos de indivíduos qualificam-se como culturas separadas na medida em que eles não aprendem socialmente uns dos outros. Na realidade, nenhum grupo cultural é ideal. Grupos se misturam, cruzam-se, sobrepõem-se. Muito frequentemente, tomamos pistas para aspectos diferentes de nossas vidas de grupos diferentes. Um indivíduo pode se identificar como uma mulher, liberal, budista, acadêmica e canadense. Por essa razão, é útil falar das dimensões da cultura: categoria de gênero, partido político, grupo profissional, religião, e assim por diante. Alguém é um membro de um grupo cultural em relação a uma dessas dimensões, à medida em que seus traços socialmente aprendidos ao longo daquela dimensão são adquiridos por conformidade àqueles membros daquele grupo. Em muitos casos, nós também acrescentamos nosso toque pessoal a traços socialmente transmitidos. Adaptamos o que aprendemos a necessidades e interesses pessoais. Combinamos traços aprendidos a partir de mais de um grupo. Culturas podem evoluir por meio de tais transformações.

Com essa caracterização geral de cultura e de pertencimento de grupo cultural, voltemo-nos à questão de quais papéis a cultura tem na moralidade. Acredito que três papéis sejam mais importantes: cultura pode ser a causa da moralidade, o efeito da moralidade e a razão para a moralidade. Como causa, cultura é a fonte primária a partir da qual conseguimos nossos valores. Aprendemos a ser morais por nos submetermos a uma educação moral, o que

envolve estarmos emocionalmente condicionados às pessoas ao nosso redor, incluindo cuidadores, papéis modelos, e pares. Os valores que adquirimos são predominantemente herdados dos valores de outros. Assim como uma cultura pode transmitir crenças religiosas ou habilidades vocacionais, pode transmitir valores. Se não fôssemos expostos a uma cultura (ou culturas), não adquiriríamos os valores resultantes em nós, e também não poderíamos ter quaisquer valores morais. A esse respeito, a moralidade pode ser considerada como um construto cultural.

Como efeito, as culturas podem ser sustentadas, pelo menos em parte, por meio de valores morais. Nós nos afiliamos àqueles que compartilham nossos valores, e evitamos aqueles que não os seguem. Portanto, a moralidade une os indivíduos. Outros construtos culturais podem construir alguns vínculos também, tais como crenças religiosas compartilhadas, mas esses fatores frequentemente funcionam juntamente com a moralidade. Se pessoas que compartilham uma perspectiva religiosa discordassem moralmente, elas talvez não pudessem continuar a interagir, e as crenças religiosas poderiam evoluir para tradições culturais distintas. Pontos de vista morais compartilhados permitem um tipo de coesão que é especialmente estável, porque a moralidade tem um impacto no comportamento.

Isso está relacionado ao terceiro papel da cultura na moralidade. É importante perguntar para que serve a moralidade. Por que nós temos valores morais? A resposta óbvia é que a moralidade emerge como um sistema de regras para fazer com que as pessoas funcionem coletivamente de modos estáveis e produtivos. Temos moralidade para construir um grupo social coerente. Valores morais nos levam a cooperar e nos previnem de prejudicar os membros de nossas comunidades. Cultura é a *raison d'être* para a moralidade. Robinson Crusoe não teria necessidade dela. A esse respeito, moralidade tem uma natureza social (ou pelo menos um *telos* social).

Esses fatos sobre a cultura são importantes para se ter em mente quando pensamos sobre o relativismo. A teoria relativista que eu tenho defendido é uma variante do que Dreier (1990) chama de "relativismo do falante". O conteúdo de um julgamento moral depende dos valores da pessoa que faz aquele julgamento. Cada

um de nós é o juiz definitivo de nossos próprios valores. Penso que o relativismo do falante está correto, estritamente falando, mas acho que o rótulo é enganoso. Prefiro pensar na moralidade como intersubjetiva. Há duas razões para isso. Primeiro, se eu estiver certo sobre a semântica, alguns termos morais, na verdade, pressupõem que há outros indivíduos que compartilham nossos pontos de vista morais. Em particular, devendo o Relativismo Agente em relação a julgamentos de dever, e, nesse ponto de vista, eu não posso dizer verdadeiramente que alguém deveria fazer algo sem pressupor que essa outra pessoa compartilha meus valores. Segundo, e de modo mais profundo, a moralidade é cultural em todas as formas que eu descrevi. Se meus valores são informados pela cultura, promovem cultura, e passam a existir para servir a fins culturais, então meus valores são muito suscetíveis de ser compartilhados por outros. Além disso, quando eu me engajo no discurso moral, eu o faço como um membro da cultura em um esforço de coordenar meu comportamento com outros. A maioria dos debates morais é provavelmente conduzida entre pessoas que compartilham valores morais básicos, e nós assumimos, tipicamente, valores compartilhados quando nos engajamos no debate.

Por essas razões, penso que podemos falar sobre comunidades de moralizadores, e os limites de tais grupos frequentemente coincidirão com os limites de grupos que são considerados pré-teoricamente como culturas. Isso é especialmente provável dentro de sociedade em pequena escala, nas quais há maior homogeneidade e menos oportunidade para interação com pessoas que têm conhecimentos históricos diferentes. Conforme os grupos se tornam maiores e mais complexos, há frequentemente grande variação, mas isso não impede generalizações. Podemos até mesmo falar sobre os valores de uma nação. Seria um erro falar sobre os valores romanos se isso significa os valores compartilhados por todos os cidadãos de Roma. Mas faz sentido falar sobre os valores promovidos por aqueles no poder em Roma. Havia esforços ativos para unificar o Império Romano por meio da promoção de valores morais. Do mesmo modo ocorria para os Astecas, que eram etnicamente diversos, mas unificados por um conjunto compartilhado de instituições e práticas que provavelmente resultaram em convergência moral considerável.

Dentro do mundo contemporâneo ocidental, encontramos exemplos similares. Nos Estados Unidos, há um discurso público sobre liberdade que se designa a promover a visão de que a democracia é um valor básico. Patriotismo é também promovido como básico. Se alguém nos Estados Unidos é acusado de não ser patriota, a resposta comum é rejeitar a acusação, e não rejeitar a afirmação de que o patriotismo tem um valor intrínseco. Em alguns países europeus, há um esforço concentrado em substituir o ideal de patriotismo pelo ideal do cosmopolitismo (ou, pelo menos, o da lealdade à Comunidade Europeia). Apesar de tais campanhas homogeneizantes, os países continuam a promover valores distintos dentro de suas próprias fronteiras. Na França, as pessoas comumente apelam a valores de República: liberdade, igualdade e fraternidade. Esses valores não são acalentados por todos na França, e eles certamente têm defensores fora da França, mas eles podem ser especialmente suscetíveis a servirem como normas fundadoras em debates morais entre os franceses. Podemos falar de “valores franceses” como praxe para valores que são altamente promovidos na França, e assim especialmente prováveis de serem adotados por aqueles que têm sido aculturados na França. É simplista assumir que qualquer nação pode ser associada a um conjunto de valores simples, mas falar de valores predominantes não é problemático. Por essa razão, podemos, às vezes, falar livremente sobre a moralidade francesa ou a moralidade asteca. Isso nem sempre é apropriado. Julgamentos morais devem, com frequência, ser relativizados para grupos cujos limites falham em coincidir com qualquer nação em particular. Contudo, seria um erro negligenciar a extensão em que os membros de uma sociedade tendem a se sujeitar. De fato, a conformidade entre grupos (e por isso valores de grupo) tem um tipo de prioridade sobre valores individuais, na medida em que um dado indivíduo é suscetível a ter adquirido valores de uma coletividade social. Se negligenciarmos a moralidade das culturas, perderemos *insights* sobre a moralidade dos indivíduos, mesmo se a moralidade cultural for apenas uma média entre indivíduos agregados.

No que se segue, eu falarei com frequência sobre os valores de um grupo cultural particular. Eu não quero dizer que isso seja um ponto de partida da forma individualista de relativismo que eu endosso, mas sim uma indicação do fato de que

indivíduos frequentemente compartilham valores morais dentro de grupos culturais. De fato, a melhor forma de provar que há variação na moralidade é olhar para os níveis das tendências do grupo. Se uma disputa moral surge entre dois indivíduos, é sempre possível que um deles esteja confuso sobre fatos, inarticulado ou insincero. Mas se dois grupos parecem ter valores divergentes, em uma média entre muitos indivíduos, as diferenças são difíceis de recusar. Além disso, a divergência moral aparente entre grupos é frequentemente mais dramática que a divergência entre indivíduos do mesmo grupo. As divergências de grupo são, dessa forma, a maior fonte de evidência para o relativismo descritivo, e um fato principal na construção de um caso para o relativismo metaético.

5.2 OBJEÇÕES AO RELATIVISMO

5.2.1 Relativismo Descritivo

Avaliarei agora algumas dessas objeções que têm sido levantadas contra o relativismo metaético. A maioria dessas objeções mira o relativismo diretamente, mas a primeira é uma objeção ao argumento que eu costumava defender em primeiro lugar. Lembremos que o meu argumento para o relativismo metaético tem o relativismo descritivo como uma premissa. O relativismo descritivo é normalmente tomado como óbvio, dada a aparente diversidade de crenças e práticas morais. Evidência para diversidade é geralmente delineada a partir de pesquisas interculturais. Quando analisamos o registro antropológico, parece óbvio que a moralidade varia entre culturas. Diferenças declaradas podem ser desafiadas, contudo, e, mesmo que diferenças possam ser estabelecidas, o relativismo descritivo não se confirma necessariamente. Antes de especificar essa objeção, é útil apresentar alguns exemplos. Em face disso, muitas culturas engajam-se em práticas que pessoas em culturas euro-americanas contemporâneas achariam repreensíveis.

Novamente, podemos começar com o canibalismo. Duas formas de canibalismo são geralmente distinguidas por antropólogos. O canibalismo

mortuário é a prática de comer os mortos. Na maioria dos casos, corpos mortos são queimados e, então, as cinzas são misturadas com um líquido e consumidas. Isso não é sempre feito do mesmo modo, contudo. No começo do século XX, as mulheres membros da cultura Fore, na Nova Guiné, costumavam consumir os corpos de parentes mortos após os enterrar por alguns dias. Essa prática levou a uma epidemia, chamada pelos Fore de *Kuru*, ou a morte risonha. Trata-se de uma forma de encefalopatia espongiiforme transmissível – uma forma humana de doença da vaca louca (Mead et al., 2003). O canibalismo mortuário contrasta com o canibalismo como prêmio de disputas, no qual as pessoas consomem os corpos daqueles que foram mortos ou capturados na guerra. Os Astecas foram responsáveis pelo sistema mais vasto de canibalismo como prêmio de disputas já registrado. Embora as estimativas variem, acredita-se que eles tenham sacrificado entre 20.000 e 80.000 pessoas a cada ano (Harner, 1977). Às vezes, milhares eram mortos em um único festival. As vítimas eram sacrificadas em altares arredondados no topo de grandes pirâmides. Elas teriam seus corações removidos por sacerdotes enquanto eles eram pendurados. Os corpos eram rolados para baixo pelos degraus das pirâmides, e então distribuídos entre seus captores. Seus membros eram cozidos e consumidos.

A carnificina pública em massa foi institucionalizada muitas vezes na história europeia. Os esportes sangrentos da Roma antiga são um exemplo óbvio. O Coliseu poderia comportar, sentados, 45.000 fãs ansiosos, com lugares para mais 5.000 pessoas em pé (Platner e Ashby, 1929). Na hora do almoço, os espectadores no Coliseu viriam os prisioneiros serem crucificados, queimados, ou dilacerados por animais. Depois disso, eles assistiam a gladiadores, que eram, na maioria, escravos, confrontarem-se em um combate mortal. Se um gladiador fosse derrotado, mas não mortalmente ferido, a multidão demandaria um golpe final. Após o gladiador ter sido morto, seu corpo era marcado com ferro quente e sua cabeça, esmagada com uma marreta. Ataques gladiatoriais não foram uma consequência de curta duração dos anos decadentes que precederam a queda de Roma. Eles duraram por mais ou menos 500 anos (264 a.C. - *circa* 404 d.C.). Quando os jogos foram finalmente banidos pelos imperadores cristãos, não foi porque se julgou que eles eram cruéis

demais. Wiedemann (1992) argumenta que os jogos foram banidos porque eles pressupuseram que a salvação poderia ser alcançada por meio de valor militar. Isso entrou em conflito com a escatologia que os cristãos estavam tentando promover.

Vamos passar, para variar, da violência para o sexo, ou pelo menos para o parentesco. Tendemos a ter atitudes morais em relação ao casamento. Nós valorizamos ligações monogâmicas, matrimoniais, de amor entre pares, e desvios desse ideal são geralmente considerados imorais. Outras culturas possuem outras convenções. Muitas culturas têm casamentos arranjados em vez de casamentos por amor. A poliginia é também um lugar-comum, no qual um homem tem mais de uma mulher. Em sociedades que requerem casamento monogâmico, há geralmente sistemas de concubinação institucionalizados, que podem ser funcionalmente equivalentes a um sistema poligâmico. Nessa combinação, os homens podem ter parceiras extramaritais múltiplas, e elas podem viver na casa do marido e herdar sua fortuna. A concubinação foi vastamente praticada na Europa antes de ser banida pela Igreja Cristã em seus primórdios (Goody, 1983). A poliandria é menos comum que a poliginia, mas não é inédita. É praticada, mais famosamente, no Nepal e no Tibet. Mesmo sem a poliandria, algumas culturas têm atitudes permissivas quanto à infidelidade feminina. Entre os Tlingit, as esposas podem fazer sexo com membros do clã, mas sexo fora do clã é punido com morte (Oberg, 1934). A infidelidade sexual feminina é também tolerada em algumas sociedades, especialmente se permanece dentro do grupo pré-estabelecido [*within the in-group*]. Usando uma amostra intercultural ampla, Hupka (1981; 1991) argumenta que há uma tolerância maior quanto à infidelidade em culturas que são sexualmente permissivas e não dão ênfase na propriedade privada.

Em quase todas as culturas (veja abaixo) há proibições contra o incesto, mas essas variam consideravelmente. Em culturas euro-americanas, é considerado incestuoso casar com um primo de primeiro grau, enquanto que em partes da Índia e do Paquistão e do Oriente Médio casar com um primo em primeiro grau é encorajado fortemente (Bittles, 1990). Em algumas culturas, incesto entre irmãos é punido severamente, e em outros é simplesmente desencorajado. Por exemplo, entre os Incas, o incesto é punido arrancando os olhos, enquanto que entre os

Truam, isso é simplesmente desaprovado (Thornhill, 1991). Kuper (2002) encontrou evidência de uma atitude vaga em relação ao incesto entre irmãos no século XIX na Grã-Bretanha. No registro da corte de um julgamento por adultério, apenas uma breve menção foi feita sobre o fato de que a mulher acusada de adultério o tinha cometido com seu irmão de sangue.

Há variação cultural em pontos de vista sobre que tipos de parceiros sexuais são apropriados. Em algumas culturas, a homossexualidade é fortemente condenada, e em outras é normal. No Japão feudal, por exemplo, relacionamentos românticos entre um samurai e homens jovens eram encorajados (Leupp, 1995). Entre os Etoro da Nova Guiné, todos os homens envolvem-se em sexo homossexual, e apenas se permitia a eles fazer sexo com mulheres durante cem dias por ano (Kelly, 1977). No ocidente, as atitudes em relação à homossexualidade variavam. Em Roma e na Europa medieval, isso era geralmente tolerado, mas, no século XIII a homossexualidade passou de legal para punível com morte (Boswell, 1980).

Relações sexuais entre adultos e crianças podem ser também verificadas. Em partes da Índia, por exemplo, as filhas de famílias pobres são dadas aos templos para servirem de *devadasi*, ou esposas dos deuses (Marglin, 1985). Essas meninas servem como prostitutas; é esperado delas que façam sexo com homens adultos a partir do momento em que se tornarem púberes. Entre os Sambia da Nova Guiné, os garotos ocupam-se em rituais de felação com homens adultos (Herdt, 1987). Algumas culturas parecem tolerar a bestialidade. Entre os Kuguru, há regras proibindo o intercuro sexual com a vaca de outra pessoa, mas não, evidentemente, com a sua própria (Davis e Whitten, 1987). No começo da Inglaterra moderna, em contraste, a bestialidade era frequentemente tratada como uma ofensa capital (Thomas, 1983).

O casamento é uma instituição designada para a produção e criação da prole. Como um exemplo final da variação cultural, quero considerar o infanticídio, um destino que muitos filhos ao redor do mundo devem enfrentar. Na China, o infanticídio era uma vez generalizado, e apesar de agora ser ilegal, continua a ser praticado (Coale e Banister, 1994). A evidência para o infanticídio ininterrupto vem

das taxas de nascimento por sexo. Em algumas províncias chinesas, muito mais nascimentos de meninos do que de meninas são relatados, o que sugere que bebês meninas estão sendo mortas. O infanticídio na China tem sido praticado com muito mais frequência contra meninas, porque elas dão menos suporte à família. Tradicionalmente, meninas eram especialmente incapazes de dar suporte na China porque elas estavam sujeitas ao enfaixamento dos pés, outra prática que foi condenada no Ocidente. O infanticídio é muito menos incomum que o enfaixamento dos pés. Isso tem sido praticado por todo o mundo, e, como na China, as vítimas são, em geral, garotas. Os Inuítes praticaram infanticídio até recentemente (Freeman, 1971). Os gregos e os romanos da antiguidade deixavam bebês rejeitados morrerem por exposição às intempéries do clima. De acordo com a lenda, isso foi o que aconteceu com Édipo, mas, como o destino queria, ele conseguiu sobreviver. Na cultura americana contemporânea, o infanticídio foi seriamente condenado. Em 1996, o estado de New Jersey pediu a pena de morte quando Amy Grossberg e Brian Peterson, um jovem casal de um subúrbio rico, foram condenados por matar seu recém-nascido.

Essa pesquisa de valores morais divergentes poderia ir adiante. Todos os exemplos demonstram que os membros de outras culturas engajam-se em comportamentos que consideramos depravados, chocantes ou atroz. É natural inferir de tais exemplos que os valores morais são culturalmente relacionados. O relativismo descritivo parece incontestável.

Mas o relativismo descritivo, na verdade, requer mais defesa. A tese enfrenta três objeções relacionadas. Primeiro, opositores ao relativismo defendem que alguns exemplos de variação têm sido exagerados. Tome a obsessão romana com os jogos gladiatoriais. Isso é tão diferente da obsessão contemporânea com o futebol, o boxe e filmes violentos? Ou considere o infanticídio inuíte. Em ambientes extremamente difíceis, com recursos escassos, o controle populacional pode ser necessário para salvar grupos. Nós não consideraríamos ser permitido tirar a vida de uma criança se estivessemos em tal cenário? (Rachels, 1986: cap. 2) Em alguns casos, o discurso público poderia implicar que os valores entre duas culturas diferem, mas observar práticas reais sugere convergência.

Segundo, em casos em que há variação cultural em valores morais, não está claro se as diferenças são intratáveis. O relativismo descritivo pode ser definido como a visão de que as pessoas têm sentimentos diferentes em relação às mesmas coisas. Tenho sugerido que tais diferenças existem, mas elas são, na verdade, difíceis de estabelecer. Não temos nada como um cerebrosópio para olhar diretamente dentro do cérebro e ver quais valores morais estão escondidos. Para mostrar que as pessoas têm diferentes sentimentos, precisamos contar com evidência indireta, e, em particular, precisamos demonstrar que há fundamentos para disputas morais. Uma disputa fundamental é aquela que persiste mesmo depois de todos os fatos não-morais se apresentarem. Se uma disputa moral negocia com diferentes crenças sobre fatos não morais, então a disputa diminuirá quando os fatos se apresentarem. Uma disputa que persiste uma vez que todos os fatos não-morais são acordados é mais bem explicada ao postular uma diferença em sentimentos morais. O argumento do relativismo descritivo para o relativismo metaético depende de demonstrar que há disputas morais desse tipo, porque tais disputas são a evidência principal para a tese de que o relativismo moral descritivo é verdadeiro. Mas a existência de tais disputas tem sido contestada.

É interessante notar que Hume não foi um relativista. Ele diz que disputas morais fundamentais não emergirão porque todos os seres humanos compartilham os mesmos valores nucleares em virtude de terem uma natureza humana compartilhada. Desacordos morais decorrem de disputas sobre fatos não-morais. Por exemplo, Hume (1751: Apêndice do Diálogo) chama a atenção para o fato de que os atenienses encorajaram relações homossexuais – algo a que ele e seus contemporâneos na Europa se opuseram violentamente. Para explicar essa aparente discordância moral, Hume concorda que os atenienses simplesmente tinham crenças factuais erradas. Em particular, eles acreditavam “absurdamente” que relações com o mesmo sexo poderiam ser uma fonte de “amizade, simpatia, ligação mútua e fidelidade”.

O exemplo de Hume não perturba muitos leitores hoje, mas a estratégia básica continua popular. As pessoas desafiam o relativismo descritivo argumentando que discordâncias aparentes sobre valores acabam por ser

discordâncias em relação a fatos não-morais. No capítulo 3, observei que a escravidão americana foi apoiada por afirmações sobre a inferioridade racial dos africanos. Alguém poderia dizer o mesmo sobre a oposição moral quanto ao sufrágio das mulheres. Os opositores argumentaram que a fraqueza física das mulheres faria delas vulneráveis a ferimentos em locais de votação. Eles também afirmaram que as mulheres tinham níveis mais baixos de inteligência do que os homens, e uma tendência à histeria (Mayor, 1974; Woloch, 2000). Uma vez que a falácia dessas afirmações é revelada, os argumentos contra o sufrágio desmoronam. Em outros casos, debates morais articulam-se com crenças religiosas. Alguns opositores ao aborto acreditam que a vida humana inicia-se na concepção, e eles fundamentam essa afirmação em ensinamentos religiosos. O sacrifício asteca poderia ser explicado com direções semelhantes. Os astecas acreditavam que seu universo teria sido criado quando um deus se jogou em uma grande fogueira e se tornou o sol. Então, outros deuses se sacrificaram para fazer o sol se mover pelo céu. Eles acreditavam que precisavam sacrificar seres humanos para esses deuses a fim de fazer com que o sol não parasse. Se os astecas tivessem esses fatos corrigidos, seus rituais de sacrifício teriam sido banidos. Em cada um dos exemplos, a divergência moral parece derivar de crenças não morais divergentes.

Michele Moody-Adams (1997) demonstrou recentemente que não há um só caso estabelecido de uma disputa moral fundamental. Cada exemplo aparente pode ser explicado em termos de diferenças em crenças não-morais. E isso levanta uma terceira objeção, bastante relacionada ao relativismo descritivo. Moody-Adams argumenta que há barreiras baseadas em princípios para determinar que as disputas morais existem. Se virmos membros de dois grupos culturais fazendo julgamentos morais diferentes sobre o mesmo evento, não podemos ter certeza se eles atribuem um mesmo "significado situacional" para esse evento. Quando consideramos outras culturas, é difícil, se não impossível, entender como eles constroem os fatos. Crenças sobre o mundo são determinadas holisticamente em relação ao conhecimento de teorias cosmológicas e ao conhecimento experiencial acumulado dentro de condições ambientais e culturais que são completamente

desconhecidas para nós. Podemos realmente ver o mundo do ponto de vista dos Astecas, dos Inuítes e dos Tlingit? Se não, não podemos ter certeza de que seus valores morais realmente diferem de nossos próprios.

Essas três objeções levantam sérias dúvidas sobre o relativismo descritivo. Ele é a primeira premissa no argumento central para o relativismo metaético. Se o relativismo descritivo é insustentável, o argumento desmorona. Abordarei uma objeção por vez.

Não acho que diferenças morais entre culturas tenham sido seriamente exageradas. De fato, penso que nossa inabilidade de agrupar nossos próprios valores torna difícil ver como poderiam ser valores diferentes em outras culturas. Esportes romanos sangrentos são um bom exemplo. Considere o que aconteceria se o governo italiano decidisse deixar as pessoas matarem-se umas às outras em público novamente. Haveria um atentado internacional em massa. O cinema violento é um grito contra filmes melosos. E suponha, implausivelmente, que o alarde⁵ retrocedesse, e as pessoas começassem a frequentar os esportes sangrentos com entusiasmo. Isso não significaria que compartilhamos valores com os romanos. Eles tinham pontos de vista morais complexos que foram moldados por suas situações culturais particulares. Os romanos tinham expectativas de vida muito curtas, e eles tinham uma preocupação em expandir seu império por meio de conquistas militares. O valor cardinal em Roma era *virtus*, um ideal que incluía coragem frente à morte violenta. Ter prazer em assistir violência não é equivalente a valorizar a violência do mesmo modo que os romanos o fizeram.

Pontos semelhantes podem ser levantados com relação ao infanticídio inuíte. Se um casal dos Estados Unidos ficasse preso na tundra ártica com seu bebê recém-nascido e um grupo de viajantes adultos, eles poderiam considerar o infanticídio. Mas eles veriam essa decisão com horror. Seria um dilema moral profundo e devastador. Os inuítes têm uma atitude mais causal em relação ao infanticídio (Thomas, 1983). O que consideramos como um mal necessário não seria considerado por eles desse modo. Poderíamos acabar com valores similares

⁵ N.T.: no original, "*outrage*", que indica o sentimento de raiva retrocedendo para dar lugar a uma empatia com a violência deslavada [N.T.]

se tivéssemos perdido gerações em condições árticas, mas, até que tenhamos, nossos valores serão diferentes dos deles.

E quanto à acusação de que valores morais são baseados em debates factuais (não-morais)? Aqui, eu acho que depende do caso. A escravidão americana e o movimento antissufragio foram certamente premissas em algumas afirmações factuais falsas. Mas os argumentos factuais que costumavam defender essas práticas foram frequentemente *post hoc*. Eles foram compostos para racionalizar valores não-rationais que já tinham espaço. O mesmo é verdade para justificativas religiosas. A religião asteca pode ter emergido após os sacrifícios canibais terem começado. Os registros antropológicos mostram que o canibalismo existe em culturas com crenças religiosas variadas.

Em alguns casos, é muito difícil fixar debates morais para crenças factuais. As audiências em confrontos gladiatoriais poderiam ter algumas crenças factuais que divergiam das nossas, mas seu entusiasmo por esportes sangrentos parece ter se baseado em seus valores. Da mesma maneira para atitudes divergentes em relação à bestialidade. Se viéssemos a descobrir que os animais envolvidos apreciavam fazer sexo com humanos, nós não acharíamos a prática menos abominável. A esse respeito, alguns debates morais podem ser como debates sobre gostos. A pessoa que não gosta de chocolate não é vítima de uma crença falsa. Além disso, em geral não há um modo fácil de delinear uma distinção fato/valor em debates morais. Se aqueles que são contrários ao aborto afirmam que um feto é uma pessoa, isso é em parte uma afirmação avaliativa. Como Locke poderia apontar, PESSOA é um conceito forense. Quem você é como uma pessoa depende do que você considera como constitutivo de uma pessoa, e isso é uma questão de decisão como descoberta. Também pode acontecer que o debate não dependa da questão da origem da vida humana ou de quaisquer outros fatos aparentemente não-morais. Em um estudo-piloto, questionei a pessoas se sua postura em relação ao aborto mudaria se elas mudassem seu ponto de vista acerca da questão sobre se um feto é ou não uma pessoa. Quinze por cento disseram que sua postura mudaria, 20% disseram que sua postura "poderia ser afetada", e os outros 65% disseram que sua postura permaneceria a mesma. É como se o valor precedesse as

razões que damos ao defendê-lo; se uma razão falha, nós nos servimos de outra. Assim como com estudos sobre perplexidade moral (Murphy et al., 2000), esses resultados sugerem que as crenças factuais estão frequentemente fazendo um papel *post hoc* na justificação moral.

Isso nos deixa com o desafio Moody-Adams. Como podemos saber quando um debate é avaliativo em vez de factual? Não é sempre possível que as pessoas com diferentes práticas construam eventos de modos bem diferentes? Acho que essa linha de questionamento na verdade apoia o relativismo. Moody-Adams está disposta a conceder o relativismo sobre crenças não-morais. Se as pessoas têm crenças radicalmente diferentes, por que não pensar que elas têm valores diferentes também? A menos que tenhamos razões independentes para pensar que valores são fixos e imunes à impregnação cultural, deveríamos tomar a divergência em crenças não-morais como evidência para a possibilidade de divergência moral. Além disso, há uma razão baseada em princípios para pensar que as moralidades podem divergir. Argumentei que os valores morais são baseados em sentimentos, e sentimentos podem certamente ser influenciados pelo contexto cultural. Podemos ser condicionados a sentirmo-nos bravos em relação às mesmas coisas e culpados em relação a outras.

Alguém pode abordar o desafio de Moody-Adams encontrando casos de debates morais entre grupos culturais que são bastante relacionados. Em tais casos, é difícil defender a afirmação de que os grupos constroem o mundo de formas radicalmente diferentes. Uma boa estratégia é procurar por *subculturas* – pessoas dentro das mesmas fronteiras nacionais, que ainda assim parecem ter uma perspectiva moral diferente. Deixem-me mencionar três exemplos possíveis. Primeiro, há desacordos morais óbvios entre os conservadores e os liberais americanos. Às vezes é sugerido que os liberais e os conservadores têm crenças factuais diferentes (p. ex., sobre o *status* de fetos como pessoas, ou se a economia de subsídios funciona), mas a persistência de debates políticos entre indivíduos inteligentes e bem-informados sugere que poderia haver uma divisão de valor aqui. Retomarei esse caso depois. Um segundo exemplo possível é a divisão por gênero. Homens e mulheres na mesma comunidade geralmente parecem ter valores morais

sutilmente diferentes. Achados relevantes incluem o seguinte: homens estão levemente mais preocupados com justiça do que as mulheres, e essas estão levemente mais preocupadas com cuidados (Jaffee e Hyde, 2000); homens tendem a ser mais a favor da punição violenta do que as mulheres, e essa diferença tem sido ligada a diferentes respostas nos centros emocionais do cérebro (Singer et al., 2006); e mulheres são significativamente mais deontológicas do que os homens em suas respostas a dilemas do bonde (Mikhail, 2002). Poderia ocorrer que homens e mulheres têm crenças diferentes sobre o mundo por causa de suas diferentes experiências, mas essas poderiam inculcar valores morais diferentes.

O terceiro exemplo que eu quero considerar é a violência entre homens brancos no sul dos Estados Unidos (Nisbett e Cohen, 1996). O caso merece atenção extra porque tem sido estudado em detalhes, e é especialmente difícil se fixar em crenças factuais. Homens brancos do sul são mais propensos do que os homens brancos do norte a cometer atos de violência em retaliação a ações que mostram desrespeito: p. ex., um homem flerta com a mulher de outro homem; ou um homem roça em outro homem sem se desculpar; ou um homem deprecia outro homem em público. Nisbett e Cohen argumentam que isso ocorre devido a uma diferença cultural. Brancos do sul aderem à cultura da honra. Eles vivem sob um código de conduta que licencia uma grande mostra de força quando sua honra é ofendida. Os brancos do sul são mais propensos a julgar que a retaliação violenta é apropriada nesses casos do que suas contrapartes do norte. Júris do sul dão sentenças mais leves, por exemplo, a criminosos violentos que estavam defendendo sua honra. Sulistas são também mais tolerantes a punições corporais, autodefesa violenta, posse de armas e penalidade de morte. Há evidência de que esse sistema de valor é implementado emocionalmente. Nisbett e Cohen (1996) mediram hormônios de estresse em sulistas e nortistas logo após eles terem sido roçados por um estranho que não pede desculpas (artisticamente encenado por experimentadores). Os sulistas mostraram um aumento significativo nos hormônios de estresse se comparados a uma condição de controle, e os nortistas não. Nisbett e Cohen especulam que os sulistas herdaram seus valores dos colonizadores escoceses e irlandeses. Esses colonizadores vieram da província de Ulster, da Irlanda, que foi

povoada por nativos irlandeses e por povos de descendência escocesa. Ulster tinha uma economia de pastoreio, e pouca proteção do governo central. Um pastor está sempre vulnerável a ladrões (compare com fazendeiros, cujas fazendas não podem ser roubadas). Como resultado, os pastores desenvolveram uma cultura de honra, em que a autodefesa violenta é considerada apropriada, mesmo em resposta a provocações relativamente inferiores. Sistemas de valor semelhantes foram documentados entre pastores em outras partes do mundo (Edgerton, 1971).

O caso da violência sulista alcança algumas preocupações metodológicas em relação à pesquisa intercultural. Poderíamos não saber o suficiente sobre os astecas para entender como eles constroem o mundo, mas compartilhamos uma quantidade gigantesca com pessoas em suas próprias fronteiras nacionais (sistemas políticos, educação padronizada, entretenimento etc.). Dizer que os nortistas e os sulistas constroem o mundo de formas radicalmente diferentes é exagero. As diferenças parecem ser avaliativas em vez de factuais.

A sugestão de que pode haver discordância moral dentro de uma cultura ganha maior apoio a partir de um experimento simples de pensamento. Imagine duas famílias de filósofos que se tornaram proponentes de teorias normativas diferentes. Uma família poderia criar seu filho para ser um utilitarista convicto, enquanto a família vizinha poderia criar seu filho para ser um deontologista convicto. Quando se deparam com o caso do bonde, as crianças concordariam completamente com os fatos, mas elas discordariam sobre qual o curso de ação estaria correto. Esse caso hipotético mostra que o relativismo descritivo é possível. Há razões, em face dos exemplos pesquisados aqui, para pensar que o relativismo descritivo é também real. É extremamente plausível que crianças no sul e no norte dos Estados Unidos são criadas com valores diferentes. Por exemplo, os pais do sul são mais propensos a usar punições corporais que os seus homólogos do norte, e é muito provável que os sulistas julguem que uma criança de dez anos deveria brigar com um garoto que lhe bateu (Cohen e Nisbett, 1994). As diferenças regionais na educação moral têm uma longa história, e estão documentadas na história do sul pré-guerra (Wyatt-Brown, 1982).

Tal variação na educação moral mina uma das estratégias de Hume para argumentar contra o relativismo descritivo. Ele sugere que as discordâncias morais aparentes devem ser baseadas em divergências sobre fatos não-morais porque todos os seres humanos têm naturalmente os mesmos valores morais fundamentais. Essa afirmação pressupõe que valores centrais derivam da natureza humana. Eu examinarei criticamente a evidência para tal ponto de vista no capítulo 7, mas nós já estamos em posição de refutar o argumento de Hume. Valores morais fundadores que eu tenho chamado de normas fundadoras – normas que não dependem psicologicamente do apelo a outros. Suponha, com Hume, que a natureza humana nos forneceu um estoque de normas fundamentais compartilhadas. Não é garantido que essas normas fundadoras são fixas ou exaustivas. A educação moral funciona por condicionamento emocional, e o condicionamento pode alterar disposições afetivas prévias. Assim como podemos adquirir novos modismos, medos e fantasias por meio do condicionamento, podemos adquirir novas normas fundadoras. Duas pessoas que foram condicionadas diferentemente durante seu desenvolvimento terão valores morais fundamentais diferentes.

Considere o próprio exemplo de Hume da homossexualidade. Ele diz que os atenienses apoiavam relações com o mesmo sexo porque eles acreditavam erradamente que tais relações promoviam a amizade e outras coisas valiosas. Isso não é plausível. Relações homossexuais provavelmente emergiram em Atenas como uma tendência da moda: as pessoas foram condicionadas a pensar que tais relações eram louváveis. Do mesmo modo, a homofobia mordaz de Hume foi provavelmente resultado de condicionamento, não de crenças sobre fatos não-morais. Uma comunidade pode inculcar lentamente atitudes negativas em relação a qualquer comportamento sexual simplesmente expressando aversão quando quer que esse comportamento seja mencionado. Hume repele a homossexualidade de tal forma que ele a descreve apenas por alusão. Minha suposição é de que as pessoas adquirem suas atitudes em relação à homossexualidade antes de elas poderem articular qualquer razão a serem a favor ou contra. Hume menciona que os atenienses podem ter sido induzidos a favor da homossexualidade por causa de

sua simpatia por esportes. Se isso estiver certo, então as disputas morais sobre a homossexualidade são conflitos nas normas fundadoras que têm sido inculcadas culturalmente. Se as normas fundadoras podem ser adquiridas por meio de condicionamento, então pode haver disputas morais que não dependem de crenças factuais diferentes.

Acredito que o caso do relativismo descritivo é esmagador. Se valores fundamentais podem ser adquiridos por meio de condicionamento, como o sentimentalismo construtivo acarreta, então os membros de culturas diferentes podem ter valores fundamentais diferentes. Os debates morais entre tais limites culturais não podem ser resolvidos pelo apelo a fatos não-morais. Tais debates refletem discordâncias fundamentais.

5.2.2 Coerência

Uma das objeções-padrão ao relativismo metaético é que a doutrina é incoerente (p. ex., Lyons, 1976; Cooper, 1978). O problema é fácil de ver se reexaminarmos a definição dada acima. A formulação original bastará:

Relativismo metaético: as condições de verdade de um julgamento moral dependem do contexto no qual o julgamento é formado.

Se minha tese estiver correta, o mesmo julgamento pode ser verdadeiro e falso. Considere o julgamento de que é moralmente aceitável capturar e consumir um habitante da comunidade vizinha. Chame-o de julgamento J. Um asteca pode dizer J verdadeiramente. Posso dizer "Não-J" verdadeiramente. Logo, "J e Não-J" é verdade. Mas "J e Não-J" é uma contradição. Qualquer teoria que acarreta uma contradição é falsa (cf. Lyons, 1976).

Há maneiras óbvias de contornar essa versão do problema de coerência. Na versão do relativismo à qual eu sou a favor, julgamentos morais devem ser avaliados conforme contextos, se eles têm valores de verdade de fato. O julgamento de que é moralmente aceitável capturar e consumir seus vizinhos expressa uma

proposição quando dita por um asteca, e outra proposição quando expressa por mim. Assim, não há contradição quando "J" e "Não-J" são expressas em contextos diferentes. Além do mais, os contextos são determinados pelos valores dos falantes, e nenhum falante concordaria com "J" e "Não-J" ao mesmo tempo; assim, não há contexto em que "J e Não-J" venha a ser verdadeiro.

Os opositores ao relativismo poderiam argumentar que esse método de preservar a coerência na verdade enfraquece o relativismo metaético em vez que preservá-lo. No relativismo do conteúdo isso ocorre, então as mudanças no contexto substituem os significados de termos morais. Quando tal mudança ocorre, não há garantia de que o vocabulário "moral" ainda merece ser chamado de "moral". Por exemplo, não há nada que garanta que a palavra "bom", quando falada por um membro de outra cultura, signifique bom. E, se não significa bom, temos poucas razões em dizer que os membros de outras culturas fazem julgamentos morais. E se eles não os fazem, então apenas nossos julgamentos podem se qualificar como morais, e há apenas uma moralidade, afinal.

Uma versão dessa objeção pode ser adaptada a partir de um trabalho de Donald Davidson (1974). Davidson argumenta contra a ideia de que poderia haver mais de um esquema conceitual, i.e., dois ou mais sistemas conceituais incomensuráveis para descrever o mesmo mundo. O argumento de Davidson é complexo, porque ele examina um número de modos em que os filósofos pensaram sobre a relação esquema-mundo. Aqui está uma capitulação bem simplificada. Suponha que havia dois esquemas, Esquema 1 e Esquema 2. Agora, alguém diz um enunciado no vocabulário do Esquema 2, e nós falamos o vocabulário do Esquema 1. De acordo com Davidson, um enunciado pode ser significativo apenas se tem condições de verdade, e condições de verdade devem ser especificadas usando o Esquema-T de Tarski. Um enunciado, E, é verdadeiro se e somente se P, onde P é uma tradução de E para uma metalinguagem. Assim, um enunciado do Esquema 2 pode ser considerado significativo apenas se pode ser traduzido para uma língua que entendemos. *Ex hypothesi*, o Esquema 2 não pode ser traduzido para o Esquema 1, então nunca podemos ter motivos para pensar que os enunciados do

Esquema 2 são significativos. Eles são traduzíveis para nossa própria língua (por isso, comensuráveis) ou são meros barulhos.

Cooper (1978) baseou-se nas ideias de Davidson para criticar o relativismo moral. Adaptando Cooper, podemos imaginar encontrar uma pessoa que diz que “é moralmente bom quebrar promessas”. Poderíamos inicialmente pensar que essa pessoa tem um esquema moral radicalmente diferente. Mas essa hipótese seria difícil de manter. O fato de que a pessoa aplica “bom” para coisas que consideramos ruins enfraquece qualquer confiança de que poderíamos ter entendido o que a pessoa estava dizendo. Ou “bom” significa qualquer outra coisa que bom no idioleto da pessoa, ou não significa nada. Se Cooper estiver certo, os princípios davidsonianos deveriam nos levar a duvidar da aparente evidência para o relativismo. De maneira mais forte, um esquema moral radicalmente diferente seria ininteligível. Não haveria razão de chama-lo de significativo, muito menos de uma teoria da moralidade.

O argumento davidsoniano contra o relativismo moral depende de duas suposições duvidosas. Primeiro, depende da suposição de que os enunciados da pessoa são significativos apenas se pudermos traduzi-los para o nosso próprio vocabulário. Em outras palavras, um enunciado é significativo se pudermos interpretá-lo. Eu rejeito essa suposição. Já defendi uma tese de como os conceitos adquirem seu significado que não depende da interpretação no sentido de Davidson. Nessa visão, um conceito significa o que quer que tenha sido criado para ser desencadeado. Este não é o lugar para uma defesa completa de Dretske sobre Davidson (veja, por exemplo, Sterelny, 1990). Eu indicarei apenas que o princípio de Davidson – se não pudermos interpretar um enunciado, então ele não tem nenhum significado – constitui uma forma forte de verificacionismo semântico. Isso acarreta, por exemplo, que animais não-humanos não têm estados mentais significativos porque não podemos atribuir crenças a eles usando nosso vocabulário (Davidson, 1975). Essa é uma afirmação bizarra em direito próprio, dado o fato de que o comportamento animal é previsível e, em sentido amplo, inteligível. Contudo, é ainda mais incrível quando aplicado ao caso humano. Suponha que encontremos uma cultura cujos membros usam termos bastante

parecidos com os nossos próprios termos morais, mas eles aplicam esses termos para coisas diferentes do que nós. Se não pudermos encontrar nenhuma tradução adequada (tente traduzir o *virtus* romano), somos forçados a dizer que essas pessoas estão dizendo coisas sem sentido. Então, nos resguardamos da visão de que há muitas moralidades ao propor que muitas culturas fazem barulhos sem significado onde faríamos avaliações morais. Em vez de dizer que essas pessoas têm esquemas conceituais diferentes, Davidson afirmaria que diríamos que eles não têm nenhum esquema. Isso é pior que bizarro; é um insulto.

O argumento davidsoniano contra o relativismo poderia ser defendido por alguém que rejeita a semântica davidsoniana. Alguém poderia defender que qualquer evidência para a diversidade moral é uma evidência ainda mais forte para mal-entendidos. Nessa abordagem, não precisaríamos negar que outras pessoas em outras culturas falam de modo inteligível; precisaríamos apenas questionar nossa compreensão do que eles estão dizendo. Seguindo Davidson, Cooper diz, se alguém usasse termos como “bom” e “ruim” de forma surpreendente, não teríamos razão para pensar que esses termos estavam sendo usados para expressar crenças morais. Essa é a segunda suposição dúbia que eu quero examinar.

Tenho argumentado que conceitos morais são sentimentais. Acreditar que algo é errado é ter certas disposições emocionais em relação a isso. Essas disposições são reconhecíveis. Podemos reconhecer raiva, repugnância e vergonha, por exemplo. Essas emoções nos permitem determinar, com considerável confiança, se um membro de outra cultura tornou alguma coisa moral. Suponha que verifiquemos que uma mulher em uma cultura estrangeira demonstre desprezo por mulheres que não cobrem suas cabeças; suponha que ela mostre vergonha quando a pegamos com sua cabeça descoberta; e suponha que ela demonstre frustração profunda quando sua filha se recusa a manter sua cabeça coberta. Suponha, além disso, que essas reações são comuns em sua cultura, e que há manuais de conduta e textos sacros que incitam as mulheres a manterem suas cabeças cobertas. Se alguém dessa cultura usa a palavra persa “gonâkhâr” para caracterizar mulheres que não cobrem suas cabeças, então temos razão para pensar que “gonâkhâr” é um termo moral. Temos razão para acreditar que ele é semanticamente relacionado à

palavra inglesa "wrong" [errado]. Esse é o tipo de informação que os antropólogos e os lexicógrafos persistem.

Decerto, seria difícil identificar o vocabulário moral em uma cultura se os sentimentos nela não se sobrepusessem com nossos sentimentos em algum grau. Se os membros dessa cultura em questão não pensassem que assassinato aleatório era "gonâkhâr", poderíamos questionar se aquele termo teria significância moral. Mas na maioria dos casos, as diferenças não serão tão extremas. A maioria das culturas sobrepõe em algum grau na aplicação de seus sentimentos. E mesmo quando a sobreposição é mínima, se as pessoas em outras culturas usam termos para expressar sentimentos que são reconhecivelmente similares para nossos sentimentos, temos razão para pensar que esses termos têm significância moral. Reconhecemos o vocabulário moral por sentimentos que denunciam, e podemos reconhecer sentimentos por fisiologia, relatos fenomenológicos, expressões faciais e comportamentos (incluindo punições). Quando sentimentos negativos são expressos em relação a coisas que valorizamos, podemos concluir que algumas pessoas têm valores morais que diferem dos nossos.

Essa resposta ao argumento davidsoniano pode ser explicada mais precisamente pela invocação da distinção entre conteúdo e característica, a qual eu apresentei na discussão de relativismo do conteúdo acima. Quando encontramos membros de outras culturas que aparentemente aplicam termos morais de formas surpreendentes, podemos concluir que esses termos possuem conteúdos diferentes naquelas culturas. Contudo, se os termos expressam sentimentos familiares, podemos também concluir que eles têm a mesma característica que nossos termos morais. Contanto que a característica seja a mesma, podemos seguramente concluir que esses termos têm significância moral quando usados por membros de outras culturas. Se o significado fosse esgotado pelo conteúdo, o argumento davidsoniano teria algum mérito. Não estaríamos autorizados a interpretar termos estrangeiros como termos morais se eles fossem aplicados de modo diferente de nossos termos morais. Com característica, introduzimos um novo nível de comparação semântica, e a afirmativa de que as culturas têm diferentes modalidades torna-se perfeitamente coerente. É equivalente ao ponto de

vista de que as pessoas em culturas diferentes têm sentimentos comparáveis em relação a coisas diferentes. Isso não é mais problemático do que afirmar que duas pessoas podem gostar de diferentes coisas.

A distinção conteúdo/característica nos deixa com uma pergunta. Se encontrarmos uma cultura com valores morais diferentes dos nossos, os termos morais deles são *traduzíveis* para nossos termos morais? A resposta a essa pergunta depende de como definimos tradução. Vamos definir os termos como fracamente intertraduzíveis se eles expressam as mesmas funções do contexto para o conteúdo. A palavra "errado" é uma função de sentimentos negativos e da cultura do falante para o conteúdo. Há palavras nas línguas de culturas moralmente diversas que encontram critérios de tradutibilidade. "Gonâkhâr" poderia ser um exemplo. Vamos definir os termos como fortemente intertraduzíveis se os termos possuírem os mesmos característica e conteúdo quando falados pelos falantes das respectivas línguas. O termo "eu" falha dramaticamente em se adequar a esse critério, porque não tem o mesmo conteúdo em nenhum dos dois dialetos. Os termos morais falham em se adequar esse critério também. Se dirigirmos nossos sentimentos morais de modo diferente de membros de outra cultura, não podemos traduzir seus termos para os nossos a partir desse critério forte. Nossos termos morais podem expressar sentimentos semelhantes, mas eles têm conteúdos diferentes. Se usássemos nosso vocabulário moral para parafrasear sua fala moral, o conteúdo dos enunciados mudaria. Se parafraseássemos seus enunciados usando palavras que têm o mesmo conteúdo, não poderíamos mais usar palavras que expressam nossos sentimentos, então a característica dos enunciados mudaria. Se definíssemos incomensurabilidade como a impossibilidade de uma tradução forte entre dois esquemas, então a incomensurabilidade moral é perfeitamente possível. Não há nada incoerente em relação à ideia de que vocabulários morais podem ser incomensuráveis.

5.2.3 Indexicabilidade

Ao responder ao argumento davidsoniano, apelei ao relativismo do conteúdo. Se essa análise estiver certa, então os termos morais têm significados que devem ser preenchidos pelo contexto. Essas são palavras que claramente funcionam desse modo. Exemplos incluem pronomes pessoais, tais como "eu", "você" e "ele"; pronomes demonstrativos, tais como "isso" e "aquilo"; e termos locativos, como "aqui" e "lá". Ninguém estaria surpreso em ouvir que essas palavras dependem do contexto. Sua indexicalidade é reconhecida por todos os falantes competentes da língua. A afirmação de que termos morais são indexicais é, em comparação, controversa. Seria uma descoberta surpreendente para a maioria dos usuários da língua. Não tem objeção quanto a isso. Os significados podem nos surpreender. Mas, se termos morais realmente são indexicais, então eles deveriam funcionar como casos paradigmáticos de termos indexicais. Palavras como "eu" e "aqui" têm padrões identificáveis de uso. Elas deixam pistas sutis – frequentemente sintáticas por natureza – que revelam sua natureza indexical. Podemos fazer testes para ver se uma palavra é realmente indexical. Os mesmos testes podem ser aplicados a termos morais. Quando isso é feito, contudo, termos morais parecem falhar. Eles não se comportam da mesma forma que alguém esperaria se eles realmente fossem indexicais. Isso põe um desafio ao relativista.

Apresentarei essa objeção apropriando alguns argumentos de Cappelen e Lepore (2005). Esses autores não argumentam diretamente contra o ponto de vista de que julgamentos morais são dependentes de contexto. Eles têm a função de desbancar a dependência de contexto de maneira mais ampla. Considere o enunciado "está chovendo". Alguns filósofos presumem que esse enunciado contém uma referência implícita à localização. Quando dizemos "está chovendo", invariavelmente pretendemos transmitir que está chovendo em algum lugar. A localização não está explicitamente mencionada, mas mesmo assim é parte do significado do enunciado. Alguns proponentes dessa visão dizem que "está chovendo" contém um indexical oculto. A localização está representada na estrutura sintática subjacente do enunciado, mesmo que não apareça na estrutura

de superfície. Cappelen e Lepore oferecem uma bateria de argumentos contra os indexicais ocultos. Eles afirmam que o significado de "está chovendo" não contém elementos que se referem à localização. A informação sobre localização é transmitida por pragmática, e não é parte da sintaxe do significado do enunciado.

Talvez Cappelen e Lepore estejam certos sobre "Está chovendo". Não tenho nenhum compromisso com isso. Mas o argumento deles parece causar problema para o relativismo. Se o relativismo estiver correto, as condições de verdade de afirmativas morais dependem do contexto no qual eles foram enunciados. Eu aprovo o relativismo de conteúdo acima, e os relativistas de conteúdo às vezes assumem que termos sensíveis ao contexto contêm componentes "indexicais ocultos", que fazem referência implícita a coisas como o falante. Se tais componentes indexicais ocultos existem, então alguém poderia esperar que eles fizessem uma contribuição para a sintaxe dos enunciados que os contêm, e sua presença deveria ser detectável usando uma variedade de testes sintáticos. Se um enunciado contendo um termo alegadamente sensível ao contexto falha em demonstrar evidência de indexicais ocultos usando testes sintáticos, então isso pode ser tomado como evidência contra a afirmação de que o termo é sensível ao contexto. Aplicado à moralidade, alguém poderia supor que, se termos morais realmente fossem sensíveis ao contexto, eles revelariam esse fato por meio de testes sintáticos. Cappelen e Lepore propuseram uma variedade de testes, e, na primeira passagem, pode parecer que os termos morais falham. Essas falhas poderiam ser dirigidas como evidência contra uma tese relativista de vocabulário moral.

Há duas estratégias que alguém poderia assumir em resposta a esse tipo de crítica sintática. Primeiro, alguém poderia argumentar que os testes sintáticos em questão não são de fato bons testes para o relativismo. Alguém poderia propor que alguns termos sensíveis ao contexto, incluindo termos morais, não apresentam elementos sintáticos especiais. Ao apresentar o relativismo do conteúdo, eu mencionei essa opção, e isso também é útil para os defensores do relativismo do conteúdo, que argumentam que a sensibilidade ao contexto surge em um nível de avaliação semântica que ocorre após o significado ser atribuído aos termos. Uma

dessas opções poderia funcionar. Tal relativismo é uma tese sobre verdade e referência, não uma tese sobre sintaxe ou forma lógica, então os relativistas não precisam aprovar a visão de indexical oculto. Mas há também uma segunda estratégia para defender o relativismo contra os argumentos sintáticos. Alguém poderia argumentar que os termos morais passam nos testes de indexicalidade oculta. Quero explorar essa possibilidade. Procederei considerando alguns testes que Cappelen e Lepore discutiram em seus esforços de desmerecerem a semântica sensível ao contexto. Seu teste pode ser adaptado para o caso moral, e poderiam parecer enfraquecer a tese de que os termos morais contêm indexicais ocultos. Argumentarei, contudo, que as aparências iniciais são enganosas. Pode ser demonstrado que termos morais passam nos testes. Se esse é o caso, a análise indexical oculta pode se mostrar verdadeira.

Na descrição de indexicais ocultos de termos morais, dizer "canibalismo é errado" é um caminho curto para algo como "canibalismo é errado *de acordo com os valores do avaliador ou de algum outro indivíduo(s) saliente(s)*". "Você não deveria praticar canibalismo" é uma forma curta de dizer "você não deveria praticar canibalismo *de acordo com os valores do avaliador e do receptor*".

Se a descrição do indexical oculto estivesse correta, o componente sintático que se refere ao contexto deveria ser detectável. Cappelen e Lepore sugerem alguns testes, baseados em como indexicais explícitos funcionam, e eles argumentam que casos mencionados de indexicais ocultos falham nesses testes. Adaptarei três de seus argumentos para o caso moral. Eles têm outros testes, além desses três, mas esses podem ser deixados para outra ocasião. O primeiro teste envolve referência anafórica. Com enunciados que contêm indexicais explícitos, alguém poderia acrescentar orações que contêm palavras que referem anaforicamente ao elemento indexical. Considere os exemplos seguintes de Cappelen e Lepore (2005: 76):

(1a) He is a senator. [Ele é senador.]

(1b) He's a senator who likes *himself*. [Ele é um senador que gosta *de si mesmo*.]

(2a) That's a table. [Aquilo é uma mesa.]

(2b) That's a table but *it* is not a book. [Aquilo é uma mesa, mas *ela* não é um livro.]

A palavra "*himself*" [si mesmo] em (1b) refere-se à palavra "*he*" [ele] e seleciona o mesmo indivíduo. Em (2b), a palavra "*it*" [ela] refere ao demonstrativo, "*that*" [aquela]. Se enunciados morais contêm indexicais ocultos, deveria ser possível acrescentar orações que também se referem anaforicamente a eles. Mas quando alguém tenta fazer isso, os enunciados resultam anômalos. Aqui está um exemplo:

(3a) Cannibalism is wrong. [O canibalismo é errado.]

(3b) *Cannibalism is wrong, but they do eat intelligent mammals. [*O canibalismo é errado, mas eles comem mamíferos inteligentes.]

Na visão do indexical oculto, (3a) tem um elemento oculto que pode ser explicado como "de acordo com os valores dos indivíduos salientes". A palavra "*they*" [eles] em (3b) deveria referir anaforicamente a esses indivíduos. O enunciado, contudo, soa peculiar, sugerindo que esse elemento oculto postulado de fato não existe. Enunciados morais não passam no teste de anáfora.

Ou assim parece. Acredito que a falha é meramente aparente. Acima, eu disse que enunciados morais com nenhum contexto explícito são tomados, por *default*, por se referirem ao contexto do avaliador que diz o enunciado. O pronome em terceira pessoa no segundo caso de (3b) soa estranho, porque alguém que ouve a primeira oração daquele enunciado assumiria que o contexto relevante é o contexto do falante. Em outras palavras, a anomalia aqui pode ser semântica, não sintática. Note que o enunciado soa perfeitamente bem se mudarmos a primeira pessoa.

(3c) Cannibalism is wrong, but we do eat intelligent mammals. [O canibalismo é errado, mas nós comemos mamíferos inteligentes.]

Infelizmente, (3c) não prova que "errado" passa no teste da anáfora, porque pronomes em primeira pessoa não podem ser usados anaforicamente. Mas essa

observação sugere que é muito difícil aplicar o teste da anáfora a termos morais. Se termos morais omitem a primeira pessoa quando nenhum parâmetro explícito forem apresentados, e pronomes em primeira pessoa não podem ser usados anaforicamente, o teste pode não funcionar. Talvez o mais perto que podemos chegar é imaginar um contexto não-linguístico, em que é óbvio que o padrão da autorreferência não é operativo. Imagine abrir um livro da área de um antropólogo, e ver (3b) escrito como uma entrada. Nessa situação, o enunciado não parece mais anômalo. É claro que o pronome se refere a algum outro grupo cultural. É claro, é possível que o antropólogo que escreveu o enunciado estava observando seus sujeitos no momento que ele escreveu isso, e o pronome foi apresentado demonstrativamente, não anaforicamente. Se esse fosse o caso, teríamos novamente que concluir que o teste de anáfora é difícil de aplicar no caso moral. Entretanto, quero sugerir que há uma leitura anafórica de (3b) no caso do livro da área de antropologia, que é perfeitamente inteligível. Se eu estiver certo, isso mina a controvérsia de que os enunciados morais falham no teste de anáfora.

Cappelen e Lepore têm outros testes, contudo, que poderiam ser usados para desafiar a tese do indexical oculto. Eles observam que indexicais explícitos podem ser usados para produzir enunciados que podem ser conhecidos *a priori*. Aqui está um exemplo que eles tomaram de Kaplan (ver Cappelen e Lepore, 2005: 78).

(4) I am the person who utters this sentence. [Eu sou a pessoa que diz esse enunciado.]

Se enunciados morais contêm indexicais ocultos, devemos ser capazes de tomarmos banalidades comparáveis ao tornar esses indexicais explícitos. Não é isso o que acontece. O seguinte enunciado parece paradigmaticamente *a posteriori*:

(5) If killing is wrong, then it is against my value system to kill. [Se matar é errado, então é contra o meu sistema de valores matar.]

Intuitivamente, parece que matar poderia ser errado mesmo se não for errado em meu sistema de valores. A incorreção de matar não depende de mim. Suponha,

por exemplo, que eu sou um mafioso, e eu acho que matar é permissível. Nesse caso, (5) seria falso, porque o antecedente seria verdadeiro e o conseqüente, falso.

Esse parece ser um contraexemplo poderoso para a tese do indexical oculto do discurso moral, mas o exemplo é falho. Lembremos que, do ponto de vista que estou defendendo, "errado" pode ser relativizado para o falante *ou* para algum outro indivíduo saliente. O enunciado (5) poderia parecer contingente, ou mesmo falso, porque o termo "errado" pode ser relativizado para alguma outra pessoa além do falante. Se "errado" é relativizado para você, e você condena moralmente o ato de matar, então o antecedente de (5) é verdadeiro, e permanece verdadeiro independentemente de quais sejam os meus valores. Mas suponha que nós ajustemos o contexto. Suponha que estipulamos que (5) é dito em um contexto em que a palavra "errado" é relativizada para o falante. Nesse caso, parece que podemos derivar (5) *a priori*. Por supor que eu sou o falante, e eu condeno matar; então eu diria verdadeiramente que matar é errado, e seria também verdadeiro que matar é contra os meus valores. Ou suponha que eu sou o falante e eu aplaudo o ato de matar; então (relativo a mim) é falso que matar seja errado. De qualquer modo o condicional é verdadeiro. Então, há uma leitura de (5) que pode ser conhecida *a priori*.

Esse exemplo pode ser modificado para mostrar que a visão do discurso moral passa por um teste mais aprofundado para a indexicalidade oculta. Lepore e Cappelen (no prelo) afirmam que podemos gerar contradições com indexicais explícitos que não podemos gerar com alguns indexicais ocultos declarados. Contraste esses dois enunciados:

(6) I am hungry, but the speaker of the sentence is not hungry. [Estou com fome, mas o falante do enunciado não está com fome.]

(7) It's raining, but it is not raining in Chapel Hill. [Está chovendo, mas não está chovendo em Chapel Hill.]

O enunciado (6) é uma contradição flagrante. Na tese do indexical oculto de "*It's raining*" [Está chovendo], o enunciado (7) deveria ser uma contradição também

em contextos em que o parâmetro de localização oculta é preenchido por "Chapel Hill". Nesses contextos, (7) deveria equivaler a (8):

(8) It's raining in Chapel Hill, but it is not raining in Chapel Hill. [Está chovendo em Chapel Hill, mas não está chovendo em Chapel Hill.]

Já que (7) nunca é contraditório, e (8) é sempre contraditório, eles não podem ser equivalentes. Isso vai contra o indexical oculto de "*It's raining*" [Está chovendo] (embora os proponentes dessa tese possam argumentar que (7) simplesmente parece ser não-contraditório, porque fatores pragmáticos excluem o contexto contraditório). Agora contraste o caso da chuva com o caso moral. O enunciado seguinte, que é um contraponto para (5), tem, de fato, uma leitura em que ele é contraditório:

(9) Killing is wrong, but it is not against my value system to kill. [Matar é errado, mas não é contra meu sistema de valores matar.]

O fato de (9) ter uma leitura contraditória sugere que é muito semelhante ao exemplo de indexical explícito em (6), ao contrário de (7) (assumindo que Cappelen e Lepore estejam certos em pensar que (7) nunca é contraditório). Isso é apenas o que a visão do indexical oculto de enunciados morais prediria.

Considerarei mais um argumento baseado em Cappelen e Lepore, que poderia ser levantado contra o relativismo. Com indexicais explícitos, os enunciados do seguinte tipo são perfeitamente aceitáveis:

(10) When John said, "I am hungry", he spoke truly, but I was not hungry. [Quando John disse "Estou com fome", ele falou verdadeiramente, mas eu não estava com fome.]

A palavra "*I*" [eu] muda de conteúdo com o contexto, então a expressão "*I am hungry*" [estou com fome] pode ser afirmada e negada no mesmo enunciado, com a ajuda da citação (ver Cappelen e Lepore, 2005: 106). Os relativistas assumem que

termos morais podem também mudar de conteúdo com o contexto. Logo, deve ser possível tomar enunciados morais aceitáveis que se assemelham a (10). Mas tais enunciados resultam anômalos:

(11) *When Moctezuma said, "You ought to eat prisoners", he spoke truly, but you ought not to eat prisoners. [*Quando Moctezuma disse "Você deveria comer prisioneiros", ele falou verdadeiramente, mas você não deveria comer prisioneiros.]

O enunciado (11) soa muito estranho quando comparado a (10). Isso sugere que termos morais não tomam seu significado do contexto, como o relativismo demanda.

Essa objeção é transponível. Há uma diferença importante entre termos morais e pronomes indexicais. Termos morais têm o que Stevenson (1937) chamou de nível de significado "dinâmico": eles expressam emoções. Quando usamos termos morais em certos contextos, nós não podemos deixar de expressar as emoções implicadas nos conceitos subjacentes a esses termos. Quando expressamos essas emoções, nós damos um aval. Penso que podemos atribuir afirmações morais a outras pessoas sem expressar nossas emoções ou confirmar nossos pontos de vista. Mas quando dizemos que as afirmações dos outros são verdadeiras, transmitimos avais, e não podemos confirmar as visões de moral dos outros sem sentir um compromisso emocional com esses pontos de vista, e isso reintroduz o nível de significado dinâmico. Quando usamos a palavra "verdade" com termos expressivos, nós confirmamos o significado descritivo deles e o seu significado dinâmico. Assim, o problema com (11) é pragmático.

Esse fenômeno não está restrito ao discurso moral. Considere a palavra "yummy" [gostoso], que descreve ao mesmo tempo comida e expressa um sentimento em relação a comida. Minha oposição é que não podemos atribuir a um enunciado que contém a palavra "yummy" aos outros sem confirmar o sentimento. O exemplo seguinte dá suporte a essa afirmação:

(12) *When John said, "The tripe is yummy", he spoke truly, but the tripe is not yummy. [*Quando John disse "As tripas são gostosas", ele falou verdadeiramente, mas as tripas não são gostosas.]

Aos meus ouvidos, (12) soa anômalo. Parece tão ruim quanto (11). Mas isso de forma alguma diminui a alegação de que "yummy" [gostoso] é um termo relativo que toma seu significado completo contextualmente. Eu proponho que (11) e (12) são pragmaticamente anômalos, não semanticamente contraditórios.

Para testar essa interpretação pragmática, precisamos encontrar um modo de cancelar a aprovação emocional que surge quando elogiamos as afirmações expressivas dos outros como verdadeiras. Consideremos o seguinte:

(13) When Moctezuma said, "You ought to eat prisoners", what he said was true in his value system, but I think you ought not to eat prisoners. [Quando Moctezuma disse "Você deveria comer prisioneiros", o que ele disse é verdade em seu sistema de valor, mas eu acho que você não deveria comer prisioneiros.]

Aos meus ouvidos, (13) não soa contraditório. Poderia ser falso, se o relativismo for falso, mas mostrando que requereria uma refutação do relativismo; não há refutação evidente aqui. Assumir que (13) é anômalo demandaria a questão contra o relativista.

Com efeito, (13) torna explícitos os componentes indexicais ocultos postulados. Poderia ser contestado que esse movimento, de fato, contamina a visão do indexical oculto, porque se (13) é apenas uma afirmação explícita do que está implícito em (11), os dois enunciados devem soar igualmente bem. Mas essa objeção perde o ponto que eu estava levantando na minha análise de (11). Quando não tornamos explícitos os componentes indexicais ocultos, pode haver ambiguidades em relação ao contexto de avaliação. O uso da palavra "verdade" soa como uma aprovação e implica que o contexto de avaliação é o próprio sistema de valor do falante. Naquela interpretação, (11) é contraditório. O enunciado (13) não soa contraditório, porque a implicatura (verdade como aprovação avaliativa) é cancelada. De fato, após considerar a interpretação tornada explícita em (13), o

enunciado (11) não parece mais tão ruim. Em uma discussão do relativismo moral, por exemplo, (11) seria perfeitamente apropriado. O relativista poderia usar (11) para expressar a ideia de que outras pessoas têm valores que possuem uma afirmação para a verdade, mas não tente fazer essas coisas em casa. Se tal leitura está disponível, então os termos morais passam nesse teste para a indexicalidade oculta.

A hipótese de que termos morais apresentam indexicais ocultos é consistente com a evidência linguística. Os relativistas não precisam aprovar uma visão de indexical oculto, mas, como acontece, termos morais parecem passar nos testes para a indexicalidade oculta. Esses testes tendem a confirmar o relativismo em vez de refutá-lo.

Não penso que a visão do indexical oculto tenha sido provada por nada que eu disse. Talvez os testes sintáticos sejam inadequados. Mas acho que desfruta, *prima facie*, de apoio suficiente para ser aceito como uma hipótese de trabalho. Leitores perspicazes poderiam se preocupar com isso. A visão do indexical oculto compreende alguns fenômenos linguísticos, mas é ostensivamente difícil adequar uma afirmação que eu fiz a respeito da psicologia do julgamento moral. No capítulo 4, eu chamei a atenção para o fato de algumas pessoas serem relativistas morais, e que outras são objetivistas morais. Mas como pode alguém ser um objetivista se a visão do indexical oculto está certa? Indexicalistas ocultos constroem o relativismo a partir da semântica lexical de termos morais. Usuários competentes deveriam perceber isso, e deveriam resistir ao objetivismo como resultado. O fato de que algumas pessoas parecem ser objetivistas sugere que a tese do indexical oculto está equivocada.

Essa inquietação pode ser respondida de formas diferentes. Uma resposta possível já foi mencionada. Não temos acesso explícito à estrutura de nossas entradas lexicais, então não poderíamos perceber imediatamente que estamos todos implicitamente comprometidos com o relativismo. Outra resposta é ainda mais decisiva, no entanto. A visão do indexical oculto é compatível com o objetivismo. Se "errado" significa errado-para-mim quando eu digo isso, e errado-para-você quando você diz isso, poderia haver uma única moralidade verdadeira. O

que está errado para mim poderia estar errado para você; e o que está errado para ambos poderia ser alguma propriedade independente da mente da realidade. A visão do indexical oculto é neutra quanto ao que os nossos conceitos morais designam no mundo, e, como tal, é compatível com ambas as teorias objetivista e não-objetivista da moralidade. Não é surpreendente que, então, existam diferenças individuais na resposta a questões sobre objetividade moral.

Em resumo, penso que a tese do indexical oculto poderia estar correta. Os argumentos que têm sido usados contra a indexicalidade oculta em outros domínios não se aplicam no caso moral, e o relativismo moral não pode ser desacreditado por testes sintáticos.

5.2.4 Ilegitimidade

Algumas das objeções que são levantadas contra o relativismo são técnicas. Tenho discutido a coerência do relativismo e a evidência sintática para o relativismo. Esses interesses técnicos podem ser respondidos, mas a resistência real ao relativismo deriva de uma preocupação ainda mais profunda. O relativismo é frequentemente considerado como uma doutrina pérfida. Se há moralidades múltiplas, então, parece, não há maneiras de provar que uma moralidade é melhor do que qualquer outra. Se esse é o caso, então a moralidade que adotamos não é privilegiada. Isso tem duas implicações perturbadoras. Primeiro, quando nos engajamos em disputas morais com outros, há um sentido no qual essas disputas são espúrias. Segundo, uma vez que descobrimos que nossas preferências morais não são privilegiadas, nossa confiança nessas preferências é desestabilizada. Por que continuar a adotar nossos valores morais se eles não têm uma afirmação única para a verdade?

Para tornar esse ponto vívido, pense a respeito de valores políticos. Em particular, considere o contraste entre os valores liberais ou conservadores. Como mencionei antes, acho que disputas entre liberais e conservadores não são meramente factuais. Eles derivam de uma diferença incomensurável nos valores morais. Para os conservadores, papéis sociais tradicionais, "papéis familiares",

mercados livres e fé são normas fundadoras. Para os liberais, a igualdade de pessoas indiferente de gênero ou credo, tolerância quanto à diversidade, privacidade e bem-estar social são normas fundadoras.

Lakoff (2002) argumenta que o contraste liberal/conservador deriva de dois estilos de parentalidade distintos. Os liberais recapitulam a abordagem maternal estereotípica, que enfatiza a interdependência e o cuidado. Os conservadores simulam a abordagem autoritária austera tradicionalmente associada a pais. Haidt e Joseph (2004) têm outra consideração sobre o contraste. Eles sugerem que todas as pessoas possuem um grupo de módulos inatos que governam o comportamento moral. Um desses dirige o comportamento que tem a ver com a reciprocidade, e o outro regula hierarquias de dominância. Liberais e conservadores poderiam ser distinguidos por uma diferença no grau para o qual esses módulos dirigem avaliações morais. Essas teorias podem parecer vagas, mas penso que elas podem estar se movendo para a direção certa. É altamente plausível supor que liberais e conservadores adotem normas fundadoras diferentes. De outro modo, os debates entre esses dois lados seriam muito mais fáceis de resolver, e as diferenças diminuiriam assim que os fatos surgissem. Por exemplo, poderíamos esperar que liberais e conservadores convergissem quanto aos pontos de vista relacionados à tributação, uma vez que eles acordam quanto a modelos econômicos. Duvido que tal convergência ocorra. Conservadores valorizam a autoconfiança, então eles se opõem usando taxas para apoiar o estado de bem-estar mesmo se essa for uma política economicamente sólida, e liberais avaliam a igualdade sobre a equidade, então eles tendem a favorecer políticas que envolvem a redistribuição de recursos. Desse modo, o debate sobre a política de tributos nunca termina.

Se isso estiver correto, implicações pérfidas parecem se seguir. Os debates entre liberais e conservadores são realmente espúrios. A mesma política pode estar certa a respeito de valores liberais, mas não está correta quanto a valores conservadores. Mas "certo" e "não certo" não são incompatíveis, porque os termos avaliativos têm conteúdos diferentes em contextos diferentes. Além disso, não há posição transcendental a partir da qual alguém pode decidir qual sistema de valor é melhor. Os sistemas de valor são simplesmente diferentes. Avaliar um sistema

como melhor que outro é fazer um julgamento de valor, e aquele julgamento deve ser feito em relação ao sistema de valor. O liberal dirá que conservadorismo é pior, mas isso é apenas declarar valores liberais. Um liberal intelectualmente honesto teria que admitir que não há posição a partir da qual o conservadorismo é objetivamente pior. Um liberal que chega a essa conclusão deveria considerá-la inquietante. Nós nos identificamos fortemente com nossas visões políticas. Descobrir que nossas visões não são objetivamente melhores que as visões daqueles que se opõem a nós é muito modesto. Isso tira o vigor de nossas convicções.

Isso se torna ainda mais inquietante quando consideramos o fato de que nossos valores políticos podem ser produtos de nossa educação. Liberais e conservadores tendem a pensar que membros da orientação política opositora devem ser perversos ou estúpidos. Os opositores políticos de alguém são ou vítimas de um raciocínio ruim, ou eles simplesmente não se preocupam suficientemente em apoiar as políticas que o bom raciocínio implicaria. Essa atitude em relação a opositores políticos é arrogante e confusa. Poderia acontecer que ambos os lados estejam raciocinando perfeitamente, mas começando com diferentes primeiros princípios. Mais importante, aqueles dois primeiros princípios não são por si só produtos do raciocínio, pelo menos em muitos casos. O fato de que algumas pessoas são liberais e algumas conservadoras é provavelmente explicado melhor por fatos *bibliográficos*. Liberais e conservadores são expostos a influências diferentes. Eles crescem em famílias e comunidades diferentes. Eles têm modelos de papéis e grupos de pares. Eles experienciam eventos de vida diferentes. Esses fatores tendem a moldar ideologias políticas. Suponha que você tenha crescido sob condições diferentes. Suas convicções políticas poderiam ter sido inteiramente diferentes. Esse é o tipo de pensamento que deveria tornar difícil dormir à noite. Por que deveríamos ficar tão mexidos quanto a debates políticos se nossos valores são meramente um resultado de nossa criação?

Falando estritamente, essas implicações não podem ser usadas como um argumento contra o relativismo. Você não pode refutar uma teoria descritiva ao mostrar que ela tem consequências desagradáveis. A verdade é horrível às vezes.

Mas o relativismo pode ser uma tese perniciosa – uma tese que enfraqueceria a moralidade se as pessoas percebessem que é verdadeira. Se o relativismo não pode ser refutado, então, pelo menos, talvez, deveria ser erradicado.

Isso, para mim, é o desafio mais urgente que enfrenta o relativista. A única coisa pior do que uma teoria falsa é uma teoria prejudicial. Felizmente, acho que o relativista pode dizer algo em resposta. Há cinco pontos que eu gostaria de levantar. Primeiro, há consequências positivas de descobrir que nossos valores não são privilegiados. Poderíamos decidir, por exemplo, que temos sido excessivamente precipitados em condenar as pessoas que não compartilham de nossos valores. Em uma palavra, o relativismo pode promover a tolerância. Se nossos valores não são privilegiados, então deveríamos permitir aos outros viver de acordo com valores diferentes. Quando uma comunidade secular permite que uma comunidade Amish continue a viver de modos tradicionais, os membros da comunidade secular estão demonstrando tolerância. Os Amish estão mostrando tolerância pelo secularismo ao não impor seus valores aos outros. Isso parece uma coisa muito boa. Isso é um julgamento de valor, é claro. Nós avaliamos a tolerância, então uma teoria que promove tolerância é boa.

Não é minha intenção implicar que absolutistas morais não possam ser tolerantes. Se há uma única moralidade verdadeira, ainda pode haver dúvidas quanto a se os valores que atualmente temos estão corretos, e essa incerteza abre espaço para tolerância (Rachels, 1986: cap. 2). Mas o absolutismo e o relativismo têm uma relação diferente quanto à tolerância. Os absolutistas toleram outras visões por causa da incerteza epistêmica sobre qual moralidade é correta. Os relativistas são tolerantes porque eles estão certos de que outras visões têm a mesma reivindicação de estarem certas. Notemos que, se um absolutista toma a incerteza seriamente, poderia estar perdendo o viço motivacional. Para ser tolerante, o absolutista deve pensar: “minha moralidade pode estar equivocada”. O relativista pode ser tolerante sem perder a convicção moral: “meus pontos de vista estão corretos, assim como os seus também”. Desse modo, ironicamente, a tolerância relativista poderia ser mais compatível com a convicção moral forte do que a tolerância absolutista. Além disso, quando os absolutistas tornam-se certos

de seus pontos de vista, a intolerância geralmente se segue. Assim, é discutível que o relativismo oferece uma forma mais satisfatória de tolerância, porque não força uma escolha entre a própria dúvida debilitadora e a autoconfiança arrogante. Os relativistas podem tolerar com convicção.

Reconhecidamente, o relativismo não acarreta tolerância. O relativismo pode mesmo ser usado para defender a intolerância. Suponha que você tenha um sentimento de desaprovação em relação à tolerância. Então, é verdade que seu sistema de valor está errado. Alguns relativistas tentaram argumentar que relativistas *não deveriam ser* intolerantes. David Wong (1984), por exemplo, argumenta que não deveríamos interferir nas ações dos outros a menos que possamos justificar essa interferência a eles. Já que sistemas múltiplos têm reivindicações iguais para a verdade, para Wong, não podemos justificar a imposição de nosso sistema sobre os outros. O problema com esse argumento é que o princípio de não-interferência de Wong é moral em caráter. Wong reconhece isso, e ele encontra conforto no fato de que kantianos e utilitaristas permitiriam o princípio de não-interferência. Mas o que kantianos e utilitaristas permitiriam é irrelevante. O princípio de não-interferência poderia ser verdadeiro sobre aqueles sistemas morais, mas essa verdade é relativa. Suponha que alguém acredita sinceramente que nós devemos interferir naqueles que fazem coisas que consideramos repulsivas, mesmo se não pudermos justificar nossa interferência a eles. Se a verdade moral é determinada por aprovação e desaprovação, então a pessoa que aplaude a interferência é moralmente obrigada a interferir.

Todavia, o relativismo torna a intolerância difícil de sustentar psicologicamente. Pessoas que aplaudem uma intervenção normalmente o fazem porque pensam que seus valores são privilegiados. É fácil ser intolerante se você pensar que você está em contato com uma única moralidade verdadeira. Se tal moralidade não existisse, você poderia se ver se tornando mais mente aberta em relação aos valores dos outros. Essa é uma conjectura empírica, é claro, mas eu suspeito que seria confirmada caso fosse testada. Se eu estiver certo, então o relativismo promove tolerância, mesmo que não acarrete tolerância. Isso pode

ajudar a contrabalancear a inquietação de que o relativismo tem implicações horríveis.

Algumas pessoas questionariam se a tolerância é realmente boa. Consideremos um exemplo controverso. Em muitas partes do mundo, mulheres submetem-se a um de vários procedimentos cirúrgicos que os opositores chamam de mutilação genital feminina. Usarei o termo neutro "circuncisão feminina". É estimado que três milhões de mulheres são circuncisadas anualmente (UNICEF, 2005). A prática é realizada em inúmeros países por inúmeros membros de grupos religiosos culturais. Os procedimentos cirúrgicos envolvidos variam consideravelmente. Em alguns casos, a porção exposta do clitóris é removida; em outras, o clitóris inteiro e os lábios vaginais são removidos. Algumas pessoas afirmariam que não deveríamos tolerar nenhum desses procedimentos. Quero examinar essa alegação.

Há alguns mitos sobre a circuncisão feminina. Um mito é que mulheres que se submetem ao procedimento perdem toda a resposta sexual. Isso é falso. Algumas versões do procedimento resultam em nenhuma redução à resposta sexual, como foi medido quanto a interesse em sexo, frequência de sexo, frequência de iniciação do sexo e frequência de orgasmos (Okonofua et al., 2002). Os autores desse estudo chamam a atenção para o fato de que seus achados enfraquecem um argumento padrão *em favor* da circuncisão feminina: os praticantes afirmam que o procedimento reduz o interesse feminino em sexo, e desse modo aumenta as chances de virgindade pré-marital, que eles consideram desejável. Os achados diminuem essa afirmação, mas também diminuem a afirmação feita pelos opositores – uma alegação frequentemente usada contra a circuncisão masculina – que o procedimento diminui o prazer sexual. Isso posto, é importante enfatizar que há versões do procedimento que prejudicam a função sexual. Quando o clitóris inteiro e os lábios vaginais são removidos, a resposta sexual poderia ser diminuída seriamente. Algumas mulheres também têm seus genitais cirurgicamente selados, deixando apenas uma pequena abertura para a urina e para os fluidos menstruais. Esse procedimento pode causar dor significativa, e impede que as mulheres tenham relações sexuais até que a cirurgia seja revertida (o que normalmente acontece com

o casamento). Em todos os procedimentos de circuncisão há também alguns riscos de infecção, e os procedimentos mais severos podem levar a taxas de fertilidade reduzidas (Almroth et al., 2005).

Outro mito a respeito da circuncisão feminina é que os homens impõem o procedimento às mulheres, e as mulheres são vítimas relutantes. Na maioria das culturas em que a circuncisão feminina é realizada, as mulheres evidentemente apoiam e promovem a prática. Frequentemente, as mulheres são aquelas que ensinam as garotas mais jovens sobre a importância da circuncisão, e elas, às vezes, estão envolvidas na realização da cirurgia. As razões para a circuncisão são variadas. Elas frequentemente envolvem a redução do comportamento sexual das mulheres, mas os genitais circuncisados são considerados mais atraivos, repressões são consideradas boas (assim como são em muitas culturas ocidentais), e submeter-se ao procedimento é considerado um direito importante de passagem que confere o *status* de pertencimento ao grupo dentro da comunidade. É possível que as mulheres passem esses valores a suas filhas sob a influência (ou a ameaça) dos homens que as dominam. Meu ponto é que essa é uma suposição que precisa ser examinada e pode variar de cultura para cultura, e caso a caso.

Muitos liberais ocidentais pensam que a circuncisão feminina é categoricamente errada, e têm havido campanhas no mundo todo para eliminar essa prática. Isso pode ser totalmente justificado, mas o relativismo metaético nos convida a examinar tais casos mais de perto. Talvez a campanha para eliminar a circuncisão feminina seja uma imposição de valores ocidentais sobre outras culturas (Tamir, 1996; Shweder, 2000). Muitos membros de outras culturas em questão veem esses esforços como uma ameaça a sua autonomia. Se os valores liberais ocidentais não são privilegiados, então a campanha contra a circuncisão feminina pode estar baseada em um engano (a crença falsa de que a prática é objetivamente errada). Em culturas em que as "vítimas" da circuncisão estão dispostas, e em que eles veem os resultados como superiores aos danos, poderíamos ter que aceitar que nossos valores simplesmente não se aplicam. A história muda dramaticamente em casos em que as "vítimas" estão relutantes, ou veem o procedimento como indesejável. Nessas culturas, as pessoas no poder

estão impondo valores a mulheres que não compartilham esses valores, e a intervenção pode ser apropriada (ver, por exemplo, Nussbaum, 1999: cap. 4).

A circuncisão vaginal é um caso muito difícil. Não quero sugerir que deveríamos abandonar as campanhas para torná-la ilegal. Algumas formas do procedimento são verdadeiramente apavorantes. Talvez nossos valores morais exijam que intervenhamos. O que eu quero sugerir é que o relativismo deveria nos levar a proceder com cautela. Não deveríamos nos iludir ao pensar que estamos ajudando vítimas femininas a conseguir o que elas realmente querem. Deveríamos estar abertos à possibilidade de que esforços para intervir podem estar baseados em premissas falsas. Se nos engajamos nesse exercício difícil de abrir a mente, poderíamos vir a reavaliar a intervenção em alguns casos. Afinal, somos tolerantes a outras práticas culturais que causam danos por causa de ideais culturais de beleza. Escarificação, placas labiais e extensão do pescoço vêm à mente. Também toleramos dietas extremas e subnutrição de mulheres em culturas ocidentais. Talvez um exame cuidadoso da circuncisão vaginal, informada por uma compreensão do relativismo, nos levará a tolerar a cirurgia genital no mesmo grau que toleramos placas labiais.

Levanto a questão da tolerância para mostrar que o relativismo tem seu lado positivo, mesmo que casos como a circuncisão feminina manchem um pouco esse lado. Tolerância é frequentemente muito difícil, mas um comprometimento forte ao relativismo a torna mais fácil. E o relativismo não requer que toleremos tudo. Em casos em que prejuízos são causados a vítimas relutantes, podemos protestar. Podemos raciocinar o seguinte: já que não há uma única moralidade verdadeira, evitarei impor a minha moralidade sobre os outros, porque não tem nenhum argumento contra eles; e do mesmo modo, eu esperaria que eles não impusessem sua moralidade em pessoas que não a compartilham; mas se as vítimas de suas ações não compartilham sua moralidade, então o meu raciocínio inicial para não intervir não se aplica. Os fatores que tornam problemático impor nossos valores sobre aqueles que não adotam esses valores não se aplicam no caso de vítimas relutantes. Essas vítimas relutantes de práticas que condenamos também condenam essas práticas. Quando intervimos em tais casos, estamos impondo

nossos valores sobre os perpetradores dos danos, mas nós não estamos impondo nada às vítimas. Um relativista comprometido pode escolher auxiliar as pessoas que têm valores semelhantes. A questão da circuncisão feminina poderia, em última análise, repousar na discussão do fato de as mulheres que se submetem ao procedimento estarem realmente dispostas a isso. É claro, a crítica poderia argumentar que uma lavagem cerebral é feita nessas mulheres para que considerem a prática aceitável. Mas essa objeção é difícil de defender. Nesse contexto, o termo "lavagem cerebral" é pejorativo para aculturação. Até onde eu sei, a maioria das mulheres que aceitam a circuncisão não sofreu quaisquer métodos coercivos incomuns de doutrinação. Elas adquirem seus valores da mesma forma que nós os adquirimos. Se esse for o caso, então a prática poderia ser tolerada nesses casos. O relativismo promove a tolerância em casos nos quais as "vítimas" estão realmente dispostas. Esse tipo de tolerância que o relativismo promove é uma coisa boa para os nossos padrões.

Quero me voltar a uma segunda observação em resposta ao argumento de que o relativismo é pernicioso. Antes eu disse que o relativismo enfraquece convicções. Num exame mais detalhado, acredito que essa implicação é duplamente exagerada. Os críticos do relativismo superestimam o quanto ruim é perder convicções, e eles superestimam na medida em que as convicções devem ser abandonadas se o relativismo for verdadeiro. O caso da circuncisão feminina ilustra o primeiro ponto. Uma pessoa que se convence de que o relativismo é verdadeiro poderia decidir que sua oposição à circuncisão feminina se baseia em uma forma insustentável de absolutismo moral. Se ela perde a sua convicção de que a circuncisão feminina é categoricamente errada, essa poderia não ser uma consequência terrível.

O segundo ponto pode ser ilustrado com uma analogia. Suponha que alguém é um absolutista quando a gostos. Ele acredita que chocolate é objetiva, universal e absolutamente maravilhoso. Agora imagine que sua crença é corrigida. Ele passa a reconhecer que gosto é relativo para o apreciador, e que, para algumas pessoas, chocolate é desagradável. Trivialmente, a pessoa vai certamente perder sua convicção de que chocolate é absolutamente maravilhoso, mas ela não precisa

abandonar a convicção de chocolate é maravilhoso *para si*. Do mesmo modo, se descobrimos que nossos valores morais não se aplicam a outras culturas, podemos continuar satisfeitos porque eles se aplicam a nós. Poderíamos, falsamente, acreditar que o canibalismo é absolutamente errado, mas não acreditamos que o canibalismo é errado *porque* é absolutamente errado. Para nós, é errado porque o consideramos cruel, repugnante e degradante para a vida humana. Todas essas coisas permanecem verdadeiras, mesmo que os outros não considerem o canibalismo dessa forma.

O ponto é que adotamos nossos valores porque eles são nossos valores. Valorizamos as coisas que valorizamos. A tautologia é a chave para convicção. O fato de que outros não valorizam o que nós valorizamos é inteiramente discutível, a menos que nossa valorização de algo dependa da suposição de que o valor é universal. Acho que é implausível. O crítico que assume que valores devem ser universais para valer a pena tê-los simplesmente levanta a questão contra o relativismo. Questões de gosto são um contraexemplo notório. Psicologicamente, valorar não requer universalizar. Não é óbvio que a valorização moral deveria depender de suposições absolutistas.

Na abertura desta subseção, eu disse que o relativismo tem duas implicações perniciosas: ele enfraquece convicções, e leva a debates espúrios. Eu recém falei sobre a primeira questão. A segunda pode ser tratada construindo-se alguns recursos a partir do capítulo 3. Nele, eu dei algumas explicações para por que as pessoas poderiam se engajar em debates mesmo se o relativismo for verdadeiro. No momento, eu estava ostensivamente interessado em explicar por que debates morais existem, não em mostrar que eles fazem sentido. A preocupação acerca da ilegitimidade centra-se na questão de os debates morais serem justificados. Se o relativismo é verdadeiro, os debates fundamentais não são fundamentalmente confusos? Os oponentes morais não estão apenas falando de coisas diferentes? Acho que não. Como expressivistas têm frequentemente apontado, os debates morais são, em geral, como debates práticos (Stevenson, 1937; Blackburn, 1998). Quando os interesses práticos se chocam, há uma discordância genuína, porque as regras morais ditam as ações. Como observei no

capítulo 3, as pessoas frequentemente têm um interesse investido em trazer os outros para suas visões morais. Se você realmente quer viver em uma sociedade em que as drogas são legais e eu não, então há uma diferença entre nós de que ambos temos um interesse em resolver (lembramos também dos missionários no Peru). Quando temos tais debates, estamos discordando sobre o que fazer. Os relativistas podem também mitigar preocupações sobre a ilegitimidade de prover uma teoria de como a persuasão moral realmente ocorre. As pessoas com diferentes sistemas de valor frequentemente têm valores que se sobrepõem. Então, quando debatemos com pessoas que têm moralidades diferentes, podemos encontrar um fundamento comum. Em debates com outros, podemos tentar usar valores compartilhados para fazer com que os outros repensem esses valores que não compartilhamos. Assim, debates morais não são uma brincadeira confusa; os opositores não estão falando apenas de coisas diferentes, ou tergiversando sobre como usar palavras morais. Em vez disso, eles estão tentando alterar o comportamento prático um do outro ao identificar fins compartilhados. Nesse cenário, seria ilusório dizer que debates morais são completamente espúrios. Esse é o meu terceiro ponto em defesa contra a objeção da ilegitimidade.

Meu quarto ponto diz respeito à afirmação de que o relativismo coloca todos os valores políticos em posição igual. Isso poderia não ser verdade. O relativismo é incompatível com qualquer ideologia política que dependa do argumento de que há uma moralidade verdadeira. Nem todas as ideologias requerem isso. Como ponto de partida, considere a perspectiva política que diz: deixe todos fazerem o que eles querem, desde que não ofendam os que querem outra coisa. A cláusula de violação é subordinada à cláusula de liberdade; prevenir a violação é uma forma de permitir que as pessoas façam o que elas querem. Permitir que as pessoas façam o que elas querem é uma forma de converter o relativismo relacionado a valores em perspectiva política.

Se alguém compara o conservantismo político com o liberalismo político, o liberalismo poderia sair na frente. Liberais, como conservadores, tomam certos valores como sendo fundamentais, evidentes em si mesmos e inalienáveis. Isso pode ser problemático. Mas os liberais também tendem a promover a tolerância em

um grau maior do que os conservadores. Eles se aproximam da visão política que incorpora o relativismo. A visão relativista diz que deveríamos proteger a privacidade, a liberdade de expressão e a diversidade – valores que são mais centrais para o pensamento liberal do que para o conservador. O relativismo poderia ser também compatível com o pensamento libertário, com o isolacionismo, e com outras perspectivas políticas. Meu ponto não é que qualquer visão política possa ser derivada do relativismo. O ponto é que o relativismo não é politicamente neutro. Certas políticas, formas de governo e ideologias podem ser facilmente reconciliadas com a hipótese de que não há moralidade absoluta. Outros são mais difíceis de conciliar com essa hipótese. Aqueles que tentam impor seus valores sob a suposição de que esses têm argumento exclusivo para a verdade estão cometendo um erro. Uma vez que nós aceitamos que nossos valores são limitados, podemos continuar a ter incentivos para impor nossos valores sobre outros, mas aqueles que fazem tal imposição em nome de uma verdade absoluta deveriam se sentir compelidos a corrigir seus modos uma vez que aceitam o relativismo.

Contra tudo isso, alguém poderia imaginar o que dizer sobre as pessoas que impõem seus valores aos outros *sem* assumir que há uma única moralidade verdadeira. Elas apenas querem que seus valores sejam divulgados, e não se importam que outros valores têm uma afirmação igual para verdade. O relativismo não tem recursos para convencer tais imperialistas avaliativos de que eles estão cometendo um erro. Eu disse que o relativismo não acarreta tolerância; apenas a promove psicologicamente. Mas o relativista poderia dizer qualquer coisa sobre por que deveríamos preferir a visão tolerante à intolerante, quando a segunda não se baseia em ilusões falsas de verdade universal?

Acredito que o relativista pode dizer isso tudo. Se você impuser seus valores aos outros, sabendo que seus valores não têm grande afirmação para a verdade, então você deve ter um valor moral de acordo com qual imposição seria aceitável. Você deve acreditar, por exemplo, que tem o direito de impor suas preferências sobre as pessoas que não compartilham essas preferências. Reconhecidamente, há provavelmente algumas pessoas que têm valores como esses, mas, deduzo, muitos de nós não as têm. Então, a afirmação de que o relativismo promove a tolerância é

restrita àqueles que já têm um valor moral de que impor preferências sobre os outros é ruim se as preferências não têm argumentos para a verdade. Essa restrição enfraquece o movimento do relativismo para a tolerância, mas não o diminui. Se a imposição das preferências é algo que as pessoas consideram amplamente como errado, então a descoberta de que o relativismo é verdadeiro pode promover a tolerância amplamente. Suspeito que os liberais e os conservadores tipicamente concordam que é errado impor as preferências sobre os outros se essas preferências não têm pretensão de verdade privilegiada, e, assim, o relativismo pode ter implicações para selecionar, entre as perspectivas liberal e conservadora, certas questões de política, p. ex., questões pertencentes à privacidade.

Tenho discutido a preocupação de que o relativismo moral acarreta indiferença política. Alguns valores políticos são mais justificáveis que outros. O ponto sobre valores políticos pode ser generalizado. Algumas moralidades são melhores que outras. Ao apresentar inquietações sobre as implicações do relativismo, eu disse que não há uma posição transcendental a partir da qual podemos decidir entre sistemas morais concorrentes. Acredito que esse argumento é parcialmente verdadeiro e parcialmente falso. Penso que não há uma posição transcendental a partir da qual alguém pode determinar qual das duas moralidades é *moralmente* melhor. Mas pode haver uma postura transcendental a partir da qual alguém pode ponderar sobre fundamentos não morais. Não vou explicitar a sugestão aqui, porque essa será abordada no capítulo 8. Se eu estiver certo, essa é a quinta razão para pensar que as implicações do relativismo não são tão terríveis quanto elas possam parecer inicialmente.

Para concluir, acho que o relativismo é uma doutrina coerente que goza de amplo suporte da antropologia e que poderia até deixar marcas na sintaxe de termos morais. Não se trata de uma doutrina que pode nos levar a reconsiderar algumas de nossas convicções morais, mas não é insidiosa. Os relativistas podem continuar a se importar com a moralidade.

Ao formular esses argumentos, eu tive a oportunidade de descrever alguns dos diferentes valores que podem ser encontrados em grupos culturais ao redor do

mundo. No capítulo 6, eu investigarei a origem das morais, delineando as fontes das diferenças culturais que eu discuti aqui. Essa exploração genealógica reforçará o caso para a diversidade moral, porque mostrará alguns dos mecanismos que podem levar sistemas morais a desenvolverem-se de diferentes formas. Mas os críticos poderiam pensar que estou exacerbando a diversidade, e eles podem olhar para a biologia para encontrar alguns universais morais. No capítulo 7, levantarei algumas dúvidas sobre esse projeto. A questão referente à origem das morais é interessante por si mesma. Trata-se de uma instância do debate natureza-educação. Mas está também relacionada a alguns dos temas que eu trouxe neste capítulo. Minhas indagações sobre as origens das morais exporão ameaças mais profundas à confiança moral. Deveríamos continuar a moralizar se descobríssemos que os valores morais são produtos de eventos históricos passados? No capítulo 8, eu argumentarei que tais descobertas poderiam ser úteis para atingir progresso moral.

6. A GENEALOGIA DA MORAL

De onde provêm os valores morais? Esta pergunta não é feita com bastante frequência. Os filósofos estão geralmente mais interessados em questões normativas, do que em questões descritivas. A questão acerca da proveniência de nossos valores é considerada irrelevante. Aqueles que investigam as origens da moralidade são por vezes acusados de cometer uma falácia genética. Diz-se que a origem de uma crença ou de um valor não tem relação com sua verdade. Penso que essa suposição precisa ser reconsiderada. Teorias normativas são muitas vezes concebidas para abarcar intuições atuais. Se essas intuições têm origens desprezíveis, talvez então elas não devam orientar projetos normativos. Esse é um tema central na crítica de Nietzsche à moralidade. Nietzsche procurou desestabilizar os nossos valores, ao expor seu passado. Como um sentimentalista construtivo, penso que a moralidade é criada por nós; como um relativista, penso que diferentes sociedades criam diferentes moralidades, em diferentes condições históricas. Neste capítulo, eu examino esse processo.

Tenho dois objetivos. Quero mostrar que o método genealógico pode ser usado eficazmente para investigar a origem dos valores. Isso ajuda a confirmar que algumas convicções morais são produtos da história social. As próprias análises históricas de Nietzsche eram especulativas, apaixonantes e, provavelmente, profundamente equivocadas. Genealogias mais plausíveis foram propostas por antropólogos e historiadores. Eu considerarei alguns exemplos. Meu segundo objetivo é determinar se a genealogia pode ser utilizada para suportar o ceticismo sobre valores morais atuais. Nietzsche certamente empregou a genealogia desse modo, mas penso que seu ceticismo é exagerado. Nietzsche também defendeu a visão de que devemos substituir os valores historicamente derivados por valores que são mais naturais. Avaliarei esse aspecto de seu programa no capítulo 7, quando tratarei da ética evolucionista.

6.1 NIETZSCHE SOBRE A GENEALOGIA

Na Alemanha do século XIX, a história estava muito em voga, e não é nenhuma surpresa que filósofos alemães tenham começado por enfatizar temas históricos em suas obras. O interesse de Nietzsche pela temática histórica aparece em toda sua obra e culmina em *A Genealogia da Moral* (1887). Para entender a moralidade, argumenta Nietzsche, devemos sondar a origem dos valores morais. O projeto não foi inédito. Nietzsche foi influenciado pela obra de seu amigo íntimo Paul Rée, que havia publicado um livro sobre *A Origem dos Sentimentos Morais* (1877). As ideias de Rée eram uma mistura de filosofia moral britânica, especialmente Hobbes e John Stuart Mill, e de evolução darwiniana. Nietzsche refere-se a Rée na *Genealogia* como um psicólogo britânico, apesar do fato de que Rée era um judeu pomerânio, que nunca tinha vivido na Inglaterra. De acordo com Rée, termos morais, como “bom” e “mau”, começaram como marcas para resultados positivos e negativos. Mais especificamente, os termos referem ao que é bom e ruim para a sobrevivência da espécie. Eventualmente, através da associação e do hábito, esses termos passaram a referir às ações que causam consequências positivas e negativas, ao invés das consequências propriamente ditas. Começamos a pensar na crueldade como má, por exemplo, ao invés de reconhecer que realmente são os efeitos da crueldade que contam como mau, no sentido original do termo.

Nietzsche rejeitou a explicação de Rée, e concebeu somente uma história de sua autoria. Como Rée, Nietzsche começa com uma análise do que se poderia chamar de “bom natural” e “mau natural” – o tipo de bem e mal que seria favorecido pela natureza humana, não adulterada pela história e pela cultura. Rée estava convencido de que os princípios darwinistas favoreceram o comportamento que beneficia a outros. Antecedendo os modernos eticistas evolucionários, ele considerou os interesses egoístas como compatíveis com as preocupações altruístas. Nietzsche pensava que os princípios darwinistas forneciam um quadro menos otimista do bem e do mal naturais. Está igualmente em nossa natureza buscar o poder e dominar os que são mais fracos. Nietzsche acreditava que esses

bens naturais foram cultuados na moralidade da Roma antiga. De acordo com Nietzsche, alguns dos valores que Rée enfatizou, incluindo o autossacrifício, não eram de modo algum naturais. Em vez disso, eles eram resultado de uma luta pelo poder no mundo antigo.

Em Roma, os cristãos eram oprimidos. Eles viviam na pobreza, e ressentiam-se de seus opressores romanos (Nietzsche usa a palavra francesa *ressentiment*, que também expressa sentimentos de ódio). Para lidar com sua situação, os cristãos começaram a demonizar os valores de seus opressores. Eles condenaram o poder e a dominação, assim como a riqueza, a liberdade e a saúde. Chamaram essas coisas de más. Eles também começaram a ver seu próprio estado de abatimento como bom, e fizeram da pobreza, da fraqueza e da doença uma virtude. Esses valores – celebrados no Sermão da Montanha – foram revolucionários. Eles inverteram os ideais romanos, transformando o bem romano em mal [*evil*], e o mau [*bad*] romano em bom. Nietzsche chamou isso de revolta dos escravos na moralidade. Quando os cristãos chegaram ao poder, o seu sistema moral firmou-se longa e duradouramente. Nietzsche achava que os valores da Europa do século XIX eram vestígios do ressentimento dos primeiros cristãos. Ele imaginou uma nova aurora em que os europeus de espírito livre restabeleceriam os valores do tipo que foram celebrados em Roma. No *Zarathustra*, Nietzsche constrói uma visão utópica, na qual os fracos “desaparecem”, e os fortes afirmam sua vontade de poder (Nietzsche, 1885: 30, 58).

A filosofia moral de Nietzsche é inquietante. Suas invectivas contra o cristianismo são exageradas, sua apoteose dos valores “aristocráticos” é ofensiva, e sua invocação da “besta louca” prenuncia a ascensão do nacional-socialismo na Alemanha. Na realidade, Nietzsche desprezava os antissemitas de sua época, e ele teria sido repellido pela mentalidade de rebanho dos carrascos de Hitler. Com efeito, a moralidade nazista é passível de uma crítica nietzscheana. Os nazistas foram impelidos pelo ressentimento da Europa após o tratado de Versalhes, e seu antissemitismo foi, provavelmente, uma consequência das táticas de conversão utilizadas pela Igreja primitiva. São Paulo tentou atrair pagãos convertidos, ao decretar que os cristãos não precisavam seguir a lei judaica, e São João tentou

demonizar os judeus, culpando-os pela crucificação. A substituição dos costumes judaicos levou a uma demonização dos judeus não convertidos no mundo cristão, e a campanha de difamação transformou-os em bodes expiatórios, nos tempos de adversidade. O resultado foi dois milênios de cruzadas, inquisições e pogroms. Nietzsche teria condenado os nazistas pela apropriação de uma forma de fanatismo que era realmente um vestígio de uma antiga luta pelo poder, confundindo-o com uma verdade científica. Infelizmente, os nazistas equivocadamente reverenciaram Nietzsche (junto com Platão), como um pai ideológico de seu movimento político. Isso, às vezes, desvia das características filosoficamente interessantes da *Genealogia* de Nietzsche.

As lições mais importantes de Nietzsche podem ser resumidas no que se segue. Primeiro, os valores que atualmente estimamos têm uma história. Em segundo lugar, essa história pode não ser bela. Ela pode não refletir uma progressão racional em direção a ideias mais verdadeiras ou mais benéficas. Em vez disso, essa história da moral, como a história em geral, envolve com frequência lutas pelo poder e motivos psicológicos questionáveis (ganância, ressentimento, xenofobia, para citar alguns). Em terceiro lugar, nossa cegueira em relação a essa história nos dá uma falsa sensação de segurança em nossos valores. Consideramos incontestável nossa perspectiva moral. Em quarto lugar, podemos fazer melhor. Podemos mudar radicalmente a moralidade, e adotar um sistema de valores que tem vantagens sobre os valores que fomos culturalmente levados a aceitar. Nietzsche não é um niilista moral. Ele julga que nós precisamos de uma nova moralidade – de um novo sistema de normas.

O otimismo de Nietzsche acerca do progresso moral baseia-se em dois pressupostos. O primeiro é que, quando descobirmos a historicidade de nossos valores, teremos razão e habilidade para rejeitá-los. Em segundo lugar, Nietzsche pensa que pode substituir valores historicamente construídos por valores que são, em algum sentido, naturais. Avaliarei essa ideia no capítulo 7.

Argumentarei que Nietzsche tem razão em relação à historicidade da moral, mas que tanto seu pessimismo acerca dos valores existentes quanto seu otimismo acerca dos valores futuros são equivocados.

6.2 NIETZSCHE NATURALIZADO

6.2.1 O cristianismo segundo Nietzsche

A consideração de Nietzsche sobre a origem dos valores cristãos é altamente especulativa e provavelmente equivocada. Por exemplo, o ascetismo cristão, que expressa a visão de que a pobreza é uma virtude, pode ter suas raízes em tradições filosóficas antigas na Grécia, como o estoicismo e o cinismo, bem como em movimentos religiosos antigos. Havia cultos ascéticos no Egito antigo, e o gnosticismo, que provavelmente surgiu pouco antes do alvorecer do cristianismo, foi um influente movimento que exigia a castidade e a renúncia do corpo. Acredita-se que o gnosticismo teve suas origens na Pérsia, e é improvável que seu ascetismo tenha derivado da opressão romana (Meeks, 1993).

A preocupação cristã com a caridade pode ter tido suas origens na lei judaica (ver, por exemplo, Deuteronômio 15:11). Os judeus podem ter enfatizado a caridade como resultado de opressão, mas não há nenhuma razão para pensar que o ressentimento ou a apoteose da pobreza foram um fator determinante. Com exceção de algumas seitas isoladas, os antigos judeus não eram ascetas, e a caridade é uma estratégia muito sensata para lidar com a opressão. Do ponto de vista do jogo teórico, a partilha de recursos entre membros de um grupo coeso pode fortalecer o grupo como um todo. Essa é uma explanação mais simples do que a proposta de Nietzsche sobre o ressentimento.

Essas considerações levam-me a concluir que os primeiros cristãos não atingiram a sua visão moral através do ressentimento. Sua ideologia apropriou-se de concepções anteriormente existentes. Evidentemente, Nietzsche sabia que a caridade e os movimentos ascéticos antecederam o cristianismo. Ele pode, no entanto, argumentar que o ressentimento levou os primeiros cristãos a se apropriarem desses valores, ao invés de outros. Os cristãos pobres podem ter se identificado com os ascetas, e eles podem ter visto a solicitude judaica para com os necessitados como uma marca distintiva de seus opressores romanos. A dificuldade inerente a essa hipótese é que os primeiros cristãos podem não ter sido

empobrecidos. Stark (1996) argumenta de modo convincente que o movimento foi impulsionado principalmente pela classe média e até mesmo pelos ricos.

Stark (1996) observa que os novos movimentos religiosos tendem a difundir-se por indivíduos financeiramente assegurados, que se desencantaram com sua crença anterior. Os registros sugerem que muitos dos primeiros convertidos eram financeiramente bem providos, e alguns foram mesmo membros da classe senatorial de Roma. Eles não foram vítimas da pobreza ou da opressão. De qualquer modo, a pobreza e o ressentimento não podem explicar como o cristianismo se difundiu. O cristianismo cresceu numa velocidade enorme nos três primeiros séculos, provavelmente a uma taxa de 40% por década. As pessoas haviam sido oprimidas em todo o Império Romano, mas nenhuma religião dos oprimidos havia desalojado a fé pagã do Estado. Stark acredita que a opressão tinha pouco a ver com isso. Outros fatores parecem mais plausíveis. Em primeiro lugar, Stark (1996) observa que houve duas grandes pragas nos primeiros séculos do cristianismo, que dizimaram boa parte da população romana. Isso poderia ter suscitado o desencanto com a religião romana. O cristianismo oferecia uma fonte de conforto, ao enfatizar a vida no doce além. Ademais, as práticas cristãs de cuidar dos necessitados (herdadas do judaísmo) podem ter resultado em maiores taxas de sobrevivência durante surtos de doença. Isso poderia ter aumentado a população cristã, além de conquistar novos convertidos.

Outro fator importante, argumenta Stark, é que o cristianismo oferece melhores perspectivas para as mulheres do que o paganismo. As mulheres pagãs com frequência casavam por volta dos 12 anos, e pouco se fazia para protegê-las dos machos galanteadores e do divórcio. As mulheres cristãs casavam com mais idade, e nem o divórcio, nem a promiscuidade eram permitidos. Assim, as mulheres estavam mais seguras. Além disso, o cristianismo proibia o aborto, que era, na época, muito perigoso para as mulheres e, muitas vezes, imposto pelos homens; também proibia o infanticídio, que era um método comum em Roma, de matar bebês indesejados do sexo feminino. O resultado é que as mulheres teriam sido atraídas pelo cristianismo, e, depois de convertidas, elas estariam muito mais propensas a manter a prole feminina. Isso é muito significativo, visto que as

mulheres desempenham uma função desproporcional na propagação da religião. Se as mulheres cristãs casavam com homens pagãos, elas podem ter conseguido convertê-los, ou criar seus filhos como cristãos. Stark mostra que as mulheres foram autorizadas a exercer funções eclesiásticas na Igreja primitiva, e ele especula que isso acrescentou uma atração a mais às mulheres, e mais um meio pelo qual elas poderiam propagar a fé.

Se a explicação de Stark está no caminho certo, então o cristianismo se propagou amplamente, porque dava poder às mulheres descontentes da classe média. Não era um caso em que pessoas pobres oprimidas transformavam suas fraquezas em virtudes. Ao contrário, era um caso de pessoas em confortável situação financeira, adotando um sistema de valores que permitia maior poder e liberdade. Dificilmente se trata ali de uma revolta de escravos na moralidade. Assim que o cristianismo tornou-se religião do Estado, as perspectivas para as mulheres na Igreja parecem ter diminuído. Por exemplo, as mulheres já não podiam exercer importantes funções eclesiásticas. Podemos assumir seguramente que esta transição no poder, das mulheres da classe média à dominação masculina, não foi impulsionado pelos destituídos. Os homens que assumiram o controle eram os maridos de classe média das mulheres que difundiram o cristianismo, e membros da elite regente de Roma.

Mesmo que a genealogia dos valores cristãos de Nietzsche esteja equivocada, os princípios básicos de sua abordagem podem ser defendidos. Os valores morais, sem dúvida, foram influenciados por fatores históricos, e esses fatores nem sempre são racionais ou nobres. Devemos considerar cada valor moral como um artefato cultural, com uma história à espera de ser descoberta. Ao invés de especular sobre a história da moral, podemos recorrer aos recursos da ciência social. A antropologia histórica é de valor especial. Os antropólogos geralmente se envolvem com algo parecido à genealogia nietzscheana, mas eles apoiam em dados suas histórias acerca da origem. Ilustrarei isso com alguns estudos de caso. Mas, em primeiro lugar, farei algumas observações sobre os mecanismos de transmissão cultural.

Nietzsche enfatizou as lutas pelo poder e os fatores psicológicos: revolução e ressentimento. O ressentimento não é um princípio *geral* de transmissão cultural; como acabei de sugerir, ele pode não ter sido especialmente importante para a revolução cristã na moralidade. A ideia de uma luta pelo poder é mais genérica. Mas as lutas pelo poder não lançam luz sobre os motivos pelos quais os valores são passados de geração para geração. Além disso, alguns valores surgem lentamente ao longo do tempo. Seria proveitoso começar com um quadro mais geral para pensar acerca da genealogia da moral, antes de se concentrar nos fatores que podem ter sido operatórios em casos particulares.

6.2.2 Transmissão cultural

Adotarei a concepção de que a transmissão cultural é uma função de aptidão (em relação a isso, confira especialmente Boyd e Richerson, 1985; Sperber, 1996). Precisamos identificar os fatores que tornam os produtos culturais especialmente apropriados para a difusão e para a transmissão. Concentrar-me-ei em três fatores. Estes não são, necessariamente, os únicos fatores que contribuem para a transmissão cultural, mas eles estão provavelmente entre os mais efetivos. A probabilidade de que uma crença valorativa seja culturalmente transmissível aumenta quando:

- (i) produz benefício material para os seus crentes ou para os membros de uma cultura que estão em situação de doutrinar outros;
- (ii) está situada num contexto narrativo que é fácil de aprender porque, por exemplo, ela se integra com as crenças existentes acerca da natureza do mundo ou capta a imaginação; ou
- (iii) ela tem um apelo emocional, devido ao conteúdo intrínseco da crença ou práticas concomitantes, como o condicionamento emocional ou os rituais religiosos emocionalmente intensos.

Referir-me-ei a eles como fatores materiais, narrativos e afetivos, respectivamente. Cada um deles aumenta a aptidão de uma maneira diferente. Fatores materiais estão mais intimamente vinculados à ideia de aptidão, como é usada na biologia evolutiva. Um elemento cultural, que tem benefícios materiais, pode aumentar as perspectivas de sobrevivência e de procriação. Uma crença materialmente benéfica pode levar a um aumento do número de pessoas na população que mantêm essa crença. Conhecimento tecnológico e médico são exemplos óbvios. As culturas transmitem conhecimento de como cultivar e de como tratar feridas. Mas os valores também podem ter custos e benefícios materiais. Considerem-se dois casos fora do domínio moral. A cobiça por ouro desempenhou um papel importante na expansão geográfica da sociedade europeia; ela motivou a colonização europeia do Novo Mundo e a expansão na fronteira norte-americana. Esta foi boa para os colonizadores e ruim para os povos indígenas, que foram mortos ou desalojados. Em contraste, a cobiça por seda pode ter contribuído para a queda do Império Romano. Roma acumulou um enorme déficit comercial com a China, o que tornou mais difícil financiar conquistas militares.

Por fatores narrativos, pretendo referir aos contextos ricos, coesos e verbais, nos quais os valores são, por vezes, apresentados. As narrativas podem assumir muitas formas: elas podem ser histórias (como no caso da Bíblia judaico-cristã), teorias científicas (como na justificação pseudocientífica do racismo), contos morais ou mitos. Há um número de razões pelas quais as narrativas podem facilitar a transmissão do valor. Em primeiro lugar, as pessoas são melhores em histórias de memorização do que em listas aleatórias de sentenças (por exemplo, Black e Bern, 1981). Em segundo lugar, as narrativas podem conter elementos surpreendentes que chamam a atenção e aumentam a memória. Boyer e Ramble (2001) descobriram que indivíduos na França, Gabão, Nepal eram todos melhores em recordar histórias em que expectativas físicas são violadas (por exemplo, um ser que pode andar através de paredes) ao invés de histórias em que nada de sobrenatural acontece. Boyer e Ramble concluíram que a transmissão de crenças religiosas é facilitada quando as mesmas são apresentadas em histórias

contraintuitivas. Os valores morais são transmitidos frequentemente através de mitos com elementos sobrenaturais, e isso provavelmente aumenta as possibilidades de que os valores sejam partilhados com os outros. Em terceiro lugar, os mitos e as histórias cativam as crianças, o que acarreta um processamento mais profundo e maior probabilidade de evocação e repetição. Em quarto lugar, as narrativas podem facilitar a integração das informações. Ao situar um valor em uma história complexa, o mesmo pode ser vinculado a um amplo conjunto de crenças, incluindo sistemas religiosos, cosmologias e teorias científicas. Se um valor está ligado a sistemas de crença mais amplos, pode ser difícil abandoná-lo. A revisão de crenças é um processo holístico. Abandonar um valor fundado em um sistema de crenças pode exigir o abandono de partes vastas desse sistema, o que não seria epistemologicamente permitido. Em quinto lugar, as narrativas podem contribuir para a internalização de valores, ao fornecer uma fonte de construção da identidade. As pessoas tendem a pensar em si mesmas em enredos coesos, que podem incluir informações autobiográficas, assim como informações sobre o passado histórico. Pode ser mais fácil adotar um conjunto de valores como meu, por exemplo, se eles forem apresentados como tendo sido descobertos ou dados a meu povo, em um passado distante. Em cada um desses modos, as narrativas aumentam a probabilidade de retenção ao longo do tempo.

Os fatores afetivos também facilitam a memória. As emoções aumentam a atenção, e a atenção aumenta a memória (por exemplo, MacKay *et al.* 2004). As emoções também podem funcionar como recompensa. O afeto positivo é compensatório, e um elemento cultural que induz a emoções positivas é, portanto, transmissível. Até mesmo as emoções negativas podem incrementar a transmissão. Por exemplo, se as violações de regras forem punidas, o medo de punição pode reforçar a conformidade à regra, e esse efeito pode proliferar por gerações. Emoções morais, como culpa, podem ter o mesmo efeito. Os sentimentos morais são ferramentas poderosas para a transmissão cultural, porque eles imputam sanções emocionais a violações de regras, não importa quem violou a regra (eu ou um outro; estranho ou parente). Em um estudo empírico muito inteligente, Nichols (2002) demonstrou que as normas emocionalmente

implementadas têm maior potencial para a transmissão social. Nichols compilou uma lista de regras elencadas no manual de etiqueta muito influente de Erasmo (1530), *De civilitate morum puerilium*. Ele então procurou um codificador para indicar qual dessas regras envolve uma proibição contra algo que naturalmente produziria aversão (por exemplo, normas envolvendo fluidos corporais). Outro codificador foi proposto para indicar que as regras de Erasmo ainda hoje são operantes. Essas duas medidas mostraram-se altamente correlatas. As normas reguladoras de comportamentos que causam repulsa são especialmente susceptíveis de serem suportadas.

Os fatores que produzem a transmissão cultural com frequência operam em conjunto. Mitos e rituais podem ser envolventes, e partilhar crenças pode promover sentimentos de afiliação e solidariedade. A solidariedade pode aumentar os benefícios materiais ao incrementar a cooperação, e benefícios materiais podem aumentar o afeto positivo. Com frequência, uma única crença é reforçada por múltiplos fatores. Considere a ideia cristã de paraíso. Essa crença se encaixa numa cosmologia coerente, que dá sentido ao sofrimento mundano, e traz uma recompensa emocional através da esperança. Como sugerido acima, a mensagem otimista pode ter sido atraente em tempos de peste, e pode ter tornado as pessoas mais suscetíveis para ajudar os necessitados, ao invés de fugir. Isso, por sua vez, pode ter aumentado as possibilidades de sobrevivência. Assim, a crença no paraíso pode ser narrativamente atrativa, emocionalmente sedutora e materialmente benéfica.

Invoquei a noção de aptidão, ao elencar fatores que contribuem para a transmissão cultural. "*Fitness*" é um termo darwinista (introduzido por Erasmus Darwin, e que se tornou famoso através de Charles). Ele sugere uma analogia entre a transmissão cultural e a seleção natural. Alguns autores tentaram obter muita coisa com essa analogia. Eles argumentam que os elementos culturais (crenças, práticas, instituições, artefatos, etc.) são transmitidos de modo muito semelhante ao qual os genes são transmitidos. Dawkins (1976) introduziu o termo "*meme*" para referir a esses elementos. Como os genes, os *memes* surgem através de mutações em culturas pré-existentes, eles se replicam, e sua capacidade de se propagar

depende de sua aptidão. Não obstante, eu hesito em chamar isso de um modelo evolutivo de transmissão de crenças.

Há algumas diferenças importantes entre os genes e os assim chamados *memes* (ver Sperber, 2000; Boyd e Richerson, 2001). Em primeiro lugar, à diferença da evolução biológica por seleção natural, as crenças não são necessariamente formadas por algo como mutação ao acaso. Elas podem ser formadas intencionalmente. Em segundo lugar, não está claro que a noção de replicação se aplica. A transmissão de crenças não é uma questão de réplica mecânica perfeita. As crenças são muitas vezes transmitidas intencionalmente. A mudança de crenças pode ser intencional e abrupta. Desse modo, há uma boa margem de descontinuidade no fluxo e refluxo de ideias. Em terceiro lugar, a transmissão cultural frequentemente é tanto horizontal quanto vertical. Os elementos culturais podem ser transmitidos para os indivíduos na mesma geração. A transmissão genética é tipicamente vertical (a transmissão viral de genes é uma exceção). Em quarto lugar, não há nenhuma equivalência óbvia para a distinção entre genótipo e fenótipo na transmissão cultural. Na evolução biológica, os genes operam em conjunto com o ambiente, para criar traços fenotípicos. Crenças avaliativas culturalmente transmitidas não são necessariamente os mecanismos pelos quais alguma outra coisa é produzida. Eles não são as causas subjacentes de características mais superficiais. Dennett (1995) sugere que *Romeu e Julieta* e *West Side Story* têm fenótipos semânticos diferentes e o mesmo genótipo semântico subjacente (ou "*memótipo*"), mas a analogia está longe de ser perfeita. A peça de Shakespeare foi fielmente reproduzida (por meio de várias impressões), e autores posteriores tiveram de interagir com o texto, a fim de imitá-lo. Assim, poder-se-ia muito bem dizer que a peça é semelhante a um genótipo, e a trama a um fenótipo.

Por essas razões, as analogias evolutivas não devem ser levadas muito longe. Falar de *memes* não auxilia muito. É importante relatar, no entanto, paralelos vagos entre a evolução biológica e a transmissão cultural. A propagação de entidades culturais depende, em certo sentido, do êxito. As entidades culturais devem ser transferíveis; elas são especialmente propensas a proliferar, se forem materialmente benéficas ou psicologicamente atraentes. As crenças que são

acessíveis, estimulantes ou proveitosas têm uma vantagem sobre aquelas que são misteriosas, maçantes ou destrutivas. O mesmo vale para as instituições e as práticas. Devemos reconhecer o papel da aptidão na transmissão cultural, sem forçar muito a analogia com a transmissão genética. Ao discutir a aptidão de atitudes avaliativas, não quero assumir a bagagem teórica da *memética*. Quero apenas dizer que podemos compreender por que certas atitudes acontecem, por que elas permanecem, e por que outras tendem a desaparecer.

Com esse quadro, podemos conseguir agora alguma perspectiva sobre os construtos explanatórios de Nietzsche. O ressentimento é um fator afetivo, e, como tal, ele pode conduzir à transmissão cultural. Se Nietzsche estava certo, no sentido de que os primeiros cristãos ressentiam-se dos romanos, isso ajudaria a explicar a propagação do cristianismo. Suspeito que outras emoções, assim como fatores narrativos e materiais, foram mais importantes (ver abaixo). Nietzsche também menciona as lutas pelo poder. Estas não são mencionadas na minha lista de fatores, porque são mecanismos de mudança, em vez de fatores de transmissão. Dito isso, penso que Nietzsche está certo em ressaltar a conexão entre poder e moralidade. Podemos citar várias razões.

Antes de mais nada, os que detêm o poder estão numa posição especialmente privilegiada, quando se trata de transmissão cultural. Há uma generalizada "inclinação ao prestígio": somos mais propensos a imitar as pessoas que são poderosas ou bem sucedidas. Pessoas poderosas são salientes, e imitá-las de modo cognitivamente simples pode aumentar as perspectivas para o nosso próprio sucesso (Henrich e Gil-White, 2001). Em segundo lugar, as pessoas com poder têm controle sobre as instituições utilizadas na transmissão cultural. Os poderosos controlam instituições religiosas, por exemplo, e estas são a fonte de cosmologias enraizadas localmente. Em terceiro lugar, faz sentido afirmar que todas as crenças culturalmente transmitidas *constituem* uma forma de poder. As crenças culturais desempenham um importante papel na estruturação de nossas sociedades e na organização de nossas vidas. Esta é uma ideia nietzscheana, mas foi exposta com mais força por Foucault, que equiparou o conhecimento ao poder (ver principalmente, Foucault, 1977). Não vou me debruçar sobre esses temas, mas

eles estarão em segundo plano. Minha abordagem da genealogia pode diferir da de Nietzsche em ênfase, mas os paralelos não seriam difíceis de extrapolar.

6.2.3 Por que as pessoas comem umas às outras?

A fim de ilustrar o método genealógico naturalizado, concentrar-me-ei em dois exemplos: o canibalismo e o casamento. As análises que eu vou rever são controversas, mas elas ajudam a estabelecer um ponto metodológico amplo: a história dos valores humanos pode ser investigada empiricamente. Nietzsche pode ser naturalizado.

Começemos com as atitudes em relação ao canibalismo. Poucas práticas nos horrorizam tanto quanto esta. A nossos olhos, o canibalismo não é apenas errado; é monstruoso. Mas essa atitude nem sempre foi a norma. No capítulo 5, observei que o canibalismo pode ter sido uma prática humana padrão. Antes da emergência das sociedades modernas, organizadas em estados, o canibalismo pode ter se difundido bastante. Até bem recentemente, foram registradas práticas canibais em sociedades pré-estatais em todo o globo. Na África, são exemplos de sociedades que se presume que tenham praticado o canibalismo os Mangbetu, Abadja e os Mayi-Mayi, tribos do atual Congo. Incluem-se como exemplos da América do Norte os Iroqueses, Hurões, Algonquinos, Kwakiutl e, talvez, os Anasazi. Exemplos documentados na América do Sul incluem os Akamara, Guaiaca, Tupinambás, Pano e os Yanomami. Existiram tribos em Papua Nova Guiné conhecidas por praticarem o canibalismo, incluindo os Fore, Baroi, Bimin-Kuskusmin e os Doboduras. Acredita-se que os habitantes das ilhas Fiji, na Polinésia, tenham sido canibais. Na Nova Zelândia, os Maoris têm a reputação de ter uma história de canibalismo. Peggy Reeves Sanday (1986) estima que o canibalismo é muito comum. Em uma amostra representativa de mais de 100 sociedades bem pesquisadas, abrangendo um intervalo de tempo de quase 4.000 anos, ela descobriu que 34 por cento delas se envolveram em práticas canibais.

Alguns céticos argumentam que as narrativas de canibalismo em terras estrangeiras são apenas histórias de marinheiros, concebidas para desumanizar os

membros de outras culturas (Arens, 1979). Embora concorde que algumas narrações de canibalismo devem ser vistas com suspeita, é mais insultante para outras culturas o ceticismo em relação ao canibalismo do que sua negação. A ânsia de negar a existência de canibalismo resulta de nossa própria convicção hegemônica, de que o canibalismo é inconcebivelmente terrível. As evidências de que o canibalismo tem sido amplamente praticado são extremamente fortes. Há relatos convergentes de testemunhas oculares, de múltiplas fontes críveis, incluindo depoimentos de praticantes (Harris, 1985). Há também evidências físicas, incluindo vestígios de sangue humano encontrados em recintos de templos, ossos humanos carbonizados em potes de cozimento, ossos humanos encontrados cortados da mesma maneira que ossos de animais usados como alimento, e tecido humano encontrado em excremento humano (DeGusta, 1999; Turner e Turner, 1999; Marlar *et al*, 2000; Walker, 2001). Um pouco mais controverso: há evidências de que as pessoas, em todo o mundo, têm um gene que evoluiu para nos proteger contra os riscos de saúde associados ao comer membros da nossa própria espécie (Mead *et al.*, 2003). Com vasta evidência testemunhal e com evidências físicas corroborantes, o ceticismo acerca do canibalismo parece infundado. Penso que devemos chegar a um acordo em relação ao fato de que o canibalismo era difundido, e temos de explicar esse fato surpreendente.

Há documentos evidenciando que o canibalismo assumiu diferentes formas. Algumas sociedades praticaram o *endocanibalismo mortuário*, ou seja, comiam parentes mortos. Muitas outras sociedades, no entanto, praticaram o *exocanibalismo guerreiro*, em que membros de outros grupos sociais são comidos depois de terem sido mortos em batalha. O canibalismo mortuário pode ser relativamente fácil de explicar. Consumir o corpo de um antepassado pode ser uma forma de lidar com o trauma da perda. É uma maneira de prolongar simbolicamente a vida dos mortos, e de transferir seus poderes vitais. O canibalismo mortuário é também uma maneira de lidar com o problema da destinação dos cadáveres. Ele elimina resíduos potencialmente perigosos. Mas o que dizer do canibalismo guerreiro?

Marvin Harris (1985: cap. 10) oferece uma explicação convincentemente simples, em termos de custos e benefícios. Vamos supor que duas tribos vizinhas entrem num conflito violento por causa de recursos. Uma tribo ganha a batalha. Vários membros da outra tribo são mortos, e muitos outros são capturados. Consideremos primeiro os mortos. Os vencedores poderiam deixar os mortos para trás, ou poderiam comê-los. A segunda opção muitas vezes faz sentido. Deixar os corpos para trás é um desperdício – não traz nenhum benefício. A carne humana é nutritiva. Se os vencedores tiverem de percorrer uma distância até o local da batalha, eles poderiam cortar a carne das vítimas para a viagem de regresso ou poderiam fazer uma festa tonificante de vitória no campo de batalha. Qualquer dessas opções é melhor do que simplesmente deixar a carne comestível para trás. Agora, o que acontece com os prisioneiros? Nesse caso, existem três opções. Eles poderiam liberar os cativos, permitindo-lhes buscar vingança em uma ocasião posterior. Os vencedores também poderiam escravizar seus prisioneiros, mas então teriam que alimentá-los e impedi-los de se rebelar. Os benefícios de manter escravos superam os custos apenas quando os mesmos podem ser usados para produzir rendimentos significativos, através da agricultura ou de outras tarefas. Se os vencedores não têm uma agricultura avançada, a escravidão é uma má ideia. Finalmente, eles podem matar os prisioneiros. Esta é a melhor opção para sociedades não agrárias; uma vez que os prisioneiros foram mortos, os vencedores confrontam-se com a escolha inicial: consumir os mortos ou deixá-los apodrecer. O consumo é, de novo, a escolha sensata. Além disso, os vencedores podem fazer os cativos marcharem de volta para sua aldeia e evitar o trabalho de carregar os seus corpos. Eles podem matá-los na aldeia e dividir a carne com os moradores que não puderam participar da batalha. Isso faz sentido, do ponto de vista econômico. Também pode ter um sentido estratégico. Se um grupo de pessoas é conhecido por serem canibais, os grupos vizinhos podem ser mais relutantes em envolver-se em guerra contra eles. Isso é especialmente verdadeiro se as vítimas dos canibais são sacrificadas de uma maneira dolorosa e ostensiva. Não é incomum nas sociedades canibais a tortura das vítimas e o estabelecimento de rituais elaborados em torno do consumo de carne.

O relato de Harris acerca do que leva as pessoas a praticar canibalismo resulta em previsões-chave. Ele prevê que as sociedades estarão propensas a envolver-se em canibalismo se forem suficientemente organizadas para se envolver em guerra, se elas têm motivos independentes para ir à guerra, e se têm infraestrutura insuficiente para tirar proveito dos escravos. As tribos [*chiefdoms*] têm uma boa chance para serem canibais, especialmente se há recursos limitados e competição. Os Estados, no entanto, não têm propensão para serem canibais. Normalmente, os Estados desenvolvem a agricultura, de modo a se beneficiar da posse de escravos. Mais importante ainda, os Estados são suficientemente organizados para ter sistemas de coleta de impostos ou de tributos por vastas áreas geográficas. Quando uma sociedade chega a esse ponto de desenvolvimento, há pouco incentivo para entrar em guerra com os povos vizinhos. Os vizinhos podem ser tributados em troca de proteção ou de outros recursos. Enquanto os vizinhos são muito fracos para promover a guerra contra o Estado, é melhor, em termos de custo-benefício, taxá-los do que conquistá-los e comê-los. Os contribuintes rendem mais benefícios do que os cadáveres. Naturalmente, vizinhos não têm propensão para serem complacentes, se eles acham que há um alto risco de virem a ser conquistados e comidos. Assim, o relato de Harris prevê que o canibalismo não apenas diminuirá com o surgimento do Estado; o canibalismo será efetivamente condenado. Na verdade, ele se tornará um tabu. Caso contrário, as pessoas podem ser tentadas a comer seus vizinhos. Harris argumenta que o tabu em relação a comer pessoas é como o tabu em relação a comer carne de cavalo. Historicamente, cavalos eram mais úteis como instrumentos de guerra do que como alimento. Comer um cavalo era atentar à segurança nacional. Isso pode ser esquecido num acesso de fome, naturalmente. Assim, sociedades que dependiam de cavalos para as campanhas militares estabeleceram tabus para desencorajar o consumo de carne de cavalo a bel-prazer. Para as sociedades avançadas, do mesmo modo, as pessoas em aldeias vizinhas são mais úteis como contribuintes do que como refeições. Ao tornar o comer pessoas um tabu, as sociedades aditaram uma sanção emocional contra aqueles que, de outro modo, poderiam ceder a suas tentações culinárias assassinas.

A hipótese de que o canibalismo desaparece com o surgimento do Estado coaduna-se bem com o registro antropológico (ver Sanday, 1986). Há, no entanto, um contraexemplo gritante: os astecas. Os astecas (que chamavam a si mesmos de *Mexica* ou *Tenochca*) tiveram um estado altamente organizado, com poderosas alianças políticas, e uma capacidade de coletar tributos e negociar com vizinhos. São essas as características que deveriam ter diminuído a probabilidade de canibalismo. Ainda assim, os astecas mantiveram as práticas canibais mais amplas de que se tem registro. Isso contrasta com a hipótese de Harris. Para resolver o enigma, Harris invoca o trabalho de Michael Harner (1977). Harner credita o canibalismo asteca à nutrição. Ao contrário de muitos outros estados organizados, os astecas não tinham uma agricultura bem desenvolvida. Eles cultivavam milho, feijão, tomate, pimenta e outros alimentos, mas eles não tinham arados ou animais de carga. Eles podem ter tido, assim, acesso limitado a alimentos com quantidades elevadas de proteína e de gordura. Eles obtinham proteína comendo insetos, répteis, aves aquáticas, e espirulina, mas eles não tinham animais domesticados. Os astecas também passaram por longos períodos de seca e fome. Em tais condições, pedaços de carne humana podem ter sido um suplemento dietético bem-vindo. A carne dos animais foi certamente considerada um prêmio valioso para os astecas; seria uma vergonha desperdiçar a carne humana, obtida fortuitamente em batalhas nos territórios vizinhos. Os cativos escravizados depois das batalhas não tinham valor econômico, visto que, sem técnicas de agricultura de alto rendimento, os custos para manter os escravos superavam os benefícios. Harner não supõe que os astecas consumiam carne humana o suficiente para uma diferença nutricional significativa, mas ele supõe que os astecas tiveram pouco incentivo para eliminar o canibalismo.

Os críticos de Harner e Harris argumentam que a dieta asteca era suficientemente rica para tornar supérfluo o canibalismo (Ortiz de Montellano, 1978; Sahlins, 1978, Garn, 1979). Eles também notam que as vítimas sacrificiais astecas eram comidas por nobres, por guerreiros exitosos, e por suas famílias. Não havia distribuição de carne humana para as massas. Além disso, as vítimas eram comidas durante os tempos de colheita, quando eram realizados festivais e quando

havia oferta de alimentos suficiente para ir às batalhas. As vítimas não foram comidas durante os períodos de grande escassez. Essas objeções, estritamente falando, não solapam a hipótese da nutrição, porque Harner e Harris não afirmam que o canibalismo era necessário. Eles simplesmente afirmam que os benefícios podem ter compensado os custos. Mas os críticos podem ponderar com uma queixa mais séria. Se o canibalismo fosse apenas uma questão de conseguir carne, por que ele seria tão ritualizado? Por que os astecas simplesmente não caçavam seres humanos como carne de caça e não consumiam carne sem cerimônia?

É possível reconstruir uma resposta a essa pergunta a partir da consideração materialista de Harris. Talvez os rituais elaborados de sacrifício foram concebidos para aterrorizar as pessoas que não eram leais ao império asteca. Harris prediz que os impérios não se envolverão em canibalismo, porque eles podem cobrar tributos de povos conquistados, mas algumas pessoas podem resistir à conquista. Se os astecas não dispunham de recursos militares para policiar territórios vizinhos, então eles poderiam ter incentivado a lealdade através das ameaças de sacrifício brutal. Ingham (1984) argumenta que esta é uma possibilidade plausível, porque os astecas não possuíam bons meios de transporte, exceto andar a pé. Enviar guerreiros às aldeias distantes seria muito mais caro do que assustar seus habitantes, submetendo-os através de rituais espetacularmente brutais.

Essa consideração da dominação política é plausível, mas incompleta. Ela implica que o sacrifício nada mais era do que uma demonstração de poder. Isso explica a função de sacrifício, mas não sua psicologia. Não há nenhuma razão para pensar que os astecas que participaram de rituais de sacrifício construíram suas práticas com base nesses termos estratégicos. É muito mais provável que o sacrifício possuísse significado religioso.

A mitologia asteca é banhada em sangue (Sanday, 1986; Carrasco, 1999). Um mito conta a história de uma deusa-monstro chamada Tlaltecuhli, que foi rasgada ao meio por dois outros deuses, Quetzalcoatl e Tezcatlipoca. Metade do seu corpo tornou-se os céus e a outra metade tornou-se a Terra. Acreditava-se que Tlaltecuhli serviria de alimento para a Terra somente se nutrida com uma dieta regular de corações humanos. Dizia-se que Quetzalcoatl e Tezcatlipoca estavam

em guerra constante um com o outro; suas batalhas conduziram à criação e destruição de quatro sóis. Cada sol corresponde a uma era no tempo, e cada uma delas terminou em cataclismo. Os astecas acreditavam que estavam vivendo na era do quinto sol, que também acabaria em desastre. A quinta era foi criada quando um deus menor se atirou no fogo e se tornou um quinto sol, chamado Tonatiuh. No início, Tonatiuh não podia mover-se através do céu, mas, então, muitos outros deuses imolaram a si mesmos, o que lhe conferiu o poder de se mover. Toda noite, no entanto, ele seria destruído, e depois recriado de manhã. A cabeça de Tonatiuh aparece no centro da pedra do calendário asteca, com uma língua de fora na forma de uma faca sacrificial de obsidiana. Os astecas acreditavam que as pessoas voltavam à vida na quinta era, quando Quetzalcoatl derramou o seu sangue sobre os ossos das pessoas que viveram em eras anteriores. Outra figura de destaque no panteão asteca era o deus da guerra, Huitzilopochtli, que foi algumas vezes visto como uma manifestação de Tonatiuh. Huitzilopochtli foi criado quando sua mãe Coatlicue foi fecundada por uma bola de penas que caíam do céu. Coatlicue tinha 400 filhos anteriores, que se sentiram desonrados pela gravidez milagrosa de sua mãe, e, por isso, decidiram matá-la. Quando eles atacaram, Huitzilopochtli irrompeu de dentro do ventre de sua mãe e destruiu a todos.

Esses mitos têm um tema comum. Todos envolvem a destruição de deuses pelos próprios deuses, e ciclos de vingança. Mais importante ainda, a mitologia asteca promoveu um sentido de extrema vulnerabilidade: há uma batalha das forças divinas em curso, que acabará por destruir o mundo. Os astecas evidentemente acreditavam que eram necessários sacrifícios para repelir a ruína inevitável. Sahlins (1978) e Sanday (1986) argumentam que o canibalismo deve ser entendido como parte desta estrutura narrativa elaborada.

As análises nutricionais e políticas do canibalismo não contribuem em nada para iluminar seu significado religioso. A análise nutricional é provavelmente melhor considerada como uma explicação de como o sacrifício ficou em primeiro plano, mas não porque ele persistiu. Os astecas começaram como uma tribo pobre e nômade, que se envolveu em conflitos militares com outros grupos. O canibalismo pode ter começado nessa fase pré-estatal, por razões nutricionais. Mas Harris e

Harner não explicam como ele se ritualizou. Uma possibilidade é que ele tem a ver com a psicologia da guerra. A guerra é carregada emocionalmente, e os guerreiros têm de passar por um treinamento e condicionamento psicológico. Emoções intensas, atividades repetidas, e situações de vida ou morte são a fonte do ritual religioso (Burkert, 1983). Depois de uma vitória militar, os ancestrais nômades dos astecas provavelmente sentiram alívio, alegria e gratidão. Se os cativos foram usados para consumo, eles provavelmente iriam ser comidos no contexto de uma celebração da vitória, a qual, em si mesma, é uma espécie de ritual. Essa celebração poderia marcar o sucesso do grupo em ter escapado da destruição através do sucesso militar. A cosmologia asteca provavelmente emergiu em tais condições. Provavelmente, foi atribuído um significado mítico aos rituais associados à vitória.

Uma vez criados, os mitos e rituais podem ter vida própria. Matar ritualmente um prisioneiro de guerra é uma experiência intensa e emocionalmente carregada. Rituais que têm um alto nível de intensidade emocional ou sensorial são suscetíveis de serem lembrados e repetidos (Whitehouse, 2000; McCauley e Lawson, 2002). A pompa do sacrifício asteca aumentou ao longo do tempo. Os sacerdotes, vestindo por vezes peles humanas, exibiam os corações de suas vítimas e então rolavam seus corpos por degraus de grandes pirâmides. Os mitos em torno do sacrifício provavelmente aumentaram a probabilidade de transmissão cultural. A religião asteca alimentou uma grande ansiedade em relação ao futuro – ansiedade apoiada em secas periódicas, furacões e outros desastres naturais. O canibalismo ritual visava impedir a ruína. Isso ajuda a explicar, juntamente com os benefícios políticos, por que a prática persistiu, mesmo depois de passado o período em que ela teria fornecido uma vantagem nutricional significativa.

A religião asteca é, em alguns aspectos, o oposto do cristianismo. Os deuses eram, muitas vezes, vistos como inimigos, ao invés de amigos; agradar aos deuses era uma maneira de evitar a catástrofe, e não um modo de entrar no paraíso. Em ambas as religiões, os seres humanos foram tratados como beneficiários de um sacrifício divino, mas, para os astecas, esse favor precisava ser pago com sacrifícios humanos; para os cristãos, no entanto, o sacrifício humano era proibido. Em outras palavras, a religião asteca prega uma mensagem de ansiedade, ao passo

que o cristianismo, muitas vezes, prega uma mensagem de esperança. Tanto a esperança quanto a ansiedade são instrumentos eficazes de transmissão.

Ao discutir os astecas, eu invoquei os fatores de transmissão cultural que descrevi acima. Em primeiro lugar, invoquei fatores materiais. Provavelmente, o canibalismo asteca surgiu porque era vantajoso do ponto de vista nutricional. Os astecas eram guerreiros nômades antes de constituir um estado, e o canibalismo pode ser a prática persistente para as sociedades com esse perfil. O canibalismo subsistiu mesmo depois da transição para organização estatal, porque conferia benefícios políticos, e os astecas não tinham o tipo de agricultura com boa relação custo-benefício para manter os prisioneiros de guerra como escravos por muito tempo. Mas talvez esses fatores materiais não tenham sido expressamente reconhecidos pelo povo asteca. Provavelmente, os astecas não viam o canibalismo simplesmente como um modo eficiente para pressionar aldeias conquistadas a pagar tributos. É mais provável que os astecas interpretaram essa prática ritual em seu contexto cosmológico; essas crenças religiosas, juntamente com as intensas emoções evocadas por rituais canibais pródigos foram provavelmente o principal meio através do qual a prática passou de geração em geração. O canibalismo foi sustentado por um conjunto de indutores de ansiedade próprio das crenças religiosas e por rituais emocionalmente evocativos. Esse coquetel de fatores materiais, crenças religiosas e emoção é uma força potente na formação, transmissão e manutenção de valores.

6.2.4 O Casamento e a Igreja Cristã

O caso do canibalismo ilustra a viabilidade de um programa genealógico empiricamente fundamentado. Nós podemos fazer especulações informadas sobre os fatores que propiciaram o surgimento, o êxito e o desaparecimento final de valores especificamente morais. Quero agora passar do canibalismo ao parentesco. O canibalismo é um valor moral estranho. Nós não mais o encontraremos como permitido. Na verdade, o canibalismo foi suprimido com tanto sucesso, de modo que raramente ocorre; os pensamentos sobre o canibalismo não exercem um papel

ativo regulador em nossas vidas. Com o parentesco é diferente. Nossas vidas são estruturadas em torno de normas operacionais, que regem as relações familiares. A genealogia de parentesco pode ser usada para lançar luz sobre valores que são centrais em nossas práticas cotidianas. Para ilustrar isso, quero considerar a genealogia dos valores ocidentais modernos referentes ao casamento. Os ocidentais acreditam que o casamento tem de ser monogâmico. A poligamia não é apenas ilegal no Ocidente; ela é considerada imoral. Os ocidentais também se opõem ao casamento entre parentes, mesmo além da família mais próxima. O casamento com um primo de primeiro grau é visto com repulsa. Os ocidentais também acreditam que o casamento deve ser consensual: ele deveria basear-se no amor ao invés de ser arranjado.

Esses valores não são universais. A monogamia é de fato muito menos comum do que a poligamia, e a sua disseminação global ocorreu bem recentemente. Em uma amostra representativa de 186 sociedades culturalmente independentes, Murdock (1981) constatou que apenas 27, abaixo de 15 por cento, são monogâmicas. Um pequeno número de sociedades poliândricas pode ser encontrado em pesquisas interculturais (uma mulher pode ter vários maridos), mas a grande maioria é polígina (um homem pode ter várias esposas). A distribuição desses sistemas de parentesco não é aleatória. A poliandria surgiu em ambientes com recursos muito escassos, como as terras áridas e montanhosas do Tibete e do Nepal. Os povos do Tibete e do Nepal tradicionalmente praticaram uma poliandria fraterna: uma mulher se casaria com todos os irmãos de uma família. Esse sistema tem certas vantagens (Goldstein, 1987). Suponhamos, como nesse caso, que as mulheres no Nepal têm, em média, três crianças do sexo masculino. Se três irmãos se casarem com uma mulher, e se a mulher tiver três filhos homens, e esses três homens se casarem com uma mulher que terá três crianças do sexo masculino, então a população permanecerá constante. Se, em vez disso, cada criança do sexo masculino se casar com uma mulher diferente, então o tamanho da população masculina crescerá exponencialmente. O crescimento exponencial é bom em territórios onde há espaço para expansão, mas é menos viável no Himalaia. Por ser difícil fixar nova residência em outro lugar, as famílias serão pressionadas a dividir

seus bens entre si. Se cada família no Nepal tiver uma propriedade rural, transmitida para herdeiros homens, e se todos os descendentes do sexo masculino se casarem com a mesma mulher, então a propriedade da família permanecerá intacta, em vez de ser dividida em partes menores, a cada geração. Os machos têm um forte incentivo para permanecer numa residência, porque eles gozarão de uma herança mais substancial, que lhes proporcionará maior conforto, liberdade e *status*. Em suma, a poliandria pode ser adaptável em alto grau num clima inóspito, onde a terra agricultável é limitada.

Na maioria das definições, no entanto, a poligamia tem sido a forma dominante de casamento. Isso não quer dizer que a maioria das pessoas em todo o mundo engajaram-se na poliginia; menciono apenas que a poliginia tem sido permitida na maioria das sociedades. A convergência de três fatores pode ser especialmente importante para explicar a prevalência de poliginia (Grossbard-Schechtman, 1984; ver também White e Burton, 1988): Em primeiro lugar, há poucas oportunidades de emprego para as mulheres em muitas sociedades; nessas sociedades, as mulheres não podem competir no mercado livre, e dependem dos homens assalariados. Em segundo lugar, em muitas sociedades, as mulheres e seus filhos podem participar de trabalhos domésticos rentáveis, incluindo a agricultura; nessas sociedades, é uma vantagem econômica ter várias mulheres em casa. Em terceiro lugar, há uma considerável estratificação social em muitas sociedades de modo que alguns homens são muito mais ricos do que outros; em tais sociedades, várias mulheres poderiam sobreviver da renda de um homem rico. Em suma, se as mulheres dependem dos homens e são rentáveis aos homens, e se alguns homens podem suportar várias esposas, então a poliginia é especialmente plausível. Esses três fatores têm sido extremamente comuns ao longo da história mundial, e podem ajudar a explicar porque a poliginia foi tão difundida. A poliginia era a regra no mundo antigo. Nas sociedades grega, egípcia e judaica antigas, homens ricos podiam ter várias esposas. Salomão teria tido 700 mulheres em seu harém (I Reis, 11:3). Hoje, a poliginia é menos comum no Ocidente, mas ainda é difundida no *Terceiro Mundo*.

O casamento entre primos também é bastante comum em certas partes do mundo, especialmente no Oriente Médio, na Ásia meridional e na África subsaariana. Bittles (1990) estima que 20 por cento da população mundial vive em sociedades que promovem ou permitem o casamento entre primos. O casamento entre primos era permitido na Grécia ateniense, na sociedade da Roma tardia e na Judéia (Thompson, 1967; Treggiari, 1993; Goody, 1983). No Antigo Testamento, por exemplo, Jacó casou-se com duas filhas do irmão de sua mãe (Gênesis 29:10-30). O casamento entre primos tem certas vantagens, a mais óbvia dentre elas é que permite às famílias ampliadas consolidar sua riqueza.

O consentimento é outra característica do casamento, com variação intercultural. Para nós, o casamento consensual é um ideal altamente moralizado, mas está longe de ser universal. Em muitas sociedades, os casamentos são arranjados. Isso era verdade na Roma antiga e no Israel antigo, e continua a ser verdade em partes da Índia, Japão, China e em muitas sociedades africanas. Há razões óbvias para as famílias influírem na escolha daqueles com quem seus filhos se casam. O amor é com frequência cego, e os casais jovens podem ser firmados em conjunto com outros fatores que a longo prazo trazem vantagem econômica. Quando famílias arranjam um casamento, elas podem escolher parceiros que possam garantir alianças benéficas, e impedir a mobilidade para uma posição inferior.

O fato mais impressionante sobre esta comparação intercultural é que as atitudes ocidentais modernas em relação ao casamento parecem ter emergido de sociedades com valores muito diferentes. Há escassos precedentes para os valores modernos na Roma e na Judéia. Isso sugere que os valores modernos não foram herdados de uma tradição precedente, mas sim que foram inventados. Mas, por quê? Existem várias explicações possíveis. Consideremos a teoria da monogamia proposta por Richard Alexander (1987). Ele argumenta que a poligamia foi proibida como uma forma de evitar conflitos no interior do grupo. Existe uma grande concorrência entre as mulheres nas sociedades polígamas, pois se um homem tem várias esposas, e se as proporções de gênero são iguais, então alguns homens serão condenados a uma vida de celibato. Isso pode conduzir a conflitos violentos;

a monogamia imposta diminui a violência entre os machos, ao dar aos homens mais oportunidades iguais no mercado de casamentos. Deve notar-se também que a monogamia reduz conflitos entre as coesposas, e os conflitos entre os filhos por mais herança, visto que os casamentos monogâmicos rendem menos filhos do que os casamentos políginos. A monogamia diminui a violência. Sanderson (2001) avaliou a evidência consistente na hipótese de Alexander. Ele observa que as sociedades monogâmicas são maiores, mais ricas, e mais pacíficas internamente, sugerindo que as leis da monogamia podem ter permitido às sociedades prosperar e crescer. Apenas 10 por cento das sociedades pré-estatais impuseram a monogamia, contra 46 por cento dos grandes Estados.

A teoria sobre a monogamia de Alexander é intuitivamente plausível, e isso pode ajudar a explicar por que a monogamia propagou-se tanto; mas ela não é a única explicação disponível. É importante ter em mente que diferentes explicações podem funcionar em casos diferentes, e é importante considerar cada um dos casos individualmente, antes de se decidir por uma teoria de como uma instituição social surgiu numa cultura particular. Jack Goody (1983) desenvolve uma hipótese provocativa sobre como a monogamia deslanchou no Ocidente. Os povos pré-cristãos europeus normalmente toleravam a poliginia ou, o que equivale a isso, eles permitiram que os homens tivessem concubinas, cujos descendentes poderiam herdar a riqueza. Os romanos eram oficialmente monogâmicos, mas eles permitiam que um homem tivesse uma concubina, caso sua esposa fosse estéril. Os romanos também permitiam aos casais se divorciarem e casarem de novo, o que equivale a uma "poligamia em série", haja vista que as pessoas podiam ter vários cônjuges ao longo da vida. Casar de novo foi mesmo necessário durante alguns períodos da história romana tardia. O sistema monogâmico minimizou os conflitos por herança, porque manteve baixo o tamanho da família e eliminou a concorrência entre as coesposas. Mas a tolerância para com o divórcio e com novos casamentos significava que um homem ainda sem herdeiros com sua primeira esposa poderia deixá-la, e tentar ter filhos com outra. Com a ascensão do cristianismo, argumenta Goody, as coisas mudaram. A poliginia e a consanguinidade (casamento com familiares próximos) foram proibidas pela Igreja no século IV, depois da conversão

de Constantino. Os casamentos arranjados foram desencorajados pela Igreja, e as proibições foram introduzidas no casamento entre o clero, no divórcio e na herança das viúvas (em que as viúvas poderiam casar-se com parentes próximos de seus maridos falecidos). A Igreja também condenou o sexo recreativo e a adoção. Foi preciso algum tempo para muitas dessas regras se firmarem como influência da Igreja em expansão pela Europa, mas elas levaram a uma transformação significativa das atitudes em relação ao casamento. Diante disso, as novas regras não têm muito em comum, mas um exame mais atento revela um fato surpreendente. Todas elas operam para reduzir o tamanho da família e aumentar o número de heranças sem herdeiros [*heirlessness*], as quais, argumenta Goody, foram rentáveis à Igreja.

Em um casamento polígino, há várias esposas e, portanto, mais filhos que podem herdar a riqueza. A proibição da poliginia levou a uma grande redução no número de herdeiros. A proscrição de adoção impossibilitou compensar essa diminuição. Proibições contra a consanguinidade, incluindo o casamento entre primos, reduziu o fácil acesso aos cônjuges e impediu as famílias de consolidarem a riqueza. A consolidação da riqueza familiar foi ainda mais restringida pela proscrição dos casamentos arranjados. Laços de parentesco estratégicos foram rompidos, e os indivíduos foram forçados a encontrar os seus próprios parceiros. A proscrição do divórcio impediu os casais de casar novamente e de ter mais filhos (poligamia em série), e as restrições em relação ao casamento das viúvas teve um efeito similar. A proibição do sexo recreativo acarretou menos gravidez acidental e, portanto, menos filhos. O resultado líquido foi que muitas famílias acabaram sem herdeiros. Com altas taxas de mortalidade e nenhum tratamento para a infertilidade, os casais monogâmicos poderiam facilmente ficar sem herdeiros. Goody estima que, em alguns períodos, heranças sem herdeiros podem ter chegado a 20%. Quando uma família não tinha herdeiros, a propriedade era herdada pela Igreja. Haja visto que o clero não podia se casar, uma vez que a propriedade ia para a Igreja, não havia modos de ela voltar para a população não clerical. Não fica claro, a partir da análise de Goody, se tudo isso foi uma estratégia elaborada por líderes da Igreja, ou uma mera coincidência, mas fica claro que a Igreja beneficiou-se enormemente do

aumento de heranças sem sucessores. Com essas políticas em vigor, a Igreja tornou-se uma das maiores proprietárias de terras na Europa. Começando sem nenhuma propriedade, a Igreja primitiva chegou eventualmente a possuir cerca de 25% de toda a Gália.

A hipótese de Goody é controversa (ver, por exemplo, Verdon, 1988), mas não é insensata. Ele mostra como podem ter surgido atitudes em relação ao casamento. Assim como Harris e Harner acerca do canibalismo, Goody enfatiza os fatores materiais. Regras estritas sobre quem poderia casar acarretaram benefícios diretos. Mas há uma diferença. No caso de canibalismo, os benefícios foram auferidos pelos canibais. No caso das leis de casamento, os benefícios foram acumulados pela Igreja, conforme a hipótese de Goody, não necessariamente pelos indivíduos que eram casados. Entretanto, esse é um cenário historicamente estável, porque a Igreja é um fornecedor primário de valores. Os valores que beneficiam a Igreja são susceptíveis de se difundir e de se manter, visto que a Igreja exerce um papel central na educação moral. Mas a hipótese de Goody depara-se com uma questão desafiadora. Por que as pessoas investiram nas reformas do casamento da Igreja, se essas reformas efetivamente reduziram sua capacidade de acumular riqueza? A resposta pode envolver os meios de transmissão. As atitudes cristãs em relação ao parentesco proliferaram através de interpretações dos textos bíblicos. Os ensinamentos de Jesus e a história da Queda puseram os valores em um contexto narrativo; a retórica do pecado e da impureza ajudou a tornar as uniões ilícitas emocionalmente repugnantes.

A mudança geral de atitude para com a sexualidade foi fortemente influenciada por Santo Agostinho e por São Paulo. Depois de uma juventude devassa, Agostinho passou a acreditar que o sexo devia ser usado apenas para a procriação. Ele fundamentou isso em antigas tradições filosóficas gregas, que ligou a bondade moral a funções naturais. O sexo tem a função de procriar, de modo que o mau uso do sexo é antinatural. Se algo é antinatural, é errado. São Paulo ensinou que todo o sexo era pecaminoso, pois a devoção deveria ser reservada a Deus. A monogamia era a melhor coisa depois do celibato, porque reduziu o sexo a um mínimo. Ao ligar o sexo ao pecado, Agostinho e Paulo fundaram restrições ao

casamento na teologia e no direito natural. Ao classificar certos atos sexuais como não naturais, os líderes da Igreja podem ter sido hábeis para tocar em disposições biológicas ao nojo dos paroquianos. O nojo obedece a uma lógica da contaminação. Se o sexo é considerado sujo, então o sexo com vários parceiros pode ser visto como um modo de disseminar a impureza. Ao cultivar atitudes escrupulosas para com a sexualidade, os líderes da Igreja podem ter transmitido mais facilmente normas contra a poligamia, o divórcio e novos casamentos.

O nojo pode ter sido explorado também nas proibições crescentes contra casamentos consanguíneos. As pessoas muitas vezes rejeitam a ideia de incesto; portanto, qualquer união rotulada como incestuosa pode ser considerada como repugnante. O casamento entre primos era comum no mundo antigo, mas a Igreja medieval proibiu o casamento entre primos até o sétimo grau. Até hoje, as pessoas aculturadas no Ocidente sentem-se desconfortáveis com o casamento entre primos, e a repugnância parece ser a reação dominante.

Emoções mais positivas podem ter favorecido a transmissão do casamento consensual. Os líderes da Igreja primitiva incentivaram as pessoas a tomar decisões matrimoniais com base no amor. Isso deu às pessoas maior autonomia e um retorno emocional muito significativo. A ideia do casamento consensual repercute nos ideais individualistas que emergiram no curso da história ocidental. Tradicionalmente, a Igreja considera os indivíduos livres, e a escatologia cristã baseia-se na ideia de que se pode ao chegar ao céu, fazendo as escolhas certas na vida. É ideologicamente difícil encaixar o casamento arranjado nesse quadro.

Em suma, a Igreja primitiva pode ter implementado mudanças revolucionárias nos costumes de parentesco, ao integrá-los num sistema de crenças atraente, com forte conteúdo emocional. Essas sugestões são vagas e altamente especulativas. Numa análise genealógica adequada seria examinada sistematicamente a retórica usada para disseminar as leis matrimoniais nos primeiros séculos da Igreja. Mas o ponto principal deve ficar claro. As mudanças radicais nas atitudes para com o casamento podem ter sido transmitidas através de uma combinação de fatores materiais, narrativos e afetivos. Provavelmente, pregadores e paroquianos estavam alheios ao impacto econômico das leis

matrimoniais revistas, mas as atitudes emocionais e filosóficas podem ser traçadas mediante o exame de documentos históricos. Genealogistas naturalizados podem usar esses registros (assim como epistemólogos naturalizados podem usar a pesquisa psicológica) como dados empíricos, ao desenvolver teorias de como cada um de nossos valores morais surgiu.

6.3 GENEALOGIA COMO CRÍTICA

Esses estudos de caso na história da moral são instrutivos. Eles mostram que os valores morais pode ser considerados historicamente. Para cada valor moral que possuímos, há uma história para contar sobre como chegamos a possuí-lo. Explorarei no capítulo 7 a possibilidade de que essas histórias, por vezes, têm uma dimensão biológica. Os exemplos discutidos neste capítulo dizem mais respeito ao desenvolvimento cultural. Em alguns casos, os novos valores morais emergem rapidamente; em outros, a transformação é mais lenta. Como Nietzsche sugeriu, a ascensão do cristianismo pode ter sido rápida o suficiente para constituir uma revolução. O desaparecimento do canibalismo pode ter sido relativamente lento em algumas sociedades. Se Harris estiver certo, o canibalismo tendeu a desaparecer quando as chefias mudaram para formas sofisticadas de agricultura e evoluíram para os estados. Nesses casos, o sacrifício humano e as práticas canibais podem ter sido gradualmente substituídos por sacrifícios de animais e, por fim, por "sacrifícios" simbólicos, incluindo qualquer coisa, desde a queima de efígies e totens até a Eucaristia cristã. O canibalismo asteca terminou rapidamente, porque os conquistadores espanhóis dizimaram o Império Asteca em uma única geração. Esse poder mudou, e os esforços posteriores dos missionários promoveram mudanças relativamente rápidas na América Central; mas as sociedades indígenas, muitas vezes, conseguiram manter vestígios de suas formas tradicionais de vida, quando foram convertidas para os sistemas de crenças de seus conquistadores.

Essas dinâmicas históricas são certamente interessantes em seu direito próprio. Elas nos ensinam algo acerca do que nos leva a acreditar no que fazemos, e por que outros têm crenças diferentes. Mas alguém pode justificadamente se

perguntar se elas têm alguma ramificação filosófica. O que elas provam sobre o estatuto da moralidade? Nietzsche acreditava que a análise genealógica poderia ser usada para criticar a moral. Ao expor a história dos valores cristãos, em particular, ele procurou corroer a sua credibilidade. Nesta seção, eu quero questionar se tais conclusões céticas estão asseguradas.

Começemos por considerar uma versão forte da reivindicação de que a historicidade mina a moralidade. Por "historicidade", eu me refiro a valores que emergiram como consequência de eventos históricos, e não por um processo, digamos, de intuição moral direta, de revelação, ou de dedução racional de princípios normativos. Em relação a esses valores, pode-se ser tentado a endossar o seguinte princípio genealógico:

(G1) Se uma crença num valor moral surgiu como uma consequência de eventos históricos, então o valor moral é falso.

Para os fins desta discussão, eu pensarei nos valores morais como proposições valorativas, da forma "fazer φ é moralmente mau" ou "fazer φ é moralmente bom". Assim, (G1) expressa que as crenças sobre o que é bom e mau podem ser desacreditadas simplesmente em virtude do fato de que elas surgiram através da história, ao invés, digamos, de processos racionais. Duvido que Nietzsche ou qualquer outra pessoa defenderia algo assim. (G1) comete uma falácia genética direta: é um erro tratar um defeito no modo como uma proposição tornou-se crível como uma evidência direta contra essa proposição. Suponhamos que eu ganhe na loteria, selecionando números que correspondem aos aniversários dos meus amigos e da minha família. Esse método de adivinhação não tem qualquer influência sobre os mecanismos que determinam os números vencedores (p. ex., uma engenhoca que aleatoriamente cospe esferas numeradas). Mas o meu palpite, apesar disso, é preciso.

Aqui, alguém pode ser tentado a argumentar que há uma importante diferença entre o caso de loteria e o caso da moral. No caso de loteria, há um mecanismo que independe da mente, por meio do qual os números são

selecionados. O mecanismo não depende de como formamos crenças acerca dos números que ganharão. Mas no caso da moral, eu argumentei que a verdade depende de nós. A verdade de um valor moral, nesse sentido, está ligado ao método de fixação da crença. O hiato entre o fato e a crença é muito estreito no domínio da moral. Isso distingue o caso da moral do caso da loteria, mas não torna (G1) defensável. Se a verdade depende de nós, então os fatos históricos sobre o modo pelo qual chegamos a sustentar um valor moral como verdadeiro estabelece também que o valor é verdadeiro. As análises genealógicas só poderiam minar a verdade de um valor moral, se elas revelarem que nós não acreditamos realmente naquilo que professamos crer. Se Nietzsche afirmasse que os cristãos não acreditam que a caridade é boa, ele teria sido capaz de refutar esse valor moral no interior da cultura cristã. Essa não era a sua intenção. O método genealógico tende a considerar as crenças morais como um ponto fixo, para depois explicar suas origens.

Precisamos de um princípio mais fraco. Talvez devêssemos abandonar o foco na verdade, e mudar para algo de caráter mais epistemológico:

(G2) Se a crença num valor moral surgiu como uma consequência de eventos históricos, então essa crença não é justificada.

Isso é um avanço. Ele parece evitar a falácia genética. O modo pelo qual uma crença é obtida pode sustentar-se no fato de a crença ser justificada. Por exemplo: os médicos medievais acreditavam que muitas plantas tinham poderes curativos, e eles, muitas vezes, assumiram que uma planta era especialmente capaz de curar a parte do corpo que mais se lhe assemelhava. A cabaça, por exemplo, pode ser usada para tratar um membro inchado. Essa crença não é justificada, porque se baseia numa premissa falsa. Mas suponhamos que um tipo particular de cabaça, por acaso, pode reduzir o inchaço em algum grau, e suponhamos que, ao longo do tempo, essa cabaça continua a ser usada medicinalmente, enquanto outras não o são, em virtude do seu sucesso passado. A crença de que a cabaça pode curar surgiu de um modo ruim, mas persistiu por um bom motivo, e agora é endossada

com uma justificativa. Esse tipo de exemplo mostra que uma crença pode ser inicialmente injustificada, pelo modo como surgiu, mas ela é justificada mais tarde, por exemplo, devido ao seu êxito na direção do comportamento. As origens ignóbeis são compatíveis, nessa medida, com a justificação.

(G2) enfrenta outra objeção também. Argumentei que conceitos morais referem a propriedades dependentes de resposta. Segue-se que os juízos morais (pelo menos aqueles que expressam normas fundantes) são autojustificantes. Se eu fizer um juízo de que algo está errado, e se o juízo for feito em condições epistêmicas, nas quais eu acessei com precisão minha memória de longo prazo e descobri um sentimento de desaprovação para com essa coisa, então a minha crença é justificada, porque ERRADO refere-se àquilo a que eu tenho tal sentimento. Se essa consideração estiver correta, isso se aplica mesmo se as minhas disposições para reagir são apresentadas como uma consequência de acontecimentos históricos. A justificação é insignificante se o sentimentalismo construtivo for verdadeiro.

Em resposta, pode-se admitir que os juízos morais são sempre justificados quando feitos sob boas condições epistêmicas, insistindo que eles deveriam ser abandonados. Quando é insignificante, a justificação pode ser insuficiente para suportar a convicção. Por analogia, suponhamos que você goste de um determinado tipo de música simplesmente por tê-lo ouvido durante os anos de formação de sua juventude. Hoje, você considera uma justificação trivial dizer que essa música lhe é agradável. O fato de que ela lhe apraz é um caso de simpatia, mesmo que o prazer seja o resultado de um fato contingente biográfico, que não tem nada a ver com as qualidades intrínsecas da música. Você pode decidir se a preferência deve ser abandonada, porque é produto de inculcação, e você pode tomar medidas para alterar o seu gosto musical. O seguinte princípio é sugerido, com vistas à aplicação ao domínio moral:

(G3) Se uma crença num valor moral surgiu como uma consequência de eventos históricos, então essa crença *deveria ser abandonada*.

(G3) é mais fraco que os princípios anteriores, visto que não faz reivindicações sobre a verdade ou a justificação. Mas ele não é mais defensável. A historicidade pode motivar-nos a rever valores, mas não necessariamente. Voltando ao exemplo da música, o fato de que nossas preferências são influenciadas pela exposição prévia não é razão suficiente para a sua revisão. O próprio fato de que você gosta de certa música agora – o mesmo vale para qualquer outra coisa – não importa qual a causa, é perfeitamente um bom motivo para ouvi-la. A analogia estética de (G3) implica que cada gosto adquirido deveria ser abandonado. Imagine se alguém lhe disser que você deve parar de comer chocolate, simplesmente porque a sua preferência por chocolate é uma consequência contingente de sua biografia. Essa é uma política estranha no domínio estético e, pode-se ponderar, é igualmente estranha no domínio dos valores morais.

Pode-se tentar defender (G3), argumentando que a moralidade difere da música em aspectos cruciais. Por um lado, não reconhecemos que nossos valores morais são históricos. A descoberta de suas origens revela que fomos enganados acerca deles. Mesmo se isso for verdade, não vejo qual é a importância disso. Suponhamos que eu acredite que gosto de jazz porque ele possui certas qualidades intrínsecas, mas, na realidade, eu gosto de jazz porque ouvi esse tipo de música ao longo dos anos de formação da minha vida. Se eu descobrir que minha preferência resulta desse fato biográfico, parece não se seguir que eu tenha um motivo para abandonar meu gosto por jazz. Nesse sentido, não haveria necessidade de uma nova premissa, segundo a qual o amor pelo jazz só valeu a pena se teve uma base puramente objetiva. É difícil sustentar essa premissa. Ela também é difícil de ser sustentada no caso da moral. Um valor pode ser digno de manter porque está arraigado e é difícil de mudar, porque é útil, ou, simplesmente, porque eu o valorizo.

É tentador procurar salvar (G3), traçando uma distinção entre dois tipos de origens. Alguns valores provêm de eventos que são inócuos, e outros provêm de eventos que consideramos retrospectivamente desprezíveis. Por exemplo, alguns valores podem ser conduzidos por motivos impróprios, tais como a ganância ou o ressentimento, e alguns podem ser-nos impostos por outros que têm mais poder. Nietzsche queria minar os valores cristãos, ao provar que eles foram reacionários.

Caso pudermos mostrar que um valor não é apenas histórico, mas que as suas origens são ignóbeis, talvez isso então seja uma razão para rebaixar o valor. Assim:

(G4) Se a crença num valor moral surgiu como uma consequência de eventos históricos ignóbeis, então essa crença deveria ser abandonada.

Isso está começando a ficar mais nietzscheano, mas não é muito satisfatório. Ignóbil é um termo valorativo. Nietzsche disse, por vezes, que há um ponto de vista moral objetivo, a partir do qual nós podemos avaliar a história da moral cristã, do ponto de vista natural daqueles espíritos livres que adotam abertamente a vontade de poder. Esse é um aspecto do programa de Nietzsche o qual faríamos bem em afastar. No próximo capítulo, levantarei dúvidas sobre a posição natural transcendental. Se a palavra "ignóbil" não pode ser aplicada de modo objetivo, então ela também pode ser um produto da história. Se a historicidade é um motivo para o ceticismo, então deveríamos ser céticos em relação às tentativas de minar os valores morais, dizendo que eles têm histórias moralmente duvidosas. A introdução de um termo moral (G4) pode ser autodestrutiva.

Pode-se tentar salvar (G4), argumentando que ele pode ser usado para diagnosticar a hipocrisia dos moralistas. Talvez Nietzsche queira que os cristãos percebam que seus valores derivam de fatores desprezíveis por seus próprios padrões. Se um sistema moral surgiu de um modo pelo qual ele próprio exige que nós o condenemos, então o sistema pode minar seus próprios fundamentos. Talvez. Mas o apologista do sistema moral em questão poderia simplesmente responder que o caminho para o céu está coberto de más intenções. Os fins justificam os meios. (G4) é apenas mais um exemplo de falácia genética.

Nesse ponto, um nietzscheano fiel pode fazer um movimento muito diferente. Quando Nietzsche fala sobre as origens ignóbeis do cristianismo, ele não está apenas se referindo ao passado. Ele pensa que os motivos questionáveis, os quais levaram à revolução cristã na ética, ainda sustentam os valores cristãos no presente. Os defensores contemporâneos desses valores são movidos pelo

ressentimento. Se isso for correto, então os moralizadores contemporâneos podem ser rotulados de hipócritas e, por seus próprios padrões, a hipocrisia é uma coisa ruim. Nessa interpretação, o ceticismo não repousa na historicidade como tal. Ao contrário, a análise genealógica é elaborada para revelar algo sobre os fundamentos psicológicos da moral contemporânea (cf. Gemes, 2006). Deveríamos abandonar os valores cristãos, porque eles são sustentados por emoções, que consideramos ser uma base ruim para os valores. Não há problema em gostar de chocolate porque é prazeroso, mas é ruim gostar dele porque você fica indignado com os que odeiam chocolate. Fica claro que esse é um juízo de valor, mas é um juízo que a pessoa, que atualmente gosta de chocolate, estaria disposta a fazer. Se você pode convencer o amante do chocolate que gostar de chocolate é reacionário, o mesmo pode envidar esforços para superar o gosto pelo chocolate. Penso que chegamos, aqui, mais perto da estratégia subjacente ao programa genealógico de Nietzsche. Para ele, métodos diacrônicos são instrumentos para fazer psicologia sincrônica. A genealogia é uma escada para ser jogada fora de novo.

Penso que essa abordagem da genealogia é muito limitada. Por um lado, duvido muito que os cristãos são (ou eram) motivados pelo *ressentimento*. Não precisamos de uma maquinaria psicodinâmica fantástica para explicar por que as pessoas têm os valores que elas empregam. O condicionamento cultural é suficiente. As emoções exercem um papel nesse processo, mas a sua contribuição pode ser muito mais modesta daquela do modelo de Nietzsche. Nós aprendemos valores morais através de treinamento emocional; os valores que se difundem mais facilmente são aqueles que são mais maleáveis a essas técnicas. Argumentei acima que as normas cristãs contra a poliginia, em parte, eram fáceis de transmitir, porque é fácil promover a repugnância quando ela ocorre em interações corporais. A ideia de relações sexuais plurais pode fazer a pessoa se sentir um pouco enojada se ela enquadra isso em termos de contaminação, de impureza e de violações da natureza. Se a história de poliginia revela algo sobre a psicologia contemporânea, ela mostra que algumas pessoas acham a poligamia repugnante. Acusar pessoas de basearem os seus valores morais na repugnância é muito diferente do que acusar pessoas de basearem os seus valores morais no ressentimento. A acusação

anterior não levaria necessariamente as pessoas a mudar sua moral. Um monogâmico dedicado pode ainda responder que a poliginia é errada precisamente porque é repugnante. Colocando isso de outra forma, Nietzsche assume que nós experimentaremos dissonância cognitiva quando descobrimos as emoções que subjazem a nossos valores morais; mas não vejo nenhuma razão para pensar que esse será o caso.

A genealogia, desse modo, não revelará sempre os motivos psicológicos ocultos que embaraçam os que moralizam. Mas a genealogia pode constranger os moralizadores de outro modo. Se você disser aos adversários de poliginia que suas preferências pela monogamia surgiram de um roubo imobiliário feito pela Igreja primitiva, isso poderia dar-lhes o que pensar. Essa revelação não é uma razão suficiente para rebaixar o movimento contra a poliginia – talvez haja boas razões em favor da monogamia – mas ela dá ensejo a uma reavaliação. Alguns valores morais se propagaram porque eles beneficiam aqueles que os possuem. Outros proliferaram porque beneficiam uma instituição poderosa, que tem um lugar privilegiado de influência na moralidade local. Quando descobrimos que um dos nossos valores se enquadra na segunda categoria, e não na primeira, temos razão para perguntar se esse valor é digno de ser mantido. Isso sugere o seguinte princípio:

(G5) Se a crença em um valor moral surgiu como uma consequência de acontecimentos históricos ignóbeis, então deveríamos *considerar* o abandono dessa crença.

Ao falar de casos em que a história é ignóbil, estou imaginando casos nos quais a disseminação de um valor não depende de qualquer benefício óbvio (moral ou de outro tipo) para o moralizador. Nesses casos, podemos justificadamente perguntar se o valor está nos trazendo agora algum benefício. Uma resposta negativa a isso não é uma razão suficiente para rebaixar o valor. Os valores morais podem ser mantidos mesmo que não tragam quaisquer recompensas. Alguns valores morais podem até mesmo ser custosos para algumas pessoas. Mas nos casos em que não se descobre nenhum benefício, a reavaliação é apropriada.

Podemos descobrir, em alguns casos, que um valor está fazendo mais mal do que bem; nesses casos, é desejável a mudança moral.

(G5) descreve a genealogia como um instrumento crítico, apesar de ser talvez mais rude do que o martelo de Nietzsche. Os valores não são necessariamente viciados por sua historicidade, mas as análises históricas nos dão uma ferramenta para avaliar se os nossos valores servem a funções pelas quais ainda nos preocupamos. Ao discutir (G2), eu disse que a justificação é insignificante para os valores morais, porque eles se referem a propriedades dependentes de resposta. Naquele ponto, eu estava falando sobre garantia de evidência: fatores que sustentam a verdade de uma crença. Eu acredito que fazer φ é ruim em virtude de ter um sentimento de desaprovação em relação a fazer φ ; tais sentimentos são validadores para reivindicações morais; assim, o fato de que eu acredito que fazer φ é ruim sustenta a conclusão de que fazer φ é ruim em meu sistema de valor. Mas a justificação pode ser também instrumental (Sinnott-Armstrong, 1996). Eu posso encontrar um valor que julgo ser útil de algum modo. Ele pode servir a uma função pela qual me importo. Argumentei que os valores morais, muitas vezes, emergem por causa de algum papel que desempenharam no passado. Esse papel pode me importar, ou não. A análise genealógica oferece um instrumento para ver que função os valores exerceram, e ela pode ser útil para determinar se os nossos valores morais estão nos ajudando ou nos impedindo na busca de nossos fins não morais.

Consideremos alguns exemplos. Eu sugeri que o tabu do canibalismo surgiu porque nós desenvolvemos modos de usar nossos vizinhos como contribuintes, em vez de alimentos. Provavelmente, deveríamos manter esse tabu no lugar. Permitir o canibalismo, mesmo em casos em que as pessoas morrem de causas naturais, pode conduzir a uma ladeira escorregadia. A regra seria fácil de explorar, o potencial de corrupção seria enorme, e o resultado seria insignificante. Além disso, a nossa oposição ao canibalismo tornou-se constitutiva de nossa identidade como "civilizados". É o resultado de termos atingido esse *status*. A suspensão do tabu poderia potencialmente alterar nossa identidade, de tal modo que teria consequências negativas imprevistas. Desde que o tabu do canibalismo não cause

dano grave ou não entre em conflito com os valores que estimamos, há pouca razão para abandoná-lo. Em qualquer caso, as razões históricas subjacentes ao tabu ainda persistem hoje. Somos ainda melhores em tributar os nossos vizinhos do que em comê-los.

Compare-se isso com a monogamia. Suponhamos que nossa transição da poligamia foi realmente uma conspiração da Igreja primitiva. Deveríamos parar de valorizar a monogamia? Essa é uma pergunta difícil. Sabemos que abandonar a monogamia teria custos. Nossos sistemas e instituições sociais foram estruturados em torno da monogamia, e sistemas que não são monogâmicos têm frequentemente características que nós desvalorizamos por outros motivos. Sociedades poligâmicas são geralmente políginas, e a poliginia depende de um duplo critério, que discrimina as mulheres. Além disso, a monogamia é gratificante para aqueles que a valorizam (compare: o jazz é agradável para quem gosta dele). Talvez isso seja o resultado do condicionamento social, mas também o são o gosto musical, as preferências de moda, a afeição pelo esporte, as línguas que falamos e outras coisas mais. Se nós fôssemos condicionados a apreciar a monogamia por razões duvidosas, não se segue que a monogamia é uma instituição ruim. Se não podemos encontrar algo horrivelmente errado na monogamia, e acontece que a valorizamos, podemos muito bem continuar a valorizá-la. Por outro lado, não podemos exigir a monogamia por lei, porque algumas pessoas podem ter o desejo e a capacidade de viver em relacionamentos poligâmicos não explorativos. Assume-se geralmente em nossa cultura que a poligamia é categoricamente errada. Na verdade, algumas pessoas argumentam contra o casamento gay, alegando que a permissão de uniões homossexuais nos colocaria numa ladeira escorregadia, e levaria à legalização da poligamia. Não é óbvio para mim que qualquer um desses tipos de casamento deveria ser proibido. O fato de que a monogamia e as proibições do casamento gay são produtos da história não é uma razão suficiente, àqueles que ainda as valorizam, para desistir de ambas; mas, se essas normas matrimoniais já não servem a qualquer função que nos importa, então não deveríamos impô-las aos outros. Para uma genealogia interessante das atitudes em relação ao casamento gay, veja Boswell (1994).

A conclusão é ainda mais óbvia, se nos voltarmos para o casamento entre primos. Para muitos de nós, o casamento entre primos soa repulsivo. Nós o julgamos moralmente errado e, em algum sentido, não natural. Mas agora descobrimos que o casamento entre primos é comum, e que nosso próprio desprezo por ele é um artefato da história. Esse fato nos leva a reavaliar nossa condenação moral do casamento entre primos. Desde os séculos em que ele foi banido, surgiram argumentos *post hoc* para julgar que o casamento entre primos é errado. Ensinaram-nos que filhos de primos são suscetíveis de terem defeitos de nascença. Essa crença revelou-se falsa (Bennett *et al.*, 2002). É uma lenda urbana, ordenada em favor de uma agenda moral. E, em qualquer caso, o risco de defeitos de nascença não torna moralmente errado se casar com alguém. Assim, não há nenhuma razão forte para manter as normas contra o casamento entre primos. Nós também podemos decidir que há boas razões em favor do casamento entre primos. Talvez os primos sejam mais propensos a partilhar coisas em comum, e a estabelecer relacionamentos estáveis. Se assim for, deveríamos ser receptivos a uma mudança de valor.

Vamos concluir com um exemplo mais controverso: o aborto. Eu notei anteriormente que o aborto foi condenado pela Igreja cristã primitiva. Essa posição pode ter sido herdada do judaísmo. Os judeus antigos tiveram duas boas razões para se opor ao aborto: o aborto era muito arriscado para as mulheres nos velhos tempos; a proibição do aborto aumentou a fertilidade, o que é uma virtude para uma pequena comunidade religiosa em constante risco de extinção. Posteriormente, a Igreja relaxou a sua posição antiaborto. Agostinho retornou à visão aristotélica, de que o feto pouco desenvolvido não podia sustentar uma alma. O aborto era considerado errado só depois de o feto agitar-se, quando a mãe sente o primeiro movimento em seu ventre. A posição de Agostinho sobre o aborto tornou-se Direito Canônico da Igreja, revertendo a visão anterior, de que todo aborto é errado. Uma possível razão para a mudança é que, no século V, a Igreja tinha interesse na redução do número de herdeiros nascidos de cada família. Essa foi a posição oficial da Igreja durante séculos. Em 1588, o papa Sisto V proibiu o aborto por um breve período, mas o papa Gregório XIV revogou a proibição três anos depois. A posição

antiaborto recuperou alguma popularidade no século XVII, mas foi somente no século XIX, no entanto, que a doutrina oficial da Igreja se voltou contra o aborto. Em 1869, o papa Pio IX aditou uma proibição do aborto ao Direito Canônico da Igreja que tem sido, desde então, a posição oficial da Igreja Católica Romana. Há escassos apoios bíblicos para essa posição, de modo que os cínicos mencionam segundas intenções. Uma conjectura, sustentada por ativistas dos direitos de aborto, é que a Igreja proibiu o aborto para aumentar o tamanho do exército francês. A população da França estava diminuindo, e o Imperador Napoleão III queria aumentar as taxas de natalidade em sua nação católica. Napoleão III reconheceu publicamente a infalibilidade do papa, no mesmo ano da proibição do aborto. Outra hipótese conspiratória é que a Igreja começou a opor-se novamente ao aborto, a fim de manter as mulheres em suas funções tradicionais. A proibição coincidiu com a ascensão do movimento de sufrágio feminino. Os críticos do sufrágio feminino com frequência argumentavam que as mulheres eram mal providas para votar, porque as demandas da maternidade mantinham-nas afastadas do trabalho e da escola, e promovia perigosas tendências ao pacifismo. O Vaticano proibiu o aborto no mesmo ano em que Inglaterra deu às mulheres o direito de votar em eleições locais, um direito sem precedentes na Europa Ocidental.

Não vou tentar avaliar essas análises históricas agora. Eu sou cético em relação à pretensão de que a posição antiaborto foi simplesmente um estratagema contra o sufrágio feminino; a Itália não deu às mulheres o direito de votar até 1925. Também sou cético em relação à afirmação de que o aborto foi proibido para aumentar o exército francês; o papa Pio IX pode não ter considerado esse fim vantajoso para Roma e, de qualquer modo, é incrivelmente presunçoso assumir que ele teria mudado o Direito Canônico sob pressão política. Além disso, a ideia de que a vida começa na concepção tem sido popular por 200 anos.

Em vez de responder à espinhosa questão acerca das razões pelas quais a Igreja proibiu o aborto, quero considerar uma questão hipotética. Imagine uma pessoa, chamada Smith, que se opõe ao aborto porque essa é a doutrina da Igreja. Imagine agora que Smith passa a acreditar que a proibição do aborto foi efetivamente introduzida como um instrumento deliberado numa campanha contra

o sufrágio feminino. Ao saber disso, Smith mudaria instantaneamente de ideia sobre o aborto? Provavelmente não. Os valores são difíceis de mudar. Mas Smith *deveria* reavaliar sua posição. Se Smith acreditava que o aborto era errado *porque* era uma violação de uma interpretação oficial dos textos sagrados, agora, a autoridade dessa interpretação é posta em dúvida pela descoberta de segundas intenções, então o motivo de Smith para se opor ao aborto caiu em descrédito. Ela terá que contar com outras razões ou tratar a oposição ao aborto como uma norma fundacional (básica).

A genealogia é um poderoso instrumento crítico, pois nos obriga a ver convicções influentes sob uma nova luz. Nós tendemos a pensar bastante superficialmente sobre os nossos valores morais. Tomamos nossos valores por verdades óbvias ou recebidas. Consideramos o comportamento imoral como antinatural. Quando adotamos uma postura histórica, nós alteramos nossa posição epistêmica em relação aos valores. No discurso ordinário, nós ou tomamos os valores morais como fundamentais, portanto, imunes ao questionamento, ou enquanto justificados pelo apelo a valores que são fundamentais. Quando oferecemos justificações, elas são tipicamente circulares, tênues ou estereotipadas. Qualquer filósofo experiente poderia inventar argumentos críveis para cada lado de um debate moral. Talvez um dos lados tenha sempre no final o melhor argumento, mas é incrivelmente improvável que cada moralizador esteja de posse de argumentos chocantes para cada uma de suas convicções morais.

Quando consideramos a moral historicamente, deixamos de lado as questões de justificação e engajamo-nos numa forma de autoantropologia. Nós não perguntamos: "por que eu deveria acreditar que P?", mas sim: "por que eu acredito que P?" A sugestão herética de Nietzsche é que a última questão descritiva pode ter alguma relação com a questão aparentemente normativa. Se as convicções morais tendem a repousar em intuições fundacionais ou em justificações imaturas, então a pergunta "por que eu deveria acreditar que P" é um convite à desilusão. Tendemos a recair em respostas prontas. A pergunta "por que eu acredito que P" pode expor os casos em que chegamos a uma convicção moral na ausência de razões decentes. Considero que essa é a condição *habitual*. Em muitos casos, os nossos estimados

valores morais emergiram de condições que não se aplicam mais. A genealogia pode ser usada para reavaliar criticamente esses valores. Em alguns casos, no entanto, os valores morais subsistirão à crítica genealógica. Alguns dos nossos valores conseguiram chegar até nós porque são úteis ou exitosos. Essa é uma possibilidade que Nietzsche não aprecia plenamente. Seu pessimismo é exagerado. Mas Nietzsche está certo em pensar que alguns de nossos valores morais têm histórias questionáveis. Alguns são transmitidos, não porque eles funcionam, mas porque servem àqueles que estão em condição de disseminar os valores. Quando nos deparamos com casos como esse, estamos diante de uma escolha desafiadora. Podemos manter o valor em questão ou podemos tentar superá-lo. Superar valores não é uma tarefa trivial, como veremos no capítulo 8.

Tive dois objetivos neste capítulo. Eu quis ilustrar o método genealógico. Se os valores morais são sentimentalmente construídos e culturalmente relativos, então tem de haver uma explicação de como chegamos a possuir cada um de nossos valores. A genealogia é um método para descobrir tais explicações. Meu segundo objetivo foi avaliar as implicações da genealogia. Nietzsche pensa que a genealogia pode ser usada para promover o ceticismo sobre os nossos estimados valores morais. Argumentei que ele é excessivamente pessimista; mas eu concordo que a genealogia nos ajuda a determinar quando um valor é especialmente apropriado para reavaliação. Eu disse muito pouco aqui sobre como essa reavaliação poderia ocorrer. A partir de que ponto de vista privilegiado nós podemos decidir se um valor é digno de ser mantido? Nietzsche pode ser lido como aquele que supõe haver uma posição transcendental a partir da qual nós podemos avaliar a moralidade e escolher novos valores. Ele sugere que nós podemos embasar a moralidade na natureza humana. Nisso reside o otimismo de Nietzsche. No capítulo 7, levantarei sérias dúvidas sobre isso e, no capítulo 8, defenderei uma abordagem alternativa. A moral provisória consiste em a genealogia ser um instrumento subtilizado para a crítica da moral. Tanto o pessimismo de Nietzsche sobre a moralidade atual quanto o seu otimismo sobre a moralidade natural são exagerados, mas ele certamente tem razão em pensar que as análises históricas podem ser valiosas na revisão da moral.

7. OS LIMITES DA ÉTICA EVOLUCIONISTA

7.1. NORMAS EVOLUÍDAS

7.1.1 O Mito de Nietzsche

A teoria ética de Nietzsche tem dois componentes: um negativo e outro positivo. O componente negativo é o seu programa genealógico. Sua função principal é crítica e destrutiva. Nietzsche queria minar os valores profundamente arraigados, expondo sua história sórdida e seus vis motivos subjacentes. O componente positivo é capturado no mito nietzscheano (1885) do *Übermensch* – um ser que é capaz de viver em um mundo além da dicotomia do bem e do mal que marca moralidade contemporânea. O *Übermensch* é um espírito livre que afirma a vida e abraça os valores que são, em certo sentido, mais naturais do que os dos homens modernos.

O tema dos valores naturais aparece em *A Genealogia da Moral*. Lá, Nietzsche diz que os nossos valores atuais são reativos, ou reacionários. Eles são alimentados pelo ressentimento em relação aos opressores romanos. Os valores romanos, pelo contrário, são ativos. Eles não são definidos em oposição a qualquer outra coisa. Nietzsche também usa a retórica dos valores naturais em *O Anticristo*, onde ele sugere que os antigos israelitas tinham valores que eram mais naturais do que os valores cristãos, e, portanto, melhores:

Originalmente, e acima de tudo no tempo da monarquia, Israel manteve a atitude certa das coisas, ou seja, a atitude natural. O seu *Yaweh* foi uma expressão de sua consciência de poder, a sua alegria em si mesmo, suas esperanças para si mesmo. (...) Mas (...) a sua concepção foi alterada – a concepção foi desnaturalizada. (...) A moralidade não é mais um reflexo das condições que contribuem para uma vida significativa e para desenvolvimento das pessoas; ela não é mais o instinto de vida primária; em vez disso, tornou-se abstrata, opondo-se à vida (Nietzsche, 1888: 25).

Assim, os valores iniciais de Israel, como os valores romanos, podem ser considerados como precursores para os valores que o *Übermensch* vai reeditar. A ponderação de Nietzsche é que os valores israelitas e romanos não são perfeitas manifestações da natureza humana, mas apenas uma aproximação. Em *A Vontade de Poder* escreve:

Nunca houve uma humanidade natural. A escolástica dos não-valores e dos antivalores é a regra, é o começo; o homem atinge uma natureza somente após longa luta – ele nunca “retorna” (...) Mais natural é nossa primeira sociedade, a dos ricos, a classe de lazer: caçam uns aos outros, o amor entre os sexos é um tipo de esporte em que o casamento fornece um obstáculo e uma provocação; eles se divertem entre si e vivem para o prazer; eles estimam vantagens físicas, acima de tudo, são curiosos e ousados (Nietzsche, 1883-8: 120).

O progresso moral virá, nesta leitura de Nietzsche, quando nós abandonarmos nossos valores reativos e retornarmos às formas mais antigas de prosperidade. Nietzsche sugere que podemos até mesmo superar os romanos, e desenvolver um sistema de valores que seja ainda mais natural e ahistórico. Para se mover para além do bem e do mal, podemos tomar inspiração no passado, mas isso não é, ao mesmo tempo, tanto uma forma de nostalgia quanto uma estratégia para escapar da temporalidade. Os antigos não eram melhores pelo fato de terem precedido os cristãos; eles foram melhores porque seus valores realmente expressavam a constituição fundamental como seres humanos.

Essa ideia segundo a qual alguns valores são historicamente construídos e outros são mais naturais é muito sedutora. Ela está incorporada no ideal romântico do bom selvagem, e também é fundamental para a distinção de Hume entre virtudes naturais e artificiais. Ao contrário de Nietzsche, Hume não queria rejeitar as virtudes artificiais, mas achava que as naturais vêm primeiro; ele pensava que as virtudes artificiais deveriam ser fundamentadas em virtudes naturais para ter qualquer autoridade. O tema dos valores naturais também é popular na filosofia moral recente. Os filósofos de hoje não invocam o mito do *Übermensch* de Nietzsche, mas alguns deles se mostram atraídos pela ideia de que temos um conjunto de normas

que nos vêm naturalmente. Em particular, supõe-se que algumas normas são produto da seleção natural; essas normas foram forjadas em nosso passado ancestral e continuam a desempenhar um papel consistente em guiar nosso comportamento moral. Os defensores da ética evolucionista raramente contrastam essas normas de base biológica com outras normas que não são de base biológica, entretanto, sua premissa permiti-los-ia fazer isso. Todas as variações culturais que tenho discutido ao longo dos dois últimos capítulos, todas as contingências históricas, poderiam ser apenas uma sobreposição dispensável que, na verdade, oculta um conjunto de valores que se encontram fora da história. Estes valores naturais pertenceriam a uma escala de tempo diferente, e é tentador pensar que eles são privilegiados em relação aos demais. Ao contrário de valores historicamente construídos, os valores naturalmente evoluídos passaram por um processo de seleção, eles são universais e podem ser considerados imutáveis. Desta forma, talvez o relativismo e o historicismo possam ser contornados. Talvez possamos simplesmente descartar os valores que tenho discutido nos dois últimos capítulos e nos restringir aos valores que são tecidos em nossa natureza como seres humanos. Este seria um afastamento significativo do sentimentalismo construtivista que tenho sustentado até aqui. Há espaço para um sentimentalismo evolutivo, mas o que devemos fazer com esse ponto de vista? É possível que os nossos sentimentos morais sejam naturalmente direcionados para algumas coisas em detrimento de outras? É possível que este fato natural privilegie esses valores naturais mais do aqueles que emergem de processos culturais?

Neste capítulo vou levantar algumas dúvidas sobre esta sugestão. Vou admitir que somos biologicamente propensos a ter certos tipos de valores, mas vou negar que há uma moralidade inata (para mais discussão, ver Prinz, 2007; no prelo). Nossas predisposições biológicas não têm autoridade sobre os valores que têm uma origem cultural. As predisposições biológicas podem ser aprimoradas, ou mesmo revertidas, sob a influência da cultura. Além disso, vou argumentar que nossas predisposições biológicas não se qualificam como regras morais sem que haja uma elaboração cultural. A moralidade é inteiramente artificial. Tomado literalmente, o termo “ética evolucionista” é um mito. É o Romantismo renascido

como cientificismo crasso, não mais plausível do que *Übermensch* de Nietzsche e, talvez, não menos insidioso. Antes de chegar a essa crítica, vamos ver o que a evolução tem a oferecer.

7.1.2 Uma mão lava a outra

As discussões contemporâneas, em ética evolucionista, procuram classificar as teorias que estão atualmente em voga partindo da tese defendida por Herbert Spencer no século XIX. Spencer cunhou a frase “sobrevivência do mais apto”, procurando descrever o progresso social humano em termos evolutivos. Spencer (1857) acreditava que cada um de nós tem o direito inato de procurar a felicidade, desde que isso não implique interferência no arbítrio dos outros. Deste modo, leis que exigem que ajudemos os menos afortunados são qualificadas como violações deste princípio básico, impostas pelo Estado. Spencer era inflexivelmente contra leis que impunham a caridade, porque acreditava que tais leis violavam os direitos dos ricos. Ele também acreditava que tais leis promoviam a preguiça entre as pessoas que eram perfeitamente capazes de trabalhar e, ao mesmo tempo, diminuíam a empatia natural em relação às pessoas realmente necessitadas. Spencer sugeria que devemos ajudar aqueles que estão desamparados devido a circunstâncias alheias à sua vontade, mas, em trechos menos acautelados, observa que os animais de rebanho se beneficiam do fato de que o fraco e doente seja morto por predadores, permitindo, assim, que os mais aptos escapem à morte e prosperem.

O darwinismo social de Spencer pode ser interpretado como uma teoria moral. Spencer pensa que temos um sentido moral para discernir o certo e o errado, mas o certo e o errado que percebemos têm sua base na aptidão. Spencer pressupõe que todo o mal, seja natural ou moral, envolve uma incongruência entre um organismo e seu meio ambiente. Os males sociais surgiram porque evoluímos para lidar com as condições bárbaras e competitivas que antecederam a civilização. Agora, temos de evoluir para padrões de comportamento mais civilizados, e devemos deixar que aqueles que não conseguem se adaptar caiam

fora. É errado interferir no processo de seleção natural, porque a seleção natural é intrinsecamente progressiva. O mal naturalmente desaparece.

Os Eticistas evolucionários contemporâneos repudiam Spencer. Eles não pensam que os princípios darwinianos podem ser usados para justificar a ideia de que os fortes devem prosperar, deixando para trás os fracos. A "sobrevivência do mais apto" não deve ser usada para ditar a política social. Spencer tentou transformar a seleção natural em uma lei moral. Eticistas evolucionários contemporâneos normalmente consideram a seleção natural como moralmente neutra. Em vez de fazerem afirmações normativas, eles investigam a psicologia moral que a seleção conferiu. A evolução gerou uma capacidade moral em nós. Essa capacidade moral não é baseada no triunfo do mais forte sobre o mais fraco. Pelo contrário, ela é baseada no desejo de ajudar aqueles em necessidade, mesmo às nossas próprias custas. A evolução nos fez altruístas (ver Sober e Wilson, 1998).

Em face disso, esta conclusão parece bem diferente da moralidade egoísta de Spencer. Mas o contraste não deve ser exagerado. Spencer acreditava que, na adaptação à existência social, temos de evoluir para cuidar de quem precisa. A visão de Spencer sobre a política social foi impulsionada por sua moralização dos princípios darwinianos, mas seu relato sobre a psicologia moral certamente antecipava as tendências contemporâneas. A principal diferença entre Spencer e moralistas contemporâneos é que o primeiro considerava o altruísmo como uma adaptação à civilização, enquanto que os teóricos contemporâneos encontram origens mais antigas para ele.

O comportamento altruísta está em toda parte na natureza. Os zangões estripam-se, deixando seus ferrões em animais que ameaçam a colmeia. Polvos-fêmea deixam de se alimentar enquanto protegem seus ovos, e no momento em que os ovos eclodem, a mãe está geralmente tão fraca que morre ou acaba sendo vítima de um predador natural. Um esquilo dá sinal de alarme quando um predador se aproxima, tornando-se vulnerável justamente para salvar os outros de sua espécie. Esses comportamentos são inicialmente intrigantes do ponto de vista darwiniano. Se o comportamento é impulsionado pelos genes, por que os genes levam uma criatura a sacrificar-se em proveito dos outros? Tal criatura seria menos propensa a

viver até a idade reprodutiva, seus genes nobres teriam poucas chances de replicação. Este enigma do altruísmo tem uma solução simples (Hamilton, 1964). Em cada um dos casos aqui mencionados, a criatura altruísta é sacrificada pela segurança de seus parentes próximos. Genes que levam criaturas a proteger seus parentes são propensos a prevalecer porque os parentes compartilham a mesma carga genética. Se os genes de uma abelha fazem com que ela se sacrifique em prol de uma abelha geneticamente semelhante, cópias desses genes sobreviverão nesta outra abelha. Se o polvo mãe morre em prol de sua descendência, sua prole terá uma melhor chance de alcançar a maturidade e produzir mais cópias de genes da mãe. O mesmo vale para os esquilos que ajudam seus irmãos, um esquilo é mais propenso a ajudar o outro se os dois são intimamente relacionados. (Sherman, 1977).

O mesmo é provavelmente verdadeiro para nós. Preferimos ajudar parentes do que estranhos. Hamilton argumenta que a seleção de parentesco – ajudar indivíduos que carregam genes semelhantes – está no cerne do altruísmo humano. Esta tese pode até estar certa, apesar de ser amplamente considerada como incompleta. A dificuldade da tese está no fato de que nós, por vezes, efetivamente ajudamos estranhos. Esse fato reintroduz o enigma do altruísmo. Na verdade, a seleção de parentesco pode levar-nos a ajudar estranhos acidentalmente, p. ex., se o altruísmo de base biológica utiliza mecanismos cegos que não podem distinguir parentes de não-parentes (Joyce, 2006: 21), mas é difícil acreditar que a bondade para com estranhos é apenas uma cegueira biológica. Somos bondosos para não-parentes mesmo quando sabemos que eles não são parentes, e nossa bondade para não-parentes é consideravelmente maior do que o observado na maioria das outras espécies. A bondade humana é suficientemente generalizada ao ponto de se poder até pensar que ela seria contrária à seleção natural. Se você se sacrificar por alguém que não compartilha seus genes, você não faz nada para aumentar as chances de que seus genes sobrevivam. Isto levanta um problema sério para a hipótese de que os valores morais têm uma base biológica. Afinal de contas, os valores morais, muitas vezes, obrigam-nos a ajudar os não-parentes. Podemos até

mesmo questionar como esses valores surgiram, já que eles são tão contrários aos interesses dos nossos próprios genes.

Trivers (1971) desenvolveu uma resposta a este desafio. Em vez de assumir que evoluímos mecanismos de autossacrifício para ajudar estranhos, pode-se supor que evoluímos mecanismos para ajudar estranhos em contextos de reciprocidade. Se nós evoluímos um mecanismo de reciprocidade, então, ajudar um estranho é uma boa estratégia para conseguir a ajuda deste indivíduo em uma ocasião futura. Trivers chama isso de altruísmo recíproco. Desde que não nos matem no processo, ajudar estranhos pode aumentar nossas próprias chances de sobrevivência a longo prazo, se partirmos da premissa que os favores são retribuídos de forma confiável. Para obter esse tipo de comportamento, precisamos desenvolver tanto uma tendência a ajudar os outros quanto uma tendência a retribuir quando somos ajudados. Devemos também desenvolver uma capacidade para identificar, evitar e, talvez, punir aqueles que não retribuem. Ajudar alguém que não vai te ajudar de volta é uma estratégia perdedora. Neste caso, a população de altruístas recíprocos poderia ser tomada por uma população de caronistas.

Trivers fundamenta sua tese usando as ferramentas da teoria dos jogos. Considere o dilema do prisioneiro, no qual os jogadores podem cooperar ou não. Um desses jogos envolve a ajuda recíproca (figura 7.1). Se você e eu ajudamos um ao outro, nós dois vamos sair felizes, mas nós dois vamos ter que trabalhar para a nossa felicidade, uma mão lava a outra. Se eu pudesse fazer você me ajudar, sem ter que te ajudar como contrapartida, eu ficaria ainda mais feliz. Claro, o mesmo é verdade para você. Se nenhum de nós ajudar o outro, não obteremos qualquer recompensa, mas também não vamos ter de suportar os custos da ajuda. Se você examinar este arranjo de utilidades, classificadas numericamente na figura 7.1, notará que se recusar a prestar ajuda é a estratégia dominante. De saída, a estratégia dominante é não ajudar, não importa o que o outro jogador faça. Isso nada mais é do que a garantia de que nenhum de nós vai prestar ajuda ao outro. Uma mão não lavará a outra. Assim, a célula-matriz no canto inferior direito é o resultado, se fizermos a nossa decisão de uma forma puramente racional. Este é um resultado lamentável, porque nós dois sairíamos beneficiados se ajudássemos

um ao outro. Deste modo, a célula no canto superior esquerdo é mais desejável. O dilema se coloca porque sabemos que é melhor cooperar, mas é mais racional não cooperar, porque a iniciativa da cooperação é arriscada.

Figura 7.1 Utilidades esperadas em uma versão de uma mão lava a outra do dilema do prisioneiro

		Eu	
		Eu te ajudo	Eu não te ajudo
Você	Você me ajuda	2 para mim; 2 para você	3 para mim; 0 para você
	Você não me ajuda	0 para mim, 3 para você	1 para mim; 1 para você

Com esta estrutura de incentivos, a reciprocidade nunca poderia sair do chão. Ninguém seria confiável. Mas a reciprocidade é muitas vezes benéfica. Um mundo de cooperadores é melhor do que um de conspiradores. Mas isto cria um problema. Como podemos atingir interações cooperativas dadas as recompensas de deserção? Trivers supõe que este problema foi resolvido pelos genes. Evoluímos instintos cooperativos e intolerância com os trapaceiros. Isto efetivamente aumenta a recompensa de cooperação (sentir-se bem em ajudar) e aumenta o custo de deserção (sentir-se mal em não retribuir a ajuda). Se identificarmos e punirmos desertores, o peso da balança pode pender ainda mais para o lado da cooperação. Como a evolução seleciona características que são benéficas, teria selecionado aquelas características que promovem o altruísmo recíproco por meio da reciprocidade.

Esta história sobre a evolução do altruísmo tem sido a principal inspiração por trás de trabalhos recentes sobre ética evolucionista. Ele explica como evoluímos não somente para sermos prestativos, mas também para penalizar aqueles que não se preocupam com os outros de forma suficiente. Nós evoluímos para fazer o bem e para punir o mal. Esta história também se encaixa muito bem com o quadro *emocionista* que tenho defendido. Nós certamente não deliberamos sobre o futuro de nossos genes ao escolhermos nos comportar de forma altruísta. A vantagem genética de descendência é invisível para quem age com o objetivo de manter a linhagem genética. Na verdade, estaríamos inclinados em criticar uma pessoa que escolhe tratar os outros com gentileza só porque acredita que, agindo assim, beneficiaria seus genes. As tendências altruístas são, evidentemente, implementadas por mecanismos psicológicos mais simples. Tratamos os outros bem porque temos prazer em fazê-lo, e também porque ficamos angustiados quando vemos alguém em necessidade. Sentimo-nos culpados quando não retornamos favores, e sentimos raiva quando nossas boas ações não obtêm reciprocidade. Nós amamos a cooperação e detestamos a deserção. Eticistas evolucionários acreditam que são os genes que promovem o altruísmo fazendo com que seus portadores desenvolvam emoções que levam ao comportamento altruísta (Trivers, 1971; Frank, 1988). Deste modo, não faríamos boas ações como o resultado de cálculo egoísta das compensações genéticas; ao contrário, faríamos boas ações porque amamos o bem. De acordo com especialistas em ética evolucionária, esta é a essência da moralidade. A causa principal é egoísta, mas os mecanismos psicológicos imediatos não o são.

7.1.3 Além do altruísmo

Eticistas evolucionistas têm-se centrado sobre o altruísmo de reciprocidade. Eles acreditam que os desejos que nos impelem a incorrer em custos pessoais e ajudar outros indivíduos são a fonte biológica primária da moralidade. Este enfoque, entretanto, tem limitações em dois aspectos. Em primeiro lugar, a reciprocidade pode ser um termo genérico para vários tipos diferentes de comportamentos. No

entanto, é uma boa ideia considerar comportamentos separadamente. Em segundo lugar, ainda que a reciprocidade seja um aspecto importante do comportamento moral, certamente não esgota o domínio da moralidade. Existem outras normas morais que podem ter uma base biológica. Nesta seção vou apresentar uma gama de domínios normativos que podem ser informados pela seleção natural. Na próxima seção, vou começar a olhar para estes domínios de forma mais crítica. O ponto principal aqui é que os eticistas evolucionários não deveriam se restringir ao altruísmo ao procurar as raízes biológicas da moral.

O altruísmo pode ser definido como qualquer ato no qual um organismo confere um benefício para outro indivíduo enquanto incorre em um custo. O conceito de altruísmo recíproco é baseado na ideia de que tais favores podem ser retribuídos. Neste modelo, os atos altruístas são parte de trocas sociais. Um organismo recebe benefícios às custas de outro e, posteriormente, incorre em custos, conferindo benefícios a outrem. Esta caracterização muito abstrata explica porque o altruísmo pode ter evoluído, mas obscurece o fato de que as trocas sociais não se restringem a uma única capacidade comportamental. No mínimo, o altruísmo requer comportamentos que impliquem conferir benefícios e retribuir àqueles que nos conferiram vantagens no passado. Os evolucionistas afirmam que o altruísmo também requer uma capacidade de manter o controle de quem ainda não retribuiu – a capacidade de detectar trapaceiros. Em seres avançados podem haver, também, formas mais sofisticadas de manter um equilíbrio. Quando somos recompensados por custos que incorremos, esperamos que a recompensa seja equânime. A recompensa deve ser não só proporcional ao custo incorrido como, também, proporcional aos benefícios que os outros receberam quando tiveram custos semelhantes.

Já podemos ver que o altruísmo envolve uma série de capacidades: realizar e retribuir favores, detectar trapaceiros e avaliar a equidade. Essas capacidades são complexas e variadas. A avaliação da equidade envolve um controle de quais benefícios os dois indivíduos receberam e de relacioná-los aos custos recíprocos. Brosnan e de Waal (2003) argumentam que os macacos-prego são capazes de fazer isso. Se um macaco-prego recebe um pepino em troca da realização de uma

tarefa, e, em seguida, vê um outro macaco-prego receber uma fruta mais desejável para realizar a mesma tarefa, ele vai, na próxima vez, recusar-se a aceitar o pepino. Brosnan e de Waal interpretam isso como um protesto contra uma recompensa desigual (ver abaixo). Num contexto de altruísmo recíproco, um senso de equidade poderia ser utilizado para determinar se cada de reciprocidade é ou não proporcional ao que está sendo retribuído.

A detecção de trapaceiros é um componente importante e destacado no altruísmo recíproco. Se os animais altruístas não conseguissem detectar trapaceiros, seriam facilmente suplantados por uma população de caronistas. Capacidades especiais para detecção de trapaceiro foram observadas em morcegos hematófagos (Wilkinson, 1984). Os morcegos vampiros vivem do sangue; depois de um morcego encontrar uma refeição, regurgita sangue para compartilhar com os outros. Morcegos detectam trapaceiros procurando por membros da mesma espécie que estão com barriga inchada de alimento (sangue), mas que, entretanto, não participam do ato de regurgitar.

Equidade e detecção de trapaceiro são capacidades especializadas que servem de suporte para o altruísmo recíproco que, por sua vez, é constituído basicamente pelo receber e retribuir favores. Esta não é, de toda sorte, uma capacidade única. Ao introduzir altruísmo recíproco acima, eu casualmente mencionei dois tipos muito diferentes de casos. Por um lado, existem normas que nos conduzem a ajudar aqueles que estão em necessidade. Por outro lado, temos normas que regem empreendimentos em que há uma troca de recursos. Jim Moore (1984) refere-se a tais comportamentos como ajudar e compartilhar, respectivamente. Ajudar e compartilhar são semelhantes em um nível abstrato, mas são comportamentalmente distintas em outro nível. Fundamentalmente, o compartilhamento envolve a troca de bens, enquanto a ajuda, não. Presumivelmente, esses comportamentos distintos envolvem, por sua vez, alguns mecanismos psicológicos distintos. Por um lado, eles podem ser promovidos por diferentes emoções. Quando ajudamos alguém em necessidade, estamos normalmente motivados por sentimentos de preocupação ou angústia empática. Já na atividade de compartilhar, é mais provável que estejam presentes os

sentimentos de filiação ou de confiança. Por outro lado, ajudar e compartilhar subsumem diferentes faixas de comportamentos. A classe de comportamento de compartilhamento inclui dar presentes, prestar serviços e o comércio. As provas das raízes biológicas destes comportamentos vêm de observações de campo de chimpanzés. Os chimpanzés caçadores compartilham sua comida com o resto da tropa; tosadores tosam aqueles que os tosaram, e chimpanzés machos fazem escambo de alimentos por sexo (de Waal, 1997; Stanford et al, 1994). Comportamentos de ajuda são igualmente diversos e incluem nutrir indivíduos da mesma espécie que se encontram feridos, proteger contra agressões e ajudar outrem a atingir determinados objetivos. Estes comportamentos podem, também, ser observados em chimpanzés (De Waal, 1996). Tais comportamentos específicos provavelmente evoluíram separadamente, e podemos encontrar espécies que apresentam um tipo de altruísmo, mas não o outro. Devemos resistir à tentação de tratar altruísmo recíproco como se fosse implementado por um único comportamento. Os estudiosos da evolução da moralidade devem ter o cuidado de examinar toda a gama de comportamentos que envolvem conferir benefícios a não-parentes. O altruísmo recíproco é um princípio abstrato que explica como comportamentos de autossacrifício podem ter evoluído, mas, se alguém estiver interessado na origem de normas humanas específicas, ao invés de na origem da moralidade em geral, é essencial distinguir as diferentes formas que a reciprocidade pode tomar.

Em adição à gama complexa de comportamentos associados com o altruísmo recíproco, existem comportamentos moralmente relevantes que são difíceis de analisar em termos de reciprocidade. Agora, vou considerar dois outros domínios onde as regras morais podem ser evidenciadas. Em primeiro lugar, existem regras relativas a hierarquias de dominância social. A maioria das sociedades humanas é estratificada em um ou outro grau, de modo que o poder quase nunca é igualmente compartilhado. As violações de hierarquia social são consideradas, muitas vezes, muito graves. Temos emoções que subscrevem estas relações de hierarquia social, como deferência, respeito e, talvez, o sentimento de submissão. As violações de hierarquia social são corrigidas por um dos pontos

cardeais das emoções morais: o desprezo (Rozin et al, 1999). Temos desprezo por aqueles que desrespeitam as autoridades, por aqueles que disputam um *status* social imerecidamente e por aqueles que abusam do poder.

As normas que fixam hierarquias sociais quase certamente devem algo à evolução. As hierarquias de dominação são encontradas em muitas espécies, incluindo os grandes primatas. As sociedades de gorilas são organizadas em torno de um único macho prateado que tem um harém de várias fêmeas e acesso sexual exclusivo. Em chimpanzés, os grupos têm vários machos e fêmeas, e as fêmeas, que vêm inicialmente a partir de outros grupos, acasalam-se com a maioria dos machos. Mas machos dominantes têm maior acesso às fêmeas, e, ocasionalmente, exigem acesso exclusivo por curtos períodos. As fêmeas são dominadas pelos machos, e, entre machos e fêmeas, há uma hierarquia de classificação. Entre os bonobos, a estrutura de poder é revertida com as fêmeas dominando os machos (de Waal e Lanting, 1997). Todos os macacos estabelecem e mantêm suas hierarquias de dominação pela força bruta, embora os bonobos sejam relativamente dóceis.

As normas de hierarquia envolvem padrões de deferência. Aqueles que são subordinados devem se curvar, muitas vezes literalmente, para os animais dominantes. Aos animais no topo da classificação são concedidos vários privilégios. Mas eles não necessariamente recebem tais privilégios sem dar algo em troca. A posição na hierarquia social pode interagir com norma de reciprocidade. Animais que ocupam altas posições na hierarquia social podem retribuir favores feitos por animais de baixa patente e, em chimpanzés, animais de nível superior, na verdade, compartilham mais do que os animais de baixa patente. Isso pode ajudar a sustentar a sua autoridade. Se todo mundo está em dívida com o macho alfa, a insurreição é menos provável.

Menos atenção tem sido dada para a evolução da posição do que para a evolução da reciprocidade, mas é fácil de dar um giro para histórias evolutivas sobre a hierarquia social. Quando os membros de uma espécie social perseguem as coisas que eles desejam, os conflitos surgem. Se esses conflitos persistem ou se aprofundam, a espécie vai se destruir. Para ser bem-sucedido, um grupo social precisa de mecanismos para impedir a desgraça. Esses mecanismos podem

assumir diferentes formas. Uma espécie pode evoluir no sentido de ser igualitária. Isso implicaria, presumivelmente, um resultado bastante estável, mas, por outro lado, exigiria surpreendente contenção de desejos. Outra adaptação seria manifestar a contenção quando e somente quando nossos esforços para alcançar os desejos fossem frustrados por um animal mais poderoso. Se criaturas evoluíram a capacidade de ceder quando derrotadas, em consequência, alguns animais têm o poder de dominar os outros. Um mundo de predadores mútuos pode evitar a desgraça da espécie se os animais mais fracos respeitarem a autoridade. Da mesma forma como Hobbes argumentava que um soberano pode proteger as pessoas de assassinato e evitar o caos sem fim, a deferência a um animal alfa pode impedir sociedades animais de destruir a si mesmas.

Outro lugar onde a moralidade parece ter um impulso a partir dos genes é o domínio das relações sexuais. O sexo é altamente moralizado. Normas sexuais podem ser encontradas em todo o mundo e em todo o reino animal. Vou me concentrar em normas que regulam quem pode acasalar com quem. Há também normas referentes a quando o sexo pode acontecer, e que tipos de atos sexuais são apropriados, mas estas têm sido menos bem investigadas em um contexto evolutivo. Muitas espécies têm preferências por posições sexuais, e muitas restringem o sexo para os períodos em que as fêmeas estão no cio, mas nós, ao exemplo dos bonobos, parecemos ser relativamente livres em relação a tais restrições impostas biologicamente. No entanto, nós não somos completamente livres no que diz respeito às restrições à seleção de parceiros sexuais. Dois tipos de normas parecem ser especialmente bem enraizadas. Em muitos casos, não estamos autorizados a escolher parceiros sexuais que já são sexualmente envolvidos com outra pessoa, e, entre os parceiros sexuais não envolvidos com outra pessoa, não estamos autorizados a ter relações sexuais com parentes próximos. Temos normas de infidelidade e normas de incesto.

Vamos começar com a infidelidade. Ninguém gosta de ser traído, mas raramente ocorre-nos que a preocupação com a fidelidade sexual possa ser uma resposta biologicamente programada. Por que teríamos desenvolvido o comportamento de odiar a infidelidade? Uma resposta é capturada na própria

palavra "trapaça". Um trapaceiro, resumindo, é alguém que colhe benefícios sem pagar os custos necessários. Psicólogos evolucionistas preferem ver as relações sexuais como uma forma de troca na qual normalmente, argumentam, homens e mulheres enfrentam custos e benefícios distintos (por exemplo, Buss, 2000). Os homens, por esse ponto de vista, oferecem apoio e proteção em troca de acesso sexual exclusivo. As mulheres, por sua vez, fornecem acesso sexual exclusivo em troca de apoio. (Psicólogos evolucionistas estão dispostos a aceitar que a biologia é, por vezes, politicamente incorreta). A exclusividade sexual é especialmente importante para os homens. Se uma mulher está sexualmente envolvida com alguém que não seja seu principal parceiro, sua prole pode acabar sendo geneticamente alheia ao seu principal parceiro. Se ele investe naquela prole que não é sua, terá garantido a sobrevivência dos genes do adúltero às custas da sua própria sobrevivência. Isso poderia levar à proliferação de adúlteros. Tal fato não seria uma boa notícia nem para os homens, nem para as mulheres. Os adúlteros iriam, por sua natureza, ser menos propensos a oferecer alianças de proteção estáveis e de longo prazo. Se todos os homens fossem adúlteros, as mulheres não poderiam contar com seu apoio. Isso diminuiria as taxas de sobrevivência feminina, o que seria um mal presságio para a espécie. Assim, a evolução deve selecionar os homens que evitam a infidelidade.

As normas de infidelidade são subscritas pelo ciúme romântico. Quando nós suspeitamos que nosso(a) parceiro(a) tem sido infiel, ficamos enciumados. Psicólogos evolucionistas procuram apoiar suas conjecturas acerca do fundamento biológico das normas de infidelidade examinando diferenças de gênero no que diz respeito ao ciúme. Se homens e mulheres têm diferentes motivos, quando se trata de relações sexuais, o seu padrão de ciúme deve também diferir. As mulheres buscam apoio e os homens buscam acesso sexual exclusivo. Buss et al. (1992) ponderam que os homens tendem a ser mais afetados pela infidelidade sexual do que pela infidelidade emocional. Os homens são biologicamente programados para se proteger contra o erro de investir em genes de outras pessoas. Se a parceira sexual de um homem tem relações sexuais com outros, ele não poderá ter certeza que seus filhos carregam seus genes. As mulheres nunca precisam se preocupar

com o fato de, acidentalmente, investir em genes de outra pessoa. Qualquer bebê que uma mulher tiver, certamente, será dela! Mas as mulheres precisam se preocupar em perder o apoio de seus parceiros masculinos. A infidelidade emocional representa um risco maior, a este respeito, que a infidelidade sexual. Como previsto, Buss et al. (1992) constataram que homens e mulheres, de fato, respondem de forma diferente à questão acerca do que é pior, se a infidelidade emocional ou a infidelidade sexual. Os homens são muito mais propensos do que as mulheres a dizer que a infidelidade sexual é pior do que a infidelidade emocional. Isso pode ajudar a explicar um duplo padrão generalizado. A infidelidade sexual masculina é bem mais tolerada do que a infidelidade sexual feminina; muito mais sociedades são polígamas (mais de 80 por cento) do que poliândricas (menos de 1 por cento).

A biologia provavelmente também contribui para atitudes morais em relação ao incesto. Evitar o incesto é muito comum entre os animais não-humanos, incluindo a maioria de nossos parentes próximos. Os babuínos expulsam os jovens do sexo masculino de suas tropas, forçando-os a se reproduzir em outros lugares. Um padrão similar é observado em gorilas; e os jovens chimpanzés machos também abandonam suas tropas iniciais. Goodall (1986) relata que os chimpanzés machos adolescentes, por vezes, investem em direção a suas mães, mas estas os repelem. O mesmo acontece entre os macacos. Incesto de mãe e filho é ainda mais raro entre os bonobos, que têm um comportamento notoriamente hipersexual (de Waal e Lanting, 1997). O incesto de pai e filha é mais difícil de rastrear entre macacos, porque os pais tendem a não participar na educação dos filhos e muitos macacos são sexualmente promíscuos. Os pais na espécie Gibbon têm o comportamento de acasalar com as suas filhas se outras fêmeas não estão disponíveis, mas esta não é a opção padrão (Bagemihl, 1999). Entre os seres humanos, a prevenção do incesto tornou-se um tabu completo. Em muitas culturas, o incesto é estritamente proibido.

Por que nós desenvolvemos o comportamento de evitar o incesto? Uma possível explicação consiste em evidenciar que evitar o incesto é um subproduto do comportamento de evitar a infidelidade. Incesto de pai e filha é tipicamente uma

forma de infidelidade. Quando um pai e uma filha se envolvem romanticamente, os dois indivíduos estão traindo o outro progenitor. Isso não explicaria as proibições contra incesto de irmão e irmã, no entanto. A explicação mais comum envolve os riscos genéticos da endogamia. Evidências sugerem que a prole de dois membros da família imediata são muitas vezes menos apta do que os filhos de parentes mais distantes. A endogamia permite o surgimento de traços recessivos; e traços recessivos tendem a ser mais prejudiciais do que os traços dominantes. Os traços dominantes estão sob uma mais pesada pressão de seleção, justamente porque são mais propensos a ditar a expressão fenotípica. Quando os animais com os mesmos traços recessivos se reproduzem, muitas vezes isto é uma má notícia para a prole. As espécies que evitam a consanguinidade também são mais propensas a ter uma prole com novas combinações de genes, e esta diversidade genética pode diminuir as chances de uma completa distinção e promover a adaptação. Se uma catástrofe ambiental ameaça indivíduos com uma composição genética particular, outros indivíduos com uma composição diferente têm mais probabilidades de sobreviver.

A existência de uma generalizada atitude de evitar o incesto suscita a questão de como implementá-la. Para evitar o acasalamento de parentes, precisamos de uma maneira de saber quem são eles. Para os seres humanos contemporâneos esta não é uma tarefa das mais difíceis. Mantemos não só uma estrita faixa de relações de parentesco, como também demonstramos um estrito entendimento do conceito de relação genética. Somos capazes de distinguir conceitualmente aqueles que compartilham nossos genes daqueles que não o fazem. Se evitar o incesto é uma resposta que decorre da evolução, no entanto, ela precisa ser implementada de uma forma que não requeira um grande conhecimento da ciência moderna. Westermarck (1891) propôs um mecanismo simples. Ele especulou que o ato de evitar o incesto pode ser desencadeado pela coabitação durante a infância: duas pessoas que crescem na mesma casa não terão interesse sexual uma em relação à outra. Esta proposta goza de algum apoio. Um estudo descobriu que os não-parentes criados juntos em Kibutzim israelenses foram significativamente menos propensos a se casar com uns com os outros do que

propensos a se casar com aqueles de fora da comunidade (Sefer, 1971). Efeitos similares foram encontrados no norte de Taiwan, onde há um sistema de "casamentos de infantes" (Wolf, 1970). Em vez de arranjar um casamento para seus filhos na idade adulta (um casamento de "maiores de idade"), os pais, às vezes, adotam uma filha e a tornam parte da família com a intenção de organizar um casamento entre ela e seu filho mais tarde. Casamentos de menores de idade tendem a ser menos bem-sucedidos do que os realizados entre maiores de idade; eles são três vezes mais propensos a terminar em divórcio (Wolf e Huang, 1980). Recentemente, a hipótese de Westermarck foi submetida a um teste mais direto. Lieberman et al. (2003) questionaram entre estudantes universitários qual era sua opinião acerca do casamento entre irmão e irmã. Os indivíduos que tinham vivido com irmãos do sexo oposto eram mais refratários do que aqueles que não o tinham; tal padrão se confirmou mesmo nos casos em que, além de adotados, os irmãos eram do mesmo sexo e não possuíam qualquer relação genética. O estudo descobriu que os homens que tinham vivido com as irmãs durante seus anos de adolescência eram especialmente inclinados a desaprovar o incesto entre irmãos.

Considero que algumas dessas evidências necessitam de uma reavaliação crítica. Abaixo, vou reconsiderar algumas das reivindicações que acabei de relatar; vou argumentar que as contribuições biológicas à moralidade, foram, na ocasião, erroneamente identificadas ou exageradas. Entretanto, mesmo que os resultados específicos não resistam a uma análise mais acurada, a mensagem básica é difícil negar. A seleção natural provavelmente nos forneceu uma variedade de disposições comportamentais e afetivas que contribuem para o surgimento de valores morais. Discuti três domínios principais: altruísmo (inclusive ajudar e compartilhar), hierarquia social e relação entre sexos. Esses domínios são exemplos dos três tipos de sistemas éticos identificados no trabalho antropológico de Richard Shweder e seus colegas (Shweder et al., 1997). Como vimos no capítulo 2, Shweder considera que as regras morais encontradas por todo o mundo tendem a dividir-se em três categorias. Algumas regras envolvem o domínio da autonomia: proibimos crimes contra pessoas individuais. As violações de reciprocidade geralmente se enquadram nesta categoria. A trapaça é considerada um mal contra uma pessoa. A

resposta emocional normal para a atitude dos caronistas é a raiva. Violações de hierarquia social são o que Shweder chamaria crimes contra a comunidade. As hierarquias sociais são os princípios básicos de organização das comunidades, de modo que o ato de violar uma regra hierárquica constitui-se numa ameaça a esta organização. A resposta emocional padrão para quem desrespeita a ordem social é o desprezo. Aqueles que desrespeitam as regras de um tribunal, por exemplo, são detidos por desacato. Aqueles que consideram que as autoridades vigentes são indignas do posto que ocupam dizem ter desprezo pelas mesmas. As violações das normas sexuais, por sua vez, seriam atos antinaturais. Shweder enquadra-os em sua ética da divindade, que, conforme sugeri, poderia ser melhor considerada como uma ética da ordem natural. O incesto é um "ato não natural" e é visto com repulsa. A infidelidade pode evocar nojo, bem como trazer à tona pensamentos de contaminação, mas é também um crime contra a pessoa cuja confiança é violada. Assim, a infidelidade é uma violação tanto de uma norma da natureza, quanto de uma norma de autonomia individual.

Evidentemente, esta não pretende ser uma lista exaustiva de normas que têm alguma base nos genes. Pode haver outras. De fato, há um sentido trivial em que cada norma que temos deve algo à nossa composição biológica. A pesquisa aqui destaca os casos em que a contribuição biológica para a moralidade tem sido especialmente bem investigada. O altruísmo, o que atraiu a maior atenção, não é o único domínio passível de uma análise evolutiva. Eticistas evolucionários deveriam procurar diversificar sua abordagem e se sentir confortáveis em saber que cada uma das principais classes de regras éticas, no modelo de Shweder, tem a possibilidade de ensejar algum apoio biológico.

7.2 LIMITAÇÕES DE ÉTICA EVOLUCIONÁRIA

7.2.1 O Natural implica o Bem?

Vamos supor que todas as conjecturas evolutivas na seção 7.1 estejam corretas. Regras relativas à reciprocidade, hierarquia social e acasalamento são

produtos da seleção natural. Estes são domínios centrais da moralidade humana. Se eles estão evoluindo, então a ética é, em parte, natural. Apesar de toda a variação cultural sondada nos capítulos 5 e 6, podemos encontrar um núcleo comum escrito pelo *design* darwiniano em nossos bioprogramas. Talvez esses aspectos de nossas vidas morais sejam os principais. Talvez possamos dispensar as contaminantes culturais e basear nossa moralidade inteiramente em normas naturais. Esta pode ser uma estratégia para alcançar a objetividade. As reivindicações antiobjetivistas do capítulo 4 podem ter sido prematuras.

A fim de sustentar esta tese, deveríamos admitir que as normas derivadas da natureza teriam que ter algum tipo especial de autoridade. Elas teriam que ser boas em virtude de serem naturais. Existe alguma maneira para defender essa moralização da natureza? Eticistas evolucionários não oferecem uma resposta uniforme. Eles estão divididos sobre o *status* de normas naturais. Alguns acreditam que a origem evolutiva dos valores morais pode ser usada para apoiar uma conclusão cética (por exemplo, Ruse 1991; Joyce, 2001). Ruse (1991: 506) afirma que "A moralidade não é mais do que uma ilusão coletiva causada por nossos genes." A moralidade é uma ilusão, sob este ponto de vista, porque supomos que o nosso senso de obrigação moral deriva de uma fonte mais profunda. Achamos que estamos sob algum tipo de obrigação especial de ajudar os outros, e pensamos que esta obrigação é independente dos fatos contingentes referentes à nossa própria história como espécie. Ruse especula que este pressuposto objetivista está arraigado; somos biologicamente obrigados a acreditar que as afirmações morais são objetivas e verdadeiras. Mas essa crença é uma ilusão. O fato de que a moralidade tem uma base na natureza humana explica porque nós fazemos julgamentos morais, e, uma vez que percebemos que tais julgamentos são arraigados, não precisamos nos preocupar postulando fatos morais.

Em contraste com o ceticismo de Ruse, podemos encontrar posições que levam muito a sério os valores morais com base evolucionária. Alguém poderia argumentar que qualquer norma que é o resultado da evolução é verdadeira, em algum sentido, e privilegiada sobre as normas que são culturalmente construídas. Esse tipo de pensamento é evidente em Spencer e Nietzsche, mas também pode ser

dada uma interpretação e uma defesa mais modernas. Sugerir que tanto Spencer quanto Nietzsche foram acusados de fazer uma inferência bastante inconsistente na passagem da natureza para o bem. Mas existem maneiras de dotar tal inferência com um grau de plausibilidade. Vamos supor que alguns dos nossos valores são produto da seleção natural. A seleção natural funciona através de um princípio de adequação. Os traços inatos que herdamos foram traços que se mostram bem-sucedidos em nossos antepassados. Se uma norma se desenvolveu através da seleção natural, então isso se deu em razão de conferir aptidão. Se as normas são desenvolvidas como um reforço da adaptação, então, talvez, seja o caso de dotá-las com autoridade normativa. Devemos fazer o que é bom para nós. Nesta leitura, a ética evolucionária torna a obrigação moral um caso especial de obrigação prudencial. Longe de minar a moralidade, a evolução é o que confere à moral validade indubitável. A prudência é a regra de ouro subjacente à moeda de convicções morais. As coisas estão selecionadas por suas possibilidades de retornos positivos, e, portanto, as coisas que nós naturalmente consideramos boas realmente o são, em um sentido prudencial. Não tenho certeza se alguém realmente tem defendido esta linha de argumentação, mas é evidente que ela tem algum apelo intuitivo.

Assim, uma abordagem evolutiva pode levar tanto ao ceticismo quanto ao realismo, dependendo de como ela é descrita. Realistas e céticos discordam sobre se fatos naturais implicam o bem, mas muitas vezes eles parecem concordar no que diz respeito a uma questão específica. Eles pensam que as normas lastreadas na evolução são superiores em relação a normas culturais. Ruse pensa que elas são superiores porque são fixas e inevitáveis (entretanto, veja Joyce, 2001). Os realistas darwinianos, por sua vez, acreditam que as normas que decorrem da evolução natural são superiores porque são produto da adaptação ao invés de um *fiat*.

Penso que tanto os darwinianos céticos quanto os darwinianos realistas estão enganados. Vamos começar com o ceticismo. Penso que a versão de Ruse sofre de duas falhas fatais. Primeiro, Ruse está errado em pensar que as crenças morais que derivam da evolução natural são fixas. Como vou mostrar a seguir, há

muitas razões para acreditar que podemos alterar, e até mesmo derrubar, através do processo de aculturação, intuições valorativas derivadas da evolução. Se eu estiver certo, isso mina a melhor razão do cético para pensar que as normas derivadas da evolução são prioritárias. Em segundo lugar, rejeito o argumento de Ruse que leva à conclusão de que as máximas morais são falsas. Em uma reconstrução, o raciocínio é o seguinte: a evolução nos fornece crenças sobre o que é bom; uma vez que essas crenças são explicadas, não há nenhuma razão para postular uma classe de verdades morais subjacentes a essas crenças; portanto, nossas crenças morais não se referem a fatos morais. Eu tenho problema com a segunda premissa. No capítulo 3, argumentei que os conceitos morais representam propriedades dependentes de resposta. O bem é aquilo que causa a aprovação moral. Se a evolução nos dá crenças morais, da mesma forma, produz uma classe de fatos morais.

É interessante comparar a estratégia de Ruse à estratégia genealógica de Nietzsche, que endosseí, de forma qualificada, no capítulo 6. Assim como Nietzsche, Ruse está tentando desacreditar a moralidade, mostrando que os valores morais têm uma origem histórica (não na escala da história humana, mas na escala da história evolutiva). Ao discutir Nietzsche, eu disse que análises históricas podem nos ajudar a avaliar se vale ou não a pena manter as normas, e o mesmo princípio se aplica aqui. Se uma norma decorrente da evolução não é mais instrumentalmente justificável, então poderíamos considerá-la obsoleta. Mas esse não é o caminho trilhado por Ruse, no que diz respeito à evolução, para criticar a moralidade. Ruse difere de Nietzsche em dois aspectos fundamentais. Primeiro, ele usa a história (da nossa espécie) para argumentar que os valores morais são bons para nós, ao passo que Nietzsche usa a história (da nossa cultura) para argumentar que os valores morais são ruins para nós. Em segundo lugar, Ruse está tentando mostrar que as crenças morais são falsas, enquanto que Nietzsche pode ser interpretado como um realista moral – as regras morais são construções sociais que são verdadeiras da mesma maneira que os fatos sociais, tais como fatos econômicos, são verdadeiros. Penso que Ruse está errado em ambos os aspectos. Eu já sugeri que ele está errado em pensar que as crenças morais são falsas; a

verdade moral é assegurada pelo fato de que os conceitos morais referem-se a propriedades dependentes de resposta. A suposição de que os valores morais decorrentes da evolução são, em si, bons para nós, também pode ser contestada. Ironicamente, Ruse, um cético darwiniano, compartilha esta suposição com realistas darwinianos. Então vamos voltar a essa abordagem agora.

De acordo com o realismo darwiniano, nossas crenças morais decorrentes da evolução são verdadeiras e superiores às normas de base cultural, pois surgiram através da seleção natural e esta, por sua vez, é a um aprimoramento da adaptação. Não tenho certeza se alguém tem defendido seriamente este ponto de vista. Richards (1986) defende a ideia relacionada, segundo a qual normas que decorrem da evolução são moralmente legitimadas porque nos levam a cuidar dos membros de nossas comunidades (ver também Joyce, 2006, para uma crítica convincente a esta e a outras tentativas de defender a autoridade normativa de normas que decorrem da evolução). O que tenho em mente aqui é mais geral do que isso. É a ideia de que a adaptação é um bem intrínseco, e quaisquer normas que decorrem da evolução por meio da seleção natural são, assim, concebidas para aumentar o bem. É intuitivamente plausível que esse fato legitima as normas que decorrem da evolução. Apesar de seu recurso, o realismo darwiniano colapsa quando sob escrutínio, e é instrutivo ver o porquê.

Em primeiro lugar, a evolução não otimiza. Não produz traços que são melhores, mas apenas traços que são bons o suficiente. Existem muitos sistemas de regras que podem levar ao sucesso reprodutivo. O que acontece é que as regras que herdamos através da evolução são suficientemente eficazes para o acasalamento e criação dos filhos, mas, se o mérito prudencial for o padrão, é certo que podemos fazer melhor. Existem regras não-naturais que podem aumentar o nosso desempenho. Considere a hierarquia. Podemos ter evoluído para disputar a hierarquia social por meio do confronto agressivo com nossos superiores. Também podemos ter evoluído para aceitar respeitosamente a nossa posição na hierarquia social, quando os indivíduos mais fortes tomam um pedaço maior do bolo. Não está claro que qualquer uma destas opções polares é maximamente ajustada. Talvez seria melhor eliminar a classificação social e dividir tudo em partes iguais,

independentemente dos méritos físicos. Os realistas darwinianos podem até estar corretos ao dizer que o natural implica o bom (no sentido de aptidão), mas essa descoberta realmente solapa qualquer estatuto privilegiado que as normas naturalmente evoluídas poderiam ter, porque natural, aqui, não implica melhor. Os valores culturalmente elaborados podem ser mais benéficos.

Em segundo lugar, a aptidão biológica nem sempre é vantajosa para nós enquanto indivíduos. A aptidão biológica é mais fundamentalmente definida em termos de aptidão genética: o que permite que genes possam se replicar. O que é bom para os nossos genes não é necessariamente bom para nós. Joyce (comunicação pessoal) dá o exemplo do parricídio: se você matar os seus pais, a fim de aumentar suas chances de sobrevivência, você, conseqüentemente, aumentará a aptidão de genes de seus pais, mas você certamente não estará aumentando a aptidão de seus pais.

Em terceiro lugar, mesmo que os traços evolutivos tenham aumentado a aptidão no passado, podem não ser vantajosos agora: o que era adequado para os nossos antepassados não é necessariamente adequado para nós. Como os céticos têm apontado, a evolução tende a fornecer normas que são bastante paroquiais. Se o altruísmo é impulsionado pela reciprocidade, então somos provavelmente biologicamente inclinados a ajudar apenas aqueles que podem retribuir. Podemos ter evoluído uma tendência para nos esquivar de ajudar o fraco, o pobre, ou aqueles que vivem em terras distantes. Essa política pode ter sido vantajosa no passado evolutivo, mas as sociedades têm crescido, e, para alcançar a estabilidade social em grandes grupos (ou em uma rede global de grupos), pode ser de vital importância adotar a conduta de respeitar os interesses daqueles que não podem retribuir. O principal incentivo para aderir a um grande coletivo social é que ele irá nos oferecer proteção em tempos de adversidade.

Em quarto lugar, os realistas darwinianos fundem duas modalidades do bem. Eles procuram reduzir o bem moral à bondade biológica; entretanto, adotar essa conduta é o mesmo que mudar o tópico. Pois mesmo se os bens morais são biologicamente bons (ou seja, aumentam a aptidão), podemos ainda nos questionar se são moralmente valiosos. O conceito de bondade biológica tem um sentido

diferente em relação à bondade moral. A conclusão de que ambos os sentidos são diferentes pode ser inferida a partir do fato de que o conceito padrão de bondade moral tem uma origem motivacional. Para que alguém possa julgar que algo é moralmente bom, deve ser motivado a fazê-lo. Por outro lado, podemos julgar que algo iria aumentar nosso sucesso reprodutivo (não usar o controle de natalidade, por exemplo) sem ser motivado a fazê-lo. O fato de que estes conceitos se relacionam de forma diferente pode ser inferido a partir do fato de que nossa categoria de bens morais pode ser alterada pela cultura. Talvez os nossos antepassados do Pleistoceno moralizassem somente aquelas coisas que eram reforços adaptativos. Nós, entretanto, não restringimos a moralidade desta maneira e, conseqüentemente, a análise evolutiva da origem da moral não é capaz de oferecer uma análise adequada do que a moralidade passou a incluir. Em suma, mesmo que natural implique bom, não implica moralmente bom.

O resultado de tudo isso é que não podemos extrair nenhuma inferência a partir da existência de uma norma derivada da evolução; quer sua falsidade, quer a sua verdade (compare com Sober, 1994). A evolução não confere *status* especial, positivo ou negativo, à moralidade. Assim, o projeto que consiste em distinguir normas naturais de normas culturalmente forjadas, o qual Nietzsche encorajava, pode não valer o esforço. Na próxima seção, vou argumentar que tal distinção é ainda mais polêmica do que eu já apontei até agora. Vou argumentar que não há utilidade na distinção entre normas naturais e culturais, porque a cultura molda todas as normas, mesmo quando elas têm uma base natural. Vou argumentar também que nenhuma norma derivada da evolução se qualificam como normas morais; assim, estritamente falando, não existe tal coisa como uma ética evolutiva.

7.2.2 Os primatas têm um senso moral?

Na seção 7.1, ofereci uma lista de normas derivadas da evolução natural que contribuem para a moralidade. Nesta seção, quero examinar de forma mais crítica os itens desta lista. Não vou retirar minhas afirmações sobre normas derivadas da evolução natural, mas vou reformulá-las. Vou argumentar que as normas derivadas

da evolução natural não constituem uma moralidade inata. Elas são, em vez disso, as restrições flexíveis dentro das quais a moralidade emerge. Nesta seção, vou me concentrar em pesquisas comparativas. Vou argumentar que os investigadores não estabeleceram a existência de capacidades morais em primatas não-humanos. Esta falha em encontrar tais capacidades enfraquece o caso de um senso moral baseado na evolução.

Vamos começar com o altruísmo recíproco, que supostamente subscreve as normas morais que têm a ver com a partilha e com a ajuda. A primeira coisa a notar é que o altruísmo recíproco é realmente raro em outros animais (Pusey e Packer, 1997). Ele porta pré-requisitos cognitivos bastante exigentes, o que pode ser difícil para a maioria das criaturas alcançar (Stevens e Hauser, 2004). Mas certamente há evidências de altruísmo recíproco em primatas não-humanos. Primatas e macacos retribuem, ajudam e compartilham. Os primatas têm claramente tendências evolutivas consistentes com o comportamento altruísta. Mas, como Sober e Wilson (1998) apontam, não podemos inferir motivações altruístas do comportamento altruísta. Os primatas poderiam até mesmo ser impulsionados por um desejo moral de fazer coisas boas, ou poderiam ser impulsionados por preocupações menos nobres. Eu acho que a última possibilidade é mais provável.

Há várias interpretações do comportamento altruísta que não necessitam da postulação de motivações altruístas. Em primeiro lugar, os primatas podem ser fazer coisas boas um para o outro, a partir de um desejo egoísta de obter recompensas. Por exemplo, há pouca razão para pensar que os macacos são movidos pelo altruísmo quando trocam bens, porque o doador recebe uma recompensa em troca (por exemplo, a troca de alimentos por sexo nos bonobos). Neste contexto, é importante notar que os psicopatas envolvem-se em ações de troca recíproca quando participam dos jogos do dilema do prisioneiro (Widom, 1976). Os psicopatas não são motivados, provavelmente, por uma preocupação com a equidade; eles simplesmente reconhecem que, fazendo um favor para outra pessoa, podem obter retornos mais elevados. Em segundo lugar, alguns casos de aparente altruísmo podem ocorrer sob coação. De acordo com uma teoria, a partilha entre os macacos é, na realidade, um furto tolerado; os macacos

permitiriam que membros da mesma espécie furtassem a fim de evitar assaltos violentos (Blurton-Jones, 1987). Em terceiro lugar, alguns atos de altruísmo aparente podem ser uma estratégia para alcançar popularidade quando os recursos excedentes estão disponíveis. Os machos alfa frequentemente distribuem mais alimentos do que outros macacos, e isso os ajuda a ser visto com bons olhos pelos membros do bando. Quando os macacos compartilham, eles não estão necessariamente incorrendo em perdas graves. Se os macacos fossem realmente impulsionados por preocupações altruístas, poderíamos esperar que eles fossem mais inclinados ao autossacrifício; poderíamos também esperar que eles fossem mais imparciais em seus esforços para compartilhar. Até mesmo os chimpanzés mais generosos doam para apenas cerca da metade daqueles que pedem comida (de Waal, 1996). Mais notavelmente, Silk et al. (2005) mostraram que os chimpanzés não partilham de forma confiável em condições em que a partilha não tem custo. Os autores criaram um aparelho em que um chimpanzé poderia puxar uma corda para obter uma recompensa alimentar, ou puxar uma corda diferente para obter essa mesma recompensa alimentar e entregar comida para outro chimpanzé. Os chimpanzés do estudo não mostraram preferência pela corda generosa sobre a corda egoísta, mesmo que não houvesse diferença em recompensas pessoais, e isso apesar do fato de que todos os chimpanzés viveram juntos por 15 anos.

Os primatas não-humanos demonstram, além disso, padrões de ajuda mútua estranhamente seletivos. Eles ajudam aqueles com quem formaram vínculos fortes, mas eles ocasionalmente demonstram o espetáculo bizarro da indiferença para com o sofrimento de indivíduos da mesma espécie. De Waal (1996) relata um episódio em que um reso adulto ameaçou bater em um reso bebê que gritava porque seu braço estava preso nos raios de uma roda de fiar. Embora episódico, tais exemplos levantam questões sobre se os primatas não-humanos ajudam uns aos outros por causa de uma preocupação moral sentida profundamente. Sua tendência para conferir benefícios pode ser movida por motivos que não iríamos considerar especialmente louváveis. Para provar o contrário, seria importante mostrar que os primatas ajudam regularmente uns aos outros quando há pouca

chance de reciprocidade. Isto, tanto quanto eu sei, não foi mostrado. Em contraste, os seres humanos freqüentemente ajudam os outros quando não há reciprocidade. Nós enviamos cheques de doação a organizações de caridade distantes, protestamos contra injustiças que não nos afetam e damos apoio a intervenções humanitárias em terras distantes. Às vezes, esses atos são contabilizados como "reciprocidade indireta" (Alexander, 1987), com o pensamento de que, ajudando os outros distantes, nós queremos melhorar a nossa reputação e obter benefícios em casa. Talvez; mas certamente tais comportamentos não são explicitamente motivados pelo desejo de boa aparência (ninguém vê quando você envia cheques para a Oxfam), e eles provavelmente são implementados pela preocupação moral. O fato de que nós não vemos esse tipo de comportamento regularmente em outras espécies de primatas sugere que a preocupação moral pode ser exclusivamente humana, e se é exclusivamente humana, então não podemos usar o método comparativo para mostrar que essa característica é derivada da evolução natural, mais do que culturalmente construída.

Dúvidas semelhantes podem ser levantadas sobre o sentido primaz da justiça. Vale lembrar aqui o estudo de Brosnan e de Waal (2003) em que um macaco-prego recusou um prêmio de pepino depois de ver um outro macaco receber uma uva para o mesmo trabalho. Ao invés de interpretar isso como um sentimento nascente de desigualdade social, pode-se supor que os macacos simplesmente renegam as recompensas medíocres quando vêem que melhores recompensas podem estar disponíveis. Este é um efeito bem conhecido na etologia. Tinklepaugh (1928) mostrou que os macacos recusam uma recompensa alimentar desejável de outro modo (alface) quando uma recompensa mais desejável está disponível (bananas). Será que você não deixaria um biscoito de água e sal por um pedaço de bolo de chocolate? Esta interpretação mais modesta dos resultados de Brosnan e de Waal é apoiada pelo fato de que os macacos-prego renegam sua recompensa de pepino em uma condição de controle em que, ao invés de ser dada a outro macaco, uma uva é colocada em uma gaiola vazia nas proximidades (Wynne, 2004).

Portanto, é preciso ter cautela ao interpretar o comportamento animal. Há uma tentação de atribuir motivos humanos, como até mesmo nos casos em que outras explicações estão disponíveis. Para dar mais um exemplo, de Waal (2000) argumenta que, após uma briga, os primatas tentam fazer as pazes, a fim de reparar um relacionamento danificado. Em uma reanálise dos dados, no entanto, Silk (2002) argumenta que o processo de paz é motivado por um desejo de obter recompensas imediatas, mais do que um desejo de proteger as relações de longo prazo. Se Silk está certo, os primatas são míopes e egoístas, ao contrário de profundamente preocupados com a harmonia social. Da mesma forma, os motivos subjacentes aos atos aparentemente morais podem ser totalmente amorais. Quando os primatas fazem favores uns aos outros, podem simplesmente estar tentando ganhar aliados, colher benefícios, ou evitar ser atacados. Eles podem até mesmo ser cegamente programados. Para provar a tese de que os primatas não-humanos são psicologicamente altruístas, seria necessário mostrar que eles têm um desejo de ser úteis originado de uma genuína preocupação com outros animais, especialmente animais com os quais não tenham uma relação de afinidade direta. Mas, e este é um ponto importante, mesmo se pudesse ser estabelecido que os macacos têm altruísmo genuíno, ainda assim tal prova ficam aquém da confirmação de que eles têm moralidade.

Costumamos dizer que o altruísmo genuíno é uma forma de comportamento moral. Mas a frase "comportamento moral" é ambígua. Ela pode significar tanto o comportamento que consideramos moralmente louvável ou o comportamento que seja impulsionado por avaliações morais. Suponha que os primatas ajudam uns aos outros a partir de um sentimento de genuína preocupação. Isso mostra que fazem algo moralmente louvável, e, se estivermos interessados nas origens do comportamento moralmente louvável em seres humanos, podemos extrair daí as evidências para a conclusão de que tais comportamentos são anteriores à nossa espécie. Isso nos diz algo sobre a evolução da decência moral, mas não nos diz nada, eu sustento, sobre a evolução da moralidade. A moralidade é um sistema de regras e valores utilizados para avaliar as ações, agentes e atitudes. Uma criatura poderia se comportar de maneiras nobres, sem qualquer capacidade de julgar se

suas ações são boas ou não. Se os primatas demonstrassem um comportamento moral, não só ajudariam uns aos outros a partir de um sentimento de preocupação, mas também ajudariam uns aos outros porque isso é o que eles moralmente *deveriam* fazer. Ambas são motivações distintas. Para ecoar Kant, há uma diferença entre conformidade com uma regra moral e agir sob uma regra moral. Kant pensa que o último envolve o reconhecimento de um imperativo categórico. Como um sentimentalista, rejeito essa exigência, mas acredito que há pré-requisitos psicológicos importantes para que se tenha pensamentos-dever, ou "atitudes-dever". Para um primata pensar que ele deveria compartilhar, deve haver um sentimento moral de partilha. Isso significa que ele deve se sentir culpado se não compartilhar. Também deve se sentir irritado com aqueles que não compartilham com ele. Este tipo de motivação difere do que eticistas evolutivos chamam motivação altruísta. Motivação altruísta envolve preocupação, um sentimento que é subscrito por emoções como simpatia e apego. Essas emoções não são necessárias e suficientes para julgamentos morais. A evidência para a motivação altruísta em primatas não-humanos ainda é muito fraca, dadas todas as interpretações alternativas. A evidência para a motivação moral ou atitudes-dever é ainda mais fraca.

Trivers (1971: 50) sugere que a tendência evoluída para retribuir é impulsionada, em parte, por uma tendência inata para se sentir culpado quando desertar (ver também Frank, 1988). Neste contexto teórico, é tentador interpretar reciprocidade em macacos como evidência para a conclusão de que os macacos sentem culpa, e isto, por sua vez, iria apoiar a afirmação de que eles podem ter atitudes-dever. Mas, como vimos, a caridade do macaco pode facilmente ser explicada sem apelar para culpa. De fato, há boas razões para pensar que os primatas não-humanos não têm atitudes-dever; nenhum estudo demonstrou análogos símios de culpa ou a sua prima, a vergonha. Coe e Rosenblum (1984) relatam que os macacos machos de baixa patente vão copular com as fêmeas enquanto o macho alfa está longe, e, então, eles vão ser extraordinariamente submissos quando o macho alfa retornar. De Waal (1996: 110) convida os leitores a interpretar esse comportamento submisso como um sinal de culpa, mas pode ser

melhor interpretado como medo de represálias. Se macacos de baixa patente realmente sentirem que era errado copular com as fêmeas, eles provavelmente não iriam tentar a cada oportunidade. De modo mais geral, se os primatas não-humanos sentissem culpa e vergonha, deveríamos esperar que eles exibam comportamentos de ajuda e de compartilhamento mais consistente e equilibradamente. Símios e macacos, muitas vezes, mostram surpreendente indiferença para com o bem-estar de seus pares.

Suspeito que os primatas não-humanos não têm os sentimentos de culpa e de vergonha, e, portanto, não podem ter "atitudes-dever". Além disso, eles provavelmente não têm dois outros traços psicológicos que figuram proeminentemente na moralidade humana saudável. Primeiro, os primatas provavelmente não apresentam as chamadas metaemoções. Nós não só nos sentimos culpados por violar as regras de conduta; sentimo-nos culpados por não nos sentirmos culpados. Os pais submetem seus filhos a "sentimentos de culpa" dizendo às crianças desobedientes "você deveria ter vergonha!". Não há nenhuma evidência sólida para primeira ordem de sentimentos morais em primatas, muito menos sentimentos da segunda ordem. Em segundo lugar, os primatas provavelmente não apresentam preocupações robustas do ponto de vista de "terceira pessoa". Os humanos, pelo contrário, ficam com raiva de malfeitores, mesmo que não sejam diretamente afetados pela transgressão. Nós até mesmo demonstramos preocupações com terceiros que estão fora de nossas comunidades. Nosso desempenho em tarefas que testam a compreensão da distinção moral/convencional sugere que acreditamos que as regras morais são válidas até mesmo em sociedades que não têm contato com a nossa. Bater em pessoas inocentes, por exemplo, seria errado em uma aldeia isolada, mesmo que os moradores desta aldeia aprovassem tal comportamento. Observe-se que essas preocupações de longa distância desafiam a lógica do altruísmo recíproco. As pessoas em lugares distantes não estão em posição de retribuir. Esse tipo de preocupação com os outros não diretamente relacionados conosco não é observada em primatas. Um primata pode intervir em um conflito entre dois outros

(de Waal, 2000), mas, quando isso é feito, aquele que interveio normalmente tinha alguma participação no conflito ou algum vínculo com uma das partes.

Em suma, há pouca razão para acreditar que os primatas tenham algo parecido com atitudes morais tipicamente humanas. A moralidade, portanto, pode ser uma capacidade excepcionalmente humana (comparar Waal, 1996; Hauser, 2001). Se isso estiver correto, há duas possibilidades. A capacidade moral poderia ser uma adaptação evoluída que ocorreu depois que nos separamos do ancestral que partilhamos com os chimpanzés. Ou alternativamente, a capacidade moral pode ser um subproduto de outras capacidades, que se desenvolveram para outros fins. Vou considerar brevemente essas duas opções, e vou concluir que o último caso é mais plausível.

7.2.3 A moralidade em seres humanos é realmente inata?

Existe alguma evidência sólida para a alegação de que a moralidade é inata no seres humano? Surpreendentemente, pouco trabalho foi feito sobre esta questão. Na psicologia do desenvolvimento, há um grande esforço para identificar um núcleo cognitivo em vários domínios, mas relativamente pouco esforço para determinar se há crenças morais fundamentais que podem ser inatas. Esta omissão é susceptível de mudança. Há um interesse crescente na cognição moral, e alguns autores estão começando a discutir explicitamente nativismo moral. Vou considerar quatro argumentos aqui (para mais discussão, ver Prinz, 2007; no prelo *a-b*).

O primeiro argumento é baseado em uma recente linha de pesquisa que examina o comportamento pró-social em crianças pequenas. Warneken e Tomasello (2006) queriam verificar se crianças de dezoito meses ajudariam espontaneamente adultos carentes. Em seu estudo, as crianças assistiam um adulto desempenhando uma tarefa mal sucedida (por exemplo, recuperar um objeto caído, empilhar livros, abrir um armário, e assim por diante). Em comparação a condições controladas em que os adultos não demonstravam disposição em ajudar, as crianças ajudavam espontaneamente. Por exemplo, as crianças pegavam um objeto caído e entregavam ao adulto. Chimpanzés jovens foram também testados

no estudo, e eles mostraram algum grau de utilidade no cenário cuja tarefa era recuperar um objeto caído, mas, em outras tarefas, não apresentaram tendência significativa para ajudar. Os autores sugerem que os seres humanos podem ter desenvolvido tendências do comportamento de ajuda mais robustas do que as apresentadas pelos chimpanzés. O fato de que essas tendências são evidentes aos 18 meses de idade pode ser tomado como suporte para a afirmação de que os seres humanos têm um senso moral inato.

Warneken e Tomasello não fazem uma inferência direta da tendência inata para ajudar para uma moralidade inata, e eles têm o direito de resistir a essa tentação. Como já discutido, ser útil não é evidência para quaisquer motivações pró-sociais (ajudar em razão da preocupação) ou julgamentos morais (ajudar porque a moralidade exige isso). Não há evidência de que a tendência a ajudar apresentada pelas crianças seja impulsionada por intenções nobres. Por razões semelhantes, não devemos tirar conclusões precipitadas quando lemos os estudos que mostram que as crianças consolam adultos que se apresentam aflitos. As crianças captam a aflição das pessoas ao seu redor, de modo que comportamentos de consolação podem ser um programa automático para gerenciamento de estresse. Até mesmo animais de estimação da família consolam seus donos em luto (Preston e de Waal, 2002; Zahn-Waxler et al, 1984). Os comportamentos que demonstram tendência a prestar ajuda, observados por Warneken e Tomasello, podem não ser inatos. Por volta dos 18 meses de idade, as crianças provavelmente já tiveram uma boa quantidade de experiência com os adultos fazendo pedidos: "você pode me dar o marcador?", "você pode abrir o armário?" , "você pode empilhar os livros?". Brincadeiras para crianças no primeiro ano de vida parecem, muitas vezes, envolver tarefas deste tipo. Além disso, as crianças são boas em discernir o comportamento que consiste numa tentativa; e os bebês de dezoito meses de idade já imitam espontaneamente os comportamentos dos adultos que apenas tentaram [a realização de uma tarefa] sem sucesso (Meltzoff, 1995). O que parece comportamento de prestar ajuda pode ser apenas uma instância especial de uma tendência geral das crianças pequenas para imitar e concluir ações malsucedidas.

Ao invés de ser parte de uma capacidade moral inata, esta tendência de imitar é provavelmente um mecanismo para a aprendizagem social.

Um segundo argumento em favor do nativismo moral centra-se na nossa capacidade de detectar trapaceiros. Cabe lembrar que o altruísmo recíproco levanta os problemas referentes ao caronista. Quando você é gentil com alguém, há sempre um risco de que o destinatário da sua bondade não esteja disposto a retribuir. Conseqüentemente, somos determinados para encontrar e punir os trapaceiros: aqueles que colhem benefícios sem pagar o custo. Nós não somos as únicas criaturas que têm a habilidade de detectar trapaceiros. Como observado anteriormente, muito tem sido demonstrado sobre a detecção de trapaceiro dentre os morcegos. Os morcegos-vampiros vivem de sangue, e eles podem morrer de fome depois de dois dias sem se alimentar. Para aumentar suas chances de sobrevivência, os grupos de morcegos partilham os alimentos: um morcego que tenha sido bem sucedido em encontrar uma vítima regurgita sangue para outros morcegos se alimentarem. Wilkinson (1990) observa que os morcegos são mais propensos a alimentar aqueles que já demonstraram ser bons em agir com reciprocidade. Eles são capazes de detectar morcegos gananciosos através do expediente de prestar atenção àqueles que não compartilham mas que, entretanto, estão com as barrigas distendidas de tanto comer. Não há nenhuma razão para inferir a partir deste exemplo que os mecanismos inatos de detecção de trapaceiro existem também em seres humanos. Afinal, os trapaceiros da espécie humana não podem ser detectados por sinais de percepção simples, como barrigas distendidas por exemplo. Mas é natural supor que temos a nossa própria versão de um mecanismo de detecção de trapaceiros. Em princípio, pode-se detectar trapaceiros sem fazer julgamentos morais (como provavelmente é o caso de morcegos-vampiros), mas o modo humano de detecção de trapaceiro tende a ser moralista. Quando detectamos um trapaceiro, rotulamo-lo como moralmente mau. Assim, a evidência de um mecanismo inato de detecção de trapaceiro nos seres humanos poderia ser tomado como suporte para a existência de um sentimento moral inato.

Um número considerável de psicólogos evolutivos – mais notadamente, Cosmides e Tooby (1992) e seus colegas – defende, através do estudo do raciocínio

humano, a existência de um mecanismo inato de detecção de trapaceiros. Os seres humanos são notoriamente malsucedidos no que diz respeito ao raciocínio sobre condicionais descritivos da forma: se *a* é *B*, então *a* é *C*. Em contraste, somos muito bons em raciocinar usando condicionais que envolvem trocas sociais da forma: Se *a* tiver recebido benefício *B*, em seguida, *a* deve arcar com o custo *C*. O contraste pode ser demonstrado quando consideradas condicionais que têm a mesma gramática de superfície, tendo, entretanto, ou um uso descritivo, ou uma interpretação de troca social. Aqui temos um exemplo, com base nos materiais utilizados por Stone et al. (2002). Suponha que você leia um estudo no jornal que diz que as crianças que assistem muita TV são também especialmente organizadas. Em particular, a manchete diz "se uma criança assiste mais de duas horas de TV por dia, isso implica dizer que o quarto dela está limpo e organizado." Você está tentando decidir se isso é verdade. Para descobrir, qual das seguintes opções você precisa verificar: (a) crianças que assistem mais de duas horas; (B) crianças que assistem menos de duas horas; (c) as crianças que têm quartos sujos; (d) as crianças que têm quartos limpos? Com exemplos como esse, a maioria dos sujeitos escolhe (a) e muito poucos escolhem outras opções. Significativamente, os indivíduos ignoram (c). Não percebem que precisam considerar as crianças que têm quartos sujos. Estes são os típicos casos em que o conseqüente da condicional é falso. Um descritivo condicional é falso quando, e somente quando, tem um verdadeiro antecedente e um falso conseqüente, por isso é logicamente imperativo para verificar ambos os casos, como (a), com os verdadeiros antecedentes, e casos como (c), com falsos conseqüentes, se você é incumbido da tarefa de determinar a verdade da condicional em questão. O fato de que os indivíduos negligenciam os casos de falsos antecedentes mostra que estão propensos a cometer um erro lógico.

Agora, considere um problema diferente. Você é um pai ou uma mãe, e os seus filhos não estão autorizados a assistir uma série de TV, a menos que eles limpem seus quartos. Sua regra em casa é a seguinte: "para que uma criança possa assistir mais de duas horas de TV por dia, deve, primeiro, organizar e limpar o seu quarto". Quando questionados sobre quem devem conferir para verificar se a regra

está sendo cumprida, as pessoas escolhem ambos (a) e (c). Os sujeitos reconhecem imediatamente que devem conferir as crianças cujos quartos estão sujos para se certificar de que elas não tinham assistido muita televisão. Esta é a resposta correta, e podemos obtê-la no caso de troca social, mas não é o caso quando estamos diante do uso meramente descritivo. Quando as pessoas são incumbidas de identificar infratores ("trapaceiros"), elas raciocinam de modo mais bem sucedido. Os psicólogos evolutivos explicam esse resultado dizendo que desenvolvemos um módulo dedicado a detectar trapaceiros (Cosmides, 1989; Gigerenzer e Hug, 1992).

Considero que essa explicação evolutiva não é convincente. Eu não posso rever a vasta literatura sobre o tema, mas deixe-me articular uma razão de princípio para o ceticismo. O problema básico com os experimentos que os psicólogos evolucionistas usam para contrastar condicionais sociais e descritivas é que as duas tarefas, embora superficialmente semelhantes, na verdade, são completamente diferentes. Quando instados a avaliar o descritivo condicional no experimento do jornal, sua tarefa é determinar se ele é verdadeiro. Este é um problema de confirmação. É como tentar determinar se todos os corvos são pretos. Para fazer isso de forma eficiente, uma boa ideia é examinar todos os corvos que encontramos, mas, por outro lado, é uma total perda de tempo examinar todas as coisas que não são pretas. A falácia lógica de evitar os (c)-casos acima é, na verdade, uma boa estratégia indutiva. Em contraste, quando se apresenta com o caso dos pais, assumimos o condicional como sendo verdadeiro; é uma regra da casa, e as regras são verdadeiras, mesmo que nem sempre sejam observadas. A tarefa não é confirmar ou refutar a condicional, mas sim encontrar os infratores. O infrator é, por definição, alguém que não tenha feito o que é necessário. Então, nós estamos imediatamente orientado para (c)-casos. As exigências da tarefa chamam a nossa atenção para as pessoas que não tenham feito o que lhes é exigido. Isso não confirma uma capacidade inata especializada em detectar trapaceiros. Isso apenas demonstra que nós somos bons em fazer raciocínios sobre condicionais quando as exigências das tarefas nos fazem sensível a casos que confirmam o antecedente (como em (a)) e aos casos em que negam o conseqüente (como em

(c)). Os condicionais de intercâmbio social têm esse recurso, já os condicionais descritivos, não.

A explicação alternativa para os resultados experimentais é bastante simples: há uma diferença intrínseca entre a tarefa de avaliar condicionais descritivos, caso em que estamos à procura de contraexemplos, e a tarefa de avaliar as condições de trocas sociais, caso em que assumimos que a condicional é verdadeira e procuramos identificar infratores. Nesta narrativa alternativa, não precisamos postular um domínio específico para o módulo de raciocínio sobre trapaceiros; ao contrário, postulamos um domínio geral de sistemas de raciocínio que são aplicados de forma diferente quando estamos realizando tarefas de confirmação e tarefas de identificar violação de norma. Essa alternativa prevê que as pessoas geralmente irão executar com êxito as condicionais que interpretarem como normas, mesmo quando esses condicionais não tenham nada a ver com a fraude ou a com a moralidade. Considere condicionais prudenciais. Suponha que aprendemos que passar tempo demais em frente ao televisor expõe os espectadores a níveis perigosos de radiação, a menos que eles se sentem a mais de um metro e meio de distância do aparelho. Este fato sobre riscos para a saúde pode ser expresso como uma regra condicional: "se uma pessoa assiste mais de duas horas de TV por dia, convém que ela permaneça a um metro e meio ou mais de distância do aparelho". Este condicional não é uma norma moral; é uma norma sobre a segurança. Se nossa capacidade de raciocinar sobre a detecção de trapaceiro fosse baseada na capacidade de domínio geral de raciocinar sobre todas as condicionantes normativas, deveríamos concluir que as pessoas são adeptas do raciocínio sobre condicionais prudenciais. Isso é o que os estudos têm mostrado (Manktelow and Over, 1990). Se você for solicitado ou solicitada a proteger as pessoas contra os perigos da radiação, você irá reconhecer imediatamente que sua tarefa é identificar quem está muito perto de uma TV para, então, buscar descobrir por quanto tempo essa pessoa está sentada em frente ao televisor.

Claro, *poderia ser o caso* de termos um módulo inato para o raciocínio dirigido à tarefa de identificar trapaceiros e outro módulo inato para raciocinar acerca das normas prudenciais, e estes módulos *calharem* de nos dotar com boas

capacidades iguais de raciocínio sobre condicionantes em cada um desses domínios. Isso seria uma interpretação extravagante; entretanto, é a interpretação que os psicólogos evolucionistas preferem. Para apoiar a conjectura de que existem dois módulos – um para as normas de prudência e outro para a detecção de trapaceiros – Cosmides e Tooby uniram-se a colaboradores oriundos da neurociência para estabelecer que há uma dissociação entre a capacidade de raciocinar sobre trapaceiros e a capacidade de raciocinar sobre normas prudenciais (Stone et al., 2002). Eles examinaram um paciente com lesões no córtex pré-frontal ventromedial, que tem uma séria redução na sua capacidade raciocinar sobre condicionais de trocas sociais (como é o caso das regras impostas pelos pais em casa), mas tem intacta sua capacidade de raciocinar sobre condicionais prudenciais (como é o caso de risco à saúde). Para provar que estas capacidades de raciocínio são independentes uma em relação à outra, os psicólogos evolucionistas precisariam apresentar uma dupla dissociação: precisariam encontrar pacientes com debilidade de raciocínio prudencial e com habilidade na tarefa de detectar trapaceiros. Tal paciente jamais foi encontrado; entretanto, Stone et al., sustentam que os pacientes ventromediais oferecerem um suporte preliminar para a especificidade do domínio de detecção de trapaceiro, uma vez que estabelecem uma das duas dissociações que devem, em última instância, ser encontradas se módulos separados estão trabalhando. O problema é que os pacientes com lesões ventromediais não têm um déficit seletivo na detecção de trapaceiro; eles têm muitos outros problemas. Por exemplo, eles são impulsivos, têm respostas emocionais anormais e são péssimos na detecção da *faux pas* social. Dado o seu profundo conjunto de déficits, não podemos inferir que eles têm uma deficiência no módulo de detecção de trapaceiro. O paciente do estudo de Stone et al. poderia ter um déficit no raciocínio acerca de situações sociais e, conseqüentemente, não poderia apreciar o que está envolvido na aplicação das regras sociais. Seus déficits emocionais também poderiam torná-lo inábil a apresentar as emoções necessárias à tarefa de identificar trapaceiros. Sem uma capacidade de compreender plenamente ou de se preocupar com as sutilezas do domínio social, este paciente poderia simplesmente não conseguir compreender as

demandas inerentes a esta tarefa de raciocínio. Pelo que se pode inferir, ele poderia até confundir as condicionais de trocas sociais como condicionais descritivas. Assim, o estudo de Stone et al. não prova que há recursos cognitivos dedicados à detecção de trapaceiro. Sem que tal evidência tenha sido encontrada, sem boas provas, acredito que devemos rejeitar a hipótese de que existe nos seres humanos um módulo inato de detecção de trapaceiro.

O terceiro argumento em favor do nativismo moral é apresentado por Sue Dwyer (1999). Para argumentar a favor da capacidade inata, Dwyer toma de empréstimo uma estratégia extraída da lingüística (ver também Rawls, 1971; Harman, 2000; Mikhail, 2000). O nativismo lingüístico, geralmente, apoia-se em argumentos que apelam para a insuficiência do estímulo. Os lingüistas afirmam que as crianças obedecem a regras gramaticais que não poderiam ter aprendido através da experiência, uma vez que os dados lingüísticos primários a que estão expostas (a fala dos adultos e as correções de sua fala pelos adultos) não fornecem informações suficientes para ajudar a criança a escolher entre a miríade de regras possíveis. Por paridade, Dwyer afirma que as crianças não recebem dados morais primários suficientes para explicar a sua compreensão do domínio moral. Em particular, afirma que as crianças não recebem informações suficientes por parte dos adultos para distinguir entre as regras morais e as convencionais. A distinção moral/convencional se faz presente no terceiro ano de vida e não é explicitamente ensinada. Os pais não dizem explicitamente aos filhos quais regras são do domínio da moral e quais são meramente convencionais; e mais: costumam punir as crianças indistintamente. Ou seja, as crianças são punidas por bater e morder e por violar regras de etiqueta. No entanto, de alguma forma, reconhecem que as regras morais são diferentes. No início da infância, tratam as regras morais como mais graves e menos dependentes de uma autoridade. Bater e morder, por exemplo, seria errado não importa o que alguém diz, mas colocar os cotovelos sobre a mesa não tem problema, se for dada permissão para tal. As crianças em todo o mundo extraem esse tipo de contraste entre regras, sem ser explicitamente ensinadas.

O argumento de Dwyer tem considerável plausibilidade *prima facie*, mas não creio que resista a uma análise mais acurada (ver também Nichols, 2005). Mesmo

que os pais não ensinam explicitamente a diferença entre regras morais e convencionais, existem diferenças marcantes na forma implícita como tais regras são transmitidas. Os pais tratam violações às normas morais de forma mais séria; além disso, aplicam-nas com mais rigor (Smetana, 1989; Grusec e Goodnow, 1994; Nucci e Weber, 1995). As transgressões às regras morais são quase sempre corrigidas com o uso da autoridade e com a indução à empatia, enquanto a observância às regras convencionais tem o suporte da racionalidade e apela para os padrões normativos da ordem social. As violações das regras morais também tendem a ter uma maior carga emocional, porque muitas vezes envolvem uma vítima. Se a pequena Sally morde Billy, Billy irá chorar, e os pais de Billy ficarão, no mínimo, muito chateados. Assim, nas ocasiões em que as crianças consideram que as regras morais são independentes de uma autoridade que as imponha, tal fato pode se dar porque internalizaram emocionalmente tais regras. Quando imaginam a ação de dar uma mordida, ficam aborrecidas, porque, no passado, a atitude de morder levou a intensas emoções negativas. Como resultado, as crianças afirmam que não seria certo dar uma mordida em outra pessoa, mesmo que os adultos, na condição de autoridades, afirmassem que não tem problema. Em suma, as crianças são expostas a *feedbacks* corretivos suficientes para diferenciar entre regras morais e convencionais.

Permitam-me apresentar um último argumento em favor do nativismo moral, o qual foi desenvolvido por Hauser et al. (no prelo), ecoando as ideias de Harman (2000) e estendendo o trabalho experimental desenvolvido por Mikhail (2000; 2002). Usando um questionário de internet, Hauser et al. coletaram milhares de intuições morais sobre uma variedade de diferentes casos do dilema do bonde (faço remissão ao capítulo 1). Em um cenário, os entrevistados são questionados se seria correto empurrar alguém para interromper o trajeto de um bonde a fim de salvar a vida de outras cinco pessoas que, por sua vez, estão amarradas sobre os trilhos logo adiante. Em outro caso, os indivíduos são convidados a avaliar se seria correto puxar uma alavanca que faria com que o bonde mudasse dos trilhos onde teria atingido cinco pessoas para os trilhos alternativos, onde teria atingido apenas uma pessoa. O estudo de Hauser et al. constatou que, no cenário em que uma pessoa

era empurrada sobre os trilhos, apenas 11 por cento dos indivíduos entrevistados consideraram tal atitude moralmente permissível. Já no cenário onde a tarefa consistia em puxar uma alavanca, 89 por cento dos entrevistados consideraram que o ato seria moralmente permissível. Os indivíduos entrevistados apresentaram, também, intuições robustas e consistentes em uma ampla gama de outros casos. Isso é surpreendente, pois os entrevistados nunca haviam se deparado com o caso do dilema do bonde. Se as pessoas dão respostas consistentes quando se deparam com um caso jamais visto, tal fato pode ser tomado como evidência da existência de regras inatas.

Este argumento trabalha com o mesmo princípio utilizado por Dwyer. Os autores identificam uma característica do julgamento moral que, por sua vez, é generalizada ainda que não explicitamente ensinada. Mas eu não acho que isso seja um argumento suficiente para provar que a moralidade é inata (ver também Nichols, 2005). Todos concordam que somos explicitamente ensinados a não prejudicar as outras pessoas. Se uma criança morde, bate, ou de alguma forma ataca alguém, as punições serão graves. Aprendemos também a importância das normas que impõem o dever de ajudar. Somos ensinados que o correto é ajudar as pessoas em necessidade. O comportamento que consiste em ajudar pode ter alguma base na biologia, mas os instintos de ajuda provavelmente são direcionados para parentes e descendentes próximos. Nós somos ensinados que é correto prestar ajuda aos estranhos (um ponto que eu vou voltar a seguir). Estas duas normas – evitar causar danos e ajudar os outros – entram em conflito no caso do dilema do bonde. Então a questão é: como é que vamos decidir qual norma aplicar? Tal decisão é determinada por princípios inatos?

No capítulo 1, argumentei que as intuições envolvidas no dilema do bonde são movidas por emoções. Acreditamos que seja errado empurrar alguém em direção aos trilhos, porque tal ato é um estimulante de intensa emoção negativa; por outro lado, acreditamos que não há problema em alternar os trilhos, porque essa ação parece mais benigna. Esta análise prevê que intuições no dilema do bonde vão flutuar conforme a relevância do dano. Essa previsão parece ser corroborada por outros cenários utilizados no estudo de Hauser et al. Em um

cenário, os indivíduos são convidados a considerar um caso em que você pode puxar uma alavanca que fará com que um bonde avance em direção a uns trilhos em *loop* onde um homem forte está deitado; no final de seu trajeto, a pista em *loop* conecta-se outra vez com a pista principal, de modo que o bonde iria matar cinco pessoas; entretanto, o homem forte interromperia o percurso e pararia o bonde. Neste cenário, 55 por cento dos entrevistados dizem acreditar ser permitido puxar a alavanca e mudar o bonde de trilhos. Este percentual fica mais ou menos na metade em comparação ao caso original do dilema que envolve a alavanca, porque, embora não seja um caso paradigmático de morte, o fato de que o homem forte seja usado como uma barreira humana torna sua morte bastante proeminente. No quarto caso, Hauser et al. apresentaram o cenário da pista em *loop* com uma pequena variação: o homem forte sobre a pista está em pé na frente de um objeto pesado. Os entrevistados são informados que o objeto pesado servirá como barreira, impedindo o bonde de atropelar cinco pessoas, entretanto, o homem forte morrerá no processo. Neste caso, 72 por cento dos indivíduos acredita que não há problema em puxar a alavanca, porque, eu sustento, a atenção é desviada da vítima para o objeto pesado.

A história da proeminência não pode explicar todos as intuição sobre casos do dilema do bonde, mas certamente parece explicar os que Hauser et al. apresentam, e o faz sem supor que nos orientamos por regras inatas. Da mesma maneira que podemos explicar por que as pessoas são ensinadas a ajudar e a não causar danos, podemos também explicar todas as variadas intuições sobre como essas normas interagem nos casos do dilema do bonde. A variação em nossas intuições não reflete um conjunto preciso de regras inatas – a gramática da moral – mais do que a variação reflete o quão facilmente os sentimentos adquiridos podem mudar em função da proeminência.

Uma consideração final contra todos esses argumentos nativistas deve ser trazida à baila. É tentador presumir que adquirimos ideias morais de forma fácil e sem muita instrução. Entretanto, isso seria um erro grave. Os tutores despendem uma enorme quantidade de energia para ensinar às crianças como se comportar. Hoffman (2000: 141) calcula que, entre as idades de dois e dez anos, as crianças

têm seu comportamento corrigido pelos tutores a cada intervalo de tempo que varia de seis a nove minutos! Isso equivale a cinquenta lições de conduta a cada dia. Se a moralidade fosse inata, toda essa instrução deveria ser desnecessária. Mas não é o caso. Claramente a criação está fazendo uma contribuição para a moralidade.

Em suma, penso que há um pouco de tendenciosidade em afirmar que os seres humanos têm um senso moral inato. Alguns dos argumentos principais para o nativismo moral me parecem implausível. Claro, isso não é um levantamento exaustivo de tais argumentos, então eu não tomo as considerações aqui apresentadas como prova decisiva contra o nativismo moral. Acho que a melhor maneira de derrotar nativismo neste domínio e em outros é o apresentando uma explicação alternativa. Acredito que a moralidade é um subproduto de outras capacidades. Se a moralidade pode ser explicada como um subproduto de outras capacidades, há uma certa tendenciosidade em afirmar que é inata. Vou começar o meu argumento em favor da tese do subproduto considerando algumas das capacidades cognitivas não-morais que podem desempenhar um papel importante na aquisição de moralidade.

7.3 UMA ABORDAGEM BIOCULTURAL DA MORALIDADE

7.3.1 A moralidade como um subproduto

Há muitas capacidades humanas que são quase universais, mas não inatas. Exemplos incluem arte, roupas, o uso do fogo, abrigos construídos, práticas mortuárias, religião, casamento e ferramentas complexas. Todas estas invenções distintamente humanas presumem-se resultar de capacidades cognitivas que evoluíram para outros fins. Nós não temos um módulo de vestuário; nós apenas sentimos frio e somos inteligentes o suficiente para resolver o problema fazendo coisas para cobrir nossos corpos. Nós não temos um módulo de religião; só temos tendência a atribuir mentalidade a coisas não-vivas e uma propensão para histórias que contrariam as expectativas (Boyer e Ramble, 2001). Minha conjectura é que a

moralidade se enquadra nessa categoria. É um subproduto de capacidades que não evoluíram, elas próprias, para a aquisição de regras morais.

Uma variedade de capacidades não-morais podem ser importante para a aquisição de um sentido moral. Em primeiro lugar, precisamos ter certas emoções: emoções dirigidas para outros, tais como raiva, desprezo e repugnância, bem como as emoções autodirigidas, como vergonha e culpa, o que, argumentei anteriormente, pode estar relacionada ao constrangimento e à tristeza. Em segundo lugar, é preciso a capacidade de formular regras. Durante o desenvolvimento moral, devemos *transferir* as emoções negativas que estamos condicionados a experimentar, quando nos comportamos mal, para o mau comportamento. Ao fazê-lo, geramos uma representação mental que nos predispõe a ter sentimentos negativos sobre um tipo de comportamento, independentemente de quem está realizando isso. Esta pode ser a base da nossa capacidade de ter atitudes morais em relação a terceiros, mesmo quando não estamos diretamente envolvidos. Seria interessante investigar se a formação da regra do tipo necessário é uma capacidade exclusivamente humana. Além disso, o desenvolvimento moral saudável pode exigir memória. Nós, muitas vezes, pegamos malfeitores bem depois de uma transgressão ter ocorrido. Outros animais podem ter de pegar uns aos outros em flagrante, a fim de reconhecer que eles tenham feito algo indesejável e que minimiza oportunidades para moldar o comportamento através de punição. Mais especulativamente, a moralidade pode depender da capacidade de imitar. Quando as crianças são punidas, eles não só se sentem mal sobre o que eles fizeram; eles também adquirem a disposição de punir os outros para executar as ações que lhes colocou em problemas.

Uma quinta capacidade que contribui para a moralidade humana é a "leitura da mente": a capacidade de atribuir estados mentais aos outros. Essa capacidade nos permite experimentar a empatia genuína, em oposição à angústia vicária – experimentamos o sofrimento de outra pessoa e reconhecemos que as nossas tristezas resultam das deles. Isso promove o comportamento pró-social e permite a aquisição de normas referentes a estados mentais: por exemplo, podemos pensar que é ruim causar sofrimento. Leitura da mente também nos permite adquirir

normas sobre normas: educadores morais podem reconhecer nossas emoções e nos impelir a lamentar nosso mau comportamento. Esta forma de treinamento aumenta a probabilidade de que iremos nos comportar de acordo com as normas. Nós nos comportamos de acordo com as normas porque nos sentiríamos culpados de outra forma, e a disposição de se sentir culpado tem considerável estabilidade, porque nos sentiríamos culpados se não nos sentimos culpados. Normas sobre como sentir estabelecem metaemoções e as metaemoções são como um plano de seguro extra contra o mau comportamento.

Metaemoções podem não ser essenciais para se ter um sentido moral – pessoas com autismo moralizam, apesar de uma deficiência na leitura da mente – mas eles podem desempenhar um papel fundamental no surgimento da moralidade através da evolução cultural. Usando modelos de teoria dos jogos, Henrich e Boyd (2001) argumentaram que o comportamento cooperativo generalizado não seria sustentado e transmitido ao longo de gerações, se não fosse para a punição; uma sociedade de cooperadores seria dominada pelos desertores. O mesmo ponto pode ser articulado em termos de normas morais; uma sociedade não continuaria em conformidade com as normas morais geração após geração, se os infratores não fossem punidos (Sripada e Stich, 2006). Se os membros da comunidade não tivessem punido os malfeitores, os malfeitores não alterariam o seu comportamento e [não] considerariam o delito como errado. Aqui, a noção de punição deve ser definida como qualquer ação que impõe um custo sobre aqueles que violam a norma. A punição pode ser física, mas também pode incluir táticas psicológicas, como a retirada do amor e a indução de empatia, que eu discuti no capítulo 1. Todos estes métodos levam as pessoas a estarem de acordo com as regras morais. Boyd et al. (2003) argumentam que a necessidade de punição introduz um desafio grave. É custoso punir. Pessoas que punem têm de exercer energia e colocar-se em risco. É muito tentador, portanto, tentar fugir sem punir ninguém, e esperar que os outros façam o trabalho sujo (como quando um dos pais permite que um outro pai interprete o policial mau). Isso cria um problema do caronista de segunda ordem. Caronistas de primeira ordem tentam escapar violando as regras morais, enquanto caronistas de segunda ordem tentam escapar

não cumprindo as regras de primeira ordem. Caronistas de sucesso são sempre melhores do que aqueles que pagam suas dívidas, então, uma situação que é tentadora para os caronistas pode se transformar em uma situação em que todo mundo está tentando se safar fazendo menos. Se há alguns caronistas de segunda ordem desde o início, eles podem, eventualmente, superar radicalmente os cooperadores de segunda ordem, ou seja, aqueles que punem os infratores das normas de primeira ordem. Mas, uma redução de colaboradores de segunda ordem implica uma proliferação de caronistas de primeira ordem, porque não há pessoas suficientes para manter os caronistas de primeira ordem sob controle. Isto equivale ao colapso moral, pois, nesta situação, não há número suficiente de pessoas para transmitir as regras morais. Boyd et al. argumentam que esta situação pode ser prevenida punindo os caronistas de segunda ordem; isso é chamado de metapunição. Se nós punirmos as pessoas que negligenciam a punição ao mau comportamento, então haverá um incentivo para punir, e a moralidade será capaz de sair do chão.

Eu acho que as metaemoções têm um papel central a desempenhar nesta história. Eu disse que as normas de primeira ordem são inculcadas condicionando as pessoas emocionalmente a agir de determinadas maneiras. Bem, como são inculcadas normas de punição? Uma possibilidade é que elas são inculcadas como uma classe especial de normas de primeira ordem: nós condicionamos as pessoas a envolverem-se em comportamentos de punição. Mas é importante ter em mente que a punição está sendo definida de forma ampla aqui para cobrir técnicas de condicionamento de emoções, como a retirada do amor. Se você quer treinar pessoas para serem boas em retirar amor, você *poderia* treiná-los apenas a se comportar como se eles estivessem retirando amor, *ou* você poderia treiná-los a realmente retirar amor. Em suma, você pode treinar as pessoas para ficarem decepcionadas (ou com raiva, ou com nojo, etc.) quando outras pessoas se comportam mal. A última estratégia parece ser mais direta e, portanto, mais eficaz. Se eu estiver certo, então a melhor maneira de implementar a metapunição é mostrar decepção (ou raiva, ou repulsa, etc.) para aqueles que não reagem a crimes com essas emoções negativas. Quando fazemos isso, nós efetivamente

condicionamos as pessoas a se sentir mal sobre não se sentir mal por mau comportamento. Assim, metaemoção é uma forma de metapunição porque atribui um custo emocional para quem falha em ter emoções punitivas em relação àqueles que transgridem. Metaemoções também têm a vantagem de ser fáceis de transmitir. Nós, muitas vezes, condicionamos as pessoas a ter metaemoções explicitamente punindo aqueles que têm as atitudes erradas ("você deveria ter vergonha de si mesmo!"), mas isso pode nem ser necessário. Há um desejo geral de que as pessoas têm de compartilhar as atitudes com os membros de seu grupo social. As pessoas que pensam de forma diferente são consideradas desviantes e são menos bem-sucedidas em alianças sociais. Assim, a incapacidade de ter as mesmas emoções de primeira ordem como os outros é interpretado como um defeito, e isso fomenta emoções negativas em direção a emoções não-conformistas de primeira ordem. Em outras palavras, as metaemoções são tanto explicitamente transmitidas quanto implicitamente transmitida através de um desejo generalizado para encaixar (Boyd et al. apelam para um "viés de conformidade"). Se esta história básica está correta, então metaemoções podem ser uma ferramenta essencial utilizada para implementar metapunição, e, como tal, eles podem desempenhar um papel essencial na manutenção de regras morais através de gerações. Se uma população carecia de metaemoções (por exemplo, indivíduos com autismo), então, mesmo que pudessem adquirir as normas morais, eles podem não ser capazes de manter e de transmitir essas normas de forma estável ao longo do tempo.

O resultado de tudo isso é que os seres humanos estão equipados com uma gama de ferramentas não-morais, tais como emoções, memória, formação de regra, imitação e leitura da mente, que poderiam coletivamente dar origem a uma capacidade moral. Quando as emoções são condicionadas no contexto de comportamento, sentimentos são formados e resultam regras apoiadas em afetos (compare com Nichols, 2004a). A questão sobre punição e metapunição mostra que esses recursos relativamente simples são muito poderosos. Em princípio, qualquer tipo de comportamento poderia ser submetido a uma forma de condicionamento emocional que resultaria na formação de uma regra moral (Sripada e Stich, 2006).

Neste quadro, a moralidade não é inata; é um subproduto poderoso de outras capacidades inatas.

A história que venho esboçando não diz nada sobre o conteúdo das regras morais. É uma história sobre como moralização surge, não é uma história sobre o que nós moralizamos. A pesquisa que foi feita sob a rubrica de "ética evolucionista" difere a este respeito. Essa pesquisa normalmente se refere ao conteúdo das normas. Eticistas evolucionistas enfatizam a ajuda, o compartilhamento, a reciprocidade e outras formas de comportamento. Eu objeto que esses comportamentos podem ocorrer na ausência de qualquer senso moral, e, para se qualificar como moral, eles precisam ser considerados como bom e motivado por essa consideração. Compartilhar é moral somente se nós compartilhamos *a fim de fazer a coisa certa*. Uma vez que disposições biológicas para compartilhar possam não ser implementadas por atitudes morais, a ética evolucionista pode falhar em seu esforço para demonstrar que temos regras morais inatas. Agora, no entanto, estamos em uma posição para ver o que os especialistas em ética evolutiva podem contribuir para um relato das origens da cognição moral. As regras morais não são inatas, mas, uma vez que aprendemos a tomar uma posição moral, os comportamentos enfatizados por eticistas evolucionistas são especialmente suscetíveis de ser moralizados. Nós compartilhamos, com os nossos primos primatas, uma tendência não-moral para compartilhar, para ajudar e para retribuir, e essas tendências se tornam objetos de louvor moral em nós. *Aprendemos* a ver estas coisas como boas. Nós podemos desenvolver atitudes morais em relação a outros comportamentos, mas os comportamentos enfatizados por eticistas evolucionistas são tipicamente o centro da vida moral humana.

Isso levanta uma questão. Se estamos naturalmente propensos a compartilhar, a ajudar e a retribuir, por que nós nunca moralizamos esses comportamentos? Se somos instintivamente caridosos, por que nós nunca nos preocupamos em moralizar a caridade? A moralização é uma forma de promover o comportamento, e não há necessidade de promover o que faríamos naturalmente. A resposta a esta questão pode envolver o crescimento das sociedades humanas (Boyd e Richerson, 1985). Em primatas não-humanos, partilha, ajuda, e

reciprocidade tendem a ficar restritas aos membros de pequenos grupos. Um animal vai conferir benefícios a outro somente se eles são ou parentes, ou membros associados de um bando. Sociedades humanas em uma escala muito pequena não pode exigir regras morais, porque os membros dessas sociedades são próximos o suficiente para estarem naturalmente inclinados a tratar bem uns aos outros. Enquanto o tamanho da população cresce, porém, encontramos-nos em contato com pessoas que não são amigos próximos ou familiares. Em grandes sociedades, muitas vezes, há projetos em grupo, como a construção coletiva e a agricultura, que dependem de cooperação. Mecanismos devem ser postos em prática para se certificar de que ninguém se descuide. Se comportamentos de ajuda evoluíram em contextos em que as relações de cooperação foram diádicos, os mecanismos evoluídos podem não ser suficientemente poderosos para assegurar a cooperação com grupos de estranhos. Os fatores que nos levam a ser agradáveis com os nossos vizinhos (afeto, repetidas atitudes “olho por olho”, medo de ser pego, etc) podem não se aplicar quando estamos lidando com os colaboradores mais distantes. A moralização pode ter surgido como uma técnica para garantir que as pessoas não afrouxassem enquanto as sociedades cresciam. A expansão coloca pressão sobre culturas para encontrar formas de estender a nossa amabilidade natural com estranhos. A moralização oferece uma solução. Essa sugestão se assemelha à sugestão de Hume (1739) de que o conceito de justiça surge com a necessidade de estender as simpatias naturais para grandes grupos sociais.

7.3.2 Normas Bioculturais

Se eu estiver certo, então as regras morais não são inatas. Em vez disso, eles surgem através de interações entre a biologia e a cultura. Comportamentos disposicionais de base biológica se estenderam através de enculturação, especialmente porque os grupos sociais crescem. Enculturação pode remodelar as disposições comportamentais de várias maneiras, e, em alguns casos, até mesmo substituí-las. Nos capítulos 5 e 6, vimos muitos exemplos de diversidade moral entre as culturas. O quadro apresentado neste capítulo pode aprofundar nossa

compreensão da diversidade cultural, reconhecendo que algumas de nossas regras morais são informadas por tendências biológicas a comportar-se de determinadas maneiras. Comportamentos de base biológica não são exatamente uma restrição sobre a genealogia de regras morais, porque a cultura pode substituí-los, mas eles são, muitas vezes, um ingrediente central. Podemos entender muitas normas morais humanas como variações especificamente culturais sobre os mesmos temas biológicos.

Na maioria das culturas, a bondade tende a ser maior para os membros do grupo mais próximo, que é exatamente o que predisposições biológicas preveriam, mas há uma variação considerável na forma como grupos mais próximos são definidos. No Japão medieval, havia uma rígida hierarquia social, e o *status* social cumpria um papel na determinação de quem deve ser gentil com quem. Limitações a respeito da bondade com os externos ao grupo no Japão podem ser ilustradas por uma prática chamada *tsujigiri*, onde um samurai testaria uma nova espada cortando um camponês aleatório ao meio (ver Midgley (1981) para uma discussão filosófica). Aqui, a preocupação moral é verticalmente delimitada: grupos de alto escalão mostram pouco respeito para as pessoas de baixo escalão. Preocupação moral também pode ser horizontalmente limitado. Na Nova Guiné pré-colonial, tribos vizinhas estavam, muitas vezes, em competição por recursos, e conflitos violentos eram comuns. Baseado em registros etnográficos, Wrangham (2004) calcula que de 20 a 38 por cento dos homens em tribos montanhosas morreu por homicídio. Nestes casos, não encontramos pessoas ricas matando pessoas pobres, mas sim uma tribo travando uma guerra com outra. Posição não é um fator. Às vezes, a violência ocorre entre as aldeias da mesma tribo. Os Yanomami da Amazônia vivem em aldeias que se dedicam a um ciclo interminável de brutalidade olho por olho. Uma aldeia vai invadir uma aldeia vizinha, matando homens, estuprando mulheres, e, depois, a aldeia vizinha vai traçar uma vingança (Chagnon, 1968). Os Yanomami são interessantes porque guerreiam uns contra os outros, apesar de uma cultura compartilhada. Eles definem o grupo mais próximo por localidade. Outras sociedades definem o grupo mais próximo pela cultura, apesar de distribuição geográfica maciça. O cristianismo tem tradicionalmente trabalhado

neste princípio; cristãos trabalham para ajudar os outros cristãos, mesmo que estejam do outro lado do mundo. Os cristãos também tentam ajudar os não-cristãos, mas as campanhas de caridade são, muitas vezes, conjugadas com os esforços missionários, o que aumenta a probabilidade de que a ajuda vai acabar nas mãos de outros cristãos. Neste caso, o grupo mais próximo não é definido por fronteiras da aldeia ou por classe, mas pela fé compartilhada, o que o torna muito mais abrangente. Historicamente, quando os cristãos têm encontrado grupos que não se convertem, a política da caridade é substituída por uma postura um pouco menos amigável, como na Inquisição e nas Cruzadas. Na sociedade contemporânea, a globalização, o pluralismo e o cosmopolitismo estão levando a novas concepções de membros do grupo. Estas concepções são mais abrangentes do que as concepções anteriores, na teoria, mas representam uma séria ameaça para os grupos que não querem se identificar com um coletivo global. Em suma, quando a nossa tendência biológica a ser gentil com os membros de um grupo mais próximo é convertido em uma regra moral, ela pode assumir uma ampla variedade de formas. A regra é aplicada de formas diferentes enquanto as culturas definem e renegociam as condições de participação no grupo.

Sob certas condições, a tendência biológica a ser gentil com os membros de um grupo mais próximo pode ser erodida. De acordo com Turnbull (1972), os Iks de Uganda se tornaram viciosos quando a seca e o deslocamento forçado os levou à beira de morrer da fome. Os Iks se envolveram em fraude constante, arrancaram comida da boca de um aldeão doente, e eles riram de alegria quando uma criança se queimou em um incêndio. Turnbull afirma que os Iks continuaram em seus caminhos infelizes, mesmo depois de sua situação econômica ter melhorado, como se tivessem sido corrompido pela pobreza extrema. A análise de Turnbull é controversa (ver Heine, 1985), mas não seria surpreendente se as condições culturais pudessem levar as pessoas a abandonar a tendência natural de ser gentil com seus amigos e familiares. A miséria pode tornar concorrentes aqueles que poderiam ser amigos. Uma vez que tal indiferença cruel entra no tecido de uma cultura, ela pode ser transmitida (para mais exemplos, consulte Edgerton, 1992).

Vamos passar da bondade para a equidade, onde as variações culturais também são fáceis de demonstrar. Henrich et al. (2001) descobriu isso quando pediu aos membros de diferentes culturas para jogar um jogo de ultimato. Requerem-se dois jogadores. Ao jogador 1 é dada uma quantia de dinheiro (digamos \$100) e diz-se a ele que ele tem que oferecer uma parte disso para o jogador 2. Se o jogador 2 aceita a oferta, o dinheiro é distribuído de acordo com a oferta. Mas o jogador 1 também é avisado de que nenhum jogador receberá qualquer dinheiro se o jogador 2 rejeita a oferta. Os jogadores não conhecem um ao outro, e eles são informados de que nunca vão jogar um com o outro novamente. Seria racional para o jogador 1 fazer uma oferta de \$0,01 e manter \$99,99 para si mesmo. Afinal, isto maximizaria os retornos, e o jogador 2 seria racionalmente forçado a aceitar a oferta desigual, porque \$0,01 é melhor do que não receber nada. Na verdade, porém, os jogadores no jogo geralmente oferecem muito mais – normalmente em torno de 40 por cento das ofertas do prêmio – e ofertas grosseiramente desiguais são rejeitadas. Parece que preferimos punir aqueles que fazem ofertas injustas e receber nada do que receber uma oferta injusta. Isto é transculturalmente verdadeiro, e é verdade quando o prêmio inicial é consideravelmente maior do que \$100. Todavia, Henrich et al. encontraram considerável variação transcultural. Em média, os Machiguenga do Peru faz e aceita ofertas de 26 por cento do prêmio, em vez de 40 por cento. A razão pode ser econômica. Os Machiguenga tem uma economia de corte-e-queima que não requer muita cooperação fora da família, então eles não são da prática de cooperar com estranhos. Normas que promovem a equidade nunca foram muito importante para eles. Compare esse caso com as pessoas Gnau e Au de Papua-Nova Guiné. As ofertas média feitas por essas pessoas estão perto das ofertas feitas por indivíduos estadunidenses, mas, ao contrário de nós, eles muitas vezes rejeitam ofertas que são hiperjustas (oferece mais de 50 por cento). A razão para isso parece ser a de que eles têm uma prática cultural de dar presentes, em que receber um presente exige fortemente reciprocidade. A generosa oferta pode ser um fardo, porque tem que ser recíproco. Os Gnau e Au parecem ter moralizado reciprocidade tão profundamente que falhar em retribuir é considerado uma ofensa muito grave.

Outra evidência para diferenças transculturais nas atitudes com relação à equidade pode ser encontrada através da apresentação de temas com perguntas sobre a alocação de recursos. Há um certo número de princípios que podem ser usados para determinar a forma de distribuir recursos. Pode-se dar aos necessitados, para aqueles que fizeram a maior parte do trabalho, ou para todos igualmente. Essas estratégias, a saber, necessidade, equidade e igualdade, têm sido amplamente investigada em estudos transculturais (Smith e Bond, 1998). Num estudo, Berman et al. (1985) pediram a indivíduos indianos e americanos para decidir se um bônus deve ser dado a um trabalhador rico que trabalhou muito duramente, ou a um trabalhador moderado que era carente financeiramente, ou se o bônus deve ser dividido igualmente entre eles. Metade dos indivíduos americanos disseram que a gratificação deve ser atribuída, com base na equidade, para aquele que trabalhou duramente, e apenas 14 por cento disse que o dinheiro deveria ir para o trabalhador carente. Entrevistados indianos tiveram a resposta oposta: metade atribuiu o bônus para o carente e apenas 14 por cento atribuiu com base na equidade. Esta diferença de valores pode estar relacionada ao fato de que há mais pobreza na Índia, e as pessoas, lá, consideram o cuidado com os necessitados como um valor importante. Em outro estudo, Leung e Bond (1984) descobriram que os indivíduos chineses preferem divisões iguais a divisões equitativas quando eles estão dividindo bens entre os amigos, e os indivíduos americanos preferiram equidade à igualdade. Uma explicação para esse efeito é que os americanos tendem a ter valores individualistas, os quais enfatizam a autoeficácia e o sucesso. As pessoas na China se preocupam com essas coisas também, mas elas tendem a ter uma orientação coletivista, o que significa que a harmonia do grupo, especialmente entre amigos e familiares, pode ser mais valiosa do que o sucesso individual.

Até agora, tenho focado em variações culturais na bondade e na equidade. Estas são duas das áreas em que há evidências de disposições comportamentais biologicamente preparadas, e eu argumentei que as culturas ampliam essas disposições de várias maneiras. A cultura também pode afetar a forma específica que o nosso comportamento toma. Considere a reciprocidade. Entre muitos animais

não-humanos, comportamento de reciprocidade é altamente estereotipado. Um morcego-vampiro retribui regurgitando sangue. Chimpanzés retribuem com cuidados com a aparência, com a partilha de alimentos, ou com a formação de alianças para a agressão e para a defesa. Mas o que dizer de nós? Não parece haver um padrão fixo. Podemos retribuir da mesma forma que os chimpanzés fazem, mas também de muitas outras formas. Um programa genético para a reciprocidade seria totalmente inadequado no nosso caso. Se os genes nos forneceram instruções específicas sobre como retribuir, elas seriam demasiadamente rígidas. E se os genes nos forneceram o comando genérico "retribuir!", isso seria demasiadamente vago para traduzir em ação. Regras de reciprocidade são altamente específicas, mas elas não são nem estereotipadas, nem fixas. Que presentes exigem reciprocidade? Seria bizarro reembolsar uma instituição por conferir um prêmio pelo sucesso. Quem deve retribuir? No Extremo Oriente, as dívidas podem ser passadas de geração em geração. Quanto retorno é o suficiente? Nós devolvemos o mesmo favor ou fazemos algo completamente diferente? Pode ser desagradavelmente impensável dar os mesmos presentes de aniversário que recebemos. As regras sutis e variáveis que regem a reciprocidade humana sugerem que a cultura está fazendo uma contribuição essencial. Em macacos e morcegos, a reciprocidade pode ser identificada com disposições comportamentais específicas. No nosso próprio caso, deve ser muito mais plástico.

Isto sugere que a contribuição biológica para a reciprocidade humana não é um comando que poderíamos realizar por conta própria. Se não tínhamos nada exceto a biologia para seguir em frente, o impulso de retribuir pode não ser comportamentalmente específico o suficiente para resultar em ação. Não saberíamos como retribuir. Em vez disso, a contribuição biológica pode ser mais como um andaime sobre o qual construir normas executáveis. Por exemplo, poderíamos ter nascido com um senso inato de gratidão quando nossos desejos foram realizados por outros. Um senso inato de gratidão não iria emitir diretamente em ação. Para mostrar a nossa gratidão, nós teríamos que aprender maneiras culturalmente específicos de reembolsar as nossas dívidas.

Em suma, proponho que os comportamentos de base biológica relativos à bondade, à equidade e à reciprocidade sejam culturalmente maleáveis e insuficientes para orientar o nosso comportamento sem elaboração cultural. Eu acho que a cultura faz duas contribuições para os *inputs* biológicos. Primeiro, ele converte esses comportamentos em normas morais, baseando-as em emoções morais. Em segundo lugar, toma as normas de base biológica que têm efeitos comportamentais, altamente estereotipados e limitados, em nossos primos primatas e os transforma em instruções culturalmente específicas sobre o que devemos fazer para quem. É este último aspecto de elaboração cultural que estou enfatizando nesta discussão de normas de reciprocidade.

A cultura faz uma contribuição paralela à classificação e à sexualidade. A primeira coisa a notar sobre classificação é que ela é altamente variável entre as culturas. Algumas culturas são altamente estratificadas, e algumas são igualitárias. Culturas igualitárias não são desprovidas de hierarquias de dominância social, mas indivíduos dominantes não exercem uma quantidade desproporcional de poder ou possuem mais bens. Nas sociedades igualitárias, os indivíduos dominantes (quase sempre homens) são aqueles que ganham o respeito pela sua sabedoria, habilidades, carisma, ou generosidade. Estes indivíduos de status elevado são conhecidos na antropologia como chefes. Chefes fazem o que todo mundo faz; eles simplesmente trabalham mais e doam mais. Eles não estão em condições de comandar pela força. Seus julgamentos sobre o que o grupo deve fazer são tomados como recomendações, não como comandos autoritários. Marvin Harris (1989) descreve o surgimento de estratificação enquanto as sociedades evoluem a partir de bandos igualitários em direção aos estados. Em algumas sociedades, as pessoas vão tentar superar umas às outras com façanhas de generosidade. Eles vão reunir amigos próximos e familiares para que eles possam dar festas luxuosas para os membros de suas aldeias. A pessoa que dá a festa mais luxuosa torna-se um grande homem. O grande homem é capaz de recrutar mais adeptos e seguidores, e ele pode, assim, manter seu *status* em face da concorrência. Nas sociedades onde existem abundantes fontes de alimentos não-perecíveis, o líder pode criar depósitos e redistribuir alimentos em épocas de escassez. Esses líderes

alcançam um novo nível de respeito e autoridade. Eles se tornam chefes. O seu papel como fornecedores é vital, e os moradores os recompensam com tributos, produtos especiais e serviços. Recompensas especiais são dadas para aqueles que oferecem o apoio mais fiel. Chefes não precisam mais caçar ou coletar alimentos. Eles acumulam riqueza e passam-na para seus filhos, que dizem ter o direito divino de levar adiante a liderança quando o chefe tiver morrido. Para aumentar o seu prêmio, chefes organizam exércitos, travam guerras contra vizinhos, e expandem a sociedade. Sistemas de tributo podem ser estendido para as aldeias vizinhas, impostas pelo controle militar. À medida em que aumenta o tamanho do grupo, o trabalho se diversifica, e aqueles que cumprem diferentes papéis acumulam diferentes graus de riqueza. O Estado nasce.

Em cada etapa deste processo, a classificação assume um novo significado. No final, pode haver vários níveis de dominação baseada, não na idade ou sexo, mas em recursos acumulados, alianças estratégicas, e papéis sociais. Essa estratificação é uma consequência de uma predisposição biológica a ter sistemas de classificação, mas é culturalmente transformada. As culturas também determinam os meios pelos quais alguém pode mudar de classificação. Pode ser uma questão de mérito, ou de formação, ou de casamento. Em algumas sociedades, o movimento entre as camadas sociais é praticamente impossível. Na Índia, a mudança de castas era tradicionalmente proibido. Dizia-se que as castas eram fixadas no nascimento e associadas a diferentes níveis no ciclo da reencarnação. Mudar de casta seria uma forma grotesca de pular níveis na subida para o nirvana. Esta justificação cosmológica pode ter uma história política. Apesar de controversa, acredita-se que o subcontinente indiano foi invadido pelos europeus por volta de 1500 a.C.. A população indígena foi suprimida, e a casta que conhecemos como os intocáveis foi criada. Agora, mais de dois milênios depois, testes genéticos mostram que indianos de casta superior são geneticamente mais próximos aos europeus do que indianos de castas inferiores (Bamshad et al., 2001). Isto sugere séculos de proibições muito bem sucedidas contra a mobilidade ascendente.

O sistema de castas indiano, certamente, não é inato, mas pode ser reforçado por uma tolerância natural de hierarquias de dominância. É uma instituição

biocultural, reforçada por emoções de obediência, de deferência, e de superioridade. As emoções referentes à classificação são provavelmente universais, mas suas expressões podem variar. Sentimentos de deferência são amplamente marcados pela reverência, por tirar o chapéu, por se ajoelhar, ou outras formas de abaixar o corpo, que podem ser consequências simbólicas culturalmente adaptadas dos gestos de abaixar a cabeça observados em estudos sobre deferência animal.

A cultura pode desempenhar um papel na determinação da prevalência de emoções referentes a hierarquias de dominância. Sociedades igualitárias tendem a suprimir sentimentos de superioridade. Ao estudar os Kung, Lee (1969) comprou um grande boi e ofereceu-o como um presente. Os beneficiários reclamaram que o boi estava muito magro para encher suas barrigas. Na realidade, eles dispensaram o presente como algo sem valor, porque não queriam que Lee sentisse que ele era superior (citado em Harris, 1989).

Tal como acontece com a superioridade, as culturas também parecem variar na medida em que elas incentivam sentimentos de obediência. Lembre-se dos experimentos de Milgram (1974) discutidos no capítulo 4. Indivíduos causavam choques elétricos quando eles eram convidados a fazê-lo por uma figura de autoridade. Estes resultados foram uniformes em diferentes tipos de personalidade, de gênero e de grupos sociais. Mas eles não foram uniformes culturalmente. Ao passo que os indivíduos americanos eram totalmente obedientes 65 por cento do tempo, os indivíduos alemães eram totalmente obedientes 85 por cento do tempo (Mantell, 1971). Isso se encaixa no estereótipo cultural que os alemães são propensos a obedecer ordens. Um padrão contrastante foi observado quando um estudo do estilo de Milgram foi realizado na Austrália: lá, apenas 40 por cento dos indivíduos do sexo masculino e 16 por cento dos indivíduos do sexo feminino foram totalmente obedientes (Kilham e Mann, 1974). Esta tendência para ignorar as autoridades podem reportar aos dias da Austrália como uma fronteira inexplorada, longe do braço do direito europeu.

Estes exemplos atestam a plasticidade da predisposição humana para formar classificações sociais. Há pouca evidência de que nossos genes nos dizem como organizar grupos sociais. Ao contrário, eles nos provêm com a obediência, a

capacidade de competir, e as emoções associadas ao *status*, os quais nos fazem suscetíveis à formação de hierarquia. Fatores culturais determinam quais características garantem o alto escalão e o grau de estratificação. Normas de classificação emergem através da interação biocultural. Estas normas são frequentemente moralizadas. É considerado como moralmente errado agir de forma inadequada para a sua posição na sociedade; por exemplo, nós pensamos que é errado para um adolescente para falar desrespeitosamente a um professor sem autorização. No capítulo 2, observei que temos aqueles que violam a classificação em desprezo, que é uma mistura de raiva e desgosto. Esses sentimentos morais com relação a violações de classificação são generalizados, mas eles provavelmente não são inatos. A mistura de raiva e de desgosto que usamos para fazer cumprir as normas de classificação pode ser um subproduto de como podemos transmitir tais normas na sociedade humana. Dentro de sociedades de símios, a classificação é imposta pela afirmação de poder e por favores. Macacos dominantes são mais fortes e mais generosos. Aqueles que disputam em vão por poder sentem-se deferentes, e, talvez, um triste sentimento de derrota (ver Price et al., 1994, sobre as origens evolutivas da depressão). Entre os seres humanos, a classificação não é estabelecida apenas por poder; é ensinado. Somos instruídos a respeitar a autoridade. Somos informados de que certos indivíduos (pais, líderes, anciãos, etc.) *têm direito* à sua elevada posição social. As pessoas que não obedecem a essa regra não são apenas imprudentes (como no caso do macaco); eles estão violando os direitos das pessoas no poder. Nesta conceitualização, uma violação da classificação é um crime contra outra pessoa, e crimes contra as pessoas provocam raiva. Nós também tendemos a conceitualizar nossas hierarquias de dominância, como se fossem parte da ordem natural ou sobrenatural. Considere os direitos divinos dos reis, ou o sistema de castas indiano, ou o fato mais geral de que as pessoas *nascem* em classes sociais, ou o fato de que algumas formas de autoridade são correlacionadas com os fatos biológicos (paternidade ou idade, por exemplo). Nesta conceitualização, violar a classificação é transgredir contra a natureza ou a divindade. Atos não naturais provocam nojo. Assim, tanto a raiva quanto o nojo são consequências inevitáveis de violações de

classificação, por causa da maneira que as culturas conceituam classificação. A raiva e o nojo se misturam no desprezo. Macacos têm hierarquias de classificação sem moralização, e, tanto quanto sabemos, não experimentam desprezo.

Permitam-me passar, agora, à moralidade sexual, começando com a infidelidade. Primatas diferem no grau em que as relações são consideradas como exclusivas. Gorilas têm seus haréns, mas chimpanzés e bonobos são bastante promíscuos. Animais dominantes, como quaisquer parceiros desejáveis, frequentemente são os primeiros escolhidos. Macacos e babuínos machos costumam brigar violentamente por direitos ao acesso sexual, e ligações com os subordinados são severamente punidos. Claro, a infidelidade ocorre. Macacos terão encontros secretos quando puderem, e até mesmo espécies de aves conhecidas por suas relações monogâmicas de longa duração, muitas vezes, esgueiram-se em algum romance paralelo. Não há evidência de que os animais não-humanos consideram tais comportamentos como imoral, ao invés de meramente arriscado. Os animais não-humanos não parecem exibir preocupações de terceiros quando se trata de infidelidade. Em contraste, nós condenamos os adúlteros, mesmo quando não somos diretamente afetados por suas ações. Nem sempre, nós fazemos os adúlteros usar letras escarlates ao redor de seus pescoços, mas os fazemos se sentirem envergonhados expressando publicamente a visão de que a infidelidade é terrivelmente errada. Eu acho que a atitude moral é um complemento aprendido em relação à política sexual encontrada em outros primatas. Símios e macacos lutam pelo acesso sexual, e o autorizaçãõ sexual é regido pela força, em vez de pelo direito. Os seres humanos foram além deste sistema. Criamos instituições, como o casamento, para restringir o acesso sexual, e essas instituições são baseadas em contratos, em leis, e, acima de tudo, em sentimentos. Promovemos fidelidade por moralizar isso.

Um nativista com respeito a normas de infidelidade pode opor-se a esta análise, argumentando que temos, um mecanismo psicológico inato de domínio específico para responder à infidelidade: o ciúme romântico. Mas a alegaçãõ de que o ciúme é uma emoçãõ inata pode ser posta em questão. Eu acho que o ciúme é, na verdade, uma mistura adquirida de raiva, de medo, de tristeza e de desgosto (Prinz,

2004). Todas essas emoções surgem naturalmente no contexto de infidelidade. Ficamos furiosos quando a nossa confiança é violada, com medo de enfrentar os pretendentes concorrentes, tristes com a perda potencial de alguém que amamos, e revoltados com a perspectiva de que essa pessoa tenha sido contaminado. Assim, essas emoções inevitavelmente se misturam quando somos romanticamente traídos, e usamos o termo ciumento para rotular essa mistura. Nós não precisamos de evolução para nos fornecer uma emoção que é dedicada a esta função. Nem precisamos postular as diferenças de gênero inatas para explicar por que as mulheres são mais preocupadas com a infidelidade emocional do que a infidelidade sexual. Na maioria das sociedades, a mulher ainda depende dos homens para o apoio material. Perder um parceiro masculino é, quase sempre, uma dificuldade financeira para as mulheres. Observe, porém, que isso pode não ter sido verdade no Pleistoceno. Talvez lá, como em algumas sociedades de primatas e de humanos, as crianças foram criadas por comunidades inteiras sem qualquer contribuição desproporcional daquele homem que era o progenitor. Algumas sociedades de caçadores-coletores contemporâneos têm tais acordos (Kelly, 1995). Se as sociedades humanas iniciais foram organizadas desta forma, então a suposição de que os homens são biologicamente programados para se preocupar com a paternidade incerta e as mulheres são programadas para se preocupar em manter por perto o ganha-pão é ridícula. Pode-se também ter problema com a alegação de que as mulheres se preocupam mais com a fidelidade emocional do que com o sexo (CR Harris, 2004). Esta constatação experimental pode simplesmente mostrar que as mulheres são desencorajadas de expressar interesse em sexo. Esse fato bem conhecido pode ser um produto culturalmente orientado da dominação masculina (ver, por exemplo, Hrdy (1999) sobre a evidência biológica para um forte desejo sexual feminino). Quando as mulheres recebem uma tarefa de memória cognitivamente exigente pouco antes de serem perguntadas sobre infidelidade, eles respondem mais como homens. A tarefa de memória impede de censurar sua resposta instintiva, que é que a infidelidade sexual é pior do que a infidelidade emocional (DeSteno et al., 2002). Além disso, respostas para a questão da infidelidade variam entre as culturas. Na Holanda, onde há maior liberdade sexual e

um alto grau de emprego para as mulheres, as respostas masculinas e femininas são muito semelhantes (Buunk et al., 1996). Todos dizem que a infidelidade emocional é pior do que a infidelidade sexual. Talvez os holandeses estejam escondendo suas verdadeiras respostas, ou talvez, para eles, o sexo só não seja uma grande coisa. De qualquer forma, o padrão de resultados não dá suporte à tese de que temos uma emoção especial evoluída a ponto de impor normas de infidelidade.

Isto não é negar que temos uma tendência evoluída para punir a infidelidade. Nós provavelmente a temos. Meu ponto é, em vez disso, que as regras morais contra a infidelidade e as emoções que as baseiam refletem interação biocultural. A evidência mais direta para isso vem da variação cultural nas normas de infidelidade. Na maioria das culturas, a infidelidade é proibida, mas, em muitas, os homens recebem mais liberdade. Psicólogos evolucionistas querem nos fazer crer que este duplo padrão é uma capitulação cultural do desejo inato que cada homem tem de espalhar sua semente. Se a infidelidade em macacos e aves é alguma indicação, no entanto, as mulheres igualmente cobiçam múltiplos parceiros. O duplo padrão é melhor explicado, portanto, pela dominação masculina no poder político. Um viés pró-masculino nos tribunais não é um produto da evolução mais do que um viés pró-branco. Além disso, há sociedades em que as mulheres estão autorizadas a ter parceiros extraconjugais. Por exemplo, Wood (1999) relata que, entre os nômades Gabra do leste da África, as mulheres casadas são incentivadas a ter amantes. Ironicamente, os Gabra ainda têm um duplo padrão. Homens solteiros estão autorizados a ter relações sexuais, mas quando as mulheres não casadas fazem sexo, elas são banidas. O casamento é a chave para a libertação sexual entre os Gabra. Isto é o inverso das nossas normas infidelidade.

Variação igualmente surpreendente pode ser encontrada quando se trata de incesto. Tabus contra incesto variam drasticamente. Entre os Zoroastrians, uniões sexuais dentro da família imediata teriam sido consideradas moralmente preferíveis a uniões exogâmicas (Slotkin, 1947). Os Na da China têm um sistema de sexualidade aberta antes do casamento, e não mantêm o controle de paternidade; então, é possível que uma mulher acabe na cama com seu pai biológico (Hua,

2001). Tongas homens na África Oriental dizem ter relações sexuais com suas filhas antes de caçar leões (Junod, 1962). Entre os Burundi, as mulheres dormem com seus filhos recém-casados como um tratamento para a impotência (Albert, 1963). Alega-se que relações sexuais entre mães e filhos também ocorrem de tempos em tempos no Japão moderno. Segundo relatos, que podem ser muito exagerados, algumas mães apaziguam apetites sexuais de seus filhos, de modo que eles não fiquem distraídos com o namoro com meninas de sua idade quando eles estão estudando para seus importantíssimos exames. Em 1980, as preocupações sobre o incesto entre mãe e filho atingiu um patamar bastante alto, semelhante a preocupações sobre incesto entre pai e filha em nações ocidentais (Allison, 2000).

O incesto entre irmão e irmã é, talvez, a forma mais comum do sexo culturalmente sancionada entre membros imediatos da família. Entre a realeza inca, a havaiana e a egípcia, foi, muitas vezes, prescrito. Muitos dos imperadores da era Júlio-Claudiana tiveram relações incestuosas. Calígula tinha um relacionamento sexual com sua irmã, Agripina, que se casou com seu tio, Cláudio, cuja filha se casou com seu filho, Nero. No Egito ptolemaico, até 21 por cento dos casamentos registrados entre os imigrantes gregos e romanos foram entre irmãos. Eles podem ter emulado os governantes ptolemaicos, dos quais oito, dos treze, casaram-se com suas irmãs. Também é plausível que estes cidadãos gregos e romanos foram incentivados a ter relações incestuosas, a fim de consolidar o poder. Os Ptolomeus governaram o Egito como um estado de *apartheid*, e impuseram leis que impedem gregos e romanos de se casar com egípcios nativos; se um egípcio assumia um nome estrangeiro, era uma ofensa capital. Uma vez forçados a se casar dentro de seus próprios grupos nacionais, gregos e romanos, que viviam espalhados por grandes cidades, ficaram com poucos potenciais cônjuges entre os quais escolher. Ao incentivar o incesto, os líderes minam qualquer tentação que esses imigrantes poderiam ter tido de se casar com a população indígena (Shaw, 1992).

Apesar de incomuns, esses casos lançam algumas dúvidas sobre a hipótese de Westermarck de que uma forte aversão ao incesto é desencadeada por ser levantada no mesmo lar. Essa hipótese também é desafiada pelo fato de que, em

algumas sociedades de pequena escala, aldeões que são criados em proximidade acabam se casando com outro. A evidência empírica utilizada para apoiar a hipótese de Westermarck também pode ser contestada. Primeiro, houve a elevada taxa de insucesso em casamentos entre menores de Taiwan, onde as meninas são adotadas por famílias em uma idade jovem e, então, criadas para ser as esposas dos meninos dessas famílias. O problema com esta descoberta é que a taxa de fracasso pode resultar de fatores que não têm nada a ver com evitar o incesto. Casamentos de menores são arranjados pelas famílias pobres que não podem pagar um dote e a pobreza pode levar à instabilidade conjugal; porque não há dinheiro em jogo, os pais de crianças em casamentos de menores têm menos interesse financeiro em vê-los ter sucesso (Harris, 1989: 201); casamentos de menores são considerados incomuns e inferiores aos casamentos de adultos; e casamentos de menores permitem à noiva e ao noivo exercer menos escolha ao selecionar um parceiro de casamento. A segunda linha de evidência para a hipótese de Westermarck foi provida pelas baixas taxas de casamento entre homens e mulheres que foram criados juntos em kibutzim israelenses. Em resposta, Harris (1989) aponta que a taxa de casamento entre membros do mesmo kibutz é realmente muito alto quando se inclui os casais que não foram educados na mesma classe. O fato de que homens e mulheres da mesma classe não se casaram pode ser explicado pelo fato de que os homens se casaram com as mulheres que eram, em média, três anos mais jovens do que eles. Além disso, não deve ser surpreendente se quase não houvesse casamentos entre pessoas criadas juntas em um kibutz, porque os jovens tipicamente saem de casa para servir nas forças armadas e ingressar de universidades, e, muitas vezes, não voltam; assim, as pessoas que são criadas em kibutzim vão embora no momento em que começam a procurar por cônjuges. Finalmente, houve a descoberta de Lieberman et al. de que as atitudes negativas em relação ao incesto entre irmãos são maiores para pessoas criadas com irmãos do sexo oposto. Aqui há uma explicação simples. Proibições contra o incesto entre irmãos são inaplicáveis às pessoas que não têm irmãos. Aqueles que têm irmãos são mais propensos a serem ensinados sobre a proibição (por meio de instrução explícita ou, mais provavelmente, da expressão de emoção

quando o assunto vem à tona), e eles estão sob mais pressão para internalizá-la. Adolescentes do sexo masculino que estão interessadas em sexo são especialmente suscetíveis de ficar sob a pressão social para se comportar adequadamente com suas irmãs; portanto, não é nenhuma surpresa que eles são especialmente sensíveis a respeito de incesto entre irmãos.

A hipótese de Westermarck também não consegue explicar o alcance do tabu do incesto, que, na maioria das sociedades, estende-se a parentes que não são membros da família. Na sociedade ocidental atual, as pessoas são repelidas pelo casamento entre primos, apesar de primos no Ocidente serem criados distantes uns dos outros. Existem muitas sociedades que estendem tabus do incesto à família por afinidade (relacionado somente por casamento). Até recentemente, era ilegal na Coreia casar-se com alguém com o mesmo sobrenome. Isso indica que os tabus do incesto não são meramente uma codificação de uma disposição inata para insultar aqueles com quem nós coabitamos durante o desenvolvimento. E, correlativamente, os tabus do incesto baseados na cultura não podem ser concebidos apenas para proteger contra a endogamia. Na verdade, os riscos de endogamia podem ser exagerados. Se um pequeno grupo de pessoas reproduz juntos ao longo de um período muito longo de tempo, podem eliminar genes recessivos deletérios. As pequenas tribos de caçadores-coletores (para não falar de animais de raça pura) são frequentemente muito saudáveis, apesar de uma profunda falta de diversidade genética.

Eu não pretendo inferir de tudo isso que não há nenhuma disposição biológica para evitar o incesto em seres humanos. Provavelmente há. Eu apenas não estou convencido pelas evidências quanto à hipótese de Westermarck. Uma hipótese alternativa é que, quando as crianças atingem a maturidade sexual, eles desenvolvem um interesse sexual em pessoas de fora da família. Tais preferências exógamas nos protegeriam contra os riscos potenciais, mas também confeririam um outro benefício: se tivéssemos que ficar em um grupo todas as nossas vidas, e esse grupo se abateu sobre um terrível destino através de doença, de um desastre, ou de predação, então, os genes no grupo morreriam; se, no entanto, descendentes mudassem os grupos, os genes da família seriam espalhados geograficamente, e

as perspectivas de sobrevivência aumentariam. Seja qual for a função evolutiva da exogamia, ela parece ser o padrão em nossos parentes animais mais próximos. Chimpanzés deixam comunidades natais antes de procurar parceiros sexuais. No caso dos chimpanzés comuns, são as fêmeas que saem, e no caso de bonobos, os machos. Nos seres humanos, não existe assimetria aparente no desejo de viajar [*wanderlust*] sexual. Ambos os humanos, machos e fêmeas, podem formar um desejo de sair de casa em busca de parceiros por volta do início da puberdade. Ou, talvez, tenhamos um botão inato que pode ser definido de forma diferente nas sociedades matrilocais e patrilocais.

Preferências de reprodução exógamas explicariam evitar o incesto. É uma característica típica das preferências sexuais de que você será repellido pela idéia de sexo com pessoas que você não deseja abertamente. Desinteresse gera desgosto. Se você não está atraído por alguém sexualmente, o pensamento de sexo com eles é repelente, não importando se eles são parentes ou não. Acho que a resposta do nojo associada a evitar o incesto é simplesmente um exemplo desta tendência mais geral. Se estamos biologicamente programados para ser exogâmicos, então não seremos atraídos por parentes, e, se não somos atraídos por parentes, vamos sentir nojo deles. Nesta história, a resposta do nojo não tem especialmente nada a ver com incesto. O mecanismo inato que é específico para o incesto é o pouco que compartilhamos com os macacos: uma atração para pessoas de fora do grupo. Em um aspecto, esta proposta é bastante semelhante à hipótese de Westermarck: evitar o incesto envolve uma falta de interesse sexual em pessoas com quem se coabita. Mas há algumas diferenças fundamentais. Na presente proposta, o tabu do incesto não é uma função da coabitação na primeira infância, que é um componente central para Westermarck, e não se manifesta primariamente como uma resposta de nojo dos irmãos. Ao contrário, evitar o incesto é uma consequência do desejo exogâmico. Esse desejo se manifesta como uma atração por parceiros de fora do grupo em que se cresce. Onde Westermarck focou em coabitação, a consideração exogâmica é neutra sobre como os grupos se definem (por exemplo, o grupo pode ser a casa natal ou a aldeia inteira). É possível que a cultura tenha uma mão na determinação dos limites de grupo relevantes.

Se esta história está no caminho certo, então mecanismo biológico para evitar o incesto não é uma norma moral – não é equivalente a um tabu do incesto. Uma coisa é evitar um comportamento, outra, bem diferente, é condenar esse comportamento. Com uma tendência natural para evitar o incesto, regras morais parecem ser desnecessárias. Nós não precisamos de regras morais contra comer carne podre ou contra cortar fora os nossos órgãos genitais. Não é surpreendente, então, que apenas 44 por cento das sociedades têm regras explícitas contra o sexo com parentes (Thornhill, 1991). Mas isso levanta uma questão. Por que evitar o incesto nunca foi moralizado? Pode haver vários fatores, mas uma provável possibilidade é que as regras de incesto são culturalmente construídas para incentivar a formação de alianças estratégicas com outras famílias, e para restringir a consolidação da riqueza dentro de qualquer família (Thornhill, 1991). Se é tabu dormir com parentes, e sexo está ligado ao casamento, então vai ser difícil manter a riqueza dentro de uma única família. Coerente com esta análise econômica, Thornhill aponta que sociedades com maiores graus de estratificação social tendem a ter penas mais severas para o incesto. As classes dominantes querem impedir os plebeus de consolidar a riqueza através da imposição de regras de incesto difíceis que se aplicam a todos, exceto aos membros da casa real. As classes dominantes podem tornar ainda mais difícil consolidar a riqueza através da definição de uma grande classe de relações como incestuosas, incluindo as relações com primos e parentes afins. As estratégias econômicas também podem estar na base de regras que são o inverso do tabu do incesto: proibições contra o casamento com quem é de fora do grupo. Em alguns grupos, como os beduínos, os indivíduos devem se casar com seus primos de primeiro grau. Esta é uma boa maneira de manter a riqueza dentro do grupo.

Em suma, não há razão para acreditar que os tabus do incesto têm uma origem biocultural. Nós temos uma tendência inata para evitar o incesto através de exogamia, mas esta tendência orienta o comportamento de maneiras culturalmente sensíveis. Primeiramente, a exogamia exige deixar o grupo, mas os limites de grupo em seres humanos são culturalmente definidos. Em segundo lugar, é possível que tanto homens quanto mulheres estejam dispostos a ser exogâmicos, mas, em

algumas culturas, essa disposição é expressa em apenas um sexo. Terceiro, e mais importante, evitar o incesto biológico não é uma norma moral. Ele só fica moralizado sob determinadas condições culturais, tendo a ver com a consolidação do poder e da riqueza. Quando o evitamento se torna tabu, culturas, muitas vezes, estendem a definição de família para incluir os indivíduos que não estão sujeitos a aversão biológica.

A moralização do incesto provavelmente funciona capitalizando sobre a capacidade humana do nojo. Nojo é a emoção negativa padrão para as coisas relativas ao corpo. Ele evoluiu para nos proteger de contaminação. Sexo envolve contato físico com outro corpo e transferência de fluidos corporais. Todos os fluidos corporais (exceto lágrimas) podem provocar uma repulsa física central. Esta não é uma resposta moral, mas uma resposta higiênica, fornecida pela seleção natural. Quando o incesto é moralizado, o nojo higiênico é reformulado como uma emoção moral. Através da educação moral, passamos a ver o casamento entre primos e outras relações benignas como repelentes. Com a moralização, o incesto também se torna vergonhoso. A vergonha surge quando sentimos que precisamos esconder nossos corpos, e um corpo que tenha sido contaminado por conduta sexual não sancionada é exatamente o tipo de coisa que se pode querer esconder.

Nesta discussão, tenho enfatizado a diversidade da moralidade humana, mesmo nos casos em que nossas normas morais têm alguma base na biologia. Ao longo do caminho, tenho tentado apresentar provas para duas conclusões. Primeiramente, as normas naturalmente evoluídas que eu introduzi na seção 7.1 não podem qualificar-se como normas morais em seu estado natural. Acho que a moralização, incluindo o surgimento de três tipos de sistemas éticos de Shweder, é o resultado do desenvolvimento cultural. Em segundo lugar, normas naturalmente evoluídas podem ser embelezadas, expandidas, e, talvez, reprimidas sob influência cultural. Influências culturais não estão limitadas a uma série fixa de resultados, mas estão em aberto.

Este ponto sobre o caráter aberto merece destaque especial. Neste contexto, quero desarmar uma objeção possível. Hoje em dia, todo mundo parece ser um interacionista. Nativistas morais estão, muitas vezes, dispostos a admitir que o

ambiente contribui para a moralidade; eles dizem que a maioria dos nossos fenótipos psicológicos são uma mistura de natureza e criação. Assim, pode parecer que não há diferença entre a minha opinião e a visão predominante entre os defensores da ética evolucionista. Mas as aparências enganam. Acho que eticistas evolucionistas subestimam sistematicamente as contribuições da cultura. Isto é certamente verdadeiro sobre psicólogos evolucionistas. Ao se dizerem como interacionistas, eles, muitas vezes, dão exemplos que sugerem um papel muito pequeno para o meio ambiente. Por exemplo, Cosmides e Tooby (1997) sugerem que um bom modelo de influências ambientais pode ser encontrado no bodião de cabeça azul. Estes peixes viajam em grupos com um macho e várias fêmeas. Se o macho morre, a maior fêmea se transforma em um macho. Isso parece um caso extraordinário de alteração no fenótipo ambientalmente impulsionada. O ambiente causa uma mudança de sexo! O problema é que isso é realmente apenas um caso de determinismo genético de flexão de gênero. Genes do bodião funcionam como um interruptor. Dependendo do ambiente, o interruptor alterna muito previsivelmente de masculino para feminino. O meio ambiente não está acrescentando nada de substancial. Está simplesmente provocando um resultado que foi pré-determinado pelos genes. As influências ambientais em que estou interessado são decididamente não-bodião. Variações na moralidade não se restringem a um conjunto fixo de configurações pré-estabelecidas em um interruptor interno. O ambiente tem um impacto muito mais substantivo sobre nós do que sobre o bodião de cabeça azul. O caráter aberto da moralidade sugere que as culturas em que vivemos, na verdade, contribuem para o conteúdo de nossas regras morais, em vez de selecionar a partir de um conjunto de regras que são pré-codificadas nos genes. A cultura não se limita a ativar programas inatos; ele reescreve o nosso *software* de moral.

Se eu estiver certo, então é um exagero dizer que as regras morais são produtos da evolução. Normas evoluídas podem ser pensadas como diretrizes esquemáticas, as quais são preenchidas através da interação com nossos ambientes culturais. Podemos ter evoluído para ser gentil com os membros do grupo mais próximo, mas a identidade dos membros do grupo precisa ser

trabalhada culturalmente. Podemos ter evoluído para respeitar as autoridades, mas o grau de estratificação e os critérios para o *status* precisam funcionar culturalmente. Podemos ter evoluído para evitar o sexo com parentes, mas o âmbito da proibição e a gravidade das violações precisam ser trabalhados culturalmente. A cultura nos diz como aplicar nossas normas naturais, e, em alguns casos (no incesto Zoroastriano? Na reciprocidade Ik?), ela até mesmo as substitui. Alguém pode colocar a questão dizendo que os seres humanos instintivamente tomam suas pistas da cultura. A cultura também converte nossas normas biológicas em normas morais – normas que se baseiam em sentimentos e são estendidas a terceiros. A evolução não é a fonte da moralidade; é, na melhor das hipóteses, um suporte através do qual nós podemos fazer nossa ascensão moral.

7.3.3 Conclusão: contra a natureza

E se a gente pudesse, de alguma forma, largar todas as influências da cultura e voltar para nossos valores pré-morais – os valores que teríamos em nosso estado natural? Para fazer isso, teríamos de eliminar o casamento e a propriedade, teríamos de abandonar as preocupações sobre justiça global, e teríamos que tolerar o castigo corporal e, talvez, algum canibalismo e estupro oportunistas. Nós ainda nos comportaríamos de algumas maneiras que poderiam ser descritas como respeitáveis. Nossos valores naturais nos levariam a cuidar uns dos outros se tivéssemos sido cuidados. Respeitaríamos a autoridade. Evitaríamos ter relações sexuais com os nossos irmãos. Nossas vidas provavelmente seriam bastante previsíveis. Ao contrário do *Übermensch* nietzscheano que cria a si mesmo – uma fantasia inventada sob o feitiço do Romantismo –, seríamos provavelmente muito parecidos com os chimpanzés.

Talvez possamos sustentar essa existência primitiva por um tempo. Talvez populações humanas iniciais eram assim. Mas não por muito tempo. Nossos ancestrais tinham moradias, roupas e ferramentas. Eles se engajaram em batalhas entre si, organizaram expedições de caça, e criaram organizações políticas para os seus clãs. Uma vez que essas coisas existem, os valores naturais já não podem

servir de guias simples para a ação. Precisamos decidir o que constitui a classificação, que bens dar quando nós retribuimos, e que laços românticos estão fora dos limites.

Será que estamos em piores condições para isso? Devemos nos esforçar para voltar à savana mítica? A resposta é óbvia. Por um lado, não podemos voltar a um estado puro, desinformado pelo conhecimento e pelas tecnologias que temos acumulado ao longo de milênios de evolução cultural. Por outro lado, não há nenhuma maneira de justificar a alegação de que o nosso estado natural é melhor simplesmente em virtude de ser mais natural. Nosso estado natural promove a construção de instituições e valores sociais. Nossos valores são agora, e sempre serão, produtos tanto da natureza quanto da educação. A ética evolucionista, esse equívoco, fornece uma visão furada da vida moral. É fascinante expor os grandes limites biológicos em que a moral cresce, mas uma investigação desses limites é grosseiramente incompleta e totalmente desprovido de implicações normativas.

8. PROGRESSO MORAL

8.1 O PROGRESSO É POSSÍVEL?

8.1.1 Dois tipos de niilismo

Nietzsche é, às vezes, chamado de niilista moral, e não sem justificativa. Ele queria enfraquecer a moralidade cristã, mas ele não queria eliminar toda a moralidade. Ele queria substituir valores existentes por valores melhores. A fórmula de Nietzsche para o desenvolvimento foi de identificar valores que são ativos, ao invés de reativos, ou naturais, ao invés de historicamente construídos. Eu levantei algumas dúvidas sobre este projeto no capítulo 7. Acredito que toda moralidade é construída. Isso interdita algumas avenidas para o progresso moral. Realmente, pode fazer parecer que o progresso é impossível. Se todo sistema moral é construção cultural, e nenhum deles tem melhor argumento sobre a verdade absoluta, quais fundamentos temos para dizer que qualquer sistema é melhor do que outro? Eu, brevemente, levantei essa questão no capítulo 6. Aqui, eu vou ter um pouco mais a dizer.

Para colocar o desafio do progresso mais precisamente, vamos distinguir dois tipos de niilismo moral. De acordo com o primeiro, todos os sistemas morais são vazios: os argumentos que eles oferecem são falsos. Teorias do erro da moralidade são niilistas neste sentido, incluindo as teorias de Mackie e Ruse. Moralidade é uma ilusão. Tenho argumentado contra as teorias do erro (capítulo 3). Acredito que reivindicações morais podem ser verdadeiras. Conceitos morais são indexicais que se referem a propriedades dependentes de resposta. O bem é aquilo que é um objeto de aprovação de um observador moral, e o mal é um objeto de desaprovação. Desde que eles sejam moralmente aprovados ou desaprovados, asseguramos a existência de fatos morais. O problema não é que fatos morais são falsos, mas que muitas afirmações morais são verdadeiras.

Há provavelmente milhares de moralidades ao redor do mundo, e um número ilimitado de moralidades meramente possíveis. Cada um destes tem uma afirmação

sobre verdade. Se argumentos morais são relativizados por sistemas de valores culturalmente inculcados em observadores morais, então qualquer argumento sobre o que moralidade demanda pode vir a ser verdade. Sempre há um sistema moral possível onde alguma ação é permissível, não importa o quão hediondo possa ser relativo ao seu próprio sistema. O bem de uma cultura pode ser o mal de outra. Na realidade, provavelmente há muita convergência entre culturas, mas, como vimos, também há muita diversidade moral. O problema é que, embora nós tenhamos preferência por nossos próprios sistemas morais frente ao sistema moral de outros, não parece haver nenhum sentido no qual nossas preferências sejam melhores do que as dos outros. O fato delas serem *nossas* preferências não fazem delas melhores preferências, a não ser para nós mesmos.

Isso nos traz para o segundo tipo de niilismo. Acontece que nenhum sistema moral é melhor do que outro sistema moral. Afirmações morais verdadeiras são essencialmente paroquiais. Minhas afirmações morais são verdadeiras, mas não mais verdadeiras do que a negação delas quando proferida por um membro de outra comunidade. Se é este o caso, então é difícil enxergar como podemos sequer falar sobre progresso moral. Ao movermos de um estado escravagista para um não-escravagista, ao lutarmos por igualdade de gênero, ao renunciarmos tortura e imperialismo, não estamos fazendo melhor do que fazíamos antes? Nossos valores atuais parecem ter melhorado frente aos valores do passado, e a humildade moral nos levou à esperança de continuarmos evoluindo. Se verdades morais são superficiais no sentido que havia descrito elas, então mudança moral não é um movimento para frente, mas um reposicionamento lateral no espaço de opções igualmente aceitáveis. Este é um niilismo sem direção em abundância. Mas devemos nos contentar com isso?

8.1.2 Além do Bem e Mal

Acredito que o progresso moral seja possível, mas penso que não conseguiremos seguir adiante movendo além dos nossos sistemas atuais de valores e pesando alternativas. Não penso haver uma instância transcendental da

qual podemos acessar teorias morais competitivas. Não há visão de lugar nenhum. Nós devemos sempre acessar o progresso de dentro para fora. Nesse sentido, moralidade não é diferente de ciência. Na tropa familiar de Otto Neurath popularizada por Quine, revisão de teoria é como reconstruir uma jangada enquanto estamos boiando no mar. Não podemos simplesmente abandonar nossa jangada e começar uma nova; precisamos substituir as tábuas na jangada que já temos. Isto é igualmente verdadeiro no domínio avaliativo. Não podemos simplesmente abandonar nossos valores atuais. O racista não consegue imaginar amar um membro da raça insultada mais do que podemos imaginar divertir-se com tortura. Devemos reequipar nossos valores por dentro.

Comumente, quando acessamos valores alternativos usando valores que já possuímos, o resultado é rejeição. Valores alternativos são percebidos como defeituosos, corruptos, ou, de alguma forma, piores do que os nossos. Nestes casos, progresso é possível apenas em uma retrospectiva, em algum sentido, vazia. A partir do presente, sempre parecemos ter valores melhores do que tínhamos no passado pela razão trivial de que nós adotamos nossos valores presentes e não mais os nossos valores passados. Os casos interessantes são aqueles nos quais consideramos a possibilidade de um valor alternativo ser melhor do que o valor que temos atualmente. Esse tipo de comparação é progressivo, ou com o olhar para o futuro. Nós podemos ver que o futuro moral pode ser melhor que o presente moral. Mas como isso é possível de dentro? Não é necessária uma instância transcendental?

Para abordar essa questão, ajudará pensar em como debates morais são conduzidos *entre* indivíduos. Uma técnica padrão é a oposição de normas. Se você acredita que alguém deve \emptyset , e seu interlocutor tem um valor oposto, você pode tentar trazê-lo para seu lado apontando para outro valor em seu repertório que possa ser usado para construir um caso a favor de fazer \emptyset . Em outras palavras, podemos usar uma norma a fim de vetar outra. Podemos fazer algo parecido em nossos próprios casos. Podemos juntar diferentes normas sobre a mesma questão, e dar a alguma norma precedência quando estas entram em conflito. A chave para a mudança moral está em opor valores atuais uns contra outros. Vivemos um tanto

complacentes com as inconsistências nos nossos valores, mas ficamos envergonhados quando elas são trazidas à nossa atenção. Quando confrontados com uma inconsistência, somos incentivados a eliminar um dos valores conflitivos e buscar um substituto.

Oposição de normas pressupõe que nossas normas são, algumas vezes, inconsistentes. Demanda que tenhamos duas normas que nos levem a direções conflitantes. Há boas razões para pensarmos que tais conflitos são possíveis. Sistemas morais não são construídos ou adquiridos em uma sentada. Eles são remendados juntos ao longo do tempo. Normas diferentes são planejadas para servir propósitos diferentes e para administrar diferentes aspectos das nossas vidas. Há uma preocupação filosófica em descobrir uma única norma moral que dite todos os nossos julgamentos (veja Williams (1985) para uma crítica). A norma do utilitarismo e do imperativo categórico são exemplos. Mesmo o mandamento “faça o que uma pessoa virtuosa faria” é altamente reducionista. Tais propostas filosóficas são inevitavelmente falhas se as utilizarmos para descrever como a psicologia moral comum funciona. As normas que seguimos são mais concretas e variáveis: diga “obrigado”; não cuspa em público; dê boas gorjetas; seja gentil com estranhos; limpe sua bagunça; não fure filas; peça desculpa quando descontar seu mau humor em pessoas que você ama. Raramente damos um passo atrás e tentamos substituir esses ditados cotidianos por um princípio mais abrangente. Quando buscamos princípios, descobrimos que muitos se aplicam. Essas normas se aplicam ao princípio da humanidade, ao princípio do utilitarismo, e ao comportamento de pessoas virtuosas; e ainda, eles não são produtos de nenhum destes princípios. Cada norma tem sua própria história, e, conseqüentemente, sempre há a possibilidade de que duas normas venham a colidir.

Imagine alguém que acredita que é aceitável não ajudar alguém morrendo de fome em partes subdesenvolvidas do mundo. Podemos apontar que essa atitude moral está em desacordo com a injunção de ser gentil com estranhos. Frente a isso, as duas atitudes são inconsistentes. Há várias formas de reproduzir inconsistências. Por exemplo, alguém poderia revisar a visão de permissibilidade de ignorar a fome no mundo. Alternativamente, alguém poderia adicionar uma

quantidade restritivas de normas sobre estranhos: seja gentil com estranhos que você venha a encontrar. Repare que provavelmente não seria abandonada por completo a norma de ser gentil com estranhos. Essa seria uma revisão mais radical do que o conflito precisa. Isso poderia também violar outras normas que se pode ter, incluindo valores sobre traços de caráter: seja bom, seja cuidadoso, seja generoso, seja simpático e assim por diante. Podemos valorar atributos assim como valoramos ações.

Este exemplo ilustra um ponto geral sobre normas morais. Esperamos que nosso sistema moral tenha algumas das mesmas virtudes teóricas que nossas crenças. Acima de tudo, elas deveriam ser consistentes. Ao revisar, frente à inconsistência, vamos usualmente preferir revisões conservadoras. O conservadorismo se manifesta de diferentes formas. Temos como objetivo o conjunto de normas com o máximo de consistência, e não mudamos mais do que precisamos. O ajuste periférico é melhor que a revisão radical. Também tentamos evitar mudar valores que possam ser pilares para outros valores – valores que são centrais para nossa rede de avaliação (compare com Quine e Ullian, 1978). A intensidade com a qual seguramos valores conflitivos também pode ser um fator. É difícil abandonar um valor que nos faz nos sentirmos especialmente passionais.

Podemos já começar a ver um início de resposta para o problema do progresso moral. O exemplo sendo considerado ilustra uma forma de como um sistema moral pode ser melhor do que outro. Um sistema moral pode ser mais consistente. Nós preferimos sistemas consistentes à inconsistentes. Demonstrar como estamos comprometidos com sistemas morais inconsistentes podem ocasionar uma revisão. Ao pesar sistemas morais alternativos de dentro do sistema moral que já possuímos, também temos uma preferência por aqueles que são similares aos que temos. Revisões conservadoras são mais desejáveis.

Inconsistência não é a única coisa que pode motivar uma revisão moral. Há uma variedade de outros padrões pelos quais se pode acessar algum sistema de valor. Considere a lista a seguir.

Algumas normas são apoiadas por apelos a falsos conhecimentos fatuais. Opositores do sufrágio feminino argumentaram que mulheres eram muito

psicologicamente delicadas para política. Essa era provavelmente uma desculpa *post hoc* para dominação masculina, mas ao sustentar o caso contra o sufrágio em um argumento factual falso, oponentes abriram as portas para revisão moral através da correção de crença. (Ironicamente, quando o sufrágio finalmente chegou, não foi convencendo pessoas de abandonar sua premissa factual. Ao contrário, em muitos Estados, homens brancos votantes concordaram com o voto para mulheres pois acreditavam que aquelas mulheres brancas nascidas nos Estados Unidos iriam usar seu voto para diluir o impacto corrosivo dos votantes imigrantes e afro-americanos. O ponto aqui é que uma correção de crença *podia* mudar a atitude frente ao sufrágio feminino.)

Algumas normas são relativamente fáceis de seguir. Outras são mais exigentes e difíceis de impor a outros. A revisão moral pode ser possível quando descobrimos que o valor atual é muito exigente. Mudar atitudes sobre o sexo pré-marital pode ser um caso de diminuição de exigências.

Algumas normas levam a uma maior estabilidade moral. Valoramos a ausência de ansiedade, e desvalorizamos mudanças disruptivas. Normas que levam a perseguição de pessoas dentro da sociedade são mais prováveis de produzirem instabilidade.

Algumas normas criam bem-estar social; elas nos protegem da dor e aumentam nosso prazer. Proibições contra danos são os exemplos mais óbvios.

Algumas normas aumentam nossa percepção subjetiva de bem-estar; elas nos ajudam a promover uma vida realizada, para e além dos confortos básicos. Por exemplo, normas podem ser designadas para promover educação, promover uma variedade de escolhas de vida, e apoiar o desenvolvimento de relações pessoais.

Algumas normas são mais gerais do que outras. Elas dão conselhos sobre uma ampla variedade de casos. A norma "dê boas gorjetas" se aplica a condições bem especiais. Ao invés de tratá-la como uma norma fundamental, nós podemos preferir a norma "demonstre gratidão pelos serviços disponibilizados". Esta norma pode ser aplicada para derivar injunções mais específicas, incluindo a norma da gorjeta, quando combinada com o conhecimento dos costumes locais.

Algumas normas são mais universais do que outras. Elas seriam consideradas sensíveis por um grande número de pessoas, em vez de poucas. Esta é uma vantagem, porque normas morais são usadas para coordenar comportamentos entre uma larga escala populacional.

Algumas normas são menos vulneráveis à crítica genealógica. Normas sobre dar aos necessitados continua (ao contrário de Nietzsche) valorável mesmo se elas surgem sob circunstâncias indignas.

Algumas normas são mais consistentes que com nossas normas biológicas pré-morais. Elas nos permitem ajudar e dividir, instalar autoridades dignas, e proteger nossos corpos de contaminação; tais normas foram o tópico do capítulo 7.

Cada um desses pontos podem ser tratados como um padrão de avaliação. Cada um dá um sentido no qual uma norma pode ser melhor que outra. A este respeito, elas provêm ferramentas para medir progresso moral. De duas normas morais competitivas, aquela que segue melhor estes padrões pode ser julgada como uma norma melhor. Este é um argumento empírico. A lista é sujeita a alterações empíricas.

Há várias coisas a se destacar. primeiro, os padrões podem competir. Por exemplo, o sistema mais benéfico (aqueles que aumentam bem-estar social [*welfare*] e bem-estar [*well-being*]) não são necessariamente as mais fáceis de se implementar. Isso não desqualifica a utilidade destes padrões, porque eles não levam sempre a conselhos inconsistentes. Normas que são universalmente desejáveis são mais prováveis de causar estabilidade, o que é mais provável de promover bem-estar.

Em segundo lugar, todos estes padrões de avaliação são, eles mesmos, valores. Consistência, coerência com os fatos, estabilidade, fácil implementação, e conformidade às normas biológicas são todas questões com as quais nos importamos. Se não déssemos valor a essas questões, elas não seriam vistas como uma vantagem para pesagem entre normas morais. Até nossa preferência por consistência é um julgamento de valor.

Em terceiro lugar, os padrões podem variar de acordo com a cultura. Por exemplo, há algumas variações culturais para promoção de bem-estar. Quando

comparados a pessoas da cultura oriental, ocidentais derivam maior bem-estar de autossatisfação, emoções positivas, e prazer pessoal; orientais se preocupam mais com satisfazer obrigações, trabalhar por objetivos, e fazer boas ações aos outros (Diener et al., 2003; ver também Tiberius, 2003). Há inclusive diferenças culturais no nível com o qual pessoas se importam com inconsistências. No Extremo Oriente, onde a razão dialética é valorizada, a consistência é considerada menos importante.

Em quarto lugar, os padrões sob consideração não são normais morais. Consistência, estabilidade, bem-estar, e até mesmo conformidade com normas biológicas são coisas que valoramos em um sentido extramoral. Quando implantamos esses padrões na avaliação de sistemas morais estamos saindo da moralidade. Nós estamos indo além das categorias de bem e de mal. Um sistema moral que está mais adiante em uma dessas dimensões não é mais moral. Nem é mais verdadeiro. Lembre que verdades morais vêm de forma barata. As dimensões que tenho discutido são, em grande medida, pragmáticas, prudenciais e hedônicas. Um macaco ou mesmo um psicopata podem se importar com estes padrões.

Quando Nietzsche disse que deveríamos nos mover para além do bem e mal, ele quis dizer que devemos retornar aos padrões de avaliação que precedem a inversão do cristianismo às normas romanas. Não é isso que estou recomendando. Mas há uma certa verdade no conselho de Nietzsche. Quando revisamos a moralidade, devemos prestar atenção a todas as coisas que valoramos. Assim como podemos opor normas morais uma contra outras, podemos opor normas extramorais contra normas morais. Já que somos obrigados a revisar a moralidade de dentro, devemos nos favorecer de todos os recursos. Todas coisas com as quais nos importamos é potencialmente relevante na decisão de como melhorar nosso sistema moral corrente.

Usando critérios extramorais, é fácil enxergar como alguns valores morais são subótimos mesmo do ponto de vista daqueles que o endossam. Considere os Tasmanianos, algumas das muitas culturas em que os homens têm dominação extrema sobre as mulheres (Edgerton, 1992). Mulheres da Tasmânia eram tratadas como propriedades masculinas, e elas precisavam produzir a maior parte da alimentação. Essas mulheres eram provavelmente doutrinadas à aceitação de seu

status inferior, mas elas certamente reconheciam que o sistema de valores tinha consequências negativas para seu bem-estar. O sistema de exploração pode também fazer a sociedade vulnerável à instabilidade devastadora. Quando colonizadores europeus vieram, Tasmanianos trocaram várias de suas mulheres por cachorros. O resultado foi uma grave redução no suprimento de comida, já que as mulheres eram as principais provedoras, e um aumento significativo na violência, uma vez que aumentou a competitividade pelas mulheres restantes. A história da Tasmânia não é incomum. Muitas sociedades (incluindo a nossa) possuem alguns valores que aumentam a violência, promovem desigualdades, e esgotam recursos naturais (Diamond, 2004).

Os Tasmanianos foram erradicados pelos colonizadores europeus, então eles nunca tiveram a oportunidade de revisar seus valores. Mas é possível especular que, na medida em que as coisas foram de mal a pior, eles talvez tivessem inaugurado um programa de reforma moral. Mas repare que os valores na Tasmânia permaneceram os mesmos por várias gerações antes dos colonizadores chegarem. A razão para isso não é clara, mas pode ser derivada do fato de que homens estiveram, em grande medida, confortáveis com a sociedade da Tasmânia, antes de começarem a vender as mulheres. Já que os homens tinham poder e estavam confortáveis, houve pouco interesse na reforma. Essa é uma observação desanimadora, porque sugere que reformas são difíceis, na medida em que as pessoas no poder não estão sofrendo.

Esta questão surge: uma pessoa que não está sofrendo pode detectar valores disfuncionais e impulsionar uma reforma? Vamos considerar um caso teste. Suponha que Smith seja um americano branco afortunado nos meados do século dezenove, que acredita ser permissível que se tenha escravos. Smith acredita que pessoas de descendência africana são inferiores, e que a escravidão é uma boa forma de expô-los aos valores cristãos, o que os levaria ao desenvolvimento moral e os garantiria um lugar no céu. Ele também acredita que a escravidão é perfeitamente justa. Como os antigos, ele vê como um tipo de loteria. Se você tiver suficientemente azar de ser capturado e escravizado, você precisa aguentar, a não ser que você escape. Qualquer um poderia ser um escravo. Escravidão também é

necessária, Smith poderia argumentar. Precisa haver uma divisão de trabalho na sociedade, e, para garantir progresso econômico, um vasto número de pessoas precisa trabalhar sem pagamento.

Em reflexão, Smith pode reconsiderar sua atitude sobre a escravidão. Pode descobrir que sua presunção de inferioridade racial de africanos é baseada em uma ciência ruim, e teria que reconhecer que a doutrinação cristã de africanos é possível (para o bem ou para o mal) sem a escravização. Smith também pode vir a considerar que manter escravos é inconsistente com outros valores morais que ele pratica no seu dia a dia. Smith é gentil com estranhos; ele não agride seu cachorro; ele paga suas contas; ele recompensa as pessoas pelos seus méritos; ele aprecia a liberdade. É, no mínimo, questionável se estes outros valores podem ser reconciliados com a escravidão. Também é questionável se a escravidão pode ser um sistema estável. Escravos podem se revoltar. É possível manter uma economia escravocrata por séculos, mas outros arranjos sociais podem ser mais estáveis. De qualquer forma, a escravidão é difícil de ser mantida. Requer a imposição forçada de valores a pessoas que provavelmente irão resistir fortemente. Escravidão também reduz o bem-estar e o bem-estar social. Reduz profundamente a qualidade de vida dos escravos, o que entra em conflito com a universalidade, e também pode reduzir a qualidade de vida de Smith, na medida em que ele é naturalmente empático com ao sofrimento e movido pela predisposição biológica de ajudar pessoas em necessidade.

Smith também pode ganhar perspectiva ao refletir sobre a genealogia da comercialização de escravos. A escravidão na Europa pode ser traçada até a antiguidade, quando havia conflitos militares com outras nações, motivados por competição acerca dos recursos e pelo desejo por novos territórios. Este era o caso de Roma, e a escravidão em Roma não foi tornada ilegal pela Igreja; apenas evoluiu, com mudanças na economia e com o declínio do imperialismo, no sistema de contratos de servidão e de escravidão. Criminosos eram, algumas vezes, vendidos como escravos. No mundo Islâmico, a escravidão continua, e escravos africanos foram tomados como trabalhadores domésticos, concubinas e mineradores. Quando os portugueses foram para a África na busca de ouro, eles adotaram a

prática de compra de escravos para vender aos muçulmanos de lá, que controlavam muitos dos portos comerciais. Ao mesmo tempo, colonizadores no Novo Mundo escravizavam povos nativos americanos para trabalharem nas fazendas. Os povos nativos americanos provaram ser maus escravos, pois, frequentemente, rebelavam-se ou escapavam. Escravos africanos importados eram mais fáceis de serem capturados quando fugiam, porque eles se destacavam e não conheciam as terras, além de não serem mais imunes a doenças letais como a malária. Eles também vieram com habilidades impressionantes de terra natal, que tinha agricultura e ofícios mais avançados comparativamente. Na medida em que a agricultura foi progredindo nos estados do sul, na América do Sul e no Caribe, a demandas por escravos africanos aumentou. Esse escravos, inicialmente, trabalharam ao lado de fazendeiros brancos e eram tratados similarmente, mas a agricultura de fazendeiros diminuiu, e o tratamento foi se deteriorando porque os escravos eram mais baratos. Com o aumento das plantações de algodão depois da Revolução, a demanda por escravos aumentou de maneira abrupta. Apenas uma pequena percentagem de branco sulistas tinham escravos, mas alguns tiveram plantações tão grandes que os negros africanos superaram em número os brancos em muitas partes do sul [dos Estados Unidos]. Isto era usado como vantagem política, porque os escravos eram contabilizados nos censos, dando aos estados do sul um vasto número de votantes no colégio eleitoral, o que aumentava a chance do presidente dos Estados Unidos ser simpático à escravidão. Essa história mostra que houve uma evolução na escravidão: do uso de povos vencidos à captura intencional de seres humanos para serem usados como propriedade. Isso levou à uma degradação do *status* dos escravos, e aumentou a tolerância ao abuso grotesco. A escravidão continuou, pois era um negócio grande: o algodão era responsável por metade das exportações norte-americanas no período da Guerra Civil. Em 1808, a importação de escravos foi banida, e o comércio doméstico de escravos ficou restrito à americanos nativos. A escravidão já tinha sido consequência da derrota militar, mas os escravos passaram a ser criados como gado e forçados a vidas miseráveis de trabalho para que os donos de plantações pudessem enriquecer. Romanos argumentaram que a escravidão era um sistema justo, porque os escravos eram nascidos em condição

de igualdade de oportunidade e de forma justa nos campos de batalha; mas nas Américas, escravos eram criados, e nasciam em uma condição de opressão inimaginável. Argumentos pela necessidade econômica da escravidão também foram perdendo credibilidade: por volta do século dezanove, havia economias altamente rentáveis sem escravos em todo o mundo, e a América urbana tinha uma larga classe média.

Mesmo refletindo sobre essas questões, americanos do século dezanove como Smith provavelmente não se tornariam abolicionistas da noite para o dia. Ele estaria nas garras da visão de que a escravidão é moralmente permissível, e poucos dos fatos históricos recém destacados falaram diretamente sobre a moralidade da escravidão. Mas Smith provavelmente começaria a suspeitar que a escravidão é ruim em um sentido extramoral. Essa mudança de perspectiva seria um primeiro passo para um novo set de normas sobre a escravidão – valores que Smith reconheceria como melhores. Uma vez que ele começasse a formar atitudes emocionais negativas frente à escravidão em bases extramoriais, ele estaria preparado para um recondicionamento moral.

Esse exemplo fictício sobre uma conversa moral pode corresponder a alguns casos atuais. Antes do século dezanove, a escravidão era reconhecida por brancos como uma instituição progressiva que permitiria a expansão social, acumulação de riqueza, e um sistema econômico moderno pautado pelo comércio global. No século dezanove, atitudes morais mudaram, e a escravidão começou a se mostrar retrógrada (Davis, 1984). Seria bom pensar que essa transição foi ocasionada, por alguns, ao reconhecerem que a escravidão falha em preservar alguns padrões que tenho discutido, especialmente o da consistência moral. Uma evidência para melhora da consistência moral poderia ser encontrada no fato que o século dezanove assistiu um aumento drástico no esforço de pautar o racismo na ciência (Gould, 1981). Na América do Sul, líderes religiosos pregavam que as pessoas de cor eram produtos de uma criação separada (Fredrickson, 1971). Isso sugere que pessoas implicitamente reconheceram a tensão entre suas atitudes frente à escravidão e suas crenças sobre a dignidade humana. Mesmo assim, seria

enganador pensar que a oposição à escravidão foi motivada pelo tipo de processo reflexivo que estou atribuindo a Smith.

Na Inglaterra, a oposição à escravidão coincidiu com a revolução industrial, e alguns pesquisadores acreditam que essa transição econômica foi gatilho para a emergência de uma ideologia que favorecia o trabalho assalariado e a hegemonia da classe média (Davis, 1984). Nos Estados Unidos, a escravidão ficou para trás ao longo do tempo. Primeiro, houve uma proibição para importação de escravos, mas isso pode ser sido motivado por avareza e por medo (Bender, 2006). Virginia, o estado mais populoso e poderoso da região, estava produzindo cada vez menos cultivo, e os virginianos tiveram mais ganhos em vender os escravos do que os usando nas plantações. Eles sabiam que poderiam subir os preços dos escravos ao tornar ilegal sua importação. Adicionalmente, houve uma grande revolta no Haiti (a Revolução Haitiana) na virada do século XIX, e líderes de Washington estavam com medo de que os escravos importados daquela região pudessem propagar a notícia e incitar insurreições na América; eles baniram as importações logo depois. No sul, algumas oposições foram motivadas pelo desejo de aumentar o número de empregos para os brancos; escravos eram vistos como roubando os empregos dos brancos. O norte tinha uma relação complicada com a escravidão, porque os financiadores de Nova Iorque e a indústria têxtil da Nova Inglaterra se beneficiaram da produção de algodão, mas também havia um movimento abolicionista bastante ativo. O abolicionismo, aponta Davis (1984), tinha duas formas, secular e religioso. O argumento secular focava na ineficácia do trabalho escravo em comparação com o trabalho assalariado, e o abolicionismo religioso argumentava que os donos de escravos pecavam ao adotar uma posição divina de controle sobre seres humanos. Não é claro se algum destes argumentos surgiram do processo de avaliação que descrevi. A escravidão era claramente reconhecida como em desacordo com alguns valores extramoraes, mas eles não eram necessariamente os valores na minha lista. Ao contrário, a escravidão era criticada pela ineficiência econômica e pela concepção cristã de irmandade. Em suma, não apresento minha perspectiva de como valores morais podem ser mudados a partir de um retrato fiel de como os

valores do passado realmente mudaram. Ao invés disso, eu ofereço uma lista de padrões extramoraís como uma ferramenta possível para mudanças no futuro.

Os métodos extramoraís de avaliação na minha lista poderiam ser usados para iluminar algumas controvérsias atuais. É aceitável matar e comer animais? A pornografia é aceitável? É aceitável manter todo o rendimento pós-impostos, em vez de dar grande parte desta para luta contra a fome? É aceitável apoiar instituições que promovem a globalização? As proibições da poligamia são moralmente defensáveis? Muitas pessoas responderam que sim a essas questões. Elas são consistentes com valores prevalentes. Mas talvez nosso sistema de valores não seja consistente. Ou talvez outras considerações pragmáticas ou prudenciais fossem argumentos em favor da revisão moral. Não estou avançando em um veredito a nenhuma dessas questões. Estou simplesmente apontando que argumentos morais podem, geralmente, ser acessados juntos dessas dimensões que tenho considerado.

O resultado é que há formas nas quais podemos comparar argumentos morais. Não podemos realizar o progresso moral ao afirmar que alguns argumentos morais são mais verdadeiros do que outros; o que é falso aqui pode ser verdadeiro em algum outro lugar. Nem podemos fazer progresso ao dizer que certos valores são moralmente melhores do que outros valores morais; quando dizemos que um valor moral é bom, estamos simplesmente revelando que internalizamos aquele valor, ao invés de outros valores internalizados por outras pessoas. O progresso exige irmos além do nosso comprometimento moral atual. Para adotar novos valores morais, é importante irmos além do bem e do mal. "Progresso moral" não pode ser interpretado como uma transição de um set de normas para um set de normas melhor. Toda vez que nossos valores mudam, vemos nossos novos valores morais como melhores do que os do passado, mas isso é uma ilusão: já não adotamos os valores antigos, então os vemos como ruins. Para fazer um progresso que não seja ilusório, precisamos buscar valores morais que sejam melhores em um sentido extramoral.

8.2 FAZENDO PROGRESSO

8.2.1 Mal coerente

Uma preocupação surge imediatamente com a proposta de que podemos medir o progresso apelando para padrões extramoriais. Ao se deparar com a descoberta de que seus valores atuais não satisfazem algum padrão extramoral, geralmente, há vários caminhos de revisão disponíveis. Cada revisão possível é uma melhoria, na proposta que estou vendendo, se é melhor em um desses padrões. Por exemplo, se você descobrir que tem dois valores conflitantes, será uma melhoria deixar um deles, porque isso aumentará a consistência. O problema com essa proposta é que essa estratégia permite – até mesmo encoraja – mudanças que podem ser pré-teoricamente consideradas regressivas, e não progressivas.

Vamos considerar um exemplo. A maioria de nós acredita que todas as pessoas merecem oportunidades iguais, independentemente do sexo, mas há evidências empíricas de que, apesar dessa convicção moral, existe um preconceito desenfreado contra as mulheres. Em um estudo, Rudman e Kilianski (2000) demonstraram que as pessoas inconscientemente nutrem atitudes negativas em relação às mulheres em cargos de autoridade, mesmo que não estejam cientes dessas atitudes explicitamente; a imagens de mulheres em empregos de alto *status* (por exemplo, uma médica) são atribuídas palavras negativas, e a fotos de mulheres em empregos de baixo *status* (por exemplo, garçonete), palavras positivas. Em outro estudo, Steinpreis et al. (1999), enviou currículos de candidatos a emprego idênticos para professores de psicologia, mas eles colocaram um nome masculino em alguns deles e um nome feminino em outros; os professores tinham mais de 50% de probabilidade de julgar que o candidato do sexo masculino era desejável. Suponha que você descubra que você valoriza a igualdade de gênero explicitamente, mas tem um preconceito implícito contra as mulheres. Esses valores entram em conflito. Usando o padrão de consistência, você pode melhorar sua moralidade deixando um desses dois valores. Mas qual deles? Se você usar

princípios extramorais, como consistência, parece não haver nenhuma resposta definitiva para essa questão. Você aperfeiçoaria sua moralidade igualmente ao trabalhar para exorcizar sua percepção de gênero *ou* ao abandonar sua convicção de que não devemos discriminar mulheres. Intuitivamente, entretanto, esses dois princípios não são igualmente bons. Um seria um aperfeiçoamento, o outro um retrocesso.

Acredito que essa preocupação possa ser exposta olhando mais atentamente para o exemplo. Rejeito a noção de que nenhuma consideração extramoral pode nos ajudar a decidir entre o chauvinismo consistente ou a consistente igualdade de gênero. Perceba, primeiro, que o chauvinismo pode se basear em premissas falsas sobre as diferenças entre os sexos, premissas como as que mulheres têm menos aptidão em certos domínios cognitivos do que o homem. Chauvinismo também é mais difícil de universalizar do que equidade, e também pode levar para um diminuição do bem-estar (certamente para as mulheres). Análises genealógicas também podem ser usadas para criticar o chauvinismo. As sociedades do mundo diferem em como elas consideram as mulheres. As mulheres gozam de altos *status* e poder considerado em algumas partes do mundo, e um *status* muito baixo em outros. Harris (1993) argumenta que as determinantes primárias para o *status* de mulheres são, genealógicamente falando, as econômicas. Nas sociedades em que a força física é a base para a saúde material (e, portanto, para o poder e para o *status*), os homens saem na frente. Onde este não é o caso, mulheres conseguem ganhar um bom nível de igualdade. Considere um exemplo. Na África ocidental, a agricultura era tradicionalmente baseada no trabalho manual com enxada, ao invés do arado. Na Índia, era baseada no trabalho arado. Arados, garfos e os animais que os puxam são muito pesados e difíceis de manusear. Em economias de arado, os homens eram naturalmente mais capazes de produzir riqueza, e subir ao poder. Nas economias das enxadas, as mulheres podem produzir tanto quanto os homens, e conseqüentemente, as mulheres tendem a ter mais igualdade nas sociedades da África Ocidental do que na Índia. A exceção é a sociedade Kerala, no sudeste da Índia, que tem um clima mais quente e, então, é menos dependente do arado. Solo úmido pode ser trabalhado sem animais

pesados, e o sudeste da Índia é dominado por pequenas plantações de arroz. Na nossa sociedade contemporânea, nós não dependemos nem de enxadas, nem de arado. Essas restrições à riqueza estão datadas. O chauvinismo masculino pode ter se desenvolvido pelo fato de as economias europeias terem sido pautadas pelo arado no início do seu desenvolvimento. Isso resultou em homens terem mais dinheiro, mais controle do comércio, mais alfabetização, e mais poder político. Atitudes de superioridade masculina estão pautadas em fatores que já não são mais relevantes. O igualitarismo faz mais sentido na nossa economia atual do que o chauvinismo masculino.

Mas essa história levanta outra questão. Os últimos comentários implicam que as desigualdades de gênero eram apropriadas quando tínhamos uma economia de arado. Essa é uma suposição vergonhosa e, talvez, até ofensiva. Ela também ameaça minar meu argumento sobre progresso moral. O chauvinismo era apropriado quando condições econômicas eram diferentes, e agora é inapropriado. Então, sou obrigado a dizer que o movimento para maior libertação da mulher não tem sido para frente; tem sido, apenas, uma transição lateral de um bom sistema moral para outro sistema moral igualmente bom.

Essa objeção pressupõe que não há mais nada de errado com sistemas morais que toleram desigualdades de gênero. Em muitos desses sistemas, são negadas às mulheres liberdades que não tem nada a ver com sua capacidade de agricultura, e tudo a ver com homens usando seu poder econômico para garantir um maior controle. O fato de que mulheres teriam menor produção nos campos não significa que mulheres devam ser tratadas como propriedade masculinas. Nas sociedades na qual isso ocorreu, o bem-estar feminino tem sido reduzido, e as normas de respeito, aplicadas de forma inconsistente. Então, a desigualdade de gênero pode ser condenada de uma perspectiva extramoral. Dito isso, devemos diferenciar desigualdade de gênero de divisão de trabalho. Padrões extramorais nos dão ferramentas para criticar sociedades que discriminam, mas elas não têm o poder de criticar sociedades nas quais homens e mulheres possuem papéis diferentes. Uma sociedade que não oprimiu mulheres, mas simplesmente designou papéis econômicos distintos pautados nos padrões de gênero não seriam

necessariamente piores, em nenhum sentido extramoral, do que uma sociedade que que torna disponível economicamente as mesmas opções para homens e mulheres. Uma sociedade de caça e colheita, na qual os homens caçam e as mulheres fazem a colheita, não é necessariamente pior do que uma sociedade pós-industrial, na qual os empregos disponíveis são para ambos, homens e mulheres. Não acho que este é um resultado contraintuitivo. Não é óbvio que a transição de estilo de caça e colheita para o estilo de vida pós-industrial tenha uma melhora moral. Sistemas morais são ferramentas para coordenação social, e como tais, elas devem ser moldadas para a variedade de condições materiais das culturas.

Ainda assim, poderia-se objetar que eu não disse nada para garantir que um sistema de valores completamente abominável não poderia ser melhor, a partir dos padrões que considere, do que um sistema de valor mais humano. Há, no final das contas, espaço para o mal coerente. Aqueles que buscam o genocídio, geralmente, constroem sistemas morais elaborados que são mais consistentes do que as moralidades liberais que eles substituem (compare a moralidade nazista com a da República de Weimar). Em resposta, devemos nos lembrar de que coerência não é tudo. Um sistema moral coerente pode ser melhor a este respeito do que uma coleção de sistemas de valor inconsistente. Mas pode ser muito pior em outros quesitos. Os padrões requisitados não são limitados por constrangimentos lógicos. Esses, sozinhos, não poderiam distinguir benevolência de brutalidade. Nós acessamos o progresso moral ao apelar para o bem-estar, para o bem-estar social, para a universalidade e para estabilidade social. É difícil de reconciliar o genocídio com esses objetivos. Concedo que uma sociedade perversa pode não ter preocupação com nenhum destes padrões extramorais, mas, com certeza, nós nos preocupamos com essas questões, e, portanto, não consideraríamos a adoção do genocídio como um sistema de valor progressista, mesmo que tenha um boa pontuação no padrão consistência.

Podemos acessar sistemas morais ao questionar quão bem eles servem para promover um tipo de vida que consideramos desejável. Claro, pessoas podem variar naquilo que consideram desejável, e, então, naquilo que consideram progresso moral. Se valorizo relações interpessoais e você valoriza sabedoria, nossas

concepções de bem-estar são distantes. Um sistema moral que serve melhor para proliferar sabedoria ao preço da comunidade pode colocar você em vantagem, mas pareceria o contrário para mim. Se padrões de progresso são avaliativos, progresso é, em si, relativo. E tudo bem com isso. O relativismo não precisa mostrar que há padrões absolutos para o progresso. O desafio inicial era explicar como o progresso é *minimamente* possível.

Avaliações divergentes do progresso são lugares-comuns. O que os liberais enaltecem como maturidade moral, os conservadores podem considerar como regresso moral. Alguns desses debates sobre moral são intratáveis. Mas não são todos intratáveis. Quando liberais e conservadores avaliam progresso, eles podem encontrar fundamentos comuns em alguns desses padrões que tenho considerado. Nossos valores se tornaram mais consistentes? A coesão social foi aumentada? Estamos prosperando? Essas são algumas das considerações que conservadores e liberais normalmente trazem quando estão debatendo sobre políticas públicas. Em relação a um conjunto compartilhado de valores extramoral, pode haver um fato do assunto sobre o qual a perspectiva política promove o melhor sistema moral.

8.2.2 Armadilhas no caminho do progresso

Há um número de armadilhas no caminho do progresso. Às vezes, pensamos que estamos fazendo progresso quando a verdade é o oposto. Quando embarcamos na reforma moral, é crucial termos isso em mente. Irei mencionar 4 armadilhas aqui, mas com certeza existem outras.

Fiz alusão a uma das armadilhas anteriormente, quando me referi a um sentido de olhar para trás do progresso. De qualquer sistema moral, outro sistema moral parece imoral. Se olharmos para os valores mantidos no nosso padrão cultural, eles são falhos. Em alguns casos, essa impressão pode ser justificada, mas ela não é confiável. O passado aparenta ser moralmente falho quando olhado pelas lentes do presente. Isto dá origem à ilusão do progresso. Um exemplo é ilustrado na discussão de Foucault (1977) sobre o sistema penal. Ele argumenta que a transição das execuções públicas para o encarceramento não foi um avanço moral motivado

pela preocupação com os prisioneiros, mas uma nova forma de controle, derivada da reconceituação de criminais como anormais ou desviantes. Por causa da mudança ideológica ocasionada pelas ciências humanas, vemos, agora, métodos de punição do passado como profundamente imorais, ao mesmo tempo em que falhamos em ver o caráter opressivo dos métodos atuais. Essa ilusão do progresso é difícil de escapar. Muitos de nós acreditamos que agora estamos em contato com a verdade moral. Pensamos que aqueles no passado eram atrasados, e dizemos que aqueles que anseiam pelo passado são reacionários. Esse conceito seria defensável se o progresso moral fosse uma consequência inevitável da história. Mas não é.

Uma segunda armadilha surge quando deixamos os fins justificarem os meios. É possível que alguns dos apoiadores do Pol Pot tenham sido vítima de um erro hediondo quando eles assassinaram uma grande porcentagem da população do Camboja. Eles acreditavam que poderiam aumentar a harmonia e o bem-estar ao forçar a criação de um Estado comunista agrário. Matando centenas de "intelectuais" (frequentemente pessoas que apenas usavam óculos ou falavam dois idiomas) parecia um custo necessário. Esse custo poderia ser evitado se os membros do Khmer Rouge não tivessem feito um cálculo grotesco. Ao planejar sua visão utópica, eles assumiram que as pessoas achariam mais compensatória a vida pós-revolução no Camboja. Assumiram que as pessoas ficariam felizes com uma vida bucólica agrária. Mas o erro nessa pressuposição foi imediatamente manifestado. Milhares de pessoas tiveram que ser forçadas a trabalhar nas fazendas, e aqueles que resistiram foram mortos. Se a campanha do Pol Pot tivesse criado com sucesso uma utopia comunista no Camboja, seria um caso para o progresso moral? Talvez. Matar todos os seus oponentes é um atalho sangrento para estabilidade social. Mas as execuções em massa do Khmer Rouge provavelmente qualificavam como maléficas pelos padrões daqueles que o seguiam. Se aplicarmos os padrões que tenho discutido para os meios assim como para os fins, podemos ver que a brutalidade a serviço de um bem é meio que uma inconsistência moral. Da mesma forma, podemos olhar retrospectivamente para os vários atos de genocídio contra as populações indígenas das Américas e

Australásia e perguntar se as nossas vidas confortáveis justificam agora o que colonizadores fizeram no passado. Existe um sentimento coletivo de culpa associado a essas velhas atrocidades, e, embora possamos sentir que fizemos grandes progressos em termos de bem-estar e estabilidade, estamos profundamente preocupados, espero eu, pela hipocrisia moral da nossa conquista.

A terceira armadilha é que podemos nos tornar corruptos. Para ver o quão facilmente isso pode acontecer, considere um estudo clássico da psicologia social (Haney et al., 1973). Em 1971, Philip Zimbardo e seus colaboradores pagaram um pequeno grupo de estudantes da universidade para participar de um experimento de interpretação de papéis de 'prisão' em Stanford, na qual alguns interpretaram os guardas e outros os prisioneiros. Depois de seis dias, o experimento planejado para duas semanas precisou ser interrompido porque os estudantes atuando como guardas se tornaram seriamente abusivos. Eles humilharam os prisioneiros, tiraram suas roupas, fizeram-nos encenar atos sexuais, retiraram-lhes a comida, colocaram-nos em confinamento na solitária e os fizeram urinar e defecar em baldes. Aqueles mesmos estudantes teriam condenado o comportamento abusivo antes do experimento, mas a situação os levou a abandonarem seus valores. Abusos semelhantes ocorreram na prisão militar de Abu Ghraib no Iraque. Este é um lembrete soberano do fato que corrosão moral é normalmente mais fácil que avanço moral. Progresso necessita de vigilância.

A armadilha final é que sempre corremos perigo de ver a moralidade como imutável. Quando consideramos normas morais como objetivas e inatacáveis, esquecemos que o progresso moral é possível. Uma estratégia profilática é ver a moralidade instrumentalmente. É uma maneira de garantir outros objetivos, como a coesão social, bem-estar e bem-estar social. É difícil ver a moralidade instrumentalmente, porque nossos sentimentos apresentam nossos valores morais básicos para nós como se fossem intrinsecamente bons. Se nos lembrarmos que a moralidade é uma ferramenta, e não uma janela da verdade em absoluto, podemos estar abertos à possibilidade de crescimento moral. Nós seríamos sábios para recalibrar nossos sentimentos regularmente com padrões extramoriais.

Avaliar o progresso é difícil. Eu posso até ter caído nessas armadilhas neste capítulo. Com a fala da escravidão, dos direitos das mulheres e do colonialismo, eu tenho dado a impressão de que a moralidade no presente é uma melhoria em relação ao moralidade do passado. Eu impliquei que nos movemos de forma inequívoca para frente, e não enfatizei maneiras pelas quais podemos fazer melhor. Isso é preocupante porque implica uma profunda vulnerabilidade ao autoengano sobre o progresso. Quando nós raciocinamos sobre se as coisas melhoraram e se melhorar ainda mais é possível, estamos sujeitos ao viés de confirmação: uma tendência para escolher evidência que apoiará nossas convicções atuais.

Contra esta preocupação muito séria, eu só posso apelar para o trabalho duro. Nós devemos constantemente nos lembrar de que nossos valores não são reflexos de uma verdade absoluta que, como itens em um livro de texto de ciência, foram descobertos através de cuidadosa investigação ao longo do tempo. A moral é inculcada e, muitas vezes, moldada, como Nietzsche diria, por lutas de poder e por acaso. Lembrando-nos disso, a profunda contingência da moralidade é apenas um primeiro passo. Devemos também sujeitar nossos valores, incluindo aqueles que mais apreciamos, à consideração rigorosa à luz dos valores extramoriais. Por exemplo, devemos examinar nossas atitudes em relação à democracia, ao capitalismo, à liberdade de expressão, à punição, ao casamento e ao infanticídio. Conservadores e liberais devem explorar a possibilidade de que seus valores são o resultado de inculcação, e aqueles que criticam os valores dos outros devem examinar suas própria cumplicidade em contribuir para os problemas do mundo. Eu não vou embarcar nesse exercício de autoexame aqui. Há uma longa linha de reformadores radicais que tentaram afastar as pessoas da complacência e instigar revoluções morais. Alguns desses reformadores (Marx e Nietzsche, por exemplo) se encontraram com o sucesso, mas outros tiveram impacto considerável na consciência das pessoas que vivem no mundo ocidental (líderes do movimento dos direitos civis e do movimento das mulheres vêm à mente). Os relativistas que acreditam no progresso devem estar abertos à possibilidade de que valores atuais possam ser melhorados.

8.2.3 Progresso e ética normativa

Nos capítulos anteriores, expressei minhas dúvidas sobre teorias éticas normativas prevalecentes, incluindo teorias avançadas de Kant, Mill, e Aristóteles. Ironicamente, todos os três ressurgiram em meus padrões de avaliação da moralidade. Eu disse que nós queremos nossas regras morais tão universais quanto possível; este é um princípio central do programa kantiano. Eu também disse que preferimos regras que aumentam o bem-estar; isso ecoa a preocupação milliana com utilidade. Eu também invoquei a construção mais ampla do bem-estar; esta é uma palavra moderna para a noção de *eudaimonia* de Aristóteles, que se encontra no centro de sua concepção de uma vida virtuosa. A invocação dessas idéias pode lançar dúvidas das minhas primeiras observações sobre Kant, Mill e Aristóteles. As teorias que eu rapidamente descartei agora estão sendo usadas para decidir quais regras morais são melhores. Até agora, os protagonistas desta história foram Hume e Nietzsche. Seria embaraçoso descobrir que Kant, Mill e Aristóteles estão realmente executando o show. Deixe-me explicar por que esse não é o caso.

Primeiramente, é importante perceber que Hume e Nietzsche estão fazendo coisas bem diferentes que Kant, Mill, e Aristóteles. Hume está fornecendo uma descrição do que é valorizar algo moralmente, e Nietzsche está oferecendo um relato de onde vêm nossos valores atuais. Ambos os projetos são descritivos. Hume é psicológico, e Nietzsche é uma mistura de psicologia e de história. Kant, Mill e Aristóteles estão oferecendo teorias normativas: teorias sobre o que devemos valorizar. Nesse sentido, suas contribuições são compatíveis com a psicologia moral humeana e genealogia nietzscheana. Hume diz que valorizar algo moralmente é ter um sentimento de aprovação em relação a isso. Nietzsche nos diz como nossos sentimentos atuais surgiram. Kant, Mill e Aristóteles estão nos dizendo o que devemos valorizar em vez disso. Eles podem ter acreditado que nós já valorizamos o dever, a utilidade e *eudaimonia*, mas é discutível se eles precisam fazer tais reivindicações descritivas. De fato, tais afirmações descritivas não são incompatíveis com as descrições afirmativas feitas por Nietzsche e Hume.

Também é importante ver que os padrões de avaliação que eu tenho discutido são diferentes dos padrões reivindicados por Kant, Mill e Aristóteles, mesmo que haja alguma semelhança. Eu especulei que muitos moralizadores têm uma preferência por regras mais universais. Isso não implica a afirmação de Kant de que a moralidade deve basear-se em princípios de racionalidade prática e que a universalização fornece o teste crucial para a permissibilidade moral. Minha sugestão de que as pessoas querem regras que aumentam o bem-estar fica aquém da máxima de Mill de que a utilidade é o árbitro final do bem. E, ao contrário do *eudaimonia* de Aristóteles, meu apelo ao bem-estar não tem ligação essencial com a noção de virtude. Eu também observei que concepções de bem-estar podem variar; eles não são ditados pela natureza humana. Meus padrões de avaliação são, na melhor das hipóteses, sombras pálidas dessas teorias normativas elaboradas, ou, talvez, sejam as sementes que dão às teorias normativas seus apelos instintivos.

Além disso, minhas invocações de universalidade, de bem-estar social e de bem-estar não são pretendidas como normas categóricas. Eu não tenho fornecido um argumento para que estes sejam padrões absolutos que qualquer um deve aceitar. Eu estou, antes de qualquer coisa, fazendo uma afirmação descritiva, sujeita a teste empírico, de que estes são padrões pelo qual julgaríamos um conjunto de valores como sendo melhores do que outros. Ao fazê-lo, estou demonstrando que o progresso moral é inteligível dentro de uma estrutura relativista. Se eu fizesse qualquer afirmação normativa sobre esses padrões – se quero recomendá-los –, isso seria uma expressão do fato de eu aceitar esses padrões, e da esperança de que você os aceite também. Ao contrário das teorias normativas clássicas, eu acho que as afirmações normativas sempre devem ser feitas dentro de um sistema de valor. Não podemos encontrar um teste decisivo transcendental para a moralidade.

Outra diferença entre a noção do progresso moral que eu estou promovendo e teorias prevaletentes em ética normativa é que eu acho que os padrões pelos quais o progresso moral é julgado não são, eles próprios, padrões morais. É importante para o progresso que possamos sair da moralidade. Caso contrário,

nossos valores atuais sempre ditariam nossas preferências morais, e o progresso seria impossível.

Portanto, quando digo, por exemplo, que preferimos teorias morais que levam para um bem-estar maior, não estou fazendo uma reivindicação moral. Bem-estar não é nem necessário, nem suficiente para a bondade moral. Alguém que faz algo que resulta em grande felicidade não necessariamente fez uma coisa moralmente boa, e alguém pode fazer a coisa moralmente boa (ou seja, fazer algo a que nos referimos com aprovação moral) mesmo que não traga felicidade. Da mesma forma, [isso vale] para os "bens" apresentados nas teorias normativas clássicas. Eu estou afirmando que estes são bens extramorais. Eles residem fora da moralidade, embora possam ser usados para melhorar a moralidade.

Além disso, eu rejeitaria a alegação de que qualquer um dos produtos apresentados por teorias normativas clássicas são fundamentais. Nenhuma é a fonte da qual todos os outros bens fluem. Mais genericamente, não acho que nenhum dos padrões de avaliação que tenho discutido é mais fundamental do que outros. Nem acho que esses padrões de avaliação são mais fundamentais do que as regras morais que eles avaliam. Regras morais não são derivadas desses padrões. Ao avaliar teorias morais, como ao fazer ciência, devemos ser holísticos. Nós devemos manter todos os nossos valores em vista, devemos saber onde eles estão em relação à nossa rede de convicções morais, e devemos ter em mente os princípios extramorais que também importam para nós. Alguns princípios morais são mais básicos do que outros, na medida em que alguns são princípios gerais e outros são aplicações. Mas quando se trata da revisão dos princípios gerais (nossas normas de ancoragem), todas as nossas convicções, morais e não-morais, são potencialmente relevantes. E, assim como os princípios extramorais podem ser usados para revisar a moralidade, princípios morais podem pesar enquanto revisamos nossos valores extramorais.

Este holismo anda de mãos dadas com o pluralismo. Com demasiada frequência, filósofos morais procuram reduzir a moralidade a um único princípio. Fazer isso seria uma realização de tirar o fôlego. As grandes teorias normativas são incrivelmente elegantes e ambiciosas. Na verdade, elas são ambiciosas até demais.

A moralidade humana é multifacetada. Nós temos uma gama de regras diferentes que não podem ser unificadas sob nenhum princípio único. A maioria dos filósofos se concentra em regras relacionadas a danos, a justiça ou a direitos. Estas são todas construções diferentes, e são apenas uma fração da moral. Todos eles pertencem à ética da autonomia, na frase de Shweder et al. (1997), deixando de fora os domínios da comunidade e da divindade/natureza. Shweder acredita que as normas de autonomia assumiram um destaque especial nas culturas individualistas. As culturas coletivistas, como as do Extremo Oriente, podem colocar comparativamente mais ênfase nas normas comunitárias, incluindo aquelas que tem a ver com ranqueamento. Sociedades não-seculares enfatizam a comunidade também, e também colocam ênfase nas regras relativas à pureza religiosa e sexual. Regras pertinentes para a classificação e para o sexo são menos salientes para nós, mas certamente existem. Elas não podem ser capturadas sob a rubrica da autonomia sem um artifício considerável. Alguns filósofos querem nos fazer acreditar que regras de classe e de sexo não são autênticas regras morais, porque elas não podem ser justificadas apelando para proibições contra danos ou outros princípios de autonomia. Mas essa atitude é uma forma de miopia moral, trazida pelo individualismo ocidental e pela socialização em profissionais da filosofia. Domínios morais fora da autonomia caíram do radar cultural mesmo que exerçam influência considerável em nossas vidas diárias. Vivemos num mundo de classe, de autoridade e de normas de respeito. Nós formamos nossas relações mais íntimas com base em regras de parentesco e normas de propriedade sexual. A moralidade da vida cotidiana pode ter mais a ver com essas coisas do que com danos, direitos e justiça. Precisamos de teorias morais para ser mais inclusivos. Filósofos acabam sendo excessivamente paroquiais, ignorando valores morais que eles também consideram paroquiais.

O pluralismo que venho descrevendo é um pluralismo de domínios morais. Dentro desses domínios, também é provável que haja um pluralismo de regras. Por exemplo, temos múltiplos costumes sexuais (incesto, bestialidade, necrofilia, masturbação, etc.) que não podem ser reduzidos a um tabu abrangente. Cada regra tem sua própria história biocultural. Princípios amplos provavelmente serão um

resultado de regras específicas, ao invés do contrário. Além de uma pluralidade de domínios e de regras morais, defendi uma pluralidade de padrões de avaliação para a moralidade. Não há um único teste decisivo para o progresso moral. Podemos encontrar sombras pálidas de Kant, Mill e Aristóteles na ladainha de padrões pelos quais o progresso é avaliado, mas não podemos encontrar nenhuma maneira de unir essas sombras. A ambição imperialista destes autores não têm lugar no presente relato do progresso moral.

Em suma, as teorias éticas normativas clássicas não são justificadas por qualquer coisa que eu tenha dito, nem são refutadas. Estas teorias são recomendações dignas para ação. Talvez um compromisso constante com o progresso moral acabe nos levando a adotar alguns componentes das teorias clássicas. A adoção da ética kantiana, da milliana ou da aristotélica pode contar como progresso moral pelos padrões que tenho enumerado. Talvez os valores morais evoluam para incorporar lições dessas abordagens. Mas duvido que qualquer cultura tenha sucesso em incutir valores kantianos, millianos ou aristotélicos puros. A vida humana é muito variada e a moralidade é muito complexa para ser guiada por uma ética tão simplificada.

8.2.4 Devemos renunciar à moralidade?

No esquema que tenho discutido, há dois níveis de avaliação. Em um nível, temos valores morais. Estes são definidos por sentimentos morais e as emoções morais que as constituem. Em outro nível, temos valores extramoriais. Eu não falei muito sobre como eles são implementados, mas é fácil imaginar uma coleção de objetivos, de desejos, de preferências e de procedimentos de decisões teóricas que pesam certas características acima de outras. Eu tenho sugerido que valores extramoriais podem desempenhar um papel fundamental na orientação da mudança moral. E isso gera uma pergunta importante: por que se preocupar com a moralidade? Não podemos nos dar bem com nossos valores extramoriais? A camada moral não é desnecessária?

Em uma leitura, Nietzsche pode ser interpretado recomendando que abandonemos a moralidade e a substituamos por outra coisa. Eu costumo lê-lo como sugerindo que substituamos uma moralidade por uma moral superior, mas ambas interpretações encontram suporte textual. Mais surpreendentemente, pode-se ler Kant, Mill e Aristóteles recomendando uma fuga da moralidade. Se o significado de "moralidade" é capturado pelo uso cotidiano de termos moralizantes como "bom" e "mal", ou "certo" e "errado", os eticistas normativos clássicos vão recomendar uma alternativa. Eu argumentei que o vocabulário moral comum é ligado essencialmente aos sentimentos. Termos morais de louvor e de culpa expressam disposições emocionais. Os eticistas normativos tradicionais tendem a ser céticos sobre aprovação e desaprovação. Eles acham que os sentimentos são maus guias para conduta. Em vez disso, devemos agir de acordo com o dever, a maximização da utilidade, ou a virtude. Todas essas coisas são muito difíceis de fazer, e as demandas que eles colocam em nós são, muitas vezes, contraintuitivas. O fato de que pode haver um choque entre intuições morais comuns e as exigências da ética normativa é uma indicação de que a ética normativa está no negócio de usurpar a moralidade. Se esse é o caso, então figuras importantes na história da ética normativa se alinham com Nietzsche. Eles são revisionistas radicais.

Eu duvido que poderíamos simplesmente abandonar a moralidade mesmo se tentássemos. Sentimentos morais podem ser produtos inevitáveis do psicológico humano, e normas morais podem ser profundamente entranhadas. Mas suponha que pudéssemos, de alguma forma, eliminá-los. Deveríamos? Eu acredito que não. Sentimentos morais servem para um número importante de funções.

Por um lado, os sentimentos morais desempenham um papel fundamental na motivação. Quando julgamos que algo é bom, somos motivados a agir porque estamos dispostos a sentir certas emoções motivadoras. De especial importância são as emoções punitivas. A ameaça de culpa e de vergonha nos ajudam a resistir às tentações. Raiva, desprezo e nojo ajudam a regular o comportamento dos outros. Sem moralidade, regras de conduta podem se tornar desapegadas, e regras desapegadas são fáceis de ignorar. Considere a regra contra atravessar a rua sem observar as regras de cruzamento. Nós todos sabemos que esta regra existe e vem

à mente toda vez que caminhamos contra a luz, mas, para muitos de nós, não tem impacto motivacional. Regras que não são baseadas em sentimentos não são especialmente eficazes na conduta de orientação.

Em segundo lugar, a transmissão de regras é bastante facilitada por sentimentos morais. Podemos fazer nossos filhos se conformarem com as normas operativas, condicionando suas emoções. Regras baseadas em outra coisa que não sentimentos são mais difíceis de transmitir. Emoções comandam a atenção e facilitam a memória. Elas servem como punição e recompensa. Se simplesmente entregássemos aos nossos filhos livros de regras, eu suspeito que eles teriam dificuldade em aprender o que a moralidade exige deles, e ainda mais dificuldade em aprender a se conformar.

Um terceiro ponto estreitamente relacionado é que os sentimentos morais podem ser usados para internalizar regras. Ao treinar nossos sentimentos, condutas que, de outra forma, não seriam um apelo para nós começam a parecer naturais e automáticas. A moralidade se torna uma parte central de nossas autoconcepções. Nós nos identificamos com regras morais e gostamos de nos filiar a pessoas cujos valores são semelhantes aos nossos. Crenças desapaixonadas tendem a não ser internalizadas dessa maneira. Podemos perder crenças sem perder nosso senso de identidade.

Em quarto lugar, os sentimentos morais nos permitem estender a moralidade. Através da educação sentimental, o bem-estar de outras pessoas torna-se uma questão de interesse pessoal. A idéia de extensão moral se relaciona com a concepção de virtudes artificiais de Hume (Hume, 1739: III.ii). De acordo com Hume, estamos naturalmente dispostos a ser benevolentes, mas nossa benevolência natural é limitada pelo amor-próprio e reservada para aqueles que estão perto de nós, que são semelhantes a nós, ou que têm qualidades que valorizamos. Nós não somos naturalmente benevolentes com estranhos, e, portanto, não temos um preconceito natural contra roubar suas propriedades quando nos convier. Ao refletir melhor, no entanto, isso seria uma péssima ideia. Se não respeitássemos os direitos de propriedade dos outros, eles não respeitariam os nossos. Assim, acabaríamos perdendo em uma sociedade que tolerasse o roubo.

Nós nos beneficiamos das regras que protegem a propriedade (o que Hume chama de "justiça") juntamente com outras regras, como promessas, que não são naturais. Amor-próprio nos leva a criar tais regras artificiais. Nós cultivamos sentimentos morais para a justiça, porque sabemos que é do nosso interesse fazê-lo. Hume não explica totalmente como esses sentimentos são cultivados em primeiro lugar, mas ele sugere que os pais possam ensinar os filhos a amar a justiça, a condenar os que quebram promessas e assim por diante. Se Hume tivesse sido mais cínico, ele poderia ter percebido que algumas virtudes artificiais não se originam no reconhecimento de que certas regras antinaturais são benéficas. Algumas podem emergir, como Nietzsche sugeriu, através de lutas pelo poder e do acaso histórico. Mas a idéia central na explicação de Hume sobre a artificialidade das virtudes parece correta. Muitas de nossas regras envolvem a aplicação de sentimentos a coisas que naturalmente não provocariam aprovação ou desaprovação. Sentimentos são valiosos porque podem ser estendidos a novos casos.

Por estas razões, penso que seria desastroso abandonar os sentimentos morais que constituem a moralidade. Se pudéssemos substituir nossas regras apaixonadas por princípios legais, haveria consequências hediondas. Suponha que poderíamos tomar pílulas que eliminam a raiva, o desprezo, o desgosto, a culpa e a vergonha. Nós poderíamos inicialmente ficar aliviados por nos livrarmos dessas emoções desagradáveis. Mas, ao fazer isso, perderíamos nossa motivação para evitar o comportamento antissocial, iríamos prejudicar nossa capacidade para transmitir regras facilmente, eliminaríamos nossas concepções de nós mesmos como agentes morais, e arriscaríamos nos tornar indiferentes às necessidades de outros distantes. Sentimentos morais são como uma vacina que nos protege da psicopatia virulenta.

É hora de fazer um balanço. Tenho defendido o sentimentalismo construtivista, que é uma forma de teoria da sensibilidade segundo a qual os fatos morais são produtos ou nossas regras morais, as regras morais são constituídas por sentimentos, e os sentimentos são estabelecidos por meio de interações bioculturais. Porque a moralidade é uma construção, é possível que tomemos as rédeas. Podemos exercer algum controle na determinação do curso da mudança moral. Uma vez que

reconhecemos que a moralidade existe para servir nossos desejos e necessidades, podemos tentar ajustar a moral atual para que eles nos sirvam melhor. Não podemos fazer isso de uma posição transcendental. Temos que usar nossos valores atuais para nos guiar. Mas os valores morais não podem ser usados para guiar a mudança moral, porque os valores morais são autoafirmativos; nós sempre pensamos que nossas convicções atuais são nobres. Para fazer progresso moral, precisamos consultar nossos valores extramoriais. Valores extramoriais podem não iluminar o caminho para uma única e verdadeira moralidade, mas podem nos ajudar a ver como nossos valores atuais não conseguem atingir fins que nos importam muito. Pode ser tentador dispensar a moralidade e deixar que valores extramoriais guiem o comportamento sem recrutar sentimentos morais. Isso seria um erro grave. Sentimentos são mais adequados para a regulação do comportamento do que qualquer alternativa desapaixonada. Eles são a salvaguarda contra a indiferença viciosa.

REFERÊNCIAS

- Adams, R. M. (1976). Motive Utilitarianism. *Journal of Philosophy*, 73: 467–81.
- Albert, E. M. (1963). Women of Burundi: A Study of Social Values. In D. Paulme (ed.), *Women of Tropical Africa*, pp. 179–216. London: Routledge & Kegan Paul.
- Alexander, R. D. (1987). *The Biology of Moral Systems*. Hawthorn, NY: Aldine De Gruyter.
- Allison, A. (2000). *Permitted and Prohibited Desires: Mothers, Comics and Censorship in Japan*. Berkeley, CA: University of California Press.
- Almroth, L., Elmusharaf, S., El Hadi, N., Obeid, A., El Sheikh, M. A., Elfadil, S. M., and Bergström, S. (2005). Primary Infertility after Genital Mutilation in Girlhood in Sudan: A Case-Control Study. *The Lancet*, 366: 385–91.
- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th edn.). Washington, DC: American Psychiatric Association.
- Annas, J. (2003). Virtue Ethics and Social Psychology. *A Priori*, 2: 20–59.
- Anscombe, G. E. M. (1957). *Intention*. Oxford : Blackwell.
- Arens, W. F. (1979). *The Man-Eating Myth: Anthropology and Anthropophagy*. New York, NY: Oxford University Press.
- Aristotle, (350 bce/1988). *The Nicomachean Ethics*, trans. D. Ross, ed. J. L. Ackrill and J. O. Urmson, Oxford: Oxford University Press.
- Arnold, M. B. (1960). *Emotion and Personality*. New York, NY: Columbia University Press.
- Arsenio, W. F., and Lover, A. (1995). Children's Conceptions of Sociomoral Affect: Happy Victimizers, Mixed Emotions and Other Expectancies. In M. Killen and D. Hart (eds.) *Morality in Everyday Life: Developmental Perspectives*, pp. 87–128. Cambridge: Cambridge University Press.
- Audi, R. (2004). *The Good in the Right: A Theory of Intuition and Intrinsic Value*. Princeton, NJ: Princeton University Press.
- Ayer, A. J. (1952). *Language, Truth, and Logic*. New York, NY: Dover.
- Bagemihl, B. (1999). *Biological Exuberance: Animal Homosexuality and Natural Diversity*. New York, NY: St. Martin's Press.

Baier, K. (1967). Fact, Value, and Norm in Stevenson's Ethics. *Noûs*, 1: 139–60.

Baille, J. (2000). *Hume on Morality*. London: Routledge.

Bamshad, M., Kivisild, T., Watkins, W. S., Dixon, M. E., Ricker, C. E., Rao, B. B., Naidu, M. J., Prasad, R. B. V., Reddy, G. P., Rasanayagam, A., Papiha, S. S., VILLEMS, R., Redd, A. J., Hammer, M. F., Nguyen, S. V., Carroll, M. L., Batzer, M. A., and Jorde, L. B. (2001). Genetic Evidence on the Origins of Indian Caste Populations. *Genome Research*, 11: 994–1004.

Barrett, K. C., Zahn-Waxler, C., and Cole, P. M. (1993). Avoiders Versus Amenders: Implication for the Investigation of Guilt and Shame During Toddlerhood? *Cognition and Emotion*, 7: 481–505.

Baumeister, R. F., Stillwell, A. M., and Heatherton, T. F. (1994). Guilt: An Interpersonal Approach. *Psychological Bulletin*, 115: 243–67.

Bender, T. (2006). *A Nation among Nations: America's Place in World History*. New York, NY: Hill and Wang.

Bennett, R. L., Motulsky, A. G., Bittles, A., Hudgins, L., Uhrich, S., Doyle, D. L., Silvey, K., Scott, C. R., Cheng, E., McGillivray, B., Steiner, R. D., and Olsen, D. (2002). Genetic Counseling and Screening of Consanguineous Couples and their Offspring: Recommendations of the National Society of Genetic Counselors. *Journal of Genetic Counseling*, 11: 97–119.

Ben-Ze'ev, A. (2000). *The Subtlety of Emotions*. Cambridge, MA: MIT Press.

Berman, J. J., Murphy-Berman, V. A., and Singh, P. (1985). Cross-Cultural Similarities and Differences in Perceptions of Fairness. *Journal of Cross-Cultural Psychology*, 16: 55–67.

Berthoz, S., Armony, J. L., Blair, R. J. R., and Dolan, R. J. (2002). An fMRI Study of Intentional and Unintentional (Embarrassing) Violations of Social Norms. *Brain*, 125: 1696–1708.

Bittles, A. H. (1990). Consanguineous Marriage: Current Global Incidence and Its Relevance to Demographic Research. Research Report No. 90–186, Population Studies Center, University of Michigan.

Black, J. B., and Bern, H. (1981). Causal Inference and Memory for Events in Narratives. *Journal of Verbal Learning and Verbal Behavior*, 20: 267–75.

Blackburn, S. (1980). Opinions and Chances. In D. H. Mellor (ed.), *Prospects for Pragmatism*, pp. 175–96. Cambridge: Cambridge University Press.

_____. (1984). *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Oxford University Press.

_____. (1985). Errors and the Phenomenology of Value. In T. Honderich (ed.), *Morality and Objectivity*. London: Routledge & Kegan Paul.

_____. (1988). Attitudes and Contents. *Ethics*, 98: 501–17.

_____. (1998). *Ruling Passions*. Oxford: Oxford University Press.

Blair, R. J. R. (1995). A Cognitive Developmental Approach to Morality: Investigating the Psychopath. *Cognition*, 57: 1–29.

_____. (1997). Moral Reasoning and the Child with Psychopathic Tendencies. *Personality and Individual Differences*, 26: 731–9.

Blair, R. J. R., Colledge, E., Murray, L., and Mitchell, D. G. (2001). A Selective Impairment in the Processing of Sad and Fearful Expressions in Children With Psychopathic Tendencies. *Journal of Abnormal Child Psychology*, 29: 491–8.

Blair, R. J. R., Jones, L., Clark, F., and Smith, M. (1997). The Psychopathic Individual: A Lack of Responsiveness to Distress Cues? *Psychophysiology*, 34: 192–8.

Blass, T. (2004). *The Man Who Shocked the World: The Life and Legacy of Stanley Milgram*. New York, NY: Basic Books.

Blurton-Jones, N. G. (1987). Tolerated Theft: Suggestions About the Ecology and Evolution of Sharing, Hoarding, and Scrounging. *Social Science Information*, 26: 31–54.

Boswell, J. (1980). *Christianity, Social Tolerance and Homosexuality*. Chicago, IL: University of Chicago Press.

Boswell, J. (1994). *Same-Sex Unions in Pre-Modern Europe*. New York, NY: Villard Press.

Boyd, R. N. (1988). How to be a Moral Realist. In G. Sayre-Mccord (ed.), *Essays on Moral Realism*. Ithaca, NY: Cornell University Press.

Boyd, R., and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.

_____. (2001). Memes: Universal Acid or a Better Mouse Trap. In R. Aunger (ed.), *Darwinizing Culture: The Status of Memetics as a Science* (pp. 143–62). Oxford: Oxford University Press.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). The Evolution of Altruistic Punishment. *Proceedings of the National Academy of Sciences*, 100: 3531–5.

Boyer, P., and Ramble, C. (2001). Cognitive Templates for Religious Concepts: Cross-Cultural Evidence for Recall of Counter-Intuitive Representations *Cognitive Science*, 25: 535–64.

Brandt, R. B. (1954). The Definition of an 'Ideal Observer' in Ethics. *Philosophy and Phenomenological Research*, 15: 407–13. (1967).

_____. Ethical Relativism. In P. Edwards (ed.). *Encyclopedia of Philosophy*, vol. 3, pp. 75–8. New York, NY: Macmillan.

Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Brogaard, B. (no prelo). Moral Contextualism and Moral Relativism. *The Philosophical Quarterly*.

Brosnan, S. F., and De Waal, F. B. M. (2003). Monkeys Reject Unequal Pay. *Nature*, 425: 297–9.

Burkert, W. (1983). *Homo Necans: the Anthropology of Ancient Greek Sacrificial Ritual and Myth*, trans. P. Bing. Berkeley, CA: University of California Press.

Buss, A. H. (1989). Personality as Traits. *American Psychologist*, 44: 1378–88.

Buss, D. M. (2000). *The Dangerous Passion: Why Jealousy Is as Necessary as Love and Sex*. New York, NY: Simon & Schuster.

Buss, D. M., Larsen, R. J., Westen, D., and Semmelroth, J. (1992). Sex Differences in Jealousy: Evolution, Physiology, and Psychology. *Psychological Science*, 3: 251–5.

Buunk, B. P., Angleitner, A., Oubaid, V., and Buss, D. M. (1996). Sex Differences in Jealousy in Evolutionary and Cultural Perspective: Tests From the Netherlands, Germany, and the United States. *Psychological Science*, 7: 359–63.

Cappelen, H., and Lepore, E. (2005). *Insensitive Semantics*. Oxford: Blackwell.

Carlsmith, J. M., and Gross, A. E. (1969). Some Effects of Guilt on Compliance. *Journal of Personality and Social Psychology*, 11: 232–9.

Carlsmith, J. M., Darley, J. M., and Robinson, P. H. (2002). Why Do We Punish? Deterrence and Just Deserts as Motives for Punishment. *Journal of Personality and Social Psychology*, 83: 284–99.

Carrasco, D. (1999) *City of Sacrifice: The Aztec Empire and the Role of Violence in Civilization*. Boston, MA: Beacon Press.

Caughlin, J. P, Huston, T. L., and Houts, R. M. (2000). How Does Personality Matter in Marriage? Anxiety, Interpersonal Negativity and Marital Satisfaction. *Journal of Personality and Social Psychology*, 78: 326–36.

Chagnon, N. A. (1968). *Yanomam'yo: The Fierce People*. New York, NY: Holt, Rinehart and Winston.

Cleckley, H. M. (1941). *The Mask of Sanity: An Attempt to Reinterpret the So-Called Psychopathic Personality*. St Louis, MO: The C. V. Mosby Company.

Coale, A. J., and Banister, J. (1994). Five Decades of Missing Females in China. *Demography*, 31: 459–79.

Coe, C. L., and Rosenblum, L. A. (1984). Male Dominance in the Bonnet Macaque: A Malleable Relationship. In P. Barchas and S. P. Mendoza (eds.), *Social Cohesion: Essays toward a Sociophysiological Perspective*, pp. 31–64. Westport, CT: Greenwood Press.

Cohen, A. B., and Rozin, P. (2001). Religion and the Morality of Mentality. *Journal of Personality and Social Psychology*, 81: 697–710.

Cohen, D., and Nisbett, R. E. (1994). Self-Protection and the Culture of Honor: Explaining Southern Violence. *Personality and Social Psychology Bulletin*, 20/5: 551–67.

Colby, A., Kohlberg, L., Gibbs, J., and Lieberman, M. (1983). A Longitudinal Study of Moral Judgment. *Monographs of the Society for Research in Child Development*, 48, nos. 1–2.

Cooper, D. (1978). Moral Relativism. *Midwest Studies in Philosophy*, 3: 97–108.

Cosmides, L. (1989). The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task. *Cognition*, 31: 187–276.

Cosmides, L., and Tooby, J. (1992). Cognitive Adaptations for Social Exchange. In J. H. Barkow, L. Cosmides, and J. Tooby (eds.), *the Adapted Mind: Evolutionary Psychology and the Generation of Culture*, pp. 163–228. New York, NY: Oxford University Press.

Cosmides, L. and Tooby, J. (1997). *Evolutionary Psychology Primer*. [Http://www.psych.ucsb.edu/research/cep/primer.html](http://www.psych.ucsb.edu/research/cep/primer.html)

Costa, P., and McCrae, R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO-FFI): Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Crimmins, M. (1989). Having Ideas and Having the Concept. *Mind and Language*, 4: 280–94.

D'Arms, J., and Jacobson, D. (2000). Sentiment and Value. *Ethics*, 110: 722–48.

_____. (2006). Sensibility Theory and Projectivism. In D. Copp (ed.), *the Oxford Handbook of Ethical Theory*, pp. 186–218. New York: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion Reason and the Human Brain*. New York, NY: Gossett/Putnam.

Damasio, A. R., and Van Hoesen, G. W. (1983). Emotional Disturbances Associated With Focal Lesions of the Limbic Frontal Lobe. In K. M. Heilman and P. Satz (eds.), *Neuropsychology of Human Emotion*. New York, NY: Guilford Press.

Darby, B. W., and Jeffers, D. (1988). The Effects of Defendant and Juror Attractiveness on Simulated Courtroom Trial Decisions. *Social Behavior and Personality*, 16: 39–50.

Darley, J. M., and Batson, C. D. (1973). From Jerusalem to Jericho: A Study of Situational and Dispositional Variables in Helping Behavior. *Journal of Personality and Social Psychology*, 27: 100–8.

Darwall, S. (1998). Empathy, Sympathy, Care. *Philosophical Studies*, 89: 261–82.

Darwall, S., Gibbard, A., and Railton, P. (1992). Toward Fin De Siècle Ethics: Some Trends, *Philosophical Review*, 101: 115–89.

Davidson, D. (1974). On the Very Idea of a Conceptual Scheme. *Proceedings and Addresses of the American Philosophical Association*, 47: 5–20.

_____. (1975). Thought and Talk. In S. Guttenplan (ed.), *Mind and Language*, pp. 7–24. Oxford: Oxford University Press.

Davidson, R. J., and Irwin, W. (1999). The Functional Neuroanatomy of Emotion and Affective Style. *Trends in Cognitive Sciences*, 3: 11–21.

Davies, J. G. V., and Maliphant, R. (1971). Autonomic Responses of Male Adolescents Exhibiting Refractory Behavior in School. *Journal of Child Psychology and Psychiatry*, 12: 115–27.

Davis, D. B. (1984). *Slavery and Human Progress*. New York, NY: Oxford University Press.

Davis, D. L., and Whitten, R. G. (1987). The Cross-Cultural Study of Human Sexuality. *Annual Review of Anthropology*, 16: 69–98.

Dawkins, R. (1976). *The Selfish Gene*. Oxford: Oxford University Press.

De Gelder, B., Vroomen, J., Pourtois, G., and Weiskrantz, L. (1999). Non-Conscious Recognition of Affect in the Absence of Striate Cortex. *Neuroreport*, 10: 3759–63.

De Waal, F. B. M. (1996). *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.

_____. (2000). Primates: A Natural Heritage of Conflict Resolution. *Science*, 289: 586–90.

_____. (1997). The Chimpanzee's Service Economy: Food for Grooming. *Evolution and Human Behavior*, 18: 375–86.

De Waal, F. B. M., and Lanting, F. (1997) *Bonobo: The Forgotten Ape*. Berkeley, CA: University of California Press.

DeGusta, D. (1999). Fijian Cannibalism: Osteological Evidence from Navatu. *American Journal of Physical Anthropology*, 110: 215–41.

Dennett, D. C. (1991). Lovely and Suspect Qualities. In Enrique Villanueva (ed.), *Consciousness*, pp. 37–43. Atascadero, CA: Ridgeview.

_____. (1995). *Darwin's Dangerous Idea*. New York, NY: Simon and Schuster.

DeSteno, D. A., Bartlett, M. Y., Salovey, P., and Braverman, J. (2002). Sex Differences in Jealousy: Evolutionary Mechanism or Artifact of Measurement? *Journal of Personality and Social Psychology*, 83, 1103–16.

Diamond, J. (2004). *Collapse: How Societies Choose to Fail or Succeed*. New York, NY: Viking Press.

Diener, E., and Oishi, S. (2003). Are Scandinavians Happier Than Asians? Issues in Comparing Nations on Subjective Well-Being. In F. Columbus (ed.), *Politics and Economics of Asia*, vol. 15, pp. 187–219. Hauppauge, NY: Nova Science Publishers.

Diener, E., Oishi, S., and Lucas, R. E. (2003). Personality, Culture, and Subjective Well-Being: Emotional and Cognitive Evaluations of Life. *Annual Review of Psychology*, 54: 403–25.

Diener, E., Scollon, C. N., and Lucas, R. E. (2004). The Evolving Concept of Subjective Well-Being: The Multifaceted Nature of Happiness. In P. T. Costa and I. C. Siegler (eds.), *Advances in Cell Aging and Gerontology*, vol. 15, pp. 187–220. Amsterdam: Elsevier.

Dion, K., Berscheid, E., and Walster, E. (1972). What Is Beautiful Is Good. *Journal of Personality and Social Psychology*, 24: 285–90.

Doris, J. M. (1998). Persons, Situations and Virtue Ethics. *Noûs*, 32: 504–30.

_____. (2002). *Lack of Character*. Cambridge: Cambridge University Press.

Downing, F. G. (1972). Ways of Deriving “Ought” From “Is”. *The Philosophical Quarterly*, 22: 234–47.

Dreier, J. (1990). Internalism and Speaker Relativism. *Ethics*, 101: 6–26.

_____. (2006). Was Moore A Moorean? In T. Horgan and M. Timmons (eds.), *The Legacy of G. E. Moore: 100 Years of Metaethics*. Oxford: Oxford University Press.

Dretske, F. I. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.

Dwyer, S. J (1999). Moral Competence. In K. Murasugi and R. Stainton (eds.), *Philosophy and Linguistics*, pp. 169–90. Boulder, CO: Westview Press.

Edgerton, R. B. (1971). *The Individual in Cultural Adaptation: A Study of Four East African Peoples*. Berkeley, CA: University of California Press.

_____. (1992). *Sick Societies: Challenging the Myth of Primitive Harmony*. New York, NY: The Free Press.

Ehrlich, P. R. (2000). *Human Natures: Genes, Cultures, and the Human Prospect*. Washington D.C.: Island Press.

Eisenberg, N. (2000). Emotion, Regulation, and Moral Development. *Annual Review of Psychology*, 51: 665–97.

Ekman, P. (1972). Universals and Cultural Differences in Facial Expressions of Emotion. In J. Cole (ed.), *Nebraska Symposium on Motivation 1971*, pp. 207–83. Lincoln, NE: University of Nebraska Press.

Ekman, P., and Friesen, W. V. (1986). A New Pan Cultural Facial Expression of Emotion. *Motivation and Emotion*, 10: 159–68.

Erasmus, D. (1530/1985). *On Good Manners for Boys*. In B. Mcgregor (trans.) and J. Sowlands (ed.), *Collected Works of Erasmus*, vol. 25. Toronto: University of Toronto Press.

Ehrlich, P. R. (2000). *Human Natures: Genes, Cultures, and the Human Prospect*. Washington D.C.: Island Press.

- Fehr, E., and Gächter, S. (2002). Altruistic Punishment in Humans. *Nature*, 415: 137–40.
- Firth, R. (1952). Ethical Absolutism and the Ideal Observer. *Philosophy and Phenomenological Research*, 12: 317–45.
- Flanagan, O. (1991). *Varieties of Moral Personality: Ethics and Psychological Realism*. Cambridge, MA: Harvard University Press.
- Fodor, J. A. (1975). *The Language of Thought*. New York, NY: Cromwell. (1990). *A Theory of Content*. Cambridge, MA: MIT Press.
- Fogas, J. P., and Bower, G. H. (1987). Mood Effects on Personal Perception Judgments. *Journal of Personality and Social Psychology*, 51: 53–60.
- Foot, P. (1972). Morality as a System of Hypothetical Imperatives. *Philosophical Review*, 81: 305–16.
- Forsyth, D. R. (1980). A Taxonomy of Ethical Ideologies. *Journal of Personality and Social Psychology*, 39: 175–84.
- Foucault, M. (1977). *Discipline and Punish: The Birth of the Prison*, trans. A. M. Sheridan-Smith. New York, NY: Vintage.
- Fowles, D. C., (1980). The Three Arousal Model: Implications of Gray's Two-Factor Learning Theory for Heart Rate, Electrodermal Activity, and Psychopathy. *Psychophysiology*, 17: 87–104.
- Frank, R. H. (1988). *Passion within Reason: The Strategic Role of the Emotions*. New York, NY: Norton.
- Fredrickson, G. M. (1971). *The Black Image in the White Mind*. New York, NY: Harper & Row.
- Freeman, M. M. R. (1971). A Social and Ecologic Analysis of Systematic Female Infanticide among the Netsilik Eskimo. *American Anthropologist*, 73: 1011–18.
- Frei, T. (2005). *Implications of the Categoricality of the Hypothetical Imperative*. Doctoral Thesis, Department of Philosophy, University of North Carolina, Chapel Hill.
- Garn, S. M. (1979). The Noneconomic Nature of Eating People. *American Anthropologist*, 81: 903.
- Geach, P. T. (1965). Assertion. *Philosophical Review*, 74: 449–65.
- Gemes, K. (2006). We Remain of Necessity Strangers to Ourselves: The Key Message of Nietzsche's *Genealogy*. In C. Acampora (ed.), *Nietzsche's On the Genealogy of Morals: Critical Essays*. New York, NY: Rowman & Littlefield.

Gewirth, A. (1994). Is Cultural Pluralism Relevant to Moral Knowledge? *Social Philosophy & Policy*, 11: 22–43.

Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.

_____. (2003). *Thinking How to Live*. Cambridge, MA: Harvard University Press.

Gigerenzer, G., and Hug, K. (1992). Domain-Specific Reasoning: Social Contracts, Cheating, and Perspective Change. *Cognition*, 43: 127–71.

Goldman, A. I., and Sripada, C. S. (2005). Simulationist Models of Face-Based Emotion Recognition, *Cognition*, 94: 193–213.

Goldstein, M. H. (1987). When Brothers Share A Wife. *Natural History*, 96: 39–49.

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.

Goody, J. (1983). *The Development of the Family and Marriage in Europe*. Cambridge: Cambridge University Press.

Gould, S. J. (1981). *The Mismeasure of Man*. New York, NY: Norton.

Gray, J. A. (1987). *The Psychology of Fear and Stress*. Cambridge: Cambridge University Press.

Gray, N. S., Macculloch, M. J., Smith, J., Morris, M., and Snowden, R. J. (2003). Forensic Psychology: Violence Viewed By Psychopathic Murderers. *Nature*, 423: 497–8.

Greene, J. D., and Haidt, J. (2002). How (and Where) Does Moral Judgment Work? *Trends in Cognitive Sciences*, 6: 517–523.

Greene, J.D., Lindsell, D., Clarke, A. C., Nystrom, L. E., and Cohen, J.D. (no prelo). What Pushes Your Moral Buttons? Modular Myopia and the Trolley Problem.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., and Cohen, J. D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science*, 293: 2105–8.

Greenspan, P. (2000). Emotional Strategies and Rationality. *Ethics*, 110: 469–87.

Grossbard-Shechtman, A. (1984). A Theory of Allocation of Time in Markets for Labour and Marriage. *The Economic Journal*, 94: 863–82.

Grusec, J. E., and Goodnow, J. J. (1994). Impact of Parental Discipline Methods on the Child's Internalization of Values: A Reconceptualization of Current Points of View. *Developmental Psychology*, 30: 4–19.

Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108: 814–34.

_____. (2003). Elevation and the Positive Psychology of Morality. In C. L. Keyes and J. Haidt (eds.), *Flourishing: Positive Psychology and the Life Well-Lived*, pp. 275–89. Washington, DC: American Psychological Association.

Haidt, J., and Joseph, C. (2004). Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus*, 133: 55–66.

Hamilton W. D. (1964). The Evolution of Social Behavior. *Journal of Theoretical Biology*, 7: 1–52.

Haney, C., Banks, C., and Zimbardo, P. (1973). Interpersonal Dynamics in A Simulated Prison. *International Journal of Criminology & Penology*, 1: 69–97.

Hardin, C. L. (1993). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis, in: Hackett Publishing Company.

_____. (1991). *The Hare Psychopathy Checklist—Revised*. Toronto: Multi-Health Systems.

_____. (1993). *Without Conscience: The Disturbing World of the Psychopaths among Us*. New York, NY: Pocket Books.

_____. (1998). Psychopathy, Affect and Behavior. In D. J. Cooke, R. D. Hare, and A. E. Forth (eds.), *Psychopathy: Theory, Research, and Implications for Society*, pp. 105–37. Dordrecht: Kluwer.

Harman, G. (1975). Moral Relativism Defended. *Philosophical Review*, 84: 3–22.

_____. (1977). *The Nature of Morality*. Oxford: Oxford University Press.

_____. (1999). Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error. *Proceedings of the Aristotelian Society*, 99: 315–31.

_____. (2000) Moral Philosophy and Linguistics. In G. Harman, *Explaining Value*, pp. 217–25. Oxford: Oxford University Press.

_____. (2001). Virtue Ethics without Character Traits. In A. Byrne, R. Stalnaker, and R. Wedgwood (eds.), *Fact and Value: Essays on Ethics and Metaphysics for Judith Jarvis Thomson*, pp. 117–27. Cambridge, MA: MIT Press.

Harner, M. (1977). The Ecological Basis for Aztec Sacrifice. *American Ethnologist*, 4: 117–35.

Harris, C. R. (2004). The Evolution of Jealousy. *American Scientist*, 92: 62–71.
Harris, M. (1985). *Good to Eat: Riddles of Food and Culture*. New York, NY: Simon and Schuster.

_____. (1989). *Our Kind*. New York, NY: Harper Collins.

_____. (1993). The Evolution of Human Gender Hierarchies: A Trial Formulation. In B. D. Miller (ed.), *Sex and Gender Hierarchies*, pp. 57–79. Cambridge: Cambridge University Press.

Harris, M. B., Benson, S. M., and Hall, C. (1975). The Effects of Confession on Altruism. *Journal of Social Psychology*, 96: 187–92.

Hauser, M. D. (2001). *Wild Minds: What Animals Really Think*. New York, NY: Henry Holt.

Hauser, M. D., Young, L., and Cushman, F. (forthcoming). Reviving Rawls' Linguistic Analogy: Operative Principles and the Causal Structure of Moral Actions. In W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality*. Cambridge, MA: MIT Press.

Heekeren, H. R., Wartenburger, I., Schmidt, H., Schwintowski, H. P., and Villringer, A. (2003). An fMRI Study of Simple Ethical Decision-Making. *Neuroreport*, 14: 1215–19.

Heine, B. (1985). The Mountain People: Some Notes on the Ik of North-Eastern Uganda. *Africa*, 55: 3–16.

Henrich, J., and Boyd, R. (2001). Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas. *Journal of Theoretical Biology*, 208: 79–89.

Henrich, J., and Gil-White, F. (2001). The Evolution of Prestige: Freely Conferred Status as a Mechanism for Enhancing the Benefits of Cultural Transmission. *Evolution and Human Behavior*, 22: 1–32.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Gintis, H., McElreath, R., and Fehe, E. (2001). In Search of Homo Economicus: Experiments in 15 Small-Scale Societies. *American Economic Review*, 91: 73–9.

Herdt, G. (1987). *The Sambia: Ritual and Gender in New Guinea*. New York, NY: Holt, Rinehart, and Winston.

Hill, T. (1973). The Hypothetical Imperative. *The Philosophical Review*, 82: 429–50.

Hoffman M. L. (1983). Affective and Cognitive Processes in Moral Internalization. In E. T. Higgins, D. N. Ruble, and W. W. Hartup (eds.), *Social Cognition and Social Development: A Sociocultural Perspective*, pp. 236–74. Cambridge: Cambridge University Press.

_____. (1998). Varieties of Empathy-Based Guilt. In J. Bybee (ed.), *Guilt and Children*, pp. 91–112. New York, NY: Academic Press.

_____. (2000). *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press.

Horgan, T., and Timmons, M. (1990–1). New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, 16: 447–65.

House, T. H., and Milligan, W. L. (1976). Autonomic Responses to Modeled Distress in Prison Psychopaths. *Journal of Personality and Social Psychology*, 34: 556–60.

Hrdy, S. (1999). *Mother Nature: A History of Mothers, Infants, and Natural Selection*. New York, NY: Pantheon Books.

Hua, C. (2001). *A Society without Fathers or Husbands: The Na of China*, trans. A. Hustvedt. New York, NY: Zone Books.

Hume, D. (1739/1978). *A Treatise of Human Nature*, ed. P. H. Nidditch. Oxford: Oxford University Press.

_____. (1751/1998). *An Enquiry Concerning the Principles of Morals*, ed. T. L. Beauchamp. Oxford: Oxford University Press.

_____. (1757/1993). On the Standard of Taste. In *Selected Essays*, ed. S. Copley and A. Edgar, pp. 133–54. Oxford: Oxford University Press.

Hupka, R. B. (1981). Cultural Determinants of Jealousy. *Alternative Lifestyles*, 4: 310–56.

_____. (1991). The Motive for the Arousal of Romantic Jealousy: Its Cultural Origin. In P. Salovey (ed.), *the Psychology of Jealousy and Envy*, pp. 252–70. New York, NY: Guilford Press.

Hutcheson, F. (1738/1994). *An Inquiry into the Original of Our Ideas of Beauty and Virtue*. In *Philosophical Writings*, ed. R. S. Downie. London: J.M.Dent.

Ingham, J. M. (1984). Human Sacrifice at Tenochtitlan. *Comparative Studies in Society and History*, 3: 379–400.

Isen, A. M., and Levin, P. F. (1972). The Effect of Feeling Good on Helping: Cookies and Kindness. *Journal of Personality and Social Psychology*, 21: 384–8.

Jaffee, S., and Hyde, J. S. (2000). Gender Differences in Moral Orientation: A Meta-Analysis. *Psychological Bulletin*, 126: 703–26.

James, W. (1884). What is an Emotion? *Mind*, 9: 188–205.

Jang, K. L., Livesley, W. J., and Vernon, P. A. (1996). Heritability of the Big Five Dimensions and their Facets: A Twin Study. *Journal of Personality*, 64: 575–91.

Johnston, M. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society Supplementary Volume*, 63: 139–74.

Joyce, R. (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.

_____. (2006). *The Evolution of Morality*. Cambridge, MA: MIT Press.

Junod, H. A. (1962). *The Life of a South African Tribe*, vol. II: *Mental Life*. New York, NY: University Books.

Kagan, S. (1989). *The Limits of Morality*. Oxford: Oxford University Press.

Kamtekar, R. (2004). Situationism and Virtue Ethics on the Content of Our Character. *Ethics*, 114, 458–91.

Kant, I. (1785/1998). *Groundwork of the Metaphysic of Morals*, ed. Mary J. Gregor. Cambridge: Cambridge University Press.

Kaplan, D. (1989). Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*, pp. 481–563. Oxford: Oxford University Press.

Kawabata, H., and Zeki, S. (2004). Neural Correlates of Beauty. *Neurophysiology*, 91: 1699–1705.

Kelly, D., Stich, S., Haley, K., Eng, S., and Fessler, D. (2007). Harm, Affect and the Moral/Conventional Distinction. *Mind & Language*, 22: 117–31.

Kelly, R. C. (1977). *Etoro Social Structure: A Study in Structural Contradiction*. Ann Arbor, MI: University of Michigan Press.

Kelly, R. L. (1995). *The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways*. Washington D.C.: Smithsonian Books.

Kilham, W., and Mann, L. (1974). Level of Destructive Obedience as A Junction of Transmitter and Executant Roles in the Milgram Obedience Paradigm. *Journal of Personality and Social Psychology*, 29: 696–702.

- Cluegel, J. R., and Smith, E. R. (1986). *Beliefs about Inequality: Americans' Views of What Is and What Ought to Be*. New York: Aldine De Gruyter.
- Knobe, J. (2003). Intentional Action in Folk Psychology: An Experimental Investigation. *Philosophical Psychology*, 16: 309–24.
- Kochanska, G., Casey, R. J., and Fukumoto, A. (1995). Toddlers' Sensitivity to Standard Violations. *Child Development*, 66: 643–56.
- Kochanska, G., Gross, J. N., Lin, M., and Nichols, K. E. (2002). Guilt in Young Children: Development, Determinants, and Relations With A Broader System of Standards. *Child Development*, 73: 461–82.
- Kohlberg, L. (1984). *The Psychology of Moral Development: Moral Stages and the Life Cycle*. San Francisco, CA: Harper & Row.
- Kölbel, M. (2003). Faultless Disagreement. *Proceedings of the Aristotelian Society*, 104: 53–73.
- Kompa, N. (2002). The Context Sensitivity of Knowledge Ascriptions, *Grazer Philosophische Studien*, 64: 79–96.
- Korsgaard, C. M. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Kripke, S. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Krueger, R. F., Caspi, A., and Moffitt, T. E. (2000). Epidemiological Personology: The Unifying Role of Personality in Population-Based Research on Problem Behaviors. *Journal of Personality*, 68: 967–98.
- Kuper, A. (2002). Incest, Cousin Marriage, and the Origin of the Human Sciences in Nineteenth-Century England. *Past and Present*, 174: 158–83.
- Kupperman, J. (2001). The Indispensability of Character. *Philosophy*, 76: 239–50.
- Lakoff, G. (2002). *Moral Politics: How Liberals and Conservatives Think*, 2nd edn. Chicago, IL: University of Chicago Press.
- Lange, C. G. (1885/1922). *Om Sindsbevaegelser: Et Psyko-Fysiologisk Studie*. Kjøbenhavn: Jacob Lunds. Reprinted in I. A. Haupt (trans.), *the Emotions*. Baltimore, MD: Williams and Wilkins Company.
- Laserson, P. (2005). Context Dependence, Disagreement, and Predicates of Personal Taste. *Linguistics and Philosophy*, 28: 643–86.

Lazarus, R. S. (1991). *Emotion and Adaptation*. New York, NY: Oxford University Press.

Ledoux J. E. (1996). *The Emotional Brain*. New York, NY: Simon and Schuster.

Lee, R. B. (1969). Eating Christmas in the Kalahari. *Natural History*, 78: 14, 16, 18, 21–2,60–3.

Leiter, B. (2001). Objectivity, Morality, and Adjudication. In B. Leiter (ed.), *Objectivity in Law and Morals*. Cambridge: Cambridge University Press.

Lepore, E., and Cappelen, H. (forthcoming). The Myth of Unarticulated Constituents. In M. O'Rourke and C. Washington (eds.), *Essays in Honor of John Perry*. Cambridge, MA: MIT Press.

Lerner, J., Goldberg, J., and Tetlock, P. E. (1998). Sober Second Thought: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility. *Personality and Social Psychology Bulletin*, 24: 563–74.

Leung, K., and Bond, M. H. (1984). The Impact of Cultural Collectivism on Reward Allocation. *Journal of Personality and Social Psychology*, 47: 793–804.

Leupp, G. P. (1995). *Male Colors: The Construction of Homosexuality in Tokugawa Japan*. Berkeley, CA: University of California Press.

Levenson, R. W., Ekman, P., and Friesen, W. V. (1990). Voluntary Facial Action Generates Emotion-Specific Autonomic Nervous System Activity. *Psychophysiology*, 27: 363–84.

Lewis, M. (1998). Emotional Competence and Development. In D. Pushkar, W. M. Bukowski, A. E. Schwartzman, D. M. Stack, and D. R. White (eds.), *Improving Competence across the Lifespan*, pp. 27–36. New York, NY: Plenum.

Lewis, D. (1989). Dispositional Theories of Value. *Proceedings of the Aristotelian Society Supplementary Volume*, 63: 113–37.

Lieberman, D., Tooby, J., and Cosmides, L. (2003). Does Morality Have A Biological Basis? An Empirical Test of the Factors Governing Moral Sentiments Regarding Incest. *Proceedings of the Royal Society, London B*, 270: 819–26.

Locke, J. (1690/1979). *An Essay Concerning Human Understanding*, ed. P. H. Nidditch. Oxford: Oxford University Press.

Lynam, D. R., Caspi, A., Moffitt, T. E., Raine, A., Loeber, R., and Stouthamer-Loeber, M. (2005). Adolescent Psychopathy and the Big Five: Results from Two Samples. *Journal of Abnormal Child Psychology*, 33: 431–43.

- Lyons, D. (1976). Ethical Relativism and the Problem of Incoherence. *Ethics*, 86: 107–21.
- McCauley, R., and Lawson, E. T. (2002). *Bringing Ritual to Mind: Psychological Foundations of Cultural Forms*. Cambridge, MA: Cambridge University Press.
- McCullough, M., Kilpatrick, S., Emmons, R., and Larson, D. (2001) Is Gratitude a Moral Affect? *Psychological Bulletin*, 127: 249–66.
- McDougall, W. (1908). *An Introduction to Social Psychology*. London: Methuen & Co.
- McDowell, J. (1985). Values and Secondary Qualities. In T. Honderich (ed.), *Morality and Objectivity*. London: Routledge & Kegan Paul.
- _____. (1987). Projection and Truth in Ethics, Lindsay Lecture, University of Kansas. Reprinted in S. Darwall, A. Gibbard, and P. Railton (eds.), *Moral Discourse and Practice: Some Philosophical Approaches*, pp. 215–27. New York: Oxford University Press.
- MacFarlane, J. (2005). The Assessment Sensitivity of Knowledge Attributions. *Oxford Studies in Epistemology*, 1: 197–233.
- Mackay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., and Dyer, J. R. (2004). Relations between Emotion, Memory, and Attention: Evidence from Taboo Stroop, Lexical Decision, and Immediate Memory Tasks. *Memory & Cognition*, 32: 474–88.
- Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin.
- McMillen, D. L., and Austin, J. B. (1971). Effect of Positive Feedback on Compliance Following Transgression. *Psychonomic Science*, 24: 59–61.
- McNaughton, D. (1988). *Moral Vision: An Introduction to Ethics*. Oxford: Blackwell.
- Malle, B. F., and Bennett, R. E. (2002). People's Praise and Blame for Intentions and Actions: Implications of the Folk Concept of Intentionality. *Technical Reports of the Institute of Cognitive and Decision Sciences*, no. 02–2, Eugene, or.
- Manktelow, K., and Over, D. (1990). Deontic Thought and the Selection Task. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, and G. Erdos (eds.), *Lines of Thinking*, pp. 153–64. London: Wiley.
- Mantell, D. M. (1971). The Potential for Violence in Germany. *Journal of Social Issues*, 27: 101–12.
- Marglin, F. A. (1985). *Wives of the God-King: Rituals of Devadasi of Puri*. Delhi: Oxford University Press.

Marlar, R. A., Leonard, B. L., Billman, B. R., Lambert, P. M., and Marlar, J. E. (2000). Biochemical Evidence of Cannibalism at a Prehistoric Puebloan Site in Southwestern Colorado. *Nature*, 407: 74–8.

Mayor, M. (1974). Fears and Fantasies of the Anti-Suffragists. *Connecticut Review*, 7: 64–74.

Mead, S., Stumpf, M. P. H., Whitfield, J., Beck, J. A., Poulter, M., Campbell, T., Uphill, J. B., Goldstein, D., Alpers, M., Fisher, E. M. C., and Collinge, J. (2003). Balancing Selection at the Prion Protein Gene Consistent With Prehistoric Kurulike Epidemics. *Science*, 300: 640–3.

Mealey, L. (1995). The Sociobiology of Sociopathy: An Integrated Evolutionary Model. *Behavioral and Brain Sciences*, 18: 523–99.

Meeks, W. A. (1993). *The Origins of Christian Morality: The First Two Centuries*. New Haven, CT: Yale University.

Mele, A. R. (1996). Internalist Moral Cognitivism and Listlessness. *Ethics*, 106: 727–53.

Meltzoff, A. (1995). Understanding the Intentions of Others: Reenactment of Intended Acts By 18-Month-Old Children. *Developmental Psychology*, 31: 838–50.

Merritt, M. (2000). Virtue Ethics and Situationist Personality Psychology. *Ethical Theory and Moral Practice*, 3: 365–83.

Midgley, M. (1981). On Trying out One's New Sword on a Chance Wayfarer. In M. Midgley, *Heart and Mind*. London: Methuen & Co.

Mikhail, J. (2000). Rawls' Linguistic Analogy: A Study of the "Generative Grammar" Model of Moral Theory Described By John Rawls in "A Theory of Justice". Doctoral Dissertation, Department of Philosophy, Cornell University.

_____. (2002). Aspects of the Theory of Moral Cognition: Investigating Intuitive Knowledge of the Prohibition of Intentional Battery and the Principle of Double Effect. Georgetown University Law Center Public Law & Legal Theory Working Paper No. 762385.

Milgram, S. (1974). *Obedience to Authority*. New York, NY: Harper & Row.

Miller, R. S. (1996). *Embarrassment: Poise and Peril in Everyday Life*. New York, NY: Guilford Press.

Millgram, E. (1999). Moral Values and Secondary Qualities. *American Philosophical Quarterly*, 36: 253–5.

- Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.
- Mischel, W. (1968). *Personality and Assessment*. New York, NY: John J. Wiley and Sons.
- Mitchell, P. G., Tetlock, P. E., Newman, D., and Lerner, J. (2003). Experiments behind the Veil: A Hypothetical Societies Approach to the Study of Social Justice. *Political Psychology*, 24: 519–47.
- Moll, J., De Oliveira-Souza, R., Bramati, I., and Grafman, J. (2002). Functional Networks in Emotional Moral and Nonmoral Social Judgments. *Neuroimage*, 16: 696–703.
- Moll, J., De Oliveira-Souza, R., and Eslinger, P. J. (2003). Morals and the Human Brain: A Working Model. *Neuroreport*, 14: 299–305.
- Moody-Adams, M. (1997). *Fieldwork in Familiar Places: Morality, Culture and Philosophy*. Cambridge, MA: Harvard University Press.
- Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- Moore, J. (1984). The Evolution of Reciprocal Sharing. *Ethology and Sociobiology*, 5: 5–14.
- Morris, J. S., Öhman, A., and Dolan, R. J. (1999). A Subcortical Pathway to the Right Amygdala Mediating “Unseen” Fear. *Proceedings of National Academy of Science*, 96: 1680–5.
- Moscovitch, M. (1995). Confabulation. In Schacter, D. (ed.), *Memory Distortion*, pp. 226–54. Cambridge, MA: Harvard University Press.
- Murdock, G. P. (1981) *Atlas of World Cultures*. Pittsburgh, PA: University of Pittsburgh Press.
- Murphy, S., Haidt, J., and Björklund, F. (2000). Moral Dumbfounding: When Intuition Finds No Reason. Unpublished Manuscript, Department of Philosophy, University of Virginia.
- Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- Neuberg, S. L., Cialdini, R. B., Brown, S. L., Luce, C., Sagarin, B. J., and Lewis, B. P. (1997). Does Empathy Lead to Anything More Than Superficial Helping? Comment on Batson et al. (1997). *Journal of Personality and Social Psychology*, 73: 510–16.
- Nichols, S. (2002). On the Genealogy of Norms: A Case for the Role of Emotion in Cultural Evolution. *Philosophy of Science*, 69: 234–55.

_____. (2004a). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York, NY: Oxford University Press.

_____. (2004b). After Objectivity: An Empirical Study of Moral Judgment. *Philosophical Psychology*, 17: 5–28.

_____. (2005). Innateness and Moral Psychology. In P. Carruthers, S. Laurence, and S. Stich (Eds.), *The Innate Mind: Structure and Content*. New York, NY: Oxford University Press.

Nietzsche, F. W. (1883–8/1968). *The Will to Power*, trans. W. Kauffman. New York, NY: Vintage.

_____. (1885/1966). *Thus Spoke Zarathustra: A Book for None and All*, trans. W. Kauffman. New York, NY: Viking Penguin.

_____. (1887/1967). *On the Genealogy of Morals*, trans. W. Kauffman. New York, NY: Random House.

_____. (1888/1920). *The Antichrist*, trans. H. L. Mencken. New York, NY: Knopf.

Nisbett, R. E. (2003). *The Geography of Thought: How Asians and Westerners Think Differently and Why*. New York, NY: Free Press.

Nisbett, R. E., and Cohen, D. (1996). *Culture of Honor: The Psychology of Violence in the South*. Boulder, CO: Westview Press.

Nisbett, R. E., and Wilson, T. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84: 231–59.

Nucci, L. P. (2001). *Education in the Moral Domain*. Cambridge: Cambridge University Press.

Nucci, L. P., and Weber, E. (1995). Social Interactions in the Home and the Development of Young Children's Conceptions of the Personal. *Child Development*, 66: 1438–52.

Nunner-Winkler, G., and Sodian, B. (1988). Children's Understanding of Moral Emotions. *Child Development*, 59: 1323–38.

Nussbaum, M. C. (1999). *Sex and Social Justice*. Oxford: Oxford University Press.

_____. (2001). *Upheavals of Thought: The Intelligence of the Emotions*. Cambridge: Cambridge University Press.

Oberg, K. (1934). Crime and Punishment in Tlingit Society. *American Anthropologist*, 36: 145–56.

- Okonofua, F. E., Larsen, U., Oronsaye, F., Snow, R. C., and Slanger, T. E. (2002). The Association between Female Genital Cutting and Correlates of Sexual and Gynaecological Morbidity in Edo State, Nigeria. *British Journal of Obstetrics and Gynaecology*, 109: 1089–96.
- Ones, D. S., Viswesvaran, C., and Schmidt, F. L. (1993). Comprehensive Meta-Analysis of Integrity Test Validities: Findings and Implications for Personnel Selection and Theories of Job Performance. *Journal of Applied Psychology*, 78: 679–703.
- Ortiz De Montellano, B. R. (1978). Aztec Cannibalism: An Ecological Necessity? *Science*, 200: 611–17.
- Panksepp, J. (1995). The Emotional Sources of "Chills" Induced by Music. *Music Perception*, 13: 171–207.
- Patrick, C. J., Bradley, M. M., and Lang, P. J. (1993). Emotion in the Criminal Psychopath: Startle Reflex Modulation. *Journal of Abnormal Psychology*, 102: 82–92.
- Paunonen, S. V., and Jackson, D. N. (2000). What Is Beyond the Big Five? Plenty! *Journal of Personality*, 68: 821–35.
- Phan, K. L., Wager, T. D., Taylor, S. F., and Liberzon, I. (2002). Functional Neuroanatomy of Emotion: A Meta-Analysis of Emotion Activation Studies in PET and fMRI. *Neuroimage*, 16: 331–48.
- Pitcher, G. (1965). Emotion. *Mind*, 74: 324–46.
- Platner, S. B. and Ashby, T. (1929). *A Topographical Dictionary of Ancient Rome*. Oxford: Oxford University Press.
- Preston, S. D., and De Waal, F. B. M. (2002). Empathy: Its Ultimate and Proximate Bases. *Behavioral and Brain Sciences*, 25: 1–19.
- Price, J., Sloman, L., Gardner, R., Gilbert, P., and Rohde, P. (1994). The Social Competition Hypothesis of Depression. *British Journal of Psychiatry*, 164: 309–15.
- Prinz, J. J. (2000). The Duality of Content. *Philosophical Studies*, 100: 1–34.
- _____. (2002). *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA: MIT Press.
- _____. (2004). *Gut Reactions: A Perceptual Theory of Emotion*. New York, NY: Oxford University Press.

_____. (2007a). Is Morality Innate? In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 1: The Evolution of Morality*. Cambridge, MA: MIT Press.

_____. (2007b) Resisting the Linguistic Analogy. In W. Sinnott-Armstrong (ed.), *Moral Psychology, Vol. 2: The Cognitive Science of Morality*. Cambridge, MA: MIT Press.

(2009). Against Moral Nativism. In M. Bishop and D. Murphy (eds.), *Stich and his Critics*. Oxford: Blackwell.

Provine, R. R. (2000). *Laughter: A Scientific Investigation*. New York, NY: Penguin.
Pusey, A. E., and Packer, C. (1997). The Ecology of Relationships. In J. R. Krebs and N. B. Davies (eds.), *Behavioural Ecology: An Evolutionary Approach*, pp. 254–83. Oxford: Blackwell.

Putnam, H. (1975). The Meaning of "Meaning." in K. Gunderson (ed.), *Language, Mind, and Knowledge*, pp. 131–93. Minneapolis, MN: University of Minnesota Press.

_____. (1980). *Reason, Truth and History*. Cambridge, MA: Cambridge University Press.

Quine, W. V. O. (1953). Two Dogmas of Empiricism. In *From a Logical Point of View*, pp. 20–46. Cambridge, MA: Harvard University Press.

_____. (1969). Natural Kinds. In *Ontological Relativity and Other Essays*, pp. 114–38. New York, NY: Columbia University Press.

Quine, W. V. O., and Ullian, J. S. (1978). *The Web of Belief*, 2nd edn. New York, NY: Random House.

Rachels, J. (1986). *The Elements of Moral Philosophy*. Philadelphia, PA: Temple University Press.

Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rée, P. (1877/2003). Origin of the Moral Sentiments. In *Basic Writings*, trans. R. Small. Champaign, IL: University of Illinois Press.

Richard, M. (2004). Contextualism and Relativism. *Philosophical Studies*, 119: 215–42. 324 References

Richards, R. (1986). A Defense of Evolutionary Ethics. *Biology and Philosophy*, 1: 265–93.

Rizzolatti, G., and Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27: 169–92.

- Roberts, B. W., and Delvecchio, W. F. (2000). The Rank-Order Consistency of Personality from Childhood to Old Age: A Quantitative Review of Longitudinal Studies. *Psychological Bulletin*, 126: 3–25.
- Ross, L., and Nisbett, R. E. (1991). *The Person and the Situation: Perspectives of Social Psychology*. New York, NY: McGraw-Hill.
- Ross, W. D. (1930). *The Right and the Good*. Oxford: Oxford University Press.
- Rozin, P., Haidt, J., and McCauley, C. (1993). Disgust. In M. Lewis and J. Haviland (eds.), *Handbook of Emotions*, pp. 575–94. New York, NY: Guilford Press.
- Rozin, P., Lowery, L., Imada, S., and Haidt, J. (1999). The CAD Triad Hypothesis: A Mapping between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity). *Journal of Personality and Social Psychology*, 76: 574–86.
- Rozin, P., Millman, L., and Nemeroff, C. (1986). Operation of the Laws of Sympathetic Magic in Disgust and Other Domains. *Journal of Personality and Social Psychology*, 50: 703–12.
- Rozin, P., and Singh, L. (1999). The Moralization of Cigarette Smoking in the United States. *Journal of Consumer Psychology*, 8: 339–42.
- Rudman, L. A., and Kilianski, S. E. (2000). Implicit and Explicit Attitudes toward Female Authority. *Personality and Social Psychology Bulletin*, 26: 1315–28.
- Ruse, M. (1991). The Significance of Evolution. In P. Singer (ed.), *A Companion to Ethics*, pp. 500–10. Oxford: Blackwell.
- Sahlins M. (1978). Culture as Protein and Profit. *The New York Review of Books*, 25: 45–53.
- Sanday, P. R. (1986). *Divine Hunger: Cannibalism as a Cultural System*. Cambridge: Cambridge University Press.
- Sanderson, S. K. (2001). Explaining Monogamy and Polygamy in Human Societies: Comment on Kanazawa and Still. *Social Forces*, 80: 329–35.
- Sanfey, A. G., Rilling, J. A., Aronson, J. K., Nystrom, L., and Cohen, J. D. (2003). The Neural Basis of Economic Decision Making in the Ultimatum Game, *Science*, 300: 1755–7.
- Sayre-McCord, G. (1994). On Why Hume's "General Point of View" Isn't Ideal—and Shouldn't Be. *Social Philosophy & Policy*, 11: 202–28.

Schalling, D., and Rosen, A. S. (1968). Porteus Maze Differences between Psychopathic and Non-Psychopathic Criminals. *British Journal of Social and Clinical Psychology*, 7: 224–8.

Scherer, K. R. (1993). Studying the Emotion Antecedent Appraisal Process: An Expert System Approach. *Cognition and Emotion*, 7: 325–55.

Schnall, S., Haidt, J., and Clore, G. L. (2005). Disgust as Embodied Moral Judgment. Unpublished Manuscript, Department of Psychology, University of Virginia.

Schneebaum, T. (1969). *Keep the River on Your Right*. New York, NY: Grove Press.

Searle, J. R. (1964). How to Derive "Ought" From "Is". *Philosophical Review*, 73: 43–58.

Shaw, B. (1992). Explaining Incest: Brother–Sister Marriage in Graeco-Roman Egypt. *Man*, 27: 267–99.

Shepher, J. (1971). Mate-Selection among Second-Generation Kibbutz Adolescents and Adults: Incest-Avoidance and Negative Imprinting. *Archives of Sexual Behavior*, 1: 293–307.

Sherman P. (1977). Nepotism and the Evolution of Alarm Calls. *Science*, 197: 1246–53.

Shin, L. M., Dougherty, D., Macklin, M. L., Orr, S. P., Pitman, R. K., and Rauch, S. L. (2000). Activation of Anterior Paralimbic Structures during Guilt-Related Script-Driven Imagery. *Biological Psychiatry*, 48: 43–50.

Shweder, R. A. (2000). What about "Female Genital Mutilation"? And Why Culture Matters in the First Place. *Daedalus*, 129: 209–32.

Shweder, R. A., Much, N. C., Mahapatra, M., and Park, L. (1997). The "Big Three" of Morality (Autonomy, Community, Divinity), and the "Big Three" Explanations of Suffering. In P. Rozin and A. Brandt (eds.), *Morality and Health*. New York, NY: Routledge.

Silk, J. B. (2002). The Form and Function of Reconciliation in Primates. *Annual Review of Anthropology*, 31: 21–44.

Silk, J. B., Brosnan, S. F., Vonk, J., Henrich, J., Povinelli, D. J., Richardson, A. F., Lambeth, S. P., Mascaró, J., and Schapiro, S. J. (2005). Chimpanzees are Indifferent to the Welfare of Other Group Members. *Nature*, 435: 1357–9.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, 303: 1157–62.

Singer, T., Seymour, B., O'Doherty, J., Sephan, K. E., Dolan, R. J., and Frith, C. (2006). Empathic Neural Responses are Modulated by the Perceived Fairness of Others. *Nature*, 439: 466–9.

Sinnott-Armstrong, W. (1996). Moral Skepticism and Justification. In W Sinnott-Armstrong and M. Timmons (eds.), *Moral Knowledge? New Readings in Moral Epistemology*. New York, NY: Oxford University Press.

Slotkin, J. S. (1947). On a Possible Lack of Incest Regulations in Old Iran. *American Anthropologist*, 49: 612–17.

Smetana, J. G. (1981). Preschool Children's Conceptions of Moral and Social Rules. *Child Development*, 52; 1333–6.

_____. (1989). Toddlers' Social Interactions in the Context of Moral and Conventional Transgressions in the Home. *Developmental Psychology*, 25: 499–508.

Smetana, J. and Braeges, J. (1990). The Development of Toddlers' Moral and Conventional Judgments. *Merrill-Palmer Quarterly*, 36: 329–46.

Smith, A. (1759/2000). *The Theory of Moral Sentiments*. Amherst, NY: Prometheus Books.

Smith, M. (1994). *The Moral Problem*. Oxford: Blackwell.

Smith, P. B., and Bond, M. H. (1998). *Social Psychology across Cultures*. New York, NY: Allyn & Bacon.

Snarey, J. R. (1985). Cross-Cultural Universality of Social-Moral Development: A Critical Review of Kohlbergian Research. *Psychological Bulletin*, 97: 202–32.

Sober, E. (1994). Prospects for an Evolutionary Ethics. In *From A Biological Point of View*, pp. 93–113. Cambridge: Cambridge University Press.

Sober, E., and Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.

Solomon, R. C. (1976). *The Passions*. New York, NY: Doubleday.

Solomon, R. L. (1980). The Opponent-Process Theory of Acquired Motivation: The Costs of Pleasure and the Benefits of Pain. *American Psychologist*, 35: 691–712.

Spencer, H. (1857). Progress: Its Law and Causes. *The Westminster Review*, 67: 445–7; 451: 454–6; 464–5.

Sperber, D. (1996). *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.

_____. (2000). An Objection to the Memetic Approach to Culture. In R. Aunger (Ed.), *Darwinizing Culture: The Status of Memetics as A Science*, pp. 163–73. Oxford: Oxford University Press.

Spinoza, B. (1677/1994). *Ethics*. In E. Curley (trans. and ed.), *A Spinoza Reader*. Princeton, NJ: Princeton University Press.

Sreenivasan, G. (2002). Errors about Errors: Virtue Theory and Trait Attribution. *Mind*, 111: 47–68.

Sripada, C. and Stich, S. (2006). A Framework for the Psychology of Norms. In P. Carruthers, S. Laurence and S. Stich (eds.), *The Innate Mind: Culture and Cognition*, pp. 280–301. New York: Oxford University Press.

Stanford, C. B., Wallis, J., Matama, H., and Goodall, J. (1994). Patterns of Predation by Chimpanzees on Red Colobus Monkeys in Gombe National Park, 1982–1991. *American Journal of Physical Anthropology*, 94: 213–28.

Stark, R. (1996). *The Rise of Christianity: A Sociologist Reconsiders History*. Princeton, NJ: Princeton University Press.

Steinpreis, R. E., Anders, K. A., and Ritzke, D. (1999). The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study. *Sex Roles: A Journal of Research*, 41: 509–28.

Sterelny, K. (1990). *The Representational Theory of Mind*. Oxford: Blackwell.

Stevens, D., Charman, T., and Blair, R. J. R. (2001). Recognition of Emotion in Facial Expressions and Vocal Tones in Children with Psychopathic Tendencies. *Journal of Genetic Psychology*, 162: 201–11.

Stevens, J. R., and Hauser, M. D. (2004). Why Be Nice? Psychological Constraints on the Evolution of Cooperation. *Trends in Cognitive Sciences*, 8: 60–5.

Stevenson, C. L. (1937). The Emotive Meaning of Ethical Terms. *Mind*, 46: 14–31.

Stone, V., Cosmides, L., Tooby, J., Kroll, N., and Knight, R. (2002). Selective Impairment of Reasoning about Social Exchange in a Patient with Bilateral Limbic System Damage. *Proceedings of the National Academy of Sciences*, 99: 11531–6.

Strack, F., Martin, L. L., and Stepper, S. (1988). Inhibiting and Facilitating Conditions of Facial Expressions: A Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of Personality and Social Psychology*, 54: 768–77.

Sturgeon, N. L. (1985). Moral Explanations. In D. Copp and D. Zimmerman (eds.), *Morality Reason and Truth*, pp. 49–78. Totowa, NJ: Rowan and Allanheld.

Svallfors, S. (1997). Worlds of Welfare and Attitudes to Redistribution: A Comparison of Eight Western Nations. *European Sociological Review*, 13: 283–304.

Tamir, Y. (1996). Hands off Clitoridectomy: What Our Revulsion Reveals about Ourselves. *Boston Review*, 21. <http://bostonreview.net/BR21.3/Tamir.html>

Thomas, K. (1983). *Man and the Natural World: A History of the Modern Sensibility*. New York: Pantheon.

Thompson, W. E. (1967). The Marriage of First Cousins in Athenian Society. *Phoenix*, 21: 273–82.

Thomson, J. J. (1976). Killing, Letting Die, and the Trolley Problem. *The Monist*, 59: 204–17.

_____. (1997). The Right and the Good. *Journal of Philosophy*, 94: 273–98.

Thornhill, N. W. (1991). An Evolutionary Analysis of Rules Regulating Human Inbreeding and Marriage. *Behavioral and Brain Sciences*, 14: 247–93.

Tiberius, V. (2003). Cultural Differences and Philosophical Accounts of Well-Being. *Journal of Happiness Studies*, 5: 293–314.

Tinklepaugh, O. L. (1928). An Experimental Study of Representative Factors in Monkeys. *Journal of Comparative Psychology*, 8: 197–236.

Treggiari, S. (1993). *Roman Marriage*. Oxford: Oxford University Press.

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *Quarterly Review of Biology*, 46: 35–57.

Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge: Cambridge University Press.

Turnbull, C. (1972). *The Mountain People*. New York, NY: Simon and Schuster.

Turner II, C. G., and Turner, J. A. (1999). *Man Corn: Cannibalism and Violence in the Prehistoric American Southwest*. Salt Lake City, UT: University of Utah Press.

Tylor, E. B. (1871). *Primitive Culture*. New York, NY: Gordon Press.

UNICEF (2005). *Female Genital Mutilation/Cutting: A Statistical Exploration*. New York, NY: The United Nations Children's Fund.

Valdesolo, P., and DeSteno, D. (2006). Manipulations of Emotional Context Shape Moral Judgment. *Psychological Science*, 17: 476–7.

- Vartanian, O., and Goel, V. (2004). Neuroanatomical Correlates of Aesthetic Preference for Paintings. *Neuroreport*, 15: 893–7.
- Verdon, M. (1988). Virgins and Widows: European Kinship and Early Christianity. *Man*, 23: 488–505.
- Vranas, P. (2005). The Indeterminacy Paradox: Character Evaluations and Human Psychology. *Noûs*, 39: 1–42.
- Wagner, H. L. (2000). The Accessibility of the Term “Contempt” and the Meaning of the Unilateral Lip Curl. *Cognition & Emotion*, 14: 689–710.
- Walker, P. L. (2001). A Bioarchaeological Perspective on the History of Violence. *Annual Review of Anthropology*, 30: 573–96.
- Warneken, F., and Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, 311: 1301–3.
- Westermarck, E. (1891). *The History of Human Marriage*. London: Macmillan.
- _____. (1900). Remarks on the Predicates of Moral Judgment. *Mind*, 9: 184–204.
- _____. (1906). *The Origin and Development of the Moral Ideas*. London: Macmillan.
- Wheatley, T., and Haidt, J. (2005). Hypnotically Induced Disgust Makes Moral Judgments More Severe. *Psychological Science*, 16: 780–4.
- White, D. R., and Burton, M. L. (1988). Causes of Polygyny: Ecology, Economy, Kinship, and Warfare. *American Anthropologist*, 90: 871–87.
- Whitehouse, H. (2000). *Arguments and Icons: Divergent Modes of Religiosity*. Oxford: Oxford University Press.
- Widom, C. S. (1976). Interpersonal Conflict and Cooperation in Psychopaths. *Journal of Abnormal Psychology*, 85: 330–4.
- Wiedemann, T. (1992). *Emperors & Gladiators*. London: Routledge.
- Wiggins, D. (1987). A Sensible Subjectivism. In *Needs, Values, Truth: Essays in the Philosophy of Value*, pp. 185–214. Oxford: Blackwell.
- Wilkinson, G. S. (1984). Reciprocal Food Sharing in Vampire Bats. *Nature*, 308: 181–4.
- Williams, B. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

- Witter, R. A., Stock, W. A., Okum, M. A., and Harding, M. J. (1985). Religion and Subjective Well-Being in Adulthood: A Quantitative Synthesis. *Review of Religious Research*, 36: 332–42.
- Wolf, A. P. (1970). Childhood Association and Sexual Attraction: A Further Test of the Westermarck Hypothesis. *American Anthropologist*, 72: 503–15.
- Wolf, A. P., and Huang, C. S. (1980). *Marriage and Adoption in China, 1845–1945*. Stanford, CA: Stanford University Press.
- Woloch, N. (2000). *Women and the American Experience: A Concise History*, 2nd edn. New York, NY: McGraw Hill.
- Wong, D. B. (1984). *Moral Relativity*. Berkeley CA: University of California Press.
- Wood, J. C. (1999). When Men Are Women: Manhood among Gabra Nomads of East Africa. Madison, WI: University of Wisconsin Press.
- Wrangham, R. (2004). Killer Species. *Daedalus*, 133: 25–35.
- Wright, C. (1992) *Truth and Objectivity*. Oxford: Blackwell.
- Wyatt-Brown, B. (1982). *Southern Honor: Ethics and Behavior in the Old South*. New York, NY: Oxford University Press.
- Wynne, C. D. L. (2004). Fair Refusal by Capuchin Monkeys. *Nature*, 428: 140.
- Zahn-Waxler, C., and Robinson, J. (1995). Empathy and Guilt: Early Origins of Feelings of Responsibility. In J. P. Tangney and K. W. Fischer (eds.), *Self-Conscious Emotions*, pp. 143–73. New York, NY: Guilford.
- Zahn-Waxler, C., Hollenbeck, B., and Radke-Yarrow, M. (1984). The Origins of Empathy and Altruism. In M. W. Fox and L. D. Mickley (eds.), *Advances in Animal Welfare Science*, pp. 21–39. Washington D.C.: Humane Society of the United States.
- Zajonc, R. B. (1984). On the Primacy of Affect. *American Psychologist*, 39: 117–23.
- Zhong, C. B., and Liljenquist, K. (2006). Washing Away Your Sins: Threatened Morality and Physical Cleansing. *Science*, 313: 1451–2.

