Editora UFPel

O Lugar das Emoções na Ética e na Metaética

Flavio Williges
Marcelo Fischborn
David Copp
(Organizadores)

DISSERTATIO
FILOSOFIA

O LUGAR DAS EMOÇÕES NA ÉTICA E NA METAÉTICA

Série Dissertatio Filosofia

O Lugar das Emoções na Ética e na Metaética

Flavio Williges
Marcelo Fischborn
David Copp
(Organizadores)

DISSERTATIO
FILOSOFIA
Pelotas, 2018

Editora UFPel

REITORIA
Reitor: Pedro Rodrigues Curi Hallal
Vice-Reitor: Luís Isaías Centeno do Amaral
Chefe de Gabinete: Taís Ullrich Fonseca
Pró-Reitor de Graduação: Maria de Fátima Cóssio
Pró-Reitor de Pesquisa e Pós-Graduação: Flávio Fernando Demarco
Pró-Reitor de Extensão e Cultura: Francisca Ferreira Michelon
Pró-Reitor de Planejamento e Desenvolvimento: Otávio Martins Peres
Pró-Reitor Administrativo: Ricardo Hartlebem Peter
Pró-Reitor de Infraestrutura: Julio Carlos Balzano de Mattos
Pró-Reitor de Assuntos Estudantis: Mário Renato de Azevedo Jr.
Pró-Reitor de Gestão Pessoas: Sérgio Batista Christino

CONSELHO EDITORIAL DA EDITORA DA UFPEL
Presidente do Conselho Editorial: João Luis Pereira Ourique
Representantes das Ciências Agronômicas: Guilherme Albuquerque de Oliveira Cavalcanti
Representantes da Área das Ciências Exatas e da Terra: Adelir José Strieder
Representantes da Área das Ciências Biológicas: Marla Piumbini Rocha
Representante da Área das Engenharias e Computação: Darci Alberto Gatto
Representantes da Área das Ciências da Saúde: Claiton Leoneti Lencina
Representante da Área das Ciências Sociais Aplicadas: Célia Helena Castro Gonsales
Representante da Área das Ciências Humanas: Charles Pereira Pennaforte
Representantes da Área das Linguagens e Artes: Josias Pereira da Silva

EDITORA DA UFPEL
Chefia: João Luis Pereira Ourique (Editor-chefe)
Seção de Pré-produção: Isabel Cochrane (Administrativo)
Seção de Produção: Gustavo Andrade (Administrativo)
Anelise Heidrich (Revisão)
Ingrid Fabiola Gonçalves (Diagramação)
Seção de Pós-produção: Madelon Schimmelpfennig Lopes (Administrativo)
Morgana Riva (Assessoria)

NEPFIL online

CONSELHO EDITORIAL
Prof. Dr. João Hobuss (Editor-Chefe)
Prof. Dr. Juliano Santos do Carmo (Editor-Chefe)
Prof. Dr. Alexandre Meyer Luz (UFSC)
Prof. Dr. Rogério Saucedo (UFSM)
Prof. Dr. Renato Duarte Fonseca (UFSM)
Prof. Dr. Arturo Fatturi (UFFS)
Prof. Dr. Jonadas Techio (UFRGS)
Profa. Dra. Sofia Albornoz Stein (UNISINOS)
Prof. Dr. Alfredo Santiago Culleton (UNISINOS)
Prof. Dr. Roberto Hofmeister Pich (PUCRS)
Prof. Dr. Manoel Vasconcellos (UFPEL)
Prof. Dr. Marco Antônio Caron Ruffino (UNICAMP)
Prof. Dr. Evandro Barbosa (UFPEL)
Prof. Dr. Ramón del Castillo (UNED/Espanha)
Prof. Dr. Ricardo Navia (UDELAR/Uruguai)
Profa. Dra. Mónica Herrera Noguera (UDELAR/Uruguai)
Profa. Dra. Mirian Donat (UEL)
Prof. Dr. Giuseppe Lorini (UNICA/Itália)
Prof. Dr. Massimo Dell'Utri (UNISA/Itália)

COMISSÃO TÉCNICA (EDITORAÇÃO)
Prof. Dr. Lucas Duarte Silva (Diagramador)
Profa. Luana Francine Nyland (Assessoria)
Acad. Vinícius Berman (Webmaster)

DIREÇÃO DO IFISP
Prof. Dr. João Hobuss

CHEFE DO DEPARTAMENTO DE FILOSOFIA
Prof. Dr. Manoel Vasconcellos

Série Dissertatio Filosofia

A Série Dissertatio Filosofia, uma iniciativa do Núcleo de Ensino e Pesquisa em Filosofia (sob o selo editorial NEPFIL online) em parceira com a Editora da Universidade Federal de Pelotas, tem por objetivo precípuo a publicação de estudos filosóficos relevantes que possam contribuir para o desenvolvimento da Filosofia no Brasil nas mais diversas áreas de investigação. Todo o acervo é disponibilizado para download gratuitamente. Conheça alguns de nossos mais recentes lançamentos.

A Mediação Imperfeita em Paul Ricoeur
Adriane da Silva Machado Möbbs

Do Romantismo a Nietzsche: Rupturas e Transformações na Filosofia do Século IXI
Clademir Luís Araldi

Didática e o Ensino de Filosofia
Tatielle Souza da Silva

Michel Foucault: As Palavras e as Coisas
Kelin Valeirão e Sônia Schio (Orgs.)

Sobre Normatividade e Racionalidade Prática
Juliano do Carmo e João Hobuss (Orgs.)

A Companion to Naturalism
Juliano do Carmo (Organizador)

Ciência Empírica e Justificação
Rejane Xavier

A Filosofia Política na Idade Média
Sérgio Ricardo Strefling

Pensamento e Objeto: A Conexão entre Linguagem e Realidade
Breno Hax

Acesse o acervo completo em:
www.nepfil.ufpel.edu.br

© Série Dissertatio de Filosofia, 2018

Universidade Federal de Pelotas
Departamento de Filosofia
Núcleo de Ensino e Pesquisa em Filosofia
Editora da Universidade Federal de Pelotas

NEPFil online
Rua Alberto Rosa, 154 – CEP 96010-770 – Pelotas/RS

Os direitos autorais dos colaboradores estão de acordo com a Política Editorial do NEPFil online. A revisão ortográfica e gramatical foi realizada pelos autores.

Primeira publicação em 2018 por NEPFil online e Editora da UFPel.

Dados Internacionais de Catalogação

N123 O lugar das emoções na ética e na metaética.
[recurso eletrônico] Organização: Flavio Williges, Marcelo Fischborn, David Copp – Pelotas: NEPFIL Online, 2018.

435p. - (Série Dissertatio Filosofia).

Modo de acesso: Internet
<http://nepfil.ufpel.edu.br>
ISBN: 978-85-67332-53-6

1. Emoções. 2. Ética. 3. Metaética. 4. Filosofia. I. Williges, Flavio II. Fischborn, Marcelo III. Copp, David.

COD 170

creative commonsbr
BY NC ND

ABEC BRASIL
Associação Brasileira de Editores Científicos

Para maiores informações, por favor visite nosso site: www.nepfil.ufpel.edu.br

AGRADECIMENTOS

Este livro, bem como o colóquio do qual resulta, não teria sido possível sem a colaboração de inúmeras pessoas e instituições. Agradecemos primeiramente a cada um dos autores dos trabalhos aqui reunidos. E também aos professores Luiz Eva (UFPR), Marco Azevedo (UNISINOS), Flávia Carvalho Chagas (UFPEL), João Carlos Brum Torres (UCS/UFRGS), César Schirmer dos Santos (UFSM), Helder Carvalho (UFPI), Adriano Naves de Britto (UNISINOS), Denis Coitinho da Silveira (UNISINOS) e Cinara Nahra (UFRN), que, embora não tenham seus textos publicados aqui, participaram como conferencistas ou comentadores, dando um brilho especial ao evento. Agradecemos ainda aos estudantes Gabriel Garmendia, Mateus Stein, Mônica Dalcol, Rafael Chiminte e Waleska Cardoso que ajudaram na organização do colóquio, e também a todos que apresentaram comunicações e compareceram ao evento, colaborando com o seu sucesso. Agradecemos o apoio financeiro do CNPq, Fapergs, Capes, Pró-Reitoria de Pós-Graduação e Pesquisa da UFSM, Departamento e Programa de Pós-Graduação em Filosofia da UFSM e ao Centro de Ciências Sociais e Humanas.

Os organizadores

NOTA SOBRE OS AUTORES

JOHN COTTINGHAM é professor emérito de Filosofia da University of Reading. Atualmente possui uma *Professorial Research Fellow* no Heythrop College da University de Londres; é professor de Filosofia da Religião na University of Roehampton, de Londres, e professor visitante no King's College, Londres. É autor de mais de trinta livros e cento e trinta artigos em revistas e livros especializados. Website pessoal: http://www.johncottingham.co.uk/

JULIEN DEONNA é membro do Departamento de Filosofia da Universidade de Genebra. Autor de inúmeros artigos científicos e livros, trabalha especialmente com filosofia moral, filosofia da mente e em psicologia moral. Atualmente é diretor do grupo de pesquisa Thumos, associado à Universidade de Genebra. Visite o website do grupo: http://www.unige.ch/lettres/philo/recherche/research-groups/thumos/

FABRICE TERONI é membro do Departamento de Filosofia da Universidade de Genebra. Autor de inúmeros artigos científicos e livros, seu interesse é em filosofia da mente e epistemologia. Atualmente é diretor do grupo de pesquisa Thumos, associado a Universidade de Genebra. Visite o ebsite do grupo:
http://www.unige.ch/lettres/philo/recherche/research-groups/thumos/

CHRISTINE TAPPOLET é diretora do Centre de Recherche en Éthique de l'Université de Montréal (CRÉUM), do Groupe de Recherche Interuniversitaire sur la Normativité (GRIN) e presidente da Société de Philosophie Analytique (SoPhA). Autora de livros, capítulo de livros, e artigos em revistas especializadas, trabalha com ética, metaética e filosofia moral. Website: http://en.philo.umontreal.ca/repertoire-departement/vue/tappolet-christine/

ÉRICO ANDRADE é Doutor em filosofia pela Sorbonne (Paris IV). Atualmente é professor associado da Universidade Federal de Pernambuco. Tem experiência na área de filosofia moderna, com ênfase na epistemologia moderna. Há alguns anos tem se dedicado à ética e à política com enfoque em abordagens contemporâneas, sobretudo ligadas aos conceitos de empatia, jogos de linguagem moral, cuidado, identidade e vulnerabilidade.

MICHAEL SLOTE é professor de ética na University of Miami, EUA, e membro da Royal Irish Academy. Autor de inúmeros livros, entre eles *From Morality to Virtue* (1992) e *Morals From Motives* (2001). Seu interesse concentra-se em ética, psicologia moral e teoria racional da escolha.

CLÁUDIA PASSOS-FERREIRA é Doutora em Filosofia pelo Programa de Pós-Graduação em Filosofia da Universidade Federal do Rio de Janeiro (PPGF/ UFRJ). Fez Pós-Doutorado/Recém Doutor (FAPERJ) (2007-2009) no Programa de Pós-Graduação em Saúde Coletiva/ UERJ. Sua área de pesquisa atual é em temas de filosofia da mente, psicologia moral e psicologia da saúde coletiva.

JÔNADAS TECHIO é Doutor em Filosofia pela Universidade Federal do Rio Grande do Sul (UFRGS). Desenvolveu dois projetos de pesquisa pós-doutoral (UFRGS 2009-2011, bolsa CAPES/PRODOC, e Universidade de Chicago 2013-14, bolsa CAPES/Fulbright). Atualmente professor adjunto no Departamento de Filosofia da Universidade Federal do Rio Grande do Sul (UFRGS). Sua área de pesquisa atual é em temas de Metafísica, Epistemologia, Filosofia da Mente e da Linguagem, Filosofia Moderna e Contemporânea e Filosofia do Cinema.

DAVID COPP é professor emérito do Departamento de Filosofia aa Universidade da Califórnia, Davis, EUA. Foi professor visitante na Universidade de Stanford e na Universidade de Leiden, na Holanda. Autor de inúmeros livros, entre eles *Morality, Normativity, and Society* (1995) e *Morality in a Natural World* (2007), e de artigos em revistas especializadas. Sua área de pesquisa atual é em temas de filosofia moral, normatividade e teoria naturalista da moral.

ALCINO EDUARDO BONELLA é professor da Universidade Federal de Uberlândia (UFU) e pesquisador do CNPq. Possui mestrado (1995) e doutorado (2000) em filosofia pela Universidade de Campinas (UNICAMP). Fez estágios de pós-doutorado na Universidade Federal de Minas Gerais (UFMG, 2005) e na Universidade de Oxford (2005 e 2011). Sua área de pesquisa atual é em temas de metaética, ética normativa e ética aplicada (com ênfase em bioética).

LEONARDO DE MELLO RIBEIRO é Doutor em Filosofia pela University of Sheffield, Inglaterra. Atualmente é professor da Universidade Federal de Minas Gerais. Suas áreas de especialização são Metaética, Ética, Psicologia Moral, História da Filosofia Moral Britânica, Filosofia da Mente e a História da Filosofia Moderna (sobretudo David Hume).

WILSON MENDONÇA é Doutor em Filosofia pela Universidade de Konstanz, Alemanha. Atualmente professor titular de Filosofia na Universidade Federal do Rio de Janeiro. É Diretor do Centro de Ética e Filosofia da Mente, da Universidade Federal do Rio de Janiero. Sua área de pesquisa atual é em temas de filosofia da mente, filosofia moral e metaética.

IDIA LAURA FERREIRA é Doutora em Filosofia pela Universidade Federal do Rio de Janeiro. Atualmente é pesquisadora de pós-doutorado no Centro de Ética e Filosofia da Mente/IFCS-UFRJ com o projeto Racionalidade Prática e Ação Moral com apoio FAPERJ desde 2012. Sua área de pesquisa atual é em temas de moralidade, metaética, filosofia da mente e filosofia moral prática.

MARINA OSHANA é Doutora em Filosofia pela Universidade da Califórnia, Davis, EUA. Atualmente é professora emérita do Departamento de Filosofia da Universidade da Califórnia, Davis, EUA. Autora de inúmeros artigos científicos, sua área de pesquisa atual é em temas de ética e psicologia moral, em autonomia e responsabilidade moral. Website institucional: http://philosophy.ucdavis.edu/people/moshana

LÍVIA GUIMARÃES é Doutora em Filosofia pela Catholic University of America, EUA. Atualmente é professora Titular do Departamento de Filosofia da Universidade Federal de Minas Gerais. Autora de inúmeros artigos científicos, sua área de pesquisa atual é em temas de Ética, Psicologia Moral, Epistemologia e Filosofia da Religião, atuando principalmente nos seguintes temas: Hume, ceticismo, naturalismo, moralistas britânicos, psicologia moral, filosofia da religião, feminismo e ética animal.

CARLOS ADRIANO FERRAZ é Doutor em Filosofia pela Pontifícia Universidade Católica do Rio Grande do Sul. Atualmente é professor do Departamento de Filosofia da Universidade Federal de Pelotas, membro da American Philosophical Association (APA) e consultor ad hoc das agências Fulbright e Alexander von Humboldt-Stiftung/Foundation. Autor de

inúmeros artigos científicos em revistas especializadas, sua área de pesquisa atual é em temas de ética, filosofia do direito, contratualismo, direito natural e filosofia da religião.

ANDRÉ KLAUDAT é Doutor em Filosofia pela University College London, Inglaterra. Atualmente é Professor Associado da Universidade Federal do Rio Grande do Sul. Autor de inúmeros artigos científicos em revistas especializadas, sua área de pesquisa atual é em temas de ética e filosofia da mente e em autores como Hume e Kant.

FRANK THOMAS SAUTTER é Doutor em Filosofia pela Universidade Estadual de Campinas. Atualmente é professor titular da Universidade Federal de Santa Maria. Autor de inúmeros artigos científicos em revistas especializadas, sua pesquisa concentra-se em temas de filosofia moral e lógica.

Sumário

Agradecimentos IV

Nota sobre os autores V

Apresentação XI

Ricardo Bins di Napoli

1. As razões, as emoções e o bem viver 16

John Cottingham

2. Emoções 39

Julien Deonna, Fabrice Teroni e Christine Tappolet

3. A soberania dos afetos: uma crítica ao utilitarismo a partir da empatia e da psicanálise 84

Érico Andrade

4. A importância da fenomenologia: o que a amabilidade e a frieza humana podem nos ensinar sobre ética? 103

Michael Slote

5. Em defesa da empatia: uma resposta a Prinz 137

Cláudia Passos-Ferreira

6. O papel da indignação de uma perspectiva perfeccionista 177

Jônadas Techio

7. Ceticismo darwiniano sobre o realismo moral 198
David Copp
8. Moralidade, motivação e objetividade 231
Alcino Eduardo Bonella
9. Adequação sentimental, univocidade e recalcitrância 244
Leonardo de Mello Ribeiro
10. Sentimentos morais, internalismo motivacional, sentimentalismo 299
Wilson Mendonça
11. Elementos da metaética: Cognitivismo internalista e a teoria volitiva 312
Idia Laura Ferreira
12. Memória, autocompreensão e agência 325
Marina Oshana
13. Delicadeza e ternura: Hume e o feminino 349
Lívia Guimarães
14. Kant sobre o sentimento moral 369
Carlos Adriano Ferraz
15. Motivação e valor em Hume e Kant 397
André Klaudat
16. A filosofia moral de Paul Grice 416
Frank Thomas Sautter

APRESENTAÇÃO

É inegável que as emoções desempenham um papel decisivo em nossas vidas. Podemos notar isso se observarmos as relações que os seres humanos – e também alguns animais – mantêm entre si. Essas relações são amplamente mediadas por sentimentos e emoções, os quais funcionam como guias que, por um lado, aproximam as pessoas, tornam suas vidas e relações mais prazerosas e plenas de significado e, por outro, separam-nas e tornam suas vidas mais miseráveis e tristes. As emoções e sentimentos são capazes de aproximar indivíduos e fortalecer seus laços, como tem sido observado nos estudos em torno da empatia, compaixão, gratidão ou respeito, mas também podem produzir distanciamento, descrédito na própria humanidade, sobretudo quando nossas vidas são permeadas por diversas formas de medo ou por emoções como a raiva, a repulsa e o ódio.

Além de serem relevantes para a vida em geral, os sentimentos e as emoções são relevantes do ponto de vista moral, pois é por meio deles que podemos perceber ou intuir, por um lado, o dano ou o sofrimento causado a nós ou a outrem – visível através de emoções como a culpa e a indignação – ou, por outro, que as coisas deram certo – gerando em nós um sentimento de dever cumprido, de gratidão, de compaixão, de amizade, de solidariedade ou de amor. Por essas e outras características, os sentimentos atraíram a atenção de filósofos morais diversos ao longo da história da filosofia. Evidentemente, a racionalidade também parece ter um papel relevante na moralidade, pois às vezes buscamos decidir o que é o certo a fazer mediante uma reflexão distanciada, que considera as emoções como instâncias importantes, mas não decisivas, para a moralidade. Levando isso em conta, uma disputa clássica na história da filosofia se dá na oposição entre o racionalismo e o sentimentalismo. Essas duas correntes disputam qual seria o devido papel das emoções e sentimentos na deliberação moral. O racionalismo rejeita quase totalmente a importância dos sentimentos, argumentando que a coisa certa a fazer não está diretamente relacionada àquilo

que sentimos, mas à consideração imparcial e objetiva do benefício quantificável que resultará das ações ou a possibilidade de generalização das ações e escolhas para outras pessoas e contextos. Já o sentimentalismo busca afirmar que a razão desempenha um papel limitado na ação moral. David Hume, um dos adeptos dessa posição, e precursor do utilitarismo, certa vez escreveu: "Mas embora a razão, quando plenamente desenvolvida e cultivada, seja suficiente para nos fazer reconhecer a tendência útil ou nociva de atributos e ações, ela sozinha não basta para originar qualquer censura ou aprovação moral"[1].Para ele, eram os nossos sentimentos (subjetivos) de repulsa ou louvor que nos faziam distinguir entre as ações virtuosas e viciosas. A abordagem de Hume teve origem na tradição do *moral sense* do empirismo inglês, mas antes dele autores como Aristóteles e Agostinho haviam dado atenção às emoções e, mais contemporaneamente, novas correntes de pensamento filosófico e científico têm evidenciado o papel dos sentimentos e emoções na vida moral, como, por exemplo, as pesquisas sobre as decisões morais difíceis – os famosos *trolley problems* – que apresentam resultados inteiramente distintos dependendo de se as decisões ativam setores emocionais do cérebro ou setores ligados ao raciocínio e reflexão.

Os racionalistas, por sua vez, encontraram apoio em certa linha interpretativa que vai de Sócrates e Platão a Immanuel Kant, mas que por um certo viés poderia incluir também os utilitaristas Jeremy Bentham e John Stuart Mill. Há também filósofos como Iris Murdoch, A. McIntyre, M. Sandel, que, de outro modo, tendem a afastar-se de um racionalismo estrito, seja articulando uma visão moral centrada em conceitos metafísicos como o conceito do bem, seja enfatizando o vinculo entre moralidade e sociedade. McIntyre e Sandel, em particular, esboçaram concepções vinculadas à identidade pessoal como narrativa. Para eles, a pessoa e suas decisões deveriam ser sempre consideradas de um ponto de vista mais amplo do que aquele que apresenta os processos de escolha moral como guiados por princípios abstratos, o qual incluiria também uma narrativa pessoal do agente

[1] HUME, D. *Uma investigação sobre os princípios da Moral*. Trad. José Oscar de A. Marques. Campinas: EdUnicamp, 1995, p.175.

inserido socialmente. As obrigações dos indivíduos seriam, assim, geradas por seus vínculos sociais.

Os capítulos que compõem este livro são oriundos principalmente do *V Colóquio Internacional sobre Ética e Ética Aplicada: Sentimentos Morais*, realizado na Universidade Federal de Santa Maria, RS, em junho de 2012. As considerações sobre a razão e a emoção apresentadas acima dão uma ideia do contexto de discussão em que essas contribuições se inserem. Sua reunião neste volume representa um feito inédito, associado a uma certa guinada que vem acontecendo na pesquisa em Ética no Brasil, pois há muito tempo carecia-se de um tratamento mais abrangente dos sentimentos morais no contexto da filosofia brasileira – o que em outros países já vinha acontecendo desde os anos 70 e 80, principalmente no âmbito da filosofia anglo-saxônica. Sem uma divisão rígida, os capítulos estão agrupados iniciando pelos mais voltados a questões éticas, passando aos mais focados em questões metaéticas e finalizando com os que dão atenção ao pensamento sobre emoções e razão em momentos específicos da história da filosofia moral.

No capítulo 1, "As razões, as emoções e o bem viver", John Cottingham, argumenta que temos de abandonar a dicotomia entre razão e sentimentos para alcançar uma concepção satisfatória do bem viver. De acordo com ele, o bem viver pode admitir dimensões variadas, mas que giram todas em torno da noção de integração. O Capítulo 2, "Emoções", escrito originalmente como um verbete por Julien Deonna, Fabrice Teroni e Christine Tappolet, oferece uma introdução ao estudo das emoções e apresenta os conceitos e problemas mais recentes sobre a natureza e a relevância moral das emoções. No capítulo 3, "A soberania dos afetos: Uma crítica ao utilitarismo a partir da empatia e da psicanálise", Érico Andrade critica o critério moral utilitarista que diz que podemos descobrir o que é o certo a se fazer com base em um cálculo que visa minimizar o sofrimento. Em vez disso, defende o autor, motivações afetivas, especialmente motivações empáticas, e outros elementos contextuais também precisam espaço na teoria ética. O capítulo 4, "A importância da fenomenologia: O que a amabilidade e a frieza humana podem nos ensinar sobre ética?", de Michael Slote, busca mostrar o papel

que a fenomenologia, entendida em sentido amplo, pode desempenhar na ética, na epistemologia e na filosofia da mente. No caso particular da moralidade, Slote defende que atentar à experiência da moralidade nos permite compreender os elementos emocionais envolvidos na moralidade e que seriam deixados de fora de uma perspectiva meramente racionalista. No capítulo 5, "Em defesa da empatia: Uma resposta a Prinz", Claudia Passos-Ferreira argumenta, partindo principalmente de considerações da psicologia moral, que a empatia desempenha um papel central no desenvolvimento da agência moral, e que é, além disso, uma emoção moral desejável. Sua proposta é oferecida em grande medida como um contraponto à abordagem de Jesse Prinz. No capítulo 6, "O papel da indignação de uma perspectiva perfeccionista", Jônadas Techio investiga a finalidade da discussão moral, por meio da delimitação de situações em que pode ser moral e racionalmente adequado expressar indignação. A questão importa, sugere o autor, porque, dependendo da concepção adotada, a expressão de indignação poderia representar justamente uma recusa irracional à discussão moral.

A seção sobre metaética inicia-se com o capítulo 7, "Ceticismo darwiniano sobre o realismo moral", no qual David Copp argumenta que a hipótese darwiniana, segundo a qual o fenômeno da psicologia moral foi fortemente influenciado pelo processo de evolução da espécie humana, é compatível com o realismo moral, isto é, com a tese de que há fatos morais. No capítulo 8, "Moralidade, motivação e objetividade", Alcino Bonella oferece uma resposta a duas críticas comuns à concepção expressivista sobre juízos morais, defendendo que essa concepção é compatível com a objetividade moral e que também descreve adequadamente o juízo moral. O capítulo 9, de Leonardo Ribeiro, intitula-se "Adequação sentimental, univocidade e recalcitrância". O autor argumenta que uma tese neossentimentalista sobre o valor é mais bem-sucedida do que uma tese judicativista para tratar de dois grupos de problemas, a saber, o problema da univocidade e o problema da recalcitrância. No capítulo 10, "Sentimentos morais, internalismo motivacional, sentimentalismo", Wilson Mendonça analisa as relações entre certas teses sobre o lugar dos sentimentos na teoria ética e argumenta que a tese motivacional – segundo a qual sentimentos garantem a ligação entre

convicções morais e ações – não implica a tese sentimentalista mais robusta que diz que sentimentos morais são constitutivos de valores ou juízos morais. A conclusão proposta é que cada uma dessas teses precisa de apoio independente. O capítulo 11, "Elementos da metaética: Cognitivismo internalista e a teoria volitiva", de Idia Laura Ferreira, explora o papel a ser desempenhado por teorias volitivas da ação, que enfatizam noções como intenção e vontade, sem reduzi-las a noções como a crença e o desejo, na compreensão da agência moral. E no capítulo 12, "Memória, autocompreensão e agência", Marina Oshana investiga o papel da memória no autocontrole sobre as próprias ações. Particularmente, ela argumenta, certo conjunto de memórias do agente sobre si mesmo, que permitem que se reconheça ao longo do tempo e que embasam sua psicologia moral, são cruciais para sua agência prática.

O capítulo 13, "Delicadeza e ternura: Hume e o feminino", de Lívia Guimarães, destaca emoções e outros elementos de caráter feminista na obra e vida de David Hume, os quais se articulam especialmente em torno da noção de 'ternura' – uma excelência feminina, como propõe a autora. No capítulo 14, "Kant e o sentimento moral", Carlos Adriano Ferraz esclarece o papel do sentimento moral na filosofia prática kantiana, o qual, segundo ele, cumpre o papel de mostrar como a razão pode ser prática, isto é, conduzir efetivamente à ação. O capítulo 15, de André Klaudat, intitula-se "Motivação e valor em Hume e Kant". O autor analisa e compara o papel da motivação na filosofia moral de Hume e Kant, sendo que para ambos essa noção é central. Por fim, o capítulo 16, de Frank Thomas Sautter, "A filosofia moral de Paul Grice", encerra o livro analisando alguns elementos da filosofia moral de Paul Grice, incluindo as condições para uma lógica do âmbito prático, o tratamento da noção aristotélica de *eudaimonia* e a proposta de uma metafísica do valor.

Ricardo Bins di Napoli
Universidade Federal de Santa Maria

1

AS RAZÕES, AS EMOÇÕES E O BEM VIVER*

JOHN COTTINGHAM

Introdução

A filosofia pretendeu, tradicionalmente, oferecer uma caracterização do que é o bem viver para os seres humanos. Eustáquio, um escolástico contemporâneo de Descartes, observou que "o objetivo de uma filosofia completa é a felicidade humana"[1]. E o próprio Descartes, fiel a essa tradição, viu a 'árvore' da sua filosofia dando como fruto para a humanidade o que ele chamou de um "sistema moral mais perfeito"[2]; o qual foi concebido como um desenvolvimento orgânico de todo seu projeto filosófico, que nasceu de suas descobertas científicas em fisiologia e psicologia, embora estivesse também profundamente enraizado em uma metafísica que via toda a realidade como o reflexo "da luz imensa" da verdade objetiva e da bondade[3]. Em seu projeto, o sistema moral representava o "último

---

*Traduzido por Flavio Williges.

[1] "*Universae philosophiae finis est humana foelicitas*". Eustachius e Sancto Paulo, *Summa Philosophiae Quadripartita* [1609], Prefácio da Parte II.

[2] "[Por moral] entendo a mais elevada e mais perfeita ciência, porque pressupõe um conhecimento integral das outras ciências. Ora, como não é das raízes e nem do tronco das árvores que se colhe os frutos, mas apenas das extremidades dos ramos, a principal utilidade da Filosofia depende daquelas suas partes que são aprendidas em último lugar" René Descartes, Prefácio à tradução francesa de 1647 dos *Principles of philosophy*, AT IXB 14-15: CSM I 186. Nesse texto, 'AT' refere-se à edição franco-latina padrão das obras de Descartes por C. Adam & P. Tannery, *Œuvres de Descartes* (12 vols, revisada e editada, Paris: Vrin/CNRS, 1964-76); 'CSM' refere-se à tradução inglesa por J. Cottingham, R. Stoothoff and D. Murdoch, *The Philosophical writings of Descartes*, vols I and II (Cambridge: Cambridge University Press, 1985), e 'CSMK' ao vol. III, *The correspondence*, pelos mesmos tradutores e mais A. Kenny (Cambridge: Cambridge University Press, l991).

[3] "*Placet hic aliquamdiu in ipsius Dei contemplatione immorari* […] *et immensi hujus luminis pulchritudinem* […] *intueri, admirari, adorare*". ("parece-me muito a propósito deter-me algum tempo

grau da sabedoria", nos ofertando uma compreensão profunda de nossa natureza como seres encarnados e nos habilitando a viver vidas felizes.

Hoje, em nossa cultura moderna, instrumental e administrativa, tais aspirações elevadas desapareceram quase completamente ou foram substituídas por tentativas pseudo-científicas de mensurar e promover o bem-estar. A abordagem avidamente abraçada por agentes públicos é curiosamente atomista: consumidores são convidados a preencher declarações em pesquisas respondendo se eles são "muito felizes", "um pouco felizes", "neutros", "um pouco infelizes", "muito infelizes" em algum dia ou hora específica ou, ainda, indicar quão satisfeitos eles estão com algum conjunto específico de bens e serviços. Esse é o legado do "cálculo felicífico" (*felicific calculus*) de Bentham, uma concepção mecânica e degradante do bem para os seres humanos, cujo único mérito (e que, pensando bem, é um mérito duvidoso) é pretender oferecer critérios empíricos e quantificáveis para alcançar o sucesso.

Poder-se-ia pensar que o que está errado nessas abordagens tecnológicas é que elas colocam muita ênfase no fator do "bem-estar" ou, em outros termos, que operam num nível cru, voltado à mensuração da satisfação emocional, sem sujeitar as emoções à avaliação racional a partir de um padrão filosófico acerca do que é o bem viver. Mas as coisas não são tão simples assim: existem também problemas em deixar a razão filosófica sozinha determinar as condições da boa vida. Já com Platão (em alguns de seus escritos)[4] e claramente entre seus sucessores helênicos, vemos os perigos da ética 'centrada na razão' que pretende alcançar a supressão ou subordinação das emoções. Antes do engajar-se na condição humana real, em seu envolvimento confuso e turbulento com o corpo e

---

na contemplação deste Deus todo perfeito, ponderar totalmente à vontade seus maravilhosos atributos, considerar, admirar e adorar a incomparável beleza dessa imensa luz"). Descartes, *Terceira Meditação*, AT VII 52: CSM II 36. Cf. edição portuguesa: Descartes, R. *Meditações*. São Paulo: Abril Cultural, p.112.

[4] Por exemplo, a *República* [*c.* 375 BC]; ver, abaixo, a seção 6. Há, claro, outros direcionamentos na abordagem ética de Platão que dão outros resultados: ver, por exemplo, o Fedro [*c.* 370 BC], e a esclarecedora discussão dele por Martha Nussbaum, *The fragility of goodness* (Cambridge: Cambridge University Press, 1986), Ch. 7.

as paixões, essa ética centrada na razão retrocede à calma isolada e à suposta autossuficiência dos "templos altos e serenos, fortificados pelas doutrinas dos sábios"[5].

Nós precisamos de uma concepção rica da vida que vale a pena, uma concepção que nem reduza a mesma à soma (total) dos bons sentimentos ou episódios de satisfação, por um lado, e nem, por outro lado, aos privilégios da razão que tenta tornar a vida humana, em toda sua riqueza e profundidade recalcitrante, conformada a um esquema intelectualmente pré-determinado. O que poderia ser tal concepção da vida boa, do bem viver? Obviamente eu não tenho esperança de articular a sua estrutura nesse pequeno texto, mas quero sugerir certas dimensões importantes que precisariam ser acomodadas em tal concepção. Como ficará claro, essas dimensões são inteiramente interdependentes, assim como as dimensões de um sólido geométrico estão intimamente inter-relacionadas – você não pode alterar uma propriedade (por exemplo, a forma e estabilidade) sem mudar radicalmente as propriedades do todo. A reflexão acerca dessas dimensões mostra que todas giram em torno da noção de *integração:* o caráter ético e psicodinâmico, o qual será o foco de boa parte desse texto. Pretendo extrair algumas lições acerca das respectivas funções das emoções e da razão para o bem viver e argumentarei que é somente desconstruindo ou transcendendo a dicotomia padrão entre razão e emoções que poderemos nos mover em direção a um entendimento próprio da natureza da plenitude humana.

1. Algumas dimensões básicas do bem viver

Para começar, então, quais são as principais dimensões que qualquer descrição filosófica plausível da vida boa deve incorporar, caso não pretenda cair num empirismo cru da mensuração-da-satisfação, por um lado, ou no modelo

---

[5] "*Sed nil dulcius est, bene quam munita tenere*/*Edita doctrina sapientum templa serena*". Lucretius, *De rerum natura*, [*c.* 40BC], II, 7-8. Embora a abordagem epicurista exposta por Lucrécio não seja inteiramente típica do 'raciocentrismo' que tenho em mente, é interessante ver o poder exercido aqui pelo ideal do distanciamento racional; em muitos aspectos a citação poderia igualmente ter vindo de um escritor estóico. Cf. a tradução portuguesa de Lucrécio. *Da natureza*. São Paulo: Abril Cultural, 1980, p.47.

super austero do comando e controle racional, por outro? A primeira dimensão pode ser denominada, por questão de conveniência, como a dimensão *holística*. Sua importância foi parcialmente reconhecida já por Aristóteles, que insistiu que a *eudaimonia*, a satisfação humana ou felicidade, tem que ser medida *em relação à vida toda* (*over a whole lifetime*)[6]. O que ele explicitamente parece ter tido em mente é que a vida feliz pode ser destruída, perto do seu fim, por algum desastre não previsto (como no caso de Príamo)[7]. Mas eu penso que há mais a respeito dessa ideia do que uma advertência cautelar da cultura grega antiga ("não considere nenhum homem feliz antes de conhecer sua morte!")[8]. A caracterização subjacente da felicidade de Aristóteles claramente explora analogias entre o bem para o homem e o bem para outras espécies biológicas[9]. Assim, para uma planta – por exemplo, uma oliveira – a felicidade é crescer de forma lenta e estável em direção a um estado final. Assim, o homem bom, poderia ser dito, (tomando de empréstimo um texto muito diferente, mas que curiosamente captura algo do espírito de Aristóteles) é "como uma árvore plantada pelos cursos d'água, que dá seus frutos na estação e cujas folhas não caem – tudo o que faz prospera"[10]. Há um ritmo, um crescimento em direção à perfeição ou *entelechia*[11]; e isso acontece com o passar do tempo, através da completude da vida do organismo.

Isso não significa simplesmente que precisamos fazer uma soma das várias atividades satisfatórias abarcadas em setenta anos ou em algo assim como o período de duração de uma vida humana típica. Isso nem mesmo significa que nós precisamos olhar em que medida as várias atividades oferecem oportunidade para o exercício de capacidades e faculdades especificamente humanas. Aristóteles

---

[6] Ἐν βίῳ τελείῳ (en biô teleiô); Aristoteles, *Nicomachean ethics* [325 BC], Bk I, Ch. 7 (1098a18).
[7] Ver Aristoteles, *Nicomachean ethics*, Bk I, Ch. 10 (1101a8).
[8] Ésquilo, *Agamemnon* [458 BC], 939.
[9] Cf. *Nicomachean Ethics*, Bk I, Ch.7.
[10] Salmos [*c.* 600 BC], 1:3.
[11] O termo tem muitos sentidos na história da filosofia; alguns deles derivam de mal-entendidos ou distorções do pensamento de Aristóteles. Eu uso o termo aqui no primeiro sentido identificado pelo *Oxford English dictionary* (versão online) como sendo o uso primário de Aristóteles: "A realização ou expressão completa de alguma função; a condição em que uma potencialidade tornou-se uma realidade".

deixa claro que há mais coisas envolvidas numa vida feliz do que uma mera soma de atividades valiosas separadas. É necessário que haja um padrão teleológico geral[12], o que significa que sua concepção da boa vida é *inerentemente* holista. A vida de virtude não é propriamente um agregado de excelências, mas uma vida em que todas as excelências ajustam-se harmoniosamente num todo: Aristóteles insistiu famosamente que a posse de uma virtude implica a posse de todas[13]. Essa alegação aparentemente paradoxal sobre as virtudes atuando conjuntamente tem sido debatida e dissecada infindavelmente em análises críticas. Mas o ponto subjacente é, creio, bem simples: o bem viver, para Aristóteles, tem uma *unidade orgânica*. Há algo que o ser humano pretende ser – um organismo unificado, feliz, que desenvolveu suas excelências características no curso da vida[14].

Desta concepção evidentemente se segue que a boa vida tem que ser uma vida *integral.* Em contraste com a concepção de Platão na *República* (que posteriormente exerceu uma forte influência nos estóicos) – uma concepção em que as emoções são controladas e subordinadas à razão – Aristóteles deixa claro que para o ser humano virtuoso e feliz a razão e as emoções estão em harmonia. O agente virtuoso não precisa exercitar um expediente de autocontrole de "segunda mão"[15], uma vez que, para ele, as emoções já estão desenvolvidas pela boa educação e treinamento, de modo que elas concordam espontaneamente com a percepção racional do bem. A exigência de integração conecta-se com a concepção da felicidade na 'vida-inteira', pois a rede da razão e emoção não opera apenas sincronicamente, em uma dada escolha num dado momento ou num determinado dia, mas diacronicamente, ou seja, sobre a trajetória completa da vida da pessoa virtuosa. Cada vez que uma ação virtuosa é realizada, a satisfação

---

[12] Não planejar sua vida em direção a algum fim é, diz Aristóteles, "um sinal de grande tolice"; *Eudemian ethics* [*c.* 325 BC], 1214b10-11.

[13] *Nicomachean ethics,* Book VI, Ch. 13 (1145a1-2).

[14] Para um tratamento mais abrangente dessa idéia e que ultrapassa a estrutura fornecida pelo pensamento de Aristóteles, ver J. Cottingham, "Happiness, temporality, meaning". In: L. Bortolotti (ed.), *Philosophy and happiness* (London: Palgrave, 2009), p.21-36.

[15] Ver J. Cottingham, *Philosophy and the good life: Reason and the passions in greek, cartesian and psychoanalytic ethics* (Cambridge: Cambridge University Press, 1998), Capítulo 2, §2.

resultante reforça a contribuição que esta ação dá à felicidade. Observe que a satisfação advém aqui não do próprio fim ou objetivo, mas de uma espécie de selo instado pela virtude[16], um sinal que tudo está funcionando como deveria em direção ao bem do todo. E a cada vez que isso ocorre, a disposição do agente virtuoso para espontaneamente extrair prazer ao fazer as escolhas certas é aprofundada e reforçada. Com o desabrochar da vida florescem também as virtudes ou excelências que nos habilitam à felicidade e a razão e as emoções operam numa harmonia mais íntima e vibrante.

Este modelo orgânico complexo tem sido reduzido por alguns intérpretes modernos da assim chamada "ética das virtudes" a uma caricatura na forma de um suposto teste para a ação certa ou boa: "um ato é bom se e somente se é a espécie de ato que uma pessoa virtuosa realizaria"[17]. Essa formula estéril, que imita a espécie de coisa que se encontra no cálculo utilitarista, ignora tudo o que é mais importante sobre a ética aristotélica, fazendo-a parecer como se fosse uma teoria concebida para avaliar uma ação separada do contexto do desenvolvimento da vida do agente – e como resultado rouba dela tudo o que dá seu sentido ético. O que eu creio que Aristóteles está apontando, embora ele não tenha um termo explicito para isso, é para a integridade como uma ideia central para o bem viver.

A noção de integridade tem um lugar central no contexto (bem diferente) da ética judaico-cristã. Na bíblia hebraica, em um dos últimos Salmos (86) encontramos o salmista dizendo "Dê-me, oh Deus, um coração não dividido!" – uma petição por uma unidade ética e psicológica[18]. E os evangelhos cristãos falam da importância de encontrar o nosso *eu verdadeiro.* Ganhar o mundo inteiro não é suficiente para compensar a perda de si mesmo (*heautos*), diz uma passagem famosa em São Lucas (9: 25). Uns poucos capítulos depois no mesmo Evangelho,

---

[16] O prazer aperfeiçoa a atividade "como um tipo de perfeição subsequente, como a força que embeleza as flores da juventude"; *Nicomachean ethics*, Bk X, Ch. 4, 1174b33.

[17] Para formulações desse tipo, confira a investigação instrutiva de Justin OAKLEY "A virtue-ethics approach". In: H. Kuhse e P. Singer (eds), *A companion to bioethics* (Oxford: Blackwell, 1998), Cap. 10.

[18] Salmo 86 [85]:11. Em hebraico o salmista pede a Deus *yahed levavi*, que significa, literalmente, "une meu coração!" (o verbo imperativo *yahed* deriva da raíz *ehad*, que significa um).

encontramos a parábola do filho pródigo, que sai em exílio para esbanjar sua herança, mas um dia acorda e "volta a si mesmo" (*eis heauton elthôn*) (Lucas 15:17). Como o escritor dominicano Timothy Radcliffe colocou, a decisão do filho pródigo de voltar para sua casa e família é realmente a mesma de redescobrir seu eu verdadeiro, "uma vez que o exílio de sua família é um exílio de sua verdadeira identidade como filho e irmão. Ele só pode encontrar-se a si mesmo novamente através deles"[19].

A ideia aqui é que eu tenho uma 'identidade verdadeira', um eu integrado, unificado, o eu que pretendo ser, o eu que expressa tudo o que é melhor e mais distintivo sobre mim mesmo; e ainda, que o objetivo de minha vida seria desenvolver-me na direção desse eu unificado. Isso pode parecer para algumas pessoas como uma concepção idealizada e exaltada do bem viver. Mas eu creio que essa é a direção que nosso pensamento deve tomar se começarmos a refletir seriamente sobre o que está envolvido no conceito de integridade[20]. Integridade, como a etimologia sugere, tem relação com totalidade – envolve a integração plena de todo o eu.

2. Integridade

Mas o que a integridade realmente implica em termos éticos e psicológicos, e como podemos interpretá-la melhor hoje? Nossa paisagem cultural contemporânea, como Alasdair MacIntyre argumentou, parece ter perdido de vista esse valor inteiramente, substituindo-o por uma visão fragmentada da vida humana, que:

> […] divide a vida humana numa série de segmentos, cada um com suas próprias normas e modalidades de comportamento. Portanto, o trabalho é separado do lazer, a vida privada da vida pública, o institucional do pessoal. Assim, a infância e a velhice foram amputadas do resto da vida humana e transformadas em dois setores distintos. E todas essas separações foram criadas de tal forma que é a

[19] Timothy RADCLIFFE. *Why go to church: The drama of the eucharist*. London: Continuum, 2008, p.20.

[20] Os temas desenvolvidos nessa seção foram discutidos em J. COTTINGHAM. "Integrity and fragmentation". In: *Journal of Applied Philosophy* 27.1, p.2-14, 2010.

> peculiaridade de cada uma delas, e não a unidade da vida do indivíduo que se experimenta nessas partes, em cujos termos somos ensinados a pensar e sentir[21].

A fragmentação é perigosa em muitos aspectos. Primeiramente, ela corrói a *significação,* que é um ingrediente essencial da verdadeira felicidade humana. Para entender o sentido do que eu estou fazendo, preciso ter algum sentido de como minha vida precedente, desde os primeiros anos da infância, estruturou e influenciou as escolhas e objetivos que tenho agora. Supor que eu sou um ser inteiramente autônomo e independente, isolado do processo que me formou e moldou não é apenas uma perigosa demonstração de arrogância; essa suposição oculta o verdadeiro significado da minha vida, de mim mesmo, o agente supostamente soberano. E o que aplica-se ao passado também aplica-se para o futuro que ainda não se realizou. Como Charles Taylor argumentou eloquentemente, para dar sentido às nossas vidas e realmente ter uma identidade tudo que "precisamos é de uma orientação para o bem"; precisamos ter algum sentido de nossas vidas se movendo em direção ao crescimento moral e à maturidade. Segue-se disso que nossas vidas têm uma estrutura narrativa: como me desenvolvo e aprendo a partir das minhas falhas e erros; há sempre uma história para ser contada sobre como me tornei o que eu sou agora e onde minha jornada atual em relação ao aperfeiçoamento me levará. Do mesmo modo que meu sentido de onde estou no espaço físico depende de como cheguei aqui e para onde irei, assim também ocorre, Taylor argumenta, com "minha orientação no espaço moral"[22].

O segundo problema com a fragmentação, com alguém vivendo a vida em segmentos divididos, é o risco dessa ser uma estratégia psicologicamente instável: pode funcionar por um tempo, mas cedo ou tarde tensões e conflitos surgidos dos elementos desintegrados da psique irão solapar o rumo da vida com resultados potencialmente desastrosos. Os *insights* centrais de Freud e Jung (sem considerar

---

[21] Alasdair MACINTYRE, *After virtue*. London: Duckworth, 1985, cap.15, p.205; a pontuação foi modificada. Cf. MACINTYRE, A. *Depois da virtude*. Tradução de Jussara Simões. Bauru: EDUSC, 2001, p.343.

[22] Charles TAYLOR, *Sources of the self: The making of modern identity*. Cambridge: Cambridge University Press, 1989, p.46-52. Ver TAYLOR, C. *As fontes do self*. São Paulo: Loyola, 1998.

se aceitamos os detalhes específicos de suas frequentemente controversas teorias psicoterapêuticas) convergem para uma alegação ética e psicológica simples: é apenas por estar preparado para sondar os sentimentos e impulsos parcialmente submersos por trás da superfície de nossas deliberações racionais que podemos começar a descobrir o arriscado e ameaçador encanto que certos dos nossos desejos podem ter para nós – e continuarão a ter, muito embora possamos saber intelectualmente que sua importância é duvidosa e enganadora. A integridade não é uma questão de suprimir ou cancelar esses desejos teimosos pela força principal, pois esse é um procedimento altamente perigoso. Quando Freud falou dos riscos da repressão, ele estava simplesmente seguindo o rastro de uma longa série de insights, como aqueles revelados diretamente por Eurípides nas *Bacantes* (*c.* 406 BC) sobre como negamos ou suprimimos os impulsos psíquicos mais obscuros por seu risco. Em vez de tentar suprimir os impulsos ou eliminar tais elementos irracionais, os desejos, em nossa composição, a integridade é uma questão de estar preparado para reconhecê-los, para trazê-los com esforço doloroso para a superfície e tentar entendê-los. Como explica Carl Jung:

> O objetivo psicanalítico é observar as apresentações sombrias – seja na forma de imagens ou de sentimentos – que estão espontaneamente envolvidas na psique e aparecem em desordem ao homem que olha para dentro. Desse modo encontramos uma vez mais o que estava reprimido ou esquecido. Embora seja realmente doloroso, isso é, em si mesmo, um ganho, pois o que é inferior ou mesmo irrelevante pertence a mim como minha sombra e me dá substância e massa. Como posso ser substancial se eu não consigo abandonar uma sombra? Eu devo ter um lado escuro se sou pleno e, à medida que torno-me consciente de minha própria sombra, eu também lembro que sou um ser humano como qualquer outro[23].

Isso, no final das contas, é o que Próspero faz na *Tempestade* de Shakespeare, quando finalmente enfrenta o vil Calibã, o qual o tem forçado a uma

[23] Carl G. JUNG, "Problems of modern psychotherapy" ["Die Probleme der modernen Pychotherapie", 1929]. In: *Modern man in search of a soul. Essays from the 1920s and 1930s*. Tradução de C. F. Baynes (London: Routledge, l933), p.40.

submissão relutante durante a peça, e, por fim, o alcança com as palavras "Reconheço em mim essa escuridão" (*This thing of darkness I acknowledge mine*)[24].

As complexidades da psique humana, o opaco e intensamente problemático caráter de nossas motivações profundas, indicam que o significado real dos próprios bens pelos quais lutamos, os próprios planos que construímos, são frequentemente obscurecidos até mesmo pelos nossos próprios esforços e construtos. A filosofia moral contemporânea padrão tipicamente opera com uma imagem de um agente inteiramente autônomo, racional, com completo controle dos seus planos e da deliberação que determina o seu eu ou a sua trajetória de vida, os objetivos da carreira escolhida, as ambições e todo o resto. Mas os *insights* revelados pela perspectiva psicanalítica do desenvolvimento humano sugerem que mesmo para o adulto aparentemente mais racional, os eventos da infância, não inteiramente compreendidos ou assimilados, podem estar adormecidos até que eles operem nas escolhas aparentemente racionais da vida adulta. Assim a integração do eu não é exatamente um opcional extra para o bem viver. O desafio da autoconsciência é um *sine qua non* para uma existência humana valiosa, uma vez que, antes de tudo, sem esta não pode haver nenhuma confiança na ação humana racional [25].

3. Razão versus emoções

O ponto de vista que alcançamos até aqui é que qualquer teoria plausível do bem viver para os seres humanos precisa reconhecer primeiro que nós não somos meros feixes de emoções e desejos, mas criaturas que se desenvolvem e amadurecem com o tempo; e, em segundo lugar, que tal maturidade não é exatamente uma questão de permitir que certos desejos sejam desfrutados, mas requer de nós o engajamento na disciplina propriamente humana da autodescoberta – a luta em direção à espécie de autoconsciência que nos permitirá entender o significado de nossos sentimentos e como eles contribuem

[24] William SHAKESPEARE, *The Tempest* [c.1610], Act V, Scene 1.

[25] Para um tratamento mais extenso de alguns desses temas, ver J. Cottingham, *Philosophy and the good life*, Cap. 4.

para ou se afastam da felicidade humana.

Como esses resultados se relacionam com o longo debate em filosofia moral sobre as funções próprias da razão e das emoções para o bem viver? Está claro, eu creio, que eles lançam uma grande dúvida na ideia humeana de que a razão "é e deve ser escrava das paixões e não deve se ocupar com nada além de servir e obedecer a ela"[26]. Na concepção de Hume, a função da razão, por si mesma sem poder e inerte, é limitada à função puramente instrumental de planejar como os objetivos dados pelos sentimentos ou *feelings* serão implementados. O próprio Hume mostrou que esta posição é em si mesma instável, dado que (mesmo sem invocar todos os problemas sobre as emoções que são transparentes na perspectiva psicanalítica) resta claro para o senso comum que algumas de nossas emoções são perigosas e danosas. Quais paixões devem ser seguidas? O apelo caracteristicamente humeano à natureza humana, como supostamente capaz de preencher a lacuna deixada pela razão, dificilmente ajudará aqui, dado que, como o próprio Hume admite, nossa natureza é inextricavelmente uma mistura do bom e do mau: "há alguma benevolência, por pequena que seja, infundida em nosso coração, algum lampejo de afeição pelo gênero humano, alguma parcela de pomba entrelaçada, em nossa constituição, a elementos de lobo e de serpente"[27]. Assim, uma teoria moral fundada no sentimento ou *feeling* sozinha se despedaça na rocha de nossa natureza humana ambígua e conflitiva. Joseph Butler antecipou, no começo do século XVII, essa dificuldade com grande acuidade quando notou que embora possamos ter disposições naturais de gentileza e compaixão,

> uma vez que outras paixões [tal como a raiva] […] que nos desencaminham […], estão em si mesmas num grau igualmente natural e frequentemente mais prevalente […] é claro que a primeira considerada meramente como natural não pode ser mais a lei para nós do que a última [28].

---

[26] David HUME, *A Treatise of human nature* [1739–40], Bk II, part iii, section 3. Ver HUME, D. *Tratado da natureza humana*. Tradução de Déborah Danowski. São Paulo: Editora da UNESP, 2001, p.451.
[27] David HUME. *Enquiry concerning the principles of morals* [1751], Section IX, part 1, §4.
[28] Samuel BUTLER. *Fifteen sermons* [1726], Sermon II, §8.

Nossas emoções humanas naturais, em si mesmas, Butler plausivelmente conclui, não podem proporcionar a autoridade necessária (ou "normatividade" como os filósofos dizem hoje) que é exigida para fundamentar uma teoria do bem viver.

Precisamos, portanto, voltarmo-nos noutra direção, para alguma forma ética centrada na razão? Essa é a solução proposta por Spinoza, cuja posição é diametralmente oposta a de Hume:

> Apenas à medida que vivem sob a condução da razão, os homens concordam, sempre e necessariamente, em natureza. À medida que os homens são afligidos por afetos que são paixões, eles podem ser diferentes em natureza e mutuamente contrários. Diz-se, por outro lado, que os homens agem apenas à medida que vivem sob a condução da razão [29].

Há, contudo, problemas correspondentes aqui também, os quais, embora muito diferentes, são, a seu modo, exatamente tão preocupantes quando aqueles envolvidos na caracterização de Hume. Spinoza fala da 'harmonia', mas é difícil não sentir que é essa espécie de harmonia fria da qual Otaviano (o futuro Augusto César) fala no final de *Antonio e Cleópatra* de Shakespeare: "o tempo da paz universal está próximo"[30]. Do mesmo modo que, para Otaviano, destruir cruelmente todos seus adversários foi a abolição da república romana e a inauguração do governo imperial absoluto, para Spinoza não pode haver nenhuma acomodação real das paixões. Elas são, para ele, inerentemente caóticas e confusas – não há nada de cognitivo sobre elas; são simples influências causais pelas quais somos meramente 'afetados', em oposição a estar numa perspectiva que permita exercitar a agência genuína. E o remédio é a luz pura do conhecimento racional:

> Concebemos, assim, facilmente, o que pode o conhecimento claro e distinto contra os afetos e, sobretudo, aquele terceiro gênero de conhecimento;

---

[29] Benedict SPINOZA. *Ethics* [*Ethica ordine geometrico demonstrata*, c.1665], Bk IV, Prop. 35. Na tradução para o português: B. SPINOZA. *Ética*. Tradução de Tomaz Tadeu. Belo Horizonte: Autêntica, 2007. Parte IV, Prop. 35, p.301.

[30] William SHAKESPEARE, *Anthony and Cleopatra* [*c.* 1606], Act IV, Scene 6.

> quer dizer, esse conhecimento, se não os suprime inteiramente, à medida que são paixões, faz, pelo menos, com que constituam a menor parte da mente [31].

Quem está certo? Hume ou Spinoza? A história subsequente da filosofia moral gerou versões cada vez mais sofisticadas de cada uma dessas posições[32], mas se o conflito entre a razão e as emoções for concebido no modo sugerido pelo modelo humeano ou spinozista há pouca expectativa de solução. O que é preciso em vez disso, sugerirei, é uma desconstrução da suposta oposição entre razão e emoções. Dito de outro modo, precisa-se de uma espécie de *Aufhebung* hegeliana, onde tanto a tese quanto a antítese sejam reconfiguradas e incorporadas numa síntese que preserva os insights latentes em ambas[33].

4. O papel cognitivo das emoções

O primeiro passo em direção à resolução desse recuo improdutivo entre as alegações da razão e das emoções é questionar a pressuposição espinosista de que as paixões não têm nenhuma contribuição a dar à cognição e, portanto, que precisam ser, na melhor das hipóteses, completamente 'removidas' ou enfraquecidas até ficarem diminuídas e marginalizadas pelo emprego do conhecimento claro e distinto. Muitos filósofos subsequentes seguiram o caminho de Spinoza ao supor que as emoções estão "além dos limites" do conhecimento racional; em boa parte da filosofia analítica moderna, por exemplo, a verdade e o significado são considerados como vinculados às sentenças que expressam fatos descritivos literais, não-ambíguos, livres de toda contaminação emocional, enquanto os proferimentos emotivos, que envolvem todo o universo das expressões metafóricas e alusivas, são considerados água turva, pouco dignos de

---

[31] SPINOZA, *Ethics*, Bk V, Prop. 20.

[32] Para exemplos que se enquadram (muito amplamente) dentro da primeira tradição, ver, por exemplo, Simon BLACKBURN, *Ruling passions*. Oxford: Oxford University Press, 1998; e, na segunda, Christine KORSGAARD. *The sources of normativity*. Cambridge: Cambridge University Press, 1996.

[33] Ver Georg HEGEL, *Phenomenology of spirit* [*Phänomenologie des Geistes*, 1807], trans. J. B. Baillie. London: Sonnenschein, 1910, p.68 (§113).

confiança, algo para ser filtrado se desejamos precisão e acuidade em nosso conhecimento do mundo e de como devemos agir[34].

O trabalho filosófico mais recente em torno das emoções, contudo, tem começado a corrigir essa representação simplista e polarizada. O trabalho criterioso de Michael Slote em filosofia moral tem colocado a possibilidade da renovação da tradição sentimentalista em ética e sugere que os sentimentos, particularmente o sentimento de empatia, pode vir a desempenhar uma função chave na moralidade[35]. Michael Stocker tem argumentado que os estados emocionais tais como a raiva e a piedade podem ter uma função vital para dirigir e focar nossa atenção, afetando, assim, radicalmente o modo de perceber as coisas: "eles selecionam, coletam e até mesmo criam e dão sustentação a fatos correspondentes […] os quais são usados para justificar e sustentar essas emoções, o que, por sua vez, conduz a mais buscas, coleta, criação e coloração"[36].As emoções, como Mark Wynn belamente colocou, têm a função de "guiar a investigação através da constituição de padrões de saliência"[37]. Sobretudo, o trabalho precursor de Martha Nussbaum sobre as emoções tem contribuído para a atribuição de uma função vital às emoções na compreensão humana; isso equivale, com efeito, a uma crítica radical da dicotomia rígida entre as faculdades supostamente antitéticas da razão e das paixões.

Em seu artigo "Love's knowledge" que tem sido merecidamente admirado,

---

[34] Para reconhecer algumas das implicações disso para a maneira como a filosofia analítica é comumente praticada, ver J. COTTINGHAM, "What is humane philosophy and why is it at risk?", *Philosophy*, Supplement 65, p.1-23, 2009; e A. O'Hear (ed.), *Conceptions of philosophy*, Royal Institute of Philosophy series. Cambridge: Cambridge University Press, 2009.

[35] Michael SLOTE, *The ethics of care and empathy*. London/New York: Routledge, 2007. Para conhecer algumas reservas na abordagem de Slote, ver J. COTTINGHAM. "Empathy and ethics". In: *Abstracta*, Special Issue IV, p.13-19, 2010, e a réplica de Slote, p.44ff.

[36] Michael STOCKER com Elizabeth HEGEMAN, *Valuing emotions*. Cambridge: Cambridge University Press, 1996, p.94. Citado por Mark WYNN, "The relationship of religion and ethics: A comparison of Newman and contemporary philosophy of religion". In: *Heythrop Journal*, 2005.

[37] Mark WYNN, "The relationship of religion and ethics". Ver também Wynn, *Emotional experience and religious understanding: Integrating perception, conception, and feeling.* Cambridge: Cambridge University Press, 2005.

Nussbaum emprega um exemplo do *A la recherche du temps perdu* de Marcel Proust. O herói do livro vinha examinando cuidadosa e judiciosamente seus sentimentos do modo mais meticuloso possível e chega, então, à conclusão racional de que não se importa mais com Albertine. No entanto, quando a empregada Françoise dá a ele a notícia que Albertine deixou a cidade – exatamente no momento em que ele se convenceu completamente, com inteira segurança, que ele não mais a amava – um profundo, agudo e avassalador sentimento de angústia diz-lhe: você estava errado! A voz autoral reflete:

> Eu estava enganado ao pensar que poderia ver claramente dentro do meu próprio coração. Um conhecimento que as enganadoras percepções da mente não me deram foi agora trazido até mim, duro, dilacerador, estranho, como o sal cristalizado pela reação abrupta da dor [38].

Nussbaum usa esse episódio impressionante para lançar uma dúvida geral à uma certa concepção do conhecimento que tem raízes profundas em nossa tradição filosófica. De acordo com essa visão, o conhecimento (por exemplo, o conhecimento sobre si mesmo, se alguém ama alguém) "pode ser melhor capturado por um escrutínio intelectual destacado, não-emocional e exato da própria condição, conduzido da maneira que um cientista conduziria seu trabalho investigativo […] separando, analisando, classificando"[39]. Mas a descoberta feita por Marcel, o protagonista do romance, chegou a um resultado diferente:

> A descrição de Marcel do autoconhecimento não é simplesmente adversária da descrição intelectual. Ele nos diz que a descrição intelectual estava *errada*: errada sobre o conteúdo da verdade acerca de Marcel, errada sobre os métodos adequados de obter esse conhecimento, errada também acerca do tipo de experiência […] que é conhecer. E ele nos diz que tentar compreender o amor intelectualmente é um modo de não sofrer, de não amar, […] um estratagema de fuga [40].

---

[38] Marcel PROUST, *A la recherche du temps perdu* [1913-27], transl. L. Scott Moncrieff. London: Chatto & Windus, 1967, Vol. III, p.426. Citado por NUSSBAUM, 'Love's Knowledge' [1988], reimpresso em M. Nussbaum, *Love's Knowledge*. Oxford: Oxford University Press, 1990, p.265.
[39] NUSSBAUM, *Love's Knowledge*, p.262.
[40] NUSSBAUM, *Love's knowledge*, p.268-9 (grifo do autor). Essa discussão de Nussbaum foi extraída

5. Uma integração profunda

Mesmo sendo indubitavelmente vitais esses reconhecimentos do poder cognitivo das emoções, eles próprios não superam a dificuldade de acomodar as emoções dentro de uma teoria racionalmente defensável da boa vida. As emoções, como sabemos, podem ser frequentemente prejudiciais ou podem nos inclinar na direção de ações egoístas e más. E mesmo quando elas apontam para algo bom, elas podem exercer uma influência deturpadora, como Descartes vividamente notou:

> Mas amiúde a paixão nos faz julgar certas coisas melhores e mais desejáveis do que o são; pois, quando nos damos muito trabalho em adquiri-las e perdemos, entretanto, a ocasião de possuir outros bens mais verdadeiros, o gozo nos faz conhecer seus defeitos, e daí provém os desdéns, os pesares e os arrependimentos[41].

Poder-se-ia acrescentar que mesmo emoções inequivocamente benignas, tais como a empatia, podem nos desencaminhar, ao mexer com nossos sentimentos em favor de vítimas de sofrimento que se encontram próximas de nós, quando uma avaliação mais desapaixonada poderia ter sido de muito mais valor atraindo-nos para aqueles distantes[42].

A saída desse impasse, parece-me, é desenvolver um modelo mais holista e integrador para o bem viver, ao qual fiz referência no começo desse texto. Um

---

de J. COTTINGHAM, *The spiritual dimension*. Cambridge: Cambridge University Press, 2005, Cap. 1. Ver John COTTINGHAM. *A dimensão espiritual: religião, filosofia e valor humano*. Tradução de Edson Bini. São Paulo: Loyola, 2008.

[41] Carta de 1 de setembro de 1645 (AT IV 284-5: CSMK 264). Descartes começa dizendo que as paixões frequentemente "nos fazem julgar certas coisas melhores e mais desejáveis do que o são e nos fazem imaginar um prazer muito maior antes de tê-lo (o bem) do que aquele que a experiência subsequente realmente revela". A carta termina com a observação que "a verdadeira função da razão na conduta da vida é examinar e considerar sem paixão o valor de todas as perfeições, tanto do corpo quanto do espírito, que podem ser adquiridas por nossa conduta, a fim de que, de tal modo que sendo comumente obrigados a nos privar de algumas, escolhamos sempre as melhores". DESCARTES, R. *Cartas*. São Paulo: Abril Cultural, 1980, p.313-316.

[42] Tais questões são reconhecidas e examinadas por Michael Slote em *The ethics of care and empathy.*

modo de fazê-lo é sugerido pelo trabalho de Iain McGilchrist, em seu estudo notável *The master and his emissary*. McGilchrist apresentou evidências abundantes em favor da ideia que os dois hemisférios do cérebro são responsáveis por espécies muito diferentes de funções cognitivas: quando a atividade do hemisfério esquerdo domina, o resultado é uma espécie de modo analítico, atomista, de *escrutinar* o mundo, um modo que se afasta do envolvimento com o mundo; já a atividade do hemisfério direito é responsável por um modo de consciência mais empática, corporalmente mediada, um modo que é capaz de apreender o contexto amplo de nosso *envolvimento* essencial com o mundo ao nosso redor e com a vida dos outros[43]. A classificação esquemática das atividades no 'hemisfério direito do cérebro' e no 'hemisfério esquerdo do cérebro' naturalmente precisa ser cuidadosamente qualificada e McGilchrist repetidamente sublinha que, senão em todas, ao menos em boa parte de nossa atividade mental há normalmente a contribuição significativa de ambos os hemisférios. Mas a evidência neurofisiológica e comportamental para uma divisão ampla de função nos termos propostos parece, não obstante isso, ser muito persuasiva. E mesmo se não houvesse achados científicos (comprovatórios), ainda poderia haver, creio, uma força intuitiva enorme na distinção que McGilchrist traça entre

> Dois modos de ser no mundo, ambos essenciais. Um deles consiste em permitir que as coisas *se* apresentem a nós em toda sua particularidade, com toda sua maleabilidade, impermanência e com sua interconexão, como parte do todo que está, para sempre, em fluxo. Nesse mundo nós […] sentimo-nos conectados com aquilo que experimentamos, nos sentimos parte do todo e não confinados num isolamento subjetivo de um mundo que é visto como um mundo objetivo. O outro é o passo para fora do fluxo da experiência, o 'experimentar' nossa experiência de um modo especial: para *re-presentar* o mundo numa forma que é […] mais útil para a manipulação

[43] Iain MCGILCHRIST, *The master and his emissary: The divided brain and the making of the western world*. New Haven: Yale University Press, 2009. McGilchrist às vezes parece falar como se os *próprios* hemisférios fossem sujeitos de consciência, embora muitos filósofos poderiam (corretamente) se abster de atribuir estados psicológicos a cérebros ou partes de cérebros, em vez de atribuir tais estados a seres humanos tomados por inteiro.

> do mundo e do outro. Esse mundo é abstraído explicitamente, compartimentalizado, fragmentado, estático […] essencialmente sem vida. Sentimo-nos separados desse mundo, mas somos poderosos na relação com ele[44].

É, eu creio, difícil ler essa passagem sem ver uma conexão imediata com a defesa de Nussbaum da consciência emocionalmente mediada que compreende coisas que um escrutínio 'intelectual' destacado não conseguiria. Ou novamente, quando Wynn fala das emoções como "padrões constitutivos de saliência", isso parece impressionantemente consistente com a idéia de McGilchrist de uma forma de consciência mediada-pelo-hemisfério-direito que permite que as coisas sejam *apresentadas a nós*, "em toda sua particularidade encorpada", em oposição a ser representada como objetos para escrutínio e manipulação distanciada.

Agora estamos em melhores condições de ver o que está errado com as assunções da ética 'centrada na razão', discernível numa longa tradição de filósofos morais desde Platão e os estoicos até Spinoza e muitos outros: ela privilegia o tipo de consciência do "hemisfério esquerdo" que quer o recuo das emoções e o controle e manipulação delas. Na *República* de Platão, as massas turbulentas, representando as paixões, são reguladas de cima, pela razão filosófica, que foi literalmente cindida e tornada uma parte separada do Estado e/ou do indivíduo, vivendo uma espécie de vida inteiramente diferente. Mas assim como pais controladores pensam, ao acompanhar o desenvolvimento de uma criança indisciplinada, que ela não pode vicejar se for submetida sempre a avaliações e comandos, que precisa ter direito a um certo espaço para crescer e se desenvolver, do mesmo modo cada um dos múltiplos lados da nossa composição humana tem de ser propriamente reconhecidos se quisermos ser

---

[44] MCGILCHRIST, *Master and emissary*, p.93 (ligeiramente adaptada). A passagem continua: "A diferença essencial […] é que o hemisfério direito dá atenção ao Outro, aquilo que existe independente de nós mesmos e com o qual se reconhece numa relação profunda. É profundamente atraído e dá vida, vinculação e proximidade; ele existe com o Outro. Por contraste, o hemisfério esquerdo dá atenção ao mundo virtual que foi criado, que é autoconsistente, autocontido e, em última instância, desconectado do Outro, tornando-o poderoso, mas, em última análise, somente capaz de operar e conhecer a si mesmo".

felizes. Em termos jungianos, somente acertando as contas com o lado 'sombrio' podemos aspirar a totalidade.

Porém, levando em conta os múltiplos insights encontrados em maneiras muito diferentes por pensadores que citamos, de Jung a Nussbaum e indo até McGilchrist, *não* estamos querendo andar longe na direção contrária, na direção de uma ética das paixões, que privilegia o sentimento? Apesar da metáfora sugestiva, mas potencialmente enganadora de McGilchrist, que descreve os dois lados do cérebro como o 'senhor' e o 'emissário', um ser humano não deve ser identificado nem com seu hemisfério direito e nem com seu hemisfério esquerdo: nós somos meros seres cujo viço mental depende, em toda nossa existência, da interação harmoniosa entre os dois. Embora na vida de alguns indivíduos e realmente em todas as fases de nossa cultura, parece haver exemplos de desequilíbrio radical entre os hemisférios, gerando resultados desastrosos para a habilidade humana de estar inteiramente em casa no mundo, não obstante isso, como McGilchrist ressalta claramente, nossa satisfação como seres humanos sempre depende das "'funções' aparentemente separadas de cada hemisfério [ajustando-se] inteligentemente para formar em cada caso uma entidade coerente e única[45]. Assim nós precisamos de uma descrição da felicidade humana que respeite as contribuições feitas por todas nossas faculdades humanas e capacidades e que permita um modo humano de ser no mundo que nem recue para um gueto racionalista quase-autista, por um lado, nem, por outro lado, sacrifique a precisão analítica e a racionalidade crítica que tem sido crucial para o sucesso de nossa espécie.

Conclusão: humildade e os limites da filosofia

Como, em conclusão, podemos obter tal integração? Aqui, eu creio, nós beiramos os limites da filosofia – pelo menos no modo em que ela tem sido frequentemente praticada hoje, pois a filosofia analítica de nossa época tem se tornado *par excellence* uma atividade "do lado esquerdo do cérebro" – uma disciplina que premia sobretudo o distanciamento emocional, o argumento abstrato

[45] McGILCHRIST, *Master and emissary*, p.461.

e a confiança numa espécie de evidência que tem sido corretamente chamada de "evidência do espectador" (evidência acessível por qualquer observador racional imparcial, em oposição ao tipo que requer envolvimento e compromisso a fim de tornar-se manifesta)[46]. Assim concebida, a filosofia moral contemporânea tem se entrincheirado numa disciplina especializada 'rigorosa', formulando critérios cada vez mais precisos e abrangentes de conduta correta, ou proporcionando análises intrincadamente mais qualificadas dos princípios morais ou das razões para a ação, mas com o custo de abstrair mais e mais da textura de nossa experiência humana vivida e real.

É significativo aqui que a maior parte da literatura publicada em revistas de filosofia analítica, mesmo em filosofia moral, resolutamente evite o uso de linguagem metafórica, de citações literárias e algum tipo de apelo para respostas emocionais do ouvinte. Algo disso deriva da admiração (inteiramente justificada) às descobertas maravilhosas da ciência moderna, juntamente com a ideia (inteiramente enganosa) de que a filosofia deve, de todas as maneiras, imitar a linguagem e os métodos da ciência. No limite, talvez haja algo mais sinistro operando: um medo de contaminação com o mundo real, um desejo de refugiar-se no "mundo virtual que […] é autoconsistente, mas autocontido e, em última análise, separado do Outro, tornando-o poderoso, mas, em última instância, menos capaz de lidar e conhecer a si mesmo"[47].

Contudo, como a filosofia moral espera realizar sua antiga tarefa de tentar articular uma concepção do bem viver se não está preparada para delinear todas as fontes da nossa cultura, incluindo, talvez em especial, aquelas que engajam nossas formas imaginativas e emocionais de compreensão? Como Eleonore Stump brilhantemente mostrou em seu estudo magistral do problema do sofrimento humano, *Wandering in darkness*, a escrita filosófica pode abraçar uma perspectiva ampla sem, de modo algum, sacrificar a argumentação cuidadosa e precisa pela qual a tradição analítica foi (justamente) valorizada. Essas últimas virtudes são naturalmente necessárias, mas é um erro supor que elas são tudo o que é

---

[46] Acerca do termo 'evidência do espectador', ver Paul MOSER, *The elusive God*. Cambridge: Cambridge University Press, 2008, p.47.
[47] MCGILCHRIST, *Master and emissary*, p.93.

necessário. À luz da nossa discussão precedente de McGilchrist, é interessante que Stump assinale a "miopia cognitiva" como um aspecto problemático da tradição analítica – sua cegueira para as espécies de *insight* associados com o hemisfério direito do cérebro e sua tendência não justificada a "supor que as habilidades do lado esquerdo sozinhas revelarão tudo o que é interessante sobre o mundo"[48]. Argumentos intrincados e tecnicamente hábeis têm sua função, mas quando nos aproximamos de problemas como esse do sofrimento humano (e o mesmo poderia obviamente se aplicar à felicidade e satisfação humana) Stump argumenta que nós precisamos de recursos extras – em particular daqueles derivados das nossas múltiplas respostas da narrativa literária (e bíblica).

Isso não significa meramente usar textos literários e outros textos narrativos para finalidades 'ilustrativas' – uma manobra que poderia ser analiticamente 'respeitável', mas que, de certo modo, mantém o poder imaginativo de tais obras a uma distância segura. Isso seria rebaixar a função de um texto a mera figuração ou exemplo. Pois a filosofia moral, propriamente considerada, não é realmente só sobre conhecimento racional, mesmo conhecimento extraído de exemplos e casos; ela é ou deveria ser muito mais do que isso – um modo de aprofundar nosso entendimento da condição humana, em toda sua riqueza. Tal aprofundamento, naturalmente, não pode derivar da reflexão racional apenas: ele virá em parte de nosso envolvimento humano real com o processo complexo de viver e interagir com outros. Mas dado que, para seres com a capacidade humana única de pensamento e razão, a reflexão racional sempre será parte da história, seria melhor que ela fosse conduzida de um modo que não fosse 'míope'. Ela precisa estar preparada para empregar, sem embaraço ou apologia, toda a série de nossos recursos mentais e cultivar modos de consciência filosófica que estejam sintonizados com o contexto inteiro de nossas vidas, bem como com todos os elementos que a compõem.

Eu tenho registrado essas conclusões em termos de sua aplicação para a filosofia moral, mas elas também se aplicam naturalmente, *mutatis mutandis*, para

---

[48] Eleanore STUMP, *Wandering in Darkness Narrative and the Problem of Suffering*. Clarendon Press, Oxford, 2010, p.24-5.

o próprio tema do bem viver. Nossa natureza nos deu a habilidade única de nos afastarmos do mundo e mesmo dos trabalhos de nossos próprios corpos, a habilidade de analisar e, em alguma medida, controlar a realidade, nos oferecendo o sonho de virmos a ser, na impressionante e, como tardamos a reconhecer, sinistra expressão de Descartes "mestres e senhores da natureza"[49]. Mas há também essa outra habilidade, tão apreciada, de *ouvir* o mundo e a voz de nossa própria consciência interna e permitir a nós mesmos sermos abertos e responsáveis pelo que ouvimos. Isso não é, em si mesmo, uma receita para o bem viver, mas talvez seja o primeiro passo em direção a essa humildade que criaturas como nós precisam, se pretendemos realmente sermos felizes.

Isso finalmente nos conduz de volta para aquele que talvez seja o elemento mais importante do bem viver: a dimensão do sentido. Nós temos acesso ao sentido, se o argumento que estou esboçando for persuasivo, não simplesmente por flutuarmos nas partes constituintes das coisas, não somente via análise distanciada e escrutínio, mas por sermos "abertos" e "porosos" às ressonâncias[50] do todo, como uma pessoa inteira, numa vida completa, madura e que cresceu em conhecimento e amor do bem. Formulações como essas podem soar mais parecidas com um projeto espiritual do que filosófico, mas isso não deve nos incomodar, particularmente à luz do trabalho de Pierre Hadot e outros que tem mostrado como objetivos espirituais e filosóficos não estão, de nenhum modo, tão separados como nossa cultura acadêmica contemporânea gostaria de torná-los[51]. E se olharmos para a forma tradicional da disciplina espiritual, o que encontramos, eu creio, é uma estratégia de perseguição do bem que reconhece a necessidade de um engajamento corporalmente mediado, afetivamente sintonizado, de toda a pessoa para uma visão ampla do bem. Nas práticas espirituais e litúrgicas típicas o que está envolvido não é justamente reflexão e análise racional, mas a

---

[49] "Maîtres et possesseurs de la nature": DESCARTES, *Discourse on the method* [*Discours de la méthode*, 1637], Part Six (AT VI 61: CSM I 142). Ver Descartes, R. *Discurso do método*. São Paulo: Abril Cultural, 1980, p.63.

[50] NUSSBAUM, *Love's knowledge,* p.281-282.

[51] Pierre HADOT, *Philosophy as a way of life* (Cambridge, Mass.: Blackwell, 1995). Originalmente publicado como *Exercises spirituels et philosophie antique* (Paris: Etudes Augustiniennes, 1987).

participação real; e essa participação se expressa na postura corporal (de pé, de joelhos, sentado), no ouvir, ver e cheirar (cantando, celebrando, queimando incenso e outros componentes típicos da liturgia tradicional), no uso multi-regulado da linguagem e dos gestos e nos movimentos coordenados de devoção empreendidos em cooperação íntima com outros. O que está envolvido, em suma, é uma armação completa de rituais e práticas concebidas para ajudar o participante a 'transcender a si mesmo' – para abarcar com todo seu eu o domínio do significado moral e espiritual em que a vida inteira assume um propósito unificado.

Existem naturalmente incontáveis questões que surgem aqui com relação a visão de mundo religiosa que tais práticas pressupõem. Mas pelo menos isso é claro: tais estruturas e práticas são ao menos *um* modo em que as pessoas tradicionalmente buscaram plenitude humana precisamente no sentido holista e integrador que esse texto tem ressaltado. Podem, naturalmente, existir outros modos de buscar a plenitude, mas nada daquilo que é comumente oferecido pela filosofia moral contemporânea chega perto de indicar quais seriam esses modos[52].

[52] Para mais informações sobre esse ponto, ver J. COTTINGHAM, "Theism and spirituality", no prelo em V. Harrison, S. Goetz, e C. Taliaferro (eds), *The Routledge companion to theism.*

2

EMOÇÕES*

Julien Deonna, Fabrice Teroni e Christine Tappolet

Introdução

Imagine que você escute, por acaso, um desconhecido fazer chacota de um de seus melhores amigos. Naturalmente, você se sente ofendido. Uma raiva enorme rapidamente toma conta de você. Seu sistema nervoso autônomo produz uma série de alterações fisiológicas. À medida que o coração acelera, você prende a respiração. Suas sobrancelhas se erguem e você fuzila com o olhar o autor da chacota. Você cerra os punhos. Pensamentos atravessam sua mente sem parar. Como se ousou insinuar alguma coisa tão sórdida a respeito de seu amigo? E, então, você reflete sobre o comportamento que deveria adotar: afastar-se? Insultar o maledicente? Ou, sem o devido respeito às almas sensíveis, partir para os socos?

Diferente de outras emoções, a raiva levanta uma quantidade de questões filosóficas fascinantes, das quais este capítulo não pode dar senão uma pequena amostra. A primeira seção se propõe a comparar as emoções com outros fenômenos afetivos, recorrendo a três distinções: a distinção entre disposições e episódios psicológicos, a distinção entre experiências e outros estados psicológicos e também a distinção entre estados psicológicos intencionais e não-intencionais. A segunda seção põe a questão da natureza das emoções. Afirma-se comumente que as emoções são avaliações, mas como compreender essa ideia? É preciso considerar as emoções como juízos avaliativos, como percepções de

---

*Este capítulo foi publicado originalmente como um verbete da Encyclo-Philo <http://encyclo-philo.fr/emotions-a/>. Sua reprodução aqui foi autorizada pelos autores e editores. A tradução é de Katarina Peixoto e Flavio Williges.

valores ou, ainda, como uma outra coisa? Para retomar nosso exemplo, a menos que seja um tipo de percepção que não lhe permita ter consciência da ofensa ao seu amigo, a sua raiva poderia consistir em julgar a chacota ofensiva. A terceira seção se dedica ao papel epistêmico que as emoções desempenham na relação com os juízos avaliativos. A raiva fornece alguma razão para julgar a chacota ofensiva? Ela lhe dá condições de aceder ao fato correspondente? A quarta seção considera diferentes aplicações do contraste inato/adquirido no tocante às emoções. A cultura e a educação têm um papel na maneira como você fica com raiva e na sua expressão dessa raiva? Se for esse o caso, qual é o papel dessa emoção? A relação entre as emoções e a motivação é o objeto da quinta seção. Qual é a relação entre a sua raiva e o seu desejo de não ser polido e dar um soco no autor da ofensa? As emoções são suficientes para gerar ações ou somente contribuem para os atos? Enfim, a sexta seção examina as funções que as emoções podem desempenhar na ética. Mais especificamente, nela se discute a hipótese de que se pode recorrer às emoções para compreender tanto a motivação, quanto o juízo moral. Partindo desse pressuposto, a reação de raiva diante da chacota revela uma significação moral.

1. As emoções e o domínio afetivo

Para quem se interessa pela natureza das emoções, uma das primeiras questões concerne a sua similaridade com outros estados psicológicos, como os humores, os sentimentos, as paixões e os desejos, estados que podemos nos sentir tentados a agrupar no seio do domínio afetivo. Ora, os filósofos da mente têm à sua disposição um conjunto de distinções que podem muito bem ajudar a responder a essa questão. Pode-se assim distinguir os *episódios* psicológicos, que são fenômenos relativamente pontuais, das *disposições* psicológicas, mais estáveis; as *experiências*, que possuem um caráter fenomenal ou qualitativo, dos estados mais "intelectuais" (juízos, suposições), que não possuem nada disso; e os estados *intencionais*, que estão voltados a coisas outras que eles mesmos, dos estados *não-intencionais*.

Com base na primeira distinção, é possível isolar duas categorias de

fenômenos afetivos: de uma parte, aqueles que aparecem à consciência, intensificam-se, reabsorvem-se e desaparecem em um curto lapso de tempo. De outra parte, aqueles que, por serem relativamente estáveis e duráveis, caracterizam, de uma maneira mais global, o engajamento afetivo do sujeito com seu ambiente (Broad 1954; Ryle 1949). Assim, o enunciado "Paul está com raiva de seu pai" pode significar tanto que Paul vivencia atualmente um episódio de raiva dirigido ao seu pai, bem como a simples disposição a ficar com raiva de seu pai, algo que estava bastante adormecido até este instante. Contrariamente à leitura episódica desse enunciado, a leitura disposicional implica somente que Paul é suscetível de se encolerizar (caso em que "manifestará" sua disposição de raiva) nas circunstâncias que envolvam seu pai (as quais constituiriam assim "o gatilho" para essa manifestação). Mesmo que se encontre exceções (Goldie 2000; Wolheim 1999), a maior parte dos filósofos e psicólogos contemporâneos empregam o termo "emoção" para designar esses episódios. Essa prática poderia bem distanciá-los do uso ordinário, segundo o qual as emoções são sempre maneiras mais ou menos estáveis de se engajar no mundo. Para designar esses fenômenos, os filósofos utilizam às vezes uma expressão técnica: "disposição emocional". Eles consideram então as preocupações, os cuidados, os vínculos, os temperamentos, os sentimentos e os traços de caráter como disposições emocionais, em vez de emoções.

Se não é em si controversa, a distinção entre episódios e disposições levanta questões filosóficas interessantes. Em primeiro lugar, é preciso fazer outras distinções dentre as disposições afetivas e, se for este o caso, com base em que critérios (Deonna & Teroni 2009)? Assim, embora todas as duas sejam disposições afetivas, o medo de avião e o amor à pátria não convidam à mesma análise. Em segundo lugar, a maior parte, talvez todas as emoções, manifestam disposições afetivas – como o ilustra a raiva de Paul, se ela manifesta seu temperamento colérico –, o que deixa a questão de ordem da explicação abrangente em aberto. É preciso dar conta das disposições em termos de suas manifestações ou, inversamente, das emoções em termos das disposições que elas manifestam? Ademais, as disposições afetivas têm de fato uma realidade psicológica que está

além de suas manifestações? Afinal, o status psicológico da aerofobia, na ausência de toda condição suscetível de dispará-la (decolagem iminente, planejamento de uma viagem) está longe de ser evidente. Considerando isso, certos filósofos preferem se abster de tais afirmações, a fim de preservarem a possibilidade de relações interessantes (explicativas, epistêmicas ou outras) entre as emoções e as disposições afetivas que elas manifestam (Naar 2013). Esse é um ponto a que voltaremos em seguida (seção 4).

Pode-se igualmente tirar proveito da segunda distinção. Experimentar uma emoção faz uma certa diferença. As emoções estão entre o número de estados psicológicos aos quais a ideia de caráter qualitativo ou fenomenal se aplica mais naturalmente; em uma palavra, trata-se de experiências (Lambie & Marcel 2002; Leighton 1985; Pugmire 1998). Aqui, também, elas diferem das disposições afetivas: em si, o temperamento colérico de Paul não é acompanhado por qualquer ressentimento da parte dele. Mas esse critério permite igualmente distinguir as emoções de outros estados psicológicos, tais como os juízos. Com efeito, pode-se julgar que dois mais dois são quatro sem qualquer sentimento (Hatzimoyis 2007; Bayne & Montague 2011). É claro que isso não quer dizer que as emoções são as únicas experiências no seio do domínio afetivo. Como veremos, os humores também satisfazem a essa descrição.

A dimensão fenomenal das emoções também traz consigo questões difíceis. Em primeiro lugar, são todas as emoções conscientes? Parece que não (Lacewig 2007; Prinz 2004a). Afinal, quando estamos tensos, nossa atenção está focada naquilo que tememos e não no nosso próprio sentimento. A fim de caracterizar esse tipo de fenômeno, pode-se considerar que essas emoções não sejam reflexivamente conscientes – na medida em que a atenção do sujeito não está centrada nelas, ele não forma crenças a seu respeito e pode, inclusive, negar com toda boa fé tê-las sentido. Em segundo lugar, podemos viver uma emoção sem sentir o que quer que seja (Hatzimoysis 2007; Lacewig 2007)? Essa questão não concerne mais à consciência reflexiva, mas antes ao que se convencionou chamar de consciência fenomenal (Block 1995). Os filósofos mostraram-se muito mais reticentes a aceitar a possibilidade de emoções sem senti-las. Certos

psicólogos estão, contudo, prontos a aceitá-las, a fim de distinguir o fato de sentir uma emoção do simples fato de tê-la. Essa distinção é pertinente, pois há dados comportamentais e neurofisiológicos que indicam a ocorrência de uma emoção sem que nada sugira que o sujeito a sente (Damasio 2000; Feldman Barrett et al. 2005). No entanto, não é evidente que esses dados constituam razões confirmando a existência de emoções não sentidas; elas não serão, por exemplo, consideradas como tais por quem admite que a consciência fenomenal não é necessariamente acompanhada pela (possibilidade da) consciência reflexiva. E, à primeira vista, parece que sentimos bem as emoções sem nos dar conta de que é esse o caso.

Sendo assim, quer as emoções impliquem ou não um sentimento, a natureza de sua fenomenologia levanta questões fundamentais. Essa fenomenologia é suficientemente rica para justificar as fronteiras que a linguagem ordinária traça no seio do mundo emocional? Os experimentos clássicos de Schachter e Singer (1962), que parecem mostrar que os sujeitos nas mesmas condições fisiológicas estão prontos para atribuir emoções muito diferentes em função das circunstâncias nas quais se encontram, foram consideradas por muitos pesquisadores como uma razão para negá-lo. Em que pese o vínculo entre os dados recolhidos por Schachter e Singer e essa conclusão estarem longe de ser evidentes (Prinz 2004; Deonna e Teroni 2016), muitos psicólogos e filósofos os aceitaram (Lyons 1980; Russell 2003), argumentando que uma só e mesma fenomenologia é compatível com as emoções habitualmente consideradas como muito diferentes. Em seguida, mesmo admitindo que muitas emoções supõem a experiência de alterações de nossa condição corpórea (muscular, fisiológica, endócrina, etc), é preciso chegar a afirmar, com William James, que a fenomenologia de todas as emoções pode ser capturada em termos corpóreos (Goldie 2000; Prinz 2004)? Quais são as consequências para essa questão, do fato que a atenção do sujeito que sente uma emoção não é, senão raramente, dirigida para as alterações de sua condição corpórea? E, mais geralmente, como a atenção mobilizada pelas emoções influencia a experiência emocional (Lambie & Marcel 2002)? Enfim, são as emoções necessariamente caracterizadas por um

sentimento bom ou ruim, quer dizer, por uma tonalidade hedônica positiva ou negativa? Esta última questão está intimamente ligada à noção de "valência" das emoções (Charland 2005; Colombetti 2005; Teroni 2011) e encontra ecos importantes no seio da literatura contemporânea sobre prazeres e desprazeres (Aydede 2006).

A terceira distinção opõe os estados intencionais aos estados não-intencionais. Ela também pode ser empregada para traçar as fronteiras no interior do domínio afetivo. Existe notavelmente um forte contraste entre as sensações corpóreas paradigmáticas e as experiências emocionais (Leighton 1985; Bedford 1957; Pitcher 1965). As cócegas ou uma sensação de frio nas costas, por exemplo, não são voltadas ao que quer que seja, ou ao menos a algo que esteja fora do corpo. Elas não têm intencionalidade, ao menos não o tipo de intencionalidade característica das emoções. Com efeito, o medo, a cólera, a admiração e o prazer estão voltados geralmente (senão sempre) a objetos, a acontecimentos e a estados de coisas exterior: fica-se aborrecido com o pai; admira-se um pôr do sol; tem-se prazer com uma piada; lamenta-se que Jeanne não pôde vir à festa. Observe a passagem que as emoções podem fazer para objetos mais complexos: certas emoções, como a vergonha e o orgulho, são reflexivas e desempenham por isso um papel importante na nossa concepção da própria identidade (Taylor 1985; Deonna, Rodogno & Teroni 2011), assim como as "metaemoções" que sentimos a propósito de nossas próprias emoções, como quando lamentamos ter sentido raiva (Jäger & Bartsch 2006). As emoções são estados intencionais. Uma vez admitido isso, não deriva essa característica simplesmente dos pensamentos que a acompanham? Poder-se-ia assim sustentar que o seu medo de um exame consiste em um pensamento voltado para ele e acompanhado de um conjunto de sensações típicas, as quais são elas mesmas despidas de intencionalidade (Aristóteles 1994; Goldstein 2003; Whiting 2011). De maneira alternativa, poder-se-ia concebê-lo como um sentimento não-intencional de insatisfação resultante da tomada de consciência de que seu desejo de não se submeter ao exame será frustrado (Reisenzein 2012). Qualquer que seja a resposta a essa questão, veremos na seção seguinte que a intencionalidade das emoções tem um outro

aspecto, que poderia muito bem ser específico delas e que elas não parecem possuir de maneira derivada.

Mas, antes, é preciso observar isto: é porque as emoções são tipicamente dirigidas para fora do corpo que os filósofos têm, conjuntamente, sido reticentes em lhes assimilar às sensações corpóreas – rejeitando por isso uma posição correta ou incorretamente associada a William James (James 1884, 1890; Ellsworth 1994; Ratcliffe 2005; Teroni 2017). Se essa propriedade diferencia as emoções das sensações, ela as distingue igualmente dos humores. Estes últimos, temos o costume de pensar, são episódicos e possuem um caráter fenomenal (Pears 1975). Ademais, em que pese os humores serem em geral bastante duradouros, seu perfil temporal não permite distingui-los das emoções de uma maneira clara. Por exemplo, uma atenção sutil pode fazer com que cesse imediatamente seu mau humor matinal. Contrariamente às emoções, os humores não são todavia intencionalmente dirigidos para o que o quer que seja: há certamente algo que faz com que estejamos de mau humor, mas dificilmente se pode ser mal humorado a respeito de alguma coisa. Observe, aliás, que isso pode consistir em uma razão para distinguir a causa de um estado afetivo de seu objeto. De uma parte, os humores têm causas, mas não objetos. De outra, os objetos das emoções não são necessariamente sua causa: a raiva de Paul pode ser causada pelo testemunho de sua mãe e ter por objeto o seu pai (Kenny 1963).

A maneira como acabamos de traçar a distinção entre as emoções e os humores está longe de constituir unanimidade, contudo. A diferença de intencionalidade entre as emoções e os humores é categórica ou se trata apenas de uma questão de grau (Crane 1998; Kind 2014; Mendelovici 2014)? Não teriam os humores simplesmente objetos menos específicos que os das emoções, sendo o mau humor um tipo de irritação dirigida contra o mundo em sua totalidade (Goldie 2000)? Ou seria preciso entender a distinção entre os humores e as emoções em termos do perfil modal específico do objeto dos humores, por exemplo, admitindo que os humores representam a probabilidade de certas possibilidades avaliativas se realizarem? (Price 2006; Tappolet 2016)?

Nós sugerimos que as emoções são estados psicológicos episódicos,

intencionais e dotados de uma certa fenomenologia. E, quando se trata de contrastar emoções e disposições, sua breve duração foi sublinhada. Isso deixa, porém, aberta a questão da categoria ontológica a que pertencem as emoções. São elas acontecimentos, como explosões e nascimentos? O fato de que uma emoção é tipicamente uma ocorrência de breve duração poderia dar a pensar. E, aliás, as emoções têm todas uma duração nula? Algumas dentre elas – a surpresa, talvez – não são instantâneas, como o é o ato de perceber alguma coisa? Se este é o caso, isso seria compatível com seu status de experiências (Geach 1969)? Uma outra possibilidade é conceber as emoções como processos. É o que permite levar em conta o fato de que as emoções parecem se desdobrar e passar tipicamente por diferentes fases – nós sublinhamos acima que elas aparecem à consciência, intensificam-se, reabsorvem-se e desaparecem. Um perfil temporal como esse, composto de etapas que não existem todas ao mesmo tempo, mas ligadas umas às outras numa sucessão ordenada, corresponde àquele de processos paradigmáticos, como a redação de uma carta (Goldie 2000; Price 2015). Se as emoções são processos, são suscetíveis de perseverar após terem sido interrompidas? No momento em que você está aborrecido com seu chefe, você se lembra de uma piada muito engraçada. A despeito do fato de que tenha caído no riso, uma parte da sua raiva perdura. Quando você redireciona a sua atenção sobre a ofensa cometida pelo seu chefe, sua emoção de raiva retoma o seu curso? Ou é o caso de concluir disso que você tem simplesmente um humor irritadiço, que dará lugar a um novo episódio de cólera dirigido contra seu chefe, quando não estiver mais absorvido pela piada (Crane 1998; Kind 2014; Mendelovici 2014; Pears 1975; de Sousa 1987)? Após terem passado muito tempo negligenciadas, essas questões concernentes à ontologia das emoções estão hoje no coração de numerosos debates (Naar & Teroni 2017).

2. Emoções e avaliações

As emoções são estados intencionais: elas visam a alguma coisa. Como constatamos, essa característica as distingue dos humores e das sensações. Sendo assim, as emoções não são evidentemente os únicos estados intencionais.

Você pode certamente acreditar que seu chefe lhe exige horas extras de trabalho, mas pode também imaginar, crer, deplorar ou mesmo lembrar-se disso. Quer dizer assim que a intencionalidade das emoções não difere daquela de outros tipos de estados psicológicos? Certamente não. Afinal, as emoções não se contentam em visar um objeto dado: elas lançam sobre ele um olhar *positivo ou negativo*. Mais precisamente, elas avaliam seus objetos e distinguem-se umas das outras pela maneira de fazê-lo (Platão 2016). Se você tem medo que seu chefe lhe exija horas extras de trabalho, é porque você vê nisso, por exemplo, uma ameaça ao seu bem-estar, ou então pensa que a resposta negativa que lhe daria prejudicaria a sua carreira. De modo mais geral, o medo avalia o seu objeto como ameaçador ou perigoso, a raiva como ofensivo, o prazer como engraçado e a vergonha como degradante (Kenny 1963; Roberts 2003). Essa ideia é consistente com a maneira com que nos acostumamos a conceber, ordinariamente, as emoções: com efeito, se as emoções são avaliações, compreende-se que se as critique ("Tu não deverias ter medo") e se as desaprove ("Tu deverias ter mais compaixão"). Além disso, observe que as diferentes ocorrências de um mesmo tipo de emoções podem destinar-se a objetos variados, mas cuja avaliação em jogo é idêntica: quer você tenha medo de insetos, de espaços confinados, de uma prova, do modo como seu tio dirige ou da morte, você sempre avaliará o objeto de sua crença como ameaçador. Parece que é sobre esse aspecto avaliativo que a intencionalidade das emoções se distingue daquela de outros estados psicológicos – a especificidade das emoções deveria ser situada não no nível de seus objetos particulares, mas no nível do que o jargão filosófico descreve como seu objeto "formal", a saber, das propriedades avaliativas que, a um só tempo, apreendem os objetos particulares a que visam (Teroni 2007).

Caso se aceite certas concepções clássicas das emoções, no entanto, essa é uma conclusão à qual se deve resistir. Considere, por exemplo, a ideia muito simples, a que subscrevem, aparentemente, os estóicos, segundo a qual as emoções são juízos avaliativos. Segundo essa concepção, estar com raiva de um vizinho consiste em julgar que ele cometeu uma ofensa, e ter vergonha equivale a acreditar que se foi de alguma maneira aviltado (Nussbaum 2004; Solomon 1993).

Mas então, não haveria nada de muito original na intencionalidade das emoções: o liame intencional entre esse episódio de raiva e o vizinho seria idêntico aquele implicado no juízo segundo o qual ele cometeu uma ofensa, assim como com a ligação intencional entre a vergonha de quem a sente seria idêntica àquela implicada no juízo de que se foi aviltado.

Essa tentativa de reduzir a intencionalidade das emoções à do juízo enfrenta, contudo, uma série de dificuldades. Para começar, não está claro que ter a experiência de uma emoção requer a formação de um juízo avaliativo correspondente. Às vezes, sentimos uma emoção a despeito de não aderirmos a uma avaliação pertinente, como quando reagimos de uma maneira que sabemos ser inapropriada. Uma pessoa que tem fobia de aracnídeos, por exemplo, terá medo de uma aranha mesmo que a julgue inofensiva (Döring 2014; Tappolet 2012). Se sentir uma emoção não implica manifestamente que se aceite a avaliação correspondente, pode ser suficiente, para senti-la, *pensar* no objeto em termos avaliativos pertinentes (Döring 2014; Greenspan 1988; Roberts 2003; Tappolet 2012; Teroni 2007). A cólera que você sente ao encontrar o seu vizinho resumir-se-ia então ao pensamento segundo o qual ele cometeu uma ofensa. Não é afinal inverossímil admitir que um tal pensamento precise ser aceito para que você esteja com raiva de seu vizinho – pode ser até mesmo que a recorrência desses pensamentos após ter sido convencido de que esse vizinho não queria lhe fazer mal explique a persistência de sua raiva.

É preciso, no entanto, resistir a essa interpretação. Com efeito, conceber as emoções em termos de juízo ou de simples pensamentos, consiste simplesmente em deixar de lado o fato de que se trata de experiências específicas. Julgar que alguém cometeu uma ofensa não implica sentir raiva. Essa constatação se impõe de maneira talvez ainda mais evidente no caso dos pensamentos correspondentes: pode-se pensar que alguém cometeu uma ofensa sem se deixar ofuscar por ela, assim como se pode pensar que um labrador é perigoso, sem ter medo dele (Leighton 1985; Maddell 1997). Por essas razões, a existência dos juízos e pensamentos em questão não constitui uma condição suficiente para as emoções. Aliás, ela não parece tampouco constituir uma condição necessária. De fato, um

grande número de animais não-humanos sentem emoções, embora não sejam, sem dúvida, capazes de emitir juízos avaliativos ou de se ocupar de pensamentos avaliativos (Deigh 1994; Roberts 1996). Essas duas constatações jogam água no moinho da mesma conclusão, a saber, que a intencionalidade avaliativa das emoções serve para compreender com base num modelo cognitivo menos exigente e mais sensível à natureza da experiência emocional do que aquele sugerido pelas concepções em discussão.

Uma pista a ser explorada, caso se busque desenvolver um modelo como esse, consiste em recorrer à percepção, antes que ao juízo ou ao pensamento. De fato, os estados perceptivos são estados mentais paradigmáticos a respeito dos quais falamos por experiência. Seriam as emoções, assim, formas da percepção? Segundo essa ideia, um objeto pode nos atingir emocionalmente como sendo ofensivo ou perigoso da mesma maneira que pode nos atingir visualmente, como sendo escarlate; as similaridades subjacentes entre emoções e percepções seriam suficientemente profundas para justificar a análise das primeiras em termos das segundas. Essa analogia, que foi o objeto de um rigoroso tratamento por parte de Ronald de Sousa (de Sousa 1987), gerou um conjunto de abordagens perceptuais das emoções (Deonna 2006; Döring 2007; Johnston 2001; Prinz 2004a; Tappolet 2000,2014; Tye 2008; Wringe 2014; Zagzebski 2003). Segundo esses autores, uma abordagem perceptual possui numerosas vantagens. De uma parte, ela é fiel à dimensão fenomenal das emoções. De outra, ela fornece uma solução sedutora para outros problemas que as posições mais intelectualistas, descritas anteriormente, enfrentam. Em particular, ela permite pensar as fobias segundo o modelo das ilusões perceptuais. Se você percebe um bastão imerso pela metade na água, sua experiência visual será a de um bastão dobrado. Daí você não conclui que o bastão está dobrado. De maneira análoga, uma pessoa que tem aracnofobia poderá ter a experiência emocional da presença de uma aranha inofensiva, como perigosa, sem por isso julgar que ela o é (Döring 2014; Tappolet 2012). Tanto em um caso, como no outro, a experiência não impede que o sujeito tenha bom senso. Ou melhor, como o sublinha McDowell (1994), a percepção pode ser compreendida como um “convite” a aceitar um certo conteúdo, mas esse é um

convite que, às vezes, devemos recusar. Uma abordagem perceptual das emoções resolve igualmente um outro problema: o das emoções dos animais. Com efeito, a representação perceptual do perigo não parece repousar nas capacidades cognitivas de que os animais são desprovidos: se os animais não são talvez capazes de acreditar que o objeto é perigoso ou ofensivo, sua bagagem cognitiva é verdadeiramente suficiente para percebê-lo como tal.

Apesar das vantagens que acabamos de enfatizar, as abordagens perceptuais se deparam com importantes desafios, na medida em que as emoções parecem, em muitas visões, distinguir-se das percepções (Deonna & Teroni 2012; Dokic & Lemaire 2013; Salmela 2011). Primeiramente, em que pese a fenomenologia das experiências emocionais ser tão proeminente como a da percepção sensorial, ela não poderia lhe ser assimilada. Em numerosos casos, a experiência de uma emoção tem algo de particular: ela implica a experiência de uma preparação para agir de uma certa maneira sobre um objeto dado – não é fácil, por exemplo, descrever a experiência do medo sem mencionar a experiência de uma preparação para evitá-lo, ou a experiência da admiração, sem mencionar a de uma preparação para manter a sua atenção sobre o que se admira. Não está claro que o modelo perceptual seja capaz de fazer justiça a essa característica, dado que a fenomenologia da percepção é em princípio estritamente receptiva.

Em segundo lugar, as propriedades representadas por experiências perceptuais são geralmente tomadas como causalmente responsáveis: se você percebe um tigre branco, seu estado mental é, evidentemente, causado pela presença de um tigre branco. Ao contrário, não é certo que as propriedades que as emoções pretendem representar possam desempenhar um papel causal pertinente. Pois quais são suas propriedades? Se admitirmos que se trata de propriedades avaliativas, como o fazem geralmente os amigos da teoria perceptual, a ideia de um poder causal pressupõe uma forma controversa de realismo a esse respeito: as propriedades avaliativas existiriam “no mundo”, independentemente de nossa mente, e elas estariam aptas a causar em nós as respostas psicológicas. Além disso, a teoria perceptual demanda uma concepção muito liberal de percepção, pois não é certo que as propriedades avaliativas

possam ser percebidas no sentido ordinário do termo. No mínimo, a avaliação de uma teoria perceptual depende da concepção da relação perceptiva e do gênero de fenômenos perceptivos sobre os quais ela repousa. Assim, na medida em que se aceita uma concepção da relação perceptiva segundo a qual ela é constituída pela consciência direta dos dados sensíveis, ou por uma forma de realismo direto, as perspectivas de uma teoria perceptual variariam. Do mesmo modo, apelar à percepção simples, que não implica o reconhecimento (Dretske 1979), a uma forma de atenção perceptiva, ao reconhecimento fundado sobre a percepção ou ainda ao ver como (Pelser 2014; Roberts 2003) leva a abordagens muito diferentes.

Em terceiro lugar, as emoções são sempre provocadas por pensamentos que concernem a acontecimentos distantes de nós, tanto espacial, como temporalmente. Você pode assim sentir alegria com o fim de um conflito longínquo ou lamentar que Júlio César foi assassinado. Há aí um contraste claro com a percepção sensorial, na medida em que as experiências perceptivas têm objetos que pertencem ao nosso ambiente imediato. Esse contraste deve evidentemente ser posto em relação com as observações precedentes a respeito da causalidade: é por que o objeto de uma experiência perceptiva deve ser sua causa que ele tem de ser circunscrito de maneira muito mais estrita que o objeto das emoções. Está igualmente ligado a um fato que nós já sublinhamos: as emoções, ao contrário das experiências perceptivas, não têm necessariamente um objeto que as cause.

Essa terceira diferença dá a ver uma quarta diferença significativa entre os estados perceptivos e as emoções: enquanto os primeiros usufruem de uma certa autonomia, os segundos dependem sempre de outros estados mentais (Deonna & Teroni 2013; Mullingan 1998). Não se pode ficar aborrecido com sua chefe sem representá-la de uma outra maneira (via a imaginação, uma lembrança, uma crença ou uma percepção, por exemplo). A relação intencional entre uma emoção e seu objeto particular pressupõe a existência de um outro estado psicológico. Ao contrário, a experiência visual que você tem de sua chefe não altera a sua crença, uma lembrança ou um estado imaginativo acerca dela – se há aqui uma relação de dependência, ela parece ir em outra direção, na direção em que crer, lembrar-se e

imaginar pressupõem, muito frequentemente, estados perceptivos anteriores.

Pode ser, evidentemente, que uma teoria perceptual seja capaz de dissipar esses desafios. Pode ser que essas objeções repousem sobre uma concepção conservadora demais da percepção, por exemplo (a esse respeito, ver Tappolet 2016). De acordo com uma abordagem menos conservadora, uma relação perceptiva é fundamentalmente uma co-variação causal sistemática, selecionada pela evolução, entre as condições ambientais dadas e as respostas psicológicas. Esse gênero de abordagem não ficou sem eco nos debates contemporâneos em filosofia das emoções, na medida em que a maneira de aplicação dessa concepção da percepção às emoções, por Jesse Prinz, suscitou um certo entusiasmo (Prinz 2004a).

Isso posto, as objeções mencionadas deveriam talvez incitar a exploração de outras opções teóricas. Duas estratégias muito gerais se destacam: de uma parte, pode-se desenvolver novas concepções das avaliações que incorporem as emoções. Em vez de assimilá-las a estados psicológicos mais familiares, como os juízos, os pensamentos ou as experiências perceptivas, pode-se sustentar que as emoções são uma categoria *sui generis* de estados mentais, no sentido que se trata de um tipo diferente de atitudes. Poder-se-ia, por exemplo, sustentar que as atitudes emocionais constituem-se de sentimentos de preparação para a ação, direcionados a certos aspectos do ambiente (Deonna & Teroni 2012, 2015). De outra parte, pode-se pensar que os problemas enfrentados pelas abordagens discutidas acima decorrem invariavelmente da ideia segundo a qual as emoções são avaliações, as quais seria melhor renunciar. Certos filósofos contemporâneos se voltaram, por consequência, a abordagens mais tradicionais, fundadas sobre a ideia de que emoções são, antes de qualquer coisa, sentidas, e que os sentimentos que elas evocam não são passíveis de compreensão em termos intencionais (Kriegel 2014; Reisenzein 2012; Whiting 2011). A despeito da ressurgência dessas abordagens em termos de sensações brutas, o caráter essencialmente avaliativo da intencionalidade das emoções persevera no centro dos debates filosóficos contemporâneos (Salmela 2014).

3. Epistemologia e compreensão

Até aqui, nos concentramos em uma questão de ordem ontológica: qual é a natureza das emoções? Mas as emoções também oferecem questões epistemológicas. Pode-se por exemplo perguntar como elas contribuem para o conhecimento. Como vimos na seção precedente, as emoções não são manifestamente juízos avaliativos. No entanto, elas dão lugar, frequentemente, a tais juízos: se você admirou um filme, ao assisti-lo, você provavelmente julgou que ele é bom, e sentir raiva de seu vizinho em um dado instante poderia muito bem levar você a julgá-lo ofensivo ou desrespeitoso. Tudo isso é bastante consensual. O que provoca discussão é saber se as emoções apenas causam esses juízos, ou se a relação que mantêm vai além disso. Duas questões se põem, a esse respeito. Primeiro, as emoções desempenham algum papel relevante em nossa compreensão do aspecto avaliativo desses juízos, quer dizer, em nosso domínio dos conceitos avaliativos que contêm? Em segundo lugar, as emoções fornecem um padrão de justificação para esses juízos? Se damos uma resposta positiva a uma dessas questões, a conclusão que se impõe é que as emoções preenchem uma função importante no conhecimento avaliativo. Para sabê-lo, convém revisitar, sob um ângulo epistemológico, certas teorias das emoções apresentadas acima.

Observe, em primeiro lugar, que, se identificamos as emoções com juízos avaliativos, essas duas questões não fazem sentido. Com efeito, ao supor que o seu medo do cão seja idêntico ao seu juízo de que o cão é perigoso, não é claro como ele poderia contribuir para a compreensão do conceito de perigo. Afinal, emitir o juízo em questão parece pressupor essa compreensão, em vez de contribuir com ela. O medo não poderia tampouco contribuir para a justificação desse juízo: uma relação de justificação pressupõe dois termos, que são reduzidos a um só pela teoria judicativa.

Por outro lado, aderir a uma teoria perceptual permite talvez responder positivamente a essas duas questões, o que constituiria um argumento importante em seu favor. No que concerne à compreensão dos conceitos avaliativos, o modelo perceptual está em todo caso em posição de traçar um paralelo

esclarecedor. Assim como um cego não saberia o que significa para uma coisa ser colorida, uma pessoa desprovida de emoções não saberia o que significa para uma coisa ser portadora de valor. Com efeito, que conceito de prazeroso ou admirável alguém incapaz de sentir o menor prazer e a menor admiração, poderia ter (Deonna & Teroni 2012; Robert & Wood 2007)? Por essa razão, certos filósofos (Goldie 2002) sentiram-se tentados a conceber a relação que uma pessoa incapaz de ter emoções tem com valores aos moldes da relacão que a famosa Maria tinha com as cores antes de sair do seu quarto incolor (Jackson 1986). Ambos podem, sem dúvida, aprender a classificar os objetos ou as situações de uma maneira que corresponda à nossa. No entanto, se essas classificações estiverem separadas de seus fundamentos no acesso perceptivo ou afetivo às propriedades pertinentes, isso terá por consequência que esses sujeitos não serão capazes de apreender nossos conceitos de cores ou de valores.

A analogia perceptual é igualmente promissora no que concerne ao eventual papel justificativo das emoções, *vis-à-vis* os juízos avaliativos (Baier 2004; Döring 2007; Tappolet 2000). Nossas crenças são frequentemente justificadas por nossos estados perceptivos. Assim, a experiência visual da vermelhidão de um tomate bem brilhante e pouco distante de nós fornece uma razão para julgar que este tomate é de fato vermelho. Na ausência de uma razão para pensar que nossos sentidos nos enganam ou que a iluminação não é boa, estamos justificados a formar esse juízo. É tentador dizer o mesmo a respeito das emoções. Quando o medo nos apresenta o perigo de um dado cão, você dispõe de uma razão para julgar que esse cão é efetivamente perigoso. E, na ausência de uma razão contrária – você não está drogado, não é tendencioso nem afetivamente desequilibrado – você está, parece, também justificado a formar esse juízo.

O modelo perceptual pode, portanto, orgulhar-se de identificar uma fonte plausível, tanto para nossa compreensão dos conceitos avaliativos, quanto para nosso conhecimento avaliativo. No entanto, a confiabilidade das emoções não poderia ser dada como certa da mesma maneira que a da percepção. Afinal, é um truísmo que as emoções possam nos induzir a erro, que os juízos que geram no calor do momento são, via de regra, duvidosos (Brady 2013; Dokic & Lemaire

2013). Essa constatação se reflete em questões que gostamos de fazer a respeito das emoções que nossos congêneres têm e que não nos ocorrem quanto a suas experiências perceptuais (Helm 2001). Pode-se explicar o fato de Helena julgar que um Dois-Cavalos está no estacionamento dizendo que viu um Dois-Cavalos no estacionamento[1]. E essa explicação é, em todo caso, satisfatória: não somente nossos estados perceptuais nos fornecem as *razões* para subscrever esses juízos correspondentes; eles constituem um ponto de chegada natural na busca de uma justificação epistêmica. Ora, parece que as emoções diferem, a esse respeito, das percepções. Suponha que você se pergunte por que você julga uma ação injusta e responda a isso mencionando a indignação que sente diante dela. Sua emoção talvez forneça uma certa sustentação ao seu juízo, mas ela não poderia constituir uma justificação última. De fato, seria legítimo perguntar: "Certo, mas por que essa indignação?". Poder-se-ia então esperar que você estivesse fazendo referência a certas características da ação a que sua emoção reage, antes que a sua própria reação emocional (Deonna & Teroni 2014). Essas considerações são compatíveis com a tese segundo a qual as emoções fornecem justificações *prima facie* ou derrotáveis (Tappolet 2016), ou jogam, em vez disso, água no moinho da ideia segundo a qual elas funcionam como sinais de alerta que devem nos levar a interrogar-nos a respeito da presença de razões para emitir juízos avaliativos correspondentes (Brady 2013)?

Em qualquer caso, é verossímil que o papel epistêmico das emoções difira daquele dos estados perceptivos. Embora as emoções se pareçam com estados perceptivos, pelo fato de nos fornecerem razões para crer, elas se distinguem deles na medida em que respondem igualmente a razões: contrariamente às percepções, elas são ou justificadas ou injustificadas.

Esse aspecto das emoções pode conduzir a pôr em questão uma ideia apresentada acima e sobre a qual os defensores de uma abordagem perceptiva sempre insistiram. A ideia de que um sujeito sente, como muitos aerófobos, uma emoção que sabe ser injustificada, é comparável àquela em que sabe que é vítima

[1] "Dois Cavalos" é um modelo de carro antigo da fábrica de automóveis francesa Citroen.

de uma ilusão perceptiva. Se dermos como favas contadas essa ideia, não é plausível afirmar que certas críticas que se podem dirigir ao primeiro – "tu és irracional!" – não se aplicam ao segundo (Helm 2015)? Os defensores de uma teoria perceptiva ficarão tentados a sublinhar que essa diferença se deve simplesmente ao fato de que as emoções são, ao contrário dos estados perceptivos, plásticas: podemos exercer uma influência indireta no longo prazo sobre as emoções que certos tipos de situações suscitam em nós (Tappolet 2012). É por isso que se pode criticar um sujeito que nada fez para corrigir suas emoções recalcitrantes. Certos filósofos expressam, no entanto, reservas quanto à possibilidade de resolver essa dificuldade, apelando à plasticidade das emoções (Price 2015).

Além disso, se as emoções respondem a razões, é porque de fato dependem, fundamentalmente, de outros estados psicológicos. Nós o constatamos anteriormente: sempre acessamos os objetos de nossas emoções através de outros estados psicológicos – estados perceptivos, crenças ou lembranças, por exemplo. Pode-se portanto pensar que as razões a que as emoções respondem decorrem ao menos em parte do conteúdo dos estados mentais de que elas dependem. Para responder à questão "por que você tem medo desse cachorro?" você mencionará, por exemplo, o fato de que você viu os dentes e já escutou o seu rosnado, ou sua crença que os pitbulls são perigosos. Parece que estados psicológicos como as percepções, as crenças e as lembranças – estados que frequentemente se afirma terem a direção de ajuste mente-mundo (Searle 1983) são capazes de fornecer razões para as emoções.

Observe, porém, que as emoções e os juízos a que elas dão lugar também são explicadas pelo que chamamos (na seção 1) de dispositivos emocionais – os sentimentos e os traços de caráter, por exemplo – de que eles são manifestações, assim como por estados conativos, tais como os desejos (Helm 2001). Só uma pessoa caridosa sente compaixão em determinadas circunstâncias; do mesmo modo, uma pessoa que não tivesse desejado acumular uma fortuna não ficaria regozijada com a subida do preço das ações investidas. De modo mais geral, nossas respostas emocionais tornam-se comumente inteligíveis pelos sentimentos,

pelas preocupações e pelos vínculos particulares que nos ligam às pessoas, às instituições e aos outros objetos. São esses os estados psicológicos que possuem, seja os da direção de ajuste mundo-mente (como os desejos), seja aqueles para os quais falar de direção de ajuste não faz, à primeira vista, muito sentido. Ora, algumas das dificuldades fundamentais encontradas quando se busca determinar a função epistêmica das emoções dizem respeito a esses estados.

Uma primeira dificuldade diz respeito, de fato, ao papel das disposições emocionais (cuidados, preocupações, sentimentos etc.) e dos estados conativos (desejos e aspirações) em relação às emoções. Limitar-se-iam eles a explicar causalmente sua aparição ou seriam eles mesmos capazes de justificá-las? A animosidade que alimenta em relação a sua chefe justifica sua raiva ao encontrá-la (Helm 2001; Price 2015) ou apenas a explica? É difícil responder a essas questões sem subscrever a uma concepção particular dos valores que as emoções supõem representar (Deonna & Teroni 2012). Suponha que o caráter ofensivo da observação de sua chefe seja particularmente constituído pela animosidade que você alimenta em relação a ela. Fica por sua conta, então poder justificar sua raiva mencionando essa disposição. Em troca, se as propriedades avaliativas que representam as emoções são independentes das disposições emocionais e dos desejos, como o pensam os realistas, convém distinguir, cuidadosamente, a contribuição causal de sua contribuição epistêmica: disposições emocionais e desejos explicam talvez a ocorrência de emoções, mas não as justificam. Isso quer dizer que o realista não leva em conta, de maneira exaustiva, a perspectiva avaliativa do agente em questão quando se trata de avaliar as emoções? Não necessariamente, na medida em que negar que as disposições emocionais e os desejos contribuem positivamente para a justificação das emoções não impede de admitir duas coisas. Primeiramente, que esses estados psicológicos possuem uma função seletiva essencial, ao atuar como filtros sobre a sensibilidade avaliativa de um sujeito para quem contribuem definindo contornos – a covardia não justifica o medo, mas ela explica em parte por que o sujeito está sempre pronto a se inquietar com qualquer coisa (Deonna & Teroni 2012b). Em segundo lugar, que existem valores prudenciais que levam em conta a especificidade de um dado agente, suas

capacidades e suas relações – o que é perigoso para um indivíduo não o é forçosamente para um outro (Ronnøw-Rasmussen 2011).

A segunda dificuldade se põe aos que adotam uma concepção realista das propriedades a que se supõe as emoções respondam. Se se admite que essas propriedades são independentes dos dispositivos emocionais, mas que esses dispositivos são determinantes maiores de nossas emoções, pode-se suspeitar que estas últimas sejam suspeitas de alterar nossa apreciação da paisagem avaliativa (Goldie 2008). Dizendo claramente: se as emoções são concebidas sob o modelo de um tipo de mecanismo de detecção das propriedades avaliativas, será melhor que sua fiabilidade esteja à altura desse estatuto. E sua fiabilidade vai, por sua vez, depender da exata natureza das disposições que explicam seu desencadeamento. As perspectivas epistêmicas de uma concepção desse tipo vão depender de muitos fatores, em particular, da natureza exata de nossas capacidades emocionais mais simples (ver a seção seguinte) e da possibilidade de corrigi-las, visando a desenvolver uma verdadeira sensibilidade avaliativa (D'Arms & Jacobson 2010).

Em terceiro lugar: a presença de uma emoção justificada é suficiente para justificar o juízo avaliativo correspondente? Alguns filósofos ofereceram uma resposta negativa a essa questão, considerando que as condições que devem ser preenchidas para que uma emoção seja justificada são menos restritivas que aquelas que concernem aos juízos avaliativos (Greenspan 1988). Trata-se de uma posição consistente? Afinal, não está claro que um sujeito que não possua razões suficientes para julgar de maneira justificada que uma pessoa foi ofensiva possa, no entanto, sentir uma raiva justificada em relação a ela. Ou melhor, a grande dificuldade para controlar as emoções poderia sugerir que é mais fácil desculpar uma pessoa que sente uma emoção por más razões do que uma pessoa que forma um juízo tendo por base essas mesmas razões.

Em quarto lugar, o fato de que as emoções são elas mesmas suscetíveis a serem justificadas ou injustificadas é compatível com a ideia de que elas justificam os juízos avaliativos aos quais dão lugar? Dito de outra forma: as razões que justificam as emoções não são suficientes para, igualmente, justificar os juízos

avaliativos pertinentes, tornando por isso todo o retorno às emoções inútil? Duas leituras se apresentam, espontaneamente, ao espírito (Deonna & Teroni 2012). Segundo a primeira, as emoções são supérfluas do ponto de vista da justificação dos juízos, podendo eles ser justificados diretamente pelas mesmas razões que justificam as emoções. Quando você julga que a observação de seu chefe foi ofensiva, seu juízo está justificado somente se você dispõe de uma razão para fazê-lo. Mas se você dispõe com efeito de uma tal razão, esta justifica seu juízo, o qual gera em outro uma reação emocional ou não (Brady 2013). De acordo com a segunda interpretação, as emoções jogam um papel ineliminável na transmissão das razões aos juízos avaliativos, o que elas podem fazer de duas maneiras. De uma parte, pode-se pensar que as emoções nos situam em uma posição epistêmica favorável, ao chamarem nossa atenção para características de nosso ambiente que, sem elas, escapar-nos-iam (Robinson 2005; de Sousa 1987). De outra parte, poderia ser o caso que as emoções contribuam essencialmente para a nossa compreensão da dimensão avaliativa desses juízos, uma ideia que já mencionamos, em conexão com a aquisição dos conceitos avaliativos. Tanto em um caso, como em outro, as emoções poderiam muito bem se apresentarem como indispensáveis ao conhecimento desses valores.

4. Inato x adquirido

Começamos este capítulo distinguindo as emoções de outros fenômenos afetivos como os humores e os traços de caráter. Em seguida, constatamos que as emoções diferem umas das outras em função dos valores que representam: o medo representa o perigo, ao passo que a tristeza representa a perda, por exemplo. Convém, no entanto, efetuar distinções estruturais maiores no seio do domínio emocional. A questão central concerne então à existência de uma divisão entre as emoções que são relativamente impermeáveis a pressões contextuais e emoções que são mais permeáveis a essas influências.

Pode-se abordar essa questão sob o ângulo de uma distinção muito intuitiva, aquela que opõe as emoções relativamente simples ou básicas, que são, por exemplo, o medo, a raiva e a alegria, às emoções mais complexas, dentre as

quais encontramos o arrependimento, o orgulho e a nostalgia. Em que pese essa abordagem ser muito antiga, os psicólogos e os filósofos contemporâneos a levaram a sério, na tentativa de identificar as emoções que têm uma ancoragem biológica distintiva e podem ser descobertas *por meio de* expressões faciais, de circuitos cerebrais, uma função evolutiva, um papel motivacional ou esquemas de ativação fisiológica (Griffths 1997; Charland 2002; Scarantino & Griffths 2011). A viabilidade desse projeto foi posta em dúvida: não está claramente estabelecido que certas partes do domínio emocional sejam, nesse sentido, primárias (Solomon 2011). Mas os pesquisadores que estão convencidos dessa abordagem têm à sua disposição recursos interessantes para articular as relações entre emoções simples e complexas.

Dentre as numerosas maneiras de articular essas relações, a de Jesse Prinz (2004a, 2004b) é particularmente atraente. Com uma distinção que já temos em mente (aquela entre o tipo de avaliação que implica uma emoção e a variedade das coisas que ela pode tomar por objeto), pode-se ser levado a pensar que certas emoções complexas resultam de uma modulação particular de emoções mais básicas. A indignação, por exemplo, seria uma forma de raiva: enquanto tal, ela avaliaria seu objeto como uma ofensa, mas teria a particularidade de se concentrar num subconjunto de objetos ofensivos, a saber, as ações imorais realizadas pelos outros. De maneira análoga, a *Schadenfreude* e a emoção japonesa *ijirashi* poderiam ser variantes da alegria exclusivamente concentrada sobre dois tipos de objetos suscetíveis de provocar esta emoção: respectivamente, a má sorte e o sucesso dos outros. Seria preciso então compreender as influências culturais vividas por nossas emoções como uma forma de explorar nosso repertório emocional básico, a fim de conduzi-lo nas direções consideradas como importantes no seio de nossa cultura. Esse tratamento permite dar conta das variações culturais importantes nas emoções (incluindo a prevalência relativa das diversas emoções), preservando a ideia segundo a qual as vidas emocionais dos seres humanos pertencentes a diferentes culturas se constroem sobre um conjunto de capacidades emocionais partilhadas.

Pode-se, contudo, julgar superficial demais essa concepção do impacto das

forças sociais sobre as emoções. Considere, por exemplo, a maneira como a expressão das emoções varia através dos contextos sociais, culturais e históricos. Em nossas sociedades, competitivas e orientadas para o sucesso, o orgulho não é estigmatizado como o era no mundo cristão medieval. Do mesmo modo, a dor não é expressa no Irã como o é na Suíça (sobre a história das emoções, ver Plamper & Tribe 2015; Rosenwein 2016). Pode-se imaginar que a expressão das emoções contribui essencialmente para a sua identidade. Daí só falta um passo em direção à ideia segundo a qual, ao expressar as emoções, um sujeito e o grupo a que pertence não se contentam em transportar manifestações naturais de respostas biologicamente determinadas em direções específicas. Eles contribuem mais fundamentalmente para fixar a própria natureza das emoções. Essa hipótese se vê reforçada quando se constata a importância da função comunicativa da manifestação de emoções (Darwin 1872; Ekman 2003), e o fato de que essa função é às vezes conscientemente desviada para fins pessoais ou coletivos diversos (Griffiths 1997). Seria, então, mais fácil compreender as emoções à luz dos fins que buscamos, moldando-as, do que compreendê-las como simples componentes de um edifício biológico.

Essas ideias estão naturalmente de acordo com uma certa tradição, com a qual já nos deparamos quando abordamos a questão da natureza da fenomenologia emocional (seção 2). Segundo essa tradição, as emoções não têm qualquer perfil distintivo, no sentido das respostas fenomenológicas, mas também cerebrais e fisiológicas, suficiente para determinar que tipo de emoção está em jogo (Cannon 1927; Russell 2003; Schachter & Singer 1962). Se é assim, cabe ao sujeito mesmo e aos membros de sua comunidade, fazê-lo. E, evidentemente, as próprias respostas a respeito não jogam senão um papel periférico nesse processo. Segundo essa forma de construtivismo, os sujeitos devem identificar as emoções ao interpretar as reações que observam com base em seu contexto de ocorrência e nas normas culturais pertinentes (Armon-Jones 1986; Averill 1980; Barrett 2006). Assim, a resposta emocional negativa que você experimenta quando vê alguém cortejar seu parceiro pode constituir um episódio de indignação virtuosa em um certo contexto e um episódio de ciúme patético, em um outro. Determinar

qual emoção você sente não supõe tanto ser atento ao que se passa em você, mas em interpretar o seu sentimento à luz das normas pessoais ou coletivas pertinentes em seu contexto.

As variações culturais que fazem das emoções objetos – tanto no nível de seus desdobramentos, como no de sua expressão – constrangem-nos a adotar um construtivismo tão radical? Inversamente: são elas incompatíveis com um modelo que reconhece o caráter universal e inato de ao menos algumas de nossas emoções? Os debates recentes nesse domínio estão largamente estruturados em torno desses dois pólos (Scarantino 2014).

5. Emoções e ação

Frequentemente se dá como certo que as emoções nos motivam a agir – o termo "emoção" deriva, a propósito, do latim "ex" (de fora) e "movere" (mover). Assim, a raiva é sempre acompanhada de comportamentos agressivos, e o medo é comumente associado a respostas específicas, tais como a fuga ou o combate. A relação entre emoções e motivação põe uma questão central na teoria das emoções. O poder motivacional das emoções decorre de sua essência ou se trata de um fenômeno contingente? Se a raiva e o medo necessariamente nos motivam, não poderíamos sentir raiva sem sermos motivados a gritar e não poderíamos ter medo sem sair correndo em disparada.

Para enfrentar essa questão, é preciso ainda saber o que é a motivação. Concorda-se geralmente em dizer que as emoções facilitam a ação, na medida em que as modificações corporais que implicam aumentam a probabilidade de certos comportamentos – a vingança no caso da raiva e a fuga naquele que tem medo, por exemplo. Considera-se igualmente que a tonalidade hedônica das emoções influencia a motivação, o prazer sendo em geral alguma coisa que se procura, contrariamente à dor. Mas a questão toda é saber se a motivação envolvida nas emoções é constituída de disposições comportamentais específicas, que alimentam diretamente o sistema motor, ou antes estados, tais como os desejos, que, presumivelmente, não têm senão uma influência indireta sobre aquilo que fazemos.

Conforme a primeira abordagem, cada emoção nos dispõe a uma gama limitada de comportamentos. No caso da raiva, os comportamentos pertinentes estão associados à agressão e à vingança; pode se tratar de uma ameaça de ataque ou de combate. O medo nos incita a fugir, imobilizar e abater (Cannon 1927; Elster 1999; Frijda 1986). Essas disposições provocariam uma gama estreita de estímulos, tais como o da visão de um competidor ou de um predador, e se caracterizam por manifestações rápidas e automáticas, independentes do pensamento e da tomada de decisão (Cosmides & Tooby 2000; Elster 1999; Frijda 1986; Griffths 1997). Essa primeira abordagem é plausível no que diz respeito à maioria das emoções sentidas por animais não-humanos (Tappolet 2010). Todavia, se for preciso reconhecer que manifestamos ocasionalmente o mesmo tipo de comportamento que os esquilos e as marmotas quando estamos assustados, o medo provoca em nós reações bastante diferentes. É certo que nos impulsiona a fugir, mas também pode nos incitar a pedir socorro, ao menos pelo celular. É por isso que muitos filósofos consideram que a relação entre as emoções e a ação é mais indireta que do que se supõe à primeira vista.

De acordo com a segunda abordagem, as emoções implicam motivos ou desejos, os quais têm um impacto indireto sobre a ação. Elas influenciam então nossas decisões, ao contribuírem para a definição de nossos objetivos (Clore 1994; Helm 2010; Prinz 2004a). Assim, seguindo Aristóteles, alguns consideram que a raiva supõe um desejo de vingança (Aristóteles 2004). Do mesmo modo, o medo poderia implicar o desejo de evitar um dano ou uma perda. Esses desejos informam o processo de decisão do agente e não resultam na ação senão mediante esse processo. Então, dados seus outros objetivos e as limitações derivadas do contexto, os desejos em questão poderiam não ter nenhum efeito em seu comportamento. Por exemplo, nenhuma vingança seguir-se-ia da raiva que você sente de quem roubou sua bicicleta enquanto você ignora quem foi o autor do roubo. Se essa segunda abordagem fornece uma explicação plausível do impacto motivacional das emoções, resta saber como interpretá-las. As emoções são essencialmente ligadas aos desejos ou essa é uma relação mais frouxa?

De fato, é difícil conciliar essa concepção com certas emoções que

implicam aquilo que se chama "ações expressivas" (Betzler 2004; Döring; Hursthouse 1991). Em um acesso de raiva, Jane perfura, em uma foto, os olhos de sua grande rival. A emoção de Jane implica manifestamente um desejo que define um objetivo: machucar essa pessoa. O ferimento, que é o desejo em questão, encerra a explicação do comportamento de Jane, que sabe, corretamente, não estar machucando a sua rival ao danificar a foto. Segundo uma sugestão plausível de Peter Goldie, as ações expressivas tais como essa de Jane se explicam pelos anseios, concebidos como estados conativos envolvendo a imaginação do desejo satisfeito (Goldie 2000). Jane deseja certamente furar os olhos de sua rival. Mas como um comportamento desse está fora de questão, ela se contenta em fazê-lo sobre a foto, imaginando que o faz na realidade. Alternativamente, poder-se-ia explicar a ação de Jane em termos do fracasso da disposição primitiva a causar problemas, que está sempre presente na raiva: levando em conta o fato de que a foto representa a rival de Jane, ela libera essa disposição primitiva como se sua própria rival estivesse, sem dúvida, presente.

As emoções dirigidas a objetos ficcionais são igualmente problemáticas. Frequentemente sentimos raiva ou medo lendo um romance ou assistindo a um filme. Mas parece que essas emoções têm apenas um vínculo tênue com a motivação. Como o observa Kendall Walton, se você assiste a um filme no qual uma matéria verde e viscosa se espalha sobre a Terra, aniquilando toda forma de vida por onde passa, você provavelmente vai se mexer em seu assento; talvez até mesmo se mexa com incômodo; mas uma coisa, no entanto, é certa: você não sairá correndo para fora do cinema, e tampouco chamará o serviço de segurança (Walton 1978). Walton tira daí a seguinte lição: as obras de ficção não geram emoções verdadeiras, mas simples "quase-emoções". Contrariamente ao que Walton defende, é possível que esses estados sejam emoções autênticas, mesmo que não cheguem a envolver desejos e juízos avaliativos típicos de contextos não-ficcionais (Deonna 2007; Tappolet 2010). Seria necessário, então, concluir que as emoções não estão essencialmente ligadas à motivação.

As emoções que são dirigidas ao passado oferecem problemas similares. Nada nos impede de sentir raiva de um antepassado que dilapidou nosso

patrimônio familiar. Embora essa emoção *possa* em princípio ser acompanhada pelo desejo de encontrar o parente em questão e lhe dar uma sova, não está absolutamente claro que se *deva* fazê-lo. Enfim, certas emoções não parecem simplesmente conectadas à motivação. É o caso, notavelmente, da admiração, da alegria e do alívio: pode-se admirar uma paisagem, sentir alegria ou sentir alívio sem a menor disposição para agir.

O que se depreende, parece, é que as emoções não estão essencialmente conectadas à motivação. Por isso, seria errôneo minimizar a relação contingente que mantêm com os nossos comportamentos e desejos específicos: as emoções ajudam muito frequentemente a nos motivar. É, aliás, por isso que são consideradas centrais em ética.

6. Emoções e ética

Com algum exagero, talvez, poderíamos dividir a comunidade filosófica em dois campos determinados. De um lado, os racionalistas situam sua confiança na razão e deploram os efeitos deletérios que as emoções têm sobre nossos raciocínios teóricos e nossas deliberações. De outro, os sentimentalistas, que duvidam do poder da razão e insistem na importância das emoções em nossas vidas. Os racionalistas consideram a maior parte das emoções como um obstáculo à racionalidade tanto teórica como prática, reprovando a interferência das emoções sobre nossos raciocínios, como geradoras de comportamentos irracionais, imprudentes ou mesmo imorais (Platão 2009). Os sentimentalistas sustentam, ao contrário, que as emoções são essenciais a um só tempo para o bom funcionamento da razão e para a ação. Segundo os trabalhos de Ronald de Sousa (1987) e Antonio Damasio (1994), a maior parte dos teóricos contemporâneos das emoções adotou uma atitude muito favorável em relação às emoções (ver Jones 2007).

Dito isso, o conflito que opõe racionalistas a sentimentalistas está longe de ser encerrado. Essa constatação é particularmente interessante em ética, onde a oposição entre racionalismo moral e sentimentalismo moral continua a estruturar os debates. Segundo o racionalismo moral, que remonta a Kant (1785), os juízos

morais encontram seu fundamento na razão. Os sentimentalistas morais o situam, ao contrário, em nossos sentimentos. David Hume, o representante mais famoso dessa tradição, afirma assim que "A moral é […] mais propriamente sentida que julgada" (1741) e sustenta explicitamente que as distinções morais não são derivadas da razão.

Para avançar nesse debate, é preciso responder a duas questões. A primeira concerne à natureza das emoções. Só se pode manter um contraste radical entre sentimentalismo e racionalismo moral caso se tome as emoções como estados despidos de intencionalidade, que nada têm a ver com a racionalidade. Ora, como sublinhamos, parece bem mais plausível admitir que as emoções não são somente estados intencionais, mas admitem justificações (cf seções 1 e 3). A segunda questão concerne ao papel exato que as emoções desempenham em nossa vida moral. Dada a sua diversidade – pense na diferença entre raiva, culpa e compaixão, por exemplo – e a complexidade de cada episódio particular, não é surpreendente que se atribua às emoções papéis tão diferentes. Além disso, numerosos objetos que interessam à filosofia moral mantêm relações com as emoções: (i) os juízos morais (que concernem ao que os agentes deveriam fazer, ao que é bom ou mau, e por aí vai), (ii) a motivação moral e a ação, (iii) o raciocínio moral e a deliberação prática, (iv) os fatos morais (tais como o fato de que Sam não deveria mentir) e (v) as virtudes e os vícios (tais como a coragem e a covardia). Note que as emoções também desempenham uma função relativa a um certo número de bens, tais como a autonomia (Friedman 1986; Johnston 2001; Jones 2003) e o bem-estar (Deonna & Teroni 2013; Kriegel 2014; Tappolet & Rossi 2014). Aqui, nos concentraremos no papel das emoções no que diz respeito aos juízos morais e à motivação moral.

De acordo com filósofos, juízos morais seriam constituídos por emoções ou até mesmo redutíveis às emoções (Hobbes 1651; Prinz 2008). Note, de passagem, que essa abordagem é a inversão perfeita da teoria segundo a qual as emoções são simplesmente juízos avaliativos (cf. seção 2). Esse tratamento tem uma atração inegável entre os adeptos do não-cognitivismo moral, que sustentam que os juízos morais não têm por função enunciar os fatos e que são, por

consequência, desprovidos de valor de verdade. Enquanto tese sobre juízos ou estados psicológicos, o não-cognitivismo é distinto, mas parente próximo, de duas doutrinas metaéticas importantes, a saber, *o expressivismo* (ou *emotivismo*), a tese semântica segundo a qual a função dos enunciados morais consiste em expressar emoções (Ayer 1936; Gibbard 1990; Stevenson 1937) e o *projetivismo*, a tese que pretende que a moralidade não seja outra coisa que uma projeção de nossas reações emocionais nos objetos que nos circundam; o mundo enquanto tal seria desprovido de normatividade (Blackburn 1984). O expressivismo é frequentemente compreendido como uma forma de não-cognitivismo, mas convém notar que se trata de uma tese semântica, sobre a natureza dos juízos, e não uma tese psicológica. Quanto ao projetivismo, trata-se de uma tese sobre a natureza das entidades morais, e não uma tese psicológica ou semântica.

O emotivismo e o projetivismo se apoiam habitualmente em dois pressupostos relativos à natureza das emoções. Primeiro: as emoções não teriam conteúdo representacional que pudesse ser avaliado em termos de verdade ou correção; segundo: as emoções seriam essencialmente estados motivacionais, por isso seria suficiente, para dar conta da força motivacional dos juízos, enunciados ou fatos morais, estabelecer uma conexão com as emoções. Mas esses pressupostos estão na contracorrente do consenso na teoria das emoções. Como observamos acima, a maior parte dos teóricos optam por uma concepção intencional das emoções e consideram que as emoções são juízos ou percepções. Ademais, como acabamos de ver na seção precedente, está longe de ser certo que as emoções também estão intimamente ligadas às motivações, como supõem os emotivistas e os projecionistas.

Existem outras concepções da relação entre juízos morais e emoções. Uma abordagem particularmente promissora, que pode remontar a Franz Brentano (1889), busca analisar os conceitos morais em termos de emoções apropriadas. Conhecida pelo nome de "neo-sentimentalismo", essa abordagem é particularmente promissora no caso dos juízos relativos às propriedades avaliativas. Em relação aos vínculos entre emoções e valores que acabamos de explorar na seção 2, isso não chega a ser evidentemente uma surpresa. A

abordagem geral pode ser explicitada de várias maneiras, mas ela afirma sempre que as seguintes equivalências se dão em virtude de conceitos que elas implicam: *x* é apavorante se e somente se *x* é tal que é apropriado ao medo; *x* é admirável se e somente se *x* é tal que é apropriado para a admiração e *x* cometeu uma ofensa se e somente se *x* é tal que é apropriado a entrar em estado de raiva perto dele (D'Arm & Jacobson 2000; McDowell 1985; Scanlon 1998). Os adeptos dessas abordagens devem, é claro, explicar o que entendem por "apropriado", uma tarefa que se mostrou particularmente complicada. Com efeito, uma tal explicação deve evitar uma objeção importante: certas emoções são apropriadas em relação a certos objetos que não instanciam, no entanto, o valor pertinente. Por exemplo, você poderia muito bem admirar o iate de seu vizinho rico, se sonhasse em ser convidado para subir a bordo, mesmo que o iate em questão não tenha no fundo nada de admirável em si (D'Arms & Jacobson 2000; Rabinovicz & Rønnow-Rasmussen 2004). Alguns neo-sentimentalistas se contrapuseram a essa objeção, ao definirem o caráter apropriado das emoções em termos de sua correção: só é apropriado admirar o iate de seu vizinho se o barco for efetivamente admirável (Danielsson & Olson 2007; Tappolet 2011).

A principal vantagem dessa abordagem é que permite dar conta de duas características de juízos morais aparentemente em conflito, a saber, a característica cognitiva e a força motivacional. Em mais de um aspecto os juízos morais parecem-se com os juízos cognitivos mais paradigmáticos. Assim como ocorre com os juízos acerca de cores e formas, eles são considerados verdadeiros ou falsos e parecem representar os fatos. No entanto, eles aparentemente mantêm uma relação especial com a motivação. Se alguém julga que deve ajudar seu amigo, esperamos que ocorra uma ação em consequência. Se não ocorre, daí concluímos que ele sofre de uma forma de irracionalidade prática como a fraqueza da vontade (Smith 1994). Pode-se tentar explicar essa característica dos juízos morais por sua conexão com as emoções, as quais normalmente implicam motivação.

A análise dos juízos morais em termos de emoções apropriadas postula uma conexão conceitual entre os conceitos morais e os conceitos de emoções. Aí

está uma concepção que se encaixa bem com a ideia segundo a qual a compreensão dos conceitos avaliativos depende da possibilidade de sentir emoções; um conceito como *admirável* ou *repulsivo* só pode ser plenamente compreendido por aquele que tem a experiência desse gênero de reações (ver a seção 3). Outras abordagens são mais empíricas, no sentido que não visam senão relações de ordem causal entre as emoções e os juízos morais. Como notamos, as emoções são às vezes vistas como permitindo, em certas circunstâncias, a justificação de juízos avaliativos (seção 3). É admitido por todos, no entanto, que existem relações estáveis de causalidade entre emoções e juízos normativos, quer sejam avaliativos ou de ordem moral. Assim, experimentos em psicologia social sugerem que as emoções têm um impacto considerável sobre nossos juízos morais. Nossa condenação moral de uma prática é fortemente influenciada pelo nojo que nos inspira (Haidt 2001; Nichols 2004). Observemos, todavia, que essa relação não tem um sentido único: é provável que nossos juízos morais influenciem, em troca, nossas reações emocionais. O que julgamos moralmente ruim, por exemplo, tende a desenvolver em nós reações negativas.

De maneira geral, os debates recentes sugerem que nossas emoções e nossas faculdades racionais contribuem para a formação de nossos juízos morais. Resta estabelecer em quais proporções elas o fazem. Por se apoiar fundamentalmente na noção de emoção apropriada, o neo-sentimentalista permite uma avaliação das emoções em termos da sua racionalidade. Pode-se dizer o mesmo de sugestões recentes a respeito dos processos causais responsáveis pelos juízos morais. Segundo uma explicação muito popular, proposta por Shaun Nichols (2004), o mecanismo ao qual devemos nossos "juízos morais fundamentais" (relacionados à interdição de danos) depende ao mesmo tempo de nossas emoções e de nossas capacidades racionais.

Os filósofos se dedicam, do mesmo modo, ao estudo da função de emoções específicas, como a piedade, a compaixão e a simpatia (Blum 1980; Nichols 2004; Nussbaum 2001), o amor (Aristóteles 2004; Frankfurt 2004; Jollimore 2011; Velleman 1999), o respeito (Blum 1988; Darwall 2004), a vergonha (Taylor 1985; Deonna et al. 2011), culpabilidade (Bruun & Teroni 2011; Deigh 1999), o nojo

(Clark & Fessler 2014), a inveja (D'Arms & Jacobson 2005), o ressentimento e a indignação (Strawson 1962; Wallace 1994) ou mesmo a ansiedade (Kurth 2015). Mesmo que a ênfase seja posta principalmente no impacto motivacional, essas emoções levantam questões específicas quanto a sua importância moral.

Além disso, as emoções consideradas centrais variam segundo as teorias morais (Bagnoli 2011). Os deontologistas – que concebem a justeza das ações em termos de regras absolutas – atribuem um lugar central às emoções, e mais especificamente ao respeito kantiano pela lei moral (Hill 1992). Em uma perspectiva consequencialista – segundo a qual a ação justa é aquela que tem as melhores consequências – leva-se em conta, primeiramente, os sentimentos solidários como a piedade e a compaixão (Hume 1741) e as emoções que Mill (1969) chamava de "punições internas", como a vergonha e a culpabilidade. Enfim, as éticas da virtude – que se interessam menos pelo valor das ações que pelo valor dos agentes – sublinharam a importância de uma grande variedade de emoções no exercício das virtudes e dos vícios (Hursthouse 2002). A coragem, por exemplo, não implicaria tanto a ausência de medo, mas um grau apropriado de medo. Os teóricos da virtude também foram os primeiros a negar que as emoções sejam independentes de nossa vontade, para afirmar, ao contrário, que necessitam de toda a nossa atenção e até de uma forma de educação – uma ideia que remonta a Aristóteles, mas que cotinua atual (Aristóteles 1994; Kristjansson 2008; Nussbaum 2001; Roberts 2013; Taylor 1985).

Conclusão

A despeito do caráter central das emoções nos trabalhos dos grandes nomes da história da filosofia – Aristóteles, Descartes, Spinoza, Hobbes e Hume, para citar apenas alguns – o interesse dos filósofos contemporâneos pelas emoções é relativamente recente. No entanto, desde os trabalhos seminais de David Lyons (1980), Robert Gordon (1987) e Ronald de Sousa (1987), a literatura filosófica dedicada às emoções literalmente explodiu. De uma matéria considerada desprovida de importância, extremamente desordenada e desencorajadora de qualquer tentativa de teorização sistemática, as emoções se tornaram, em pouco

tempo, um dos terrenos favoritos dos filósofos, que reconhecem desde então a necessidade de aprofundar a compreensão do tema. Isso transparece claramente na variedade e riqueza dos questionamentos que estruturam as discussões filosóficas contemporâneas. Como se deve compreender as relações entre as emoções e os outros fenômenos afetivos? Qual a maneira característica das emoções nos conectarem ao nosso entorno? Se aceitamos a ideia segundo a qual as emoções mantêm uma relação estreita com as propriedades avaliativas, como modelar melhor essa relação? As emoções são capazes de preencher uma função epistemológica significativa perante juízos avaliativos? Como compreender as influências sociais, às quais elas se prendem inegavelmente? As emoções nos motivam essencialmente ou sua relação com a motivação é mais acidental? Qual a ligação entre os juízos morais e a motivação e como a natureza dessa ligação condiciona nossa concepção das relações entre as emoções e a moral? Nesse capítulo, sublinhamos em mais de uma perspectiva a rede de relações complexas em que se inscrevem essas questões. Sua pertinência para a atividade humana justifica largamente a retomada do interesse dos filósofos.

Referências

ARISTOTE. *Rhétorique* (trad. F. Midal). Paris: Agora, 2004.

_____.. *Ethique à Nicomaque* (trad. J. Tricot). Paris: Vrin, 1994.

ARMON-JONES, C. “The thesis of constructionism”. In: Harré, R. (éd.) *The Social Construction of the Emotions*. Oxford: Oxford University Press; 1986, p.32–56.

AVERILL, J. “A constructivist view of emotion”. In: Plutchik, R. & Kellermann, H. (éd.) *Emotion: Theory, Research and Experience*. Burlington: Academic Press; p.305–339, 1980.

AYDEDE, M. *Pain: New Essays on its Nature and the Methodology of its Study*. Cambridge, Mass.: MIT Press, 2006.

AYER, A. J. *Langage, vérité et logique*. Paris: Flammarion, 1936/1956.

BAGNOLI, C. (éd.). *Morality and the Emotions*. New York: Oxford University Press, 2011.

BAIER, A. “Feelings that matter”. In: Solomon, R. C. (éd.) *Thinking about Feeling: Contemporary Philosophers on Emotions*. New York: Oxford University Press, 2004, p.200-213.

BARRETT, L. F. “Emotions as natural kinds?” In: *Perspect Psychol Sci*, 1: p.28-58, 2006.

BAYNE, T. & MONTAGUE, M. (éd.). *Cognitive Phenomenology*. New York: Oxford University Press, 2011.

BEDFORD, E. “Emotions”. In: *Proc Aristotelian Soc*, 57: p.281-304, 1957.

BETZLER, M. “Expressive actions”. In: *Inquiry*, 52: p.272-292, 2004.

BLACKBURN, S. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Oxford University Press, 1984.

BLOCK, N. “On a confusion about a function of consciousness”. In: *Behavior and Brain Sciences*, 18: p.227-287, 1995.

BLUM, A. “On respect”. In: *Philos Inq*, 10: p.58-63, 1988.

_____. *Friendship, Altruism and Morality*. London: Routledge and Kegan Paul, 1980.

BRADY, M. *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press, 2013.

BRENTANO, F. *Les origines de la connaissance morale* (trad. J.-C. Gens). Paris: Gallimard, 1889/2003.

BROAD, C.D. “Emotion and sentimento”. In: Cheeney, D. (éd.) *Broad’s Critical Essays in Moral Philosophy*. New York: George Allen & Unwin, 1954.

BRUUN, O. & TERONI, F. “Shame, guilt and morality”. In: *J Moral Philos*, 8: p.221-243, 2011.

CANNON, W. B. “The James-Lange theory of emotions: A critical examination and an alternative theory”. In: *Am J Psychol*, 39: p.106-124, 1927.

CHARLAND, L. “The natural kind status of emotion”. In: *Br J Philos Sci*, 53: p.511-537, 2002.

_____. “The heat of emotion: Valence and the demarcation problem”. In: *J Consciousness Stud*, 12: p.82-102, 2005.

CLARK, J. & FESSLER, D. “The role of disgust in norms, and norms in disgust research: Why liberals shouldn’t be disgusted by moral disgust”. In: *Topoi*, 34: p.483-498, 2014.

CLORE, G. “Why emotions are felt”. In: Ekman, P. & Davidson, R. (éd.) *The Nature of Emotion*. New York: Oxford University Press; p.103-111, 1994.

COLOMBETTI, G. “Appraising valence”. In: *J Consciousness Stud*, 12: p.103-126, 2005.

COSMIDES, L. & TOOBY, J. “Evolutionary psychology and the emotions”. In: Lewis, M. & Haviland-Jones, J. (éd.). *Handbook of Emotions*. 2nd ed. New York: Guilford, 2000.

CRANE, T. “Intentionality as the mark of the mental”. In: O’Hear, A. (éd.) *Contemporary Issues in the Philosophy of Mind*. Cambridge: Cambridge University Press, 1998, p.229-251.

D’ARMS, J. & JACOBSON, D. “The moralistic fallacy: On the ‘appropriateness’ of emotions”. In: *Philos Phenomenol Res*, 61: p.65-90, 2000.

_____; _____. “Anthropocentric constraints on human value”. In: *Oxford Stud Metaethics*, 1: p.99-126, 2005.

_____; _____. “Demystifying sensibilities: Sentimental values and the instability of affect”. In: Dans P. Goldie (éd.), *The Oxford Handbook of Philosophy of Emotion*. New York: Oxford University Press, 2010, p.585-613.

DAMASIO, A. *L’erreur de Descartes: La raison des émotions*. Paris: Odile Jacob, 1994/2010.

_____. *Le sentiment même de soi: Corps, émotions, conscience*. Paris: Odile Jacob, 2000/2002.

DANIELSSON, S. & OLSON, J. Brentano and the buck-passers. *Mind*, 106: p.511-522, 2007.

DARWALL, S. “Respect and the second-person standpoint”. In: *Proc Address Am Philos Assoc*, 78, p.43-59, 2004.

DARWIN, C. *L’expression des émotions chez l’homme et chez les animaux*. CTHS: Paris, 1872/1998.

DEIGH, J. “Cognitivism in the theory of emotions”. In: *Ethics*, 104: p.824-854, 1994.

_____. “All kinds of guilt”. In: *Law Philos*, 18, p.313-325, 1999.

DEONNA, J. “Emotion, perception, and perspective”. In: *Dialectica*, 60, p.29-46, 2006.

_____. “The structure of empathy”. In: *J Moral Philos*, 4, p.99-116, 2007.

_____; RODOGNO, R.; TERONI, F. *In Defense of Shame: The Faces of an Emotion*. Oxford: Oxford University Press, 2011.

_____; TERONI, F. “Taking affective explanations to heart”. In: *Soc Sci Inform*, 48, p.359-377, 2009.

_____; _____. *The Emotions: A Philosophical Introduction*. New York: Routledge, 2012.

_____; _____. “From Justified Emotions to Justified Evaluative Judgements”. In: *Dialogue*, 51.1, p.55-77, 2012b.

_____; _____. “What Role for Emotions in Well-Being?” In: *Philosophical Topics*, 41.1, p.123-142, 2013.

_____; _____. “Getting bodily feelings into emotional experience in the right way”. In: *Emotion Review*, 2017.

DOKIC, J. & LEMAIRE, S. “Are emotions perceptions of value?” In: *Can J Philos*, p.227–247, 2013.

DÖRING, S. A. “Explaining action by emotion”. In: *Philos Q*, 53, p.214-230, 2003.

_____. “Affective perception and rational motivation”. In: *Dialectica*, 61: p.363-394, 2007.

_____. “Why recalcitrant emotions are not irrational”. In Todd, C. & Roeser, S. (éd.) *Emotion and Value*. Oxford: Oxford University Press, 2014.

DRETSKE, F. “Simple seeing”. In: Dans D. Gustafson & B. Tapscott (éd.), *Body, Mind, and Method*. Kluwer, 1979, p.1-15.

EKMAN, P. *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. New York: Times Books, 2003.

ELLSWORTH, P. “William James and Emotion: Is a Century of Fame Worth a Century of Misunderstanding?”. In: *Psychological Review*, 101.2, p.222-229, 1994.

ELSTER, J. *Alchemies of the Mind: Rationality and the Emotions*. New York: University Press, 1999.

FELDMAN BARRETT, L.; NIEDENTHAL, P.; WINKIELMAN, P. *Emotions and Consciousness*. New York: Guilford Press, 2005.

FRANKFURT, H. G. *Les raisons de l’amour* (trad. D. Dubroca & A. Pavia). Circé: Paris, 2004/2006.

FRIEDMAN, M. “Autonomy and the Split-Level Self”. In: *Southern Journal of Philosophy*, 24, p.19-35, 1986.

FRIJDA, N. *The Emotions*. Cambridge: Cambridge University Press, 1986.

GEACH, P. *God and the Soul*. Londres: Routledge and Kegan Paul, 1969.

GIBBARD, A. *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press, 1990.

GOLDIE, P. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press, 2000.

_____. “Emotions, Feelings and Intentionality”. *Phenomenology and the Cognitive Sciences*, 1, p.235-254, 2002.

_____. “Misleading emotions”. In Brun, G., Doguoglu, U. & Kuenzle, D. (éd.) *Epistemology and the Emotions*. Aldershot: Ashgate, 2008, p.149-165.

GOLDSTEIN, I. “Are emotions feelings? A further look at hedonic theories of emotions”. In: *Consciousness Emotion*, 3, p.21-33, 2003.

GORDON, R. *The Structure of Emotions: Investigations in Cognitive Psychology*. Cambridge: Cambridge University Press, 1987.

GREENSPAN, P. S. *Emotions and Reasons*. New York: Routledge, 1988.

GRIFFITHS, P. *What Emotions Really Are: The Problem of Psychological Categories*. Chicago: University of Chicago Press, 1997.

HAIDT, J. “The emotional dog and its rational tail: A social intuitionist approach to moral judgment”. In: *Psychol Rev*, 108, p.814–834, 2001.

HATZIMOYSIS, A. “The case against unconscious emotions”. In: *Analysis*, 67, p.292-299, 2007.

HELM, B. *Emotional Reason: Deliberation, Motivation, and the Nature of Value*. Cambridge: Cambridge University Press, 2001.

_____. “Emotions and motivation: reconsidering neo-Jamesian accounts”. In Goldie, P. (éd.) *Oxford Handbook of Philosophy of Emotion*. Oxford: Oxford University Press, 2010.

_____. “Emotions and recalcitrance: Reevaluating the perceptual model”. In: *Dialectica*, 69.3, p.417-433, 2015.

HILL, T. E. “*Dignity and Practical Reason in Kant’s Moral Theory*”. In: Ithaca: Cornell University Press, 1992.

HOBBES, T. *Léviathan*. Paris: Gallimard, 1651/2000.

HUME, D. *Dissertation sur les passions: Traité de la nature humaine Livre II*. Paris: Gallimard, 1741/2015.

HURSTHOUSE, R. “Arational actions”. In: *J Philos*, 88, p.57-68, 1991.

_____. *On Virtue Ethics*. Oxford: Oxford University Press, 2002.

JACKSON, F. “What Mary did not know”. In: *Journal of Philosophy*, 83, p.291-295, 1986.

JÄGER, C.; BARTSCh, A. “Meta-emotions”. In: *Grazer Philosophische Studien*, 73, p.179-204, 2006.

JAMES, W. “What is an emotion?”. In: *Mind*, 9, p.188-205, 1884.

_____. *The Principles of Psychology*. New York: Dover, 1890/1950.

JOHNSTON, M. “The authority of affect”. In: *Philos Phenomenol Res*, 53: 181-214, 2001.

JOLLIMORE, T. *Love’s Vision*. Princeton: Princeton University Press, 2011.

JONES, K. “Emotion, Weakness of Will, and the Normative Conception of Agency”. In: Hatzimoyis, A. (éd.) *Philosophy and the Emotions*. Cambridge: Cambridge University Press, 2003.

_____. “Quick and smart? Modularity and the pro-emotion consensus”. In: *Can J Philos*, 36: p.3-27, 2007.

KANT, I. *Fondements de la métaphysique des mœurs* (trad. V. Delbos). Paris: Le Livre de Poche, 1785/1993.

KENNY, A. *Action, Emotion and Will*. London: Routledge and Kegan Paul, 1963.

KIND, A. “The case against representationalism about moods”. In: Kriegel, U. (éd.) *Current Controversies in Philosophy of Mind*. New York: Routledge, 2014.

KRIEGEL, U. “Towards a new feeling theory of emotion”. In: *Eur J Philos*, 22, p.420-442, 2014.

KRISTJANSSON, K. *The Self and its Emotions*. New York: Cambridge University Press, 2008.

KURTH, C. “Moral Anxiety and Moral Agency”. In: Timmons, M. (éd.) *Oxford Studies in Normative Ethics*. Oxford: Oxford University Press, 2015, p.171-195.

LACEWING, M. “Do unconscious emotions involve unconscious feelings?” In: *Philos Psychol*, 20, p.81-104, 2007.

LAMBIE, J.; MARCEL, A. “Consciousness and the variety of emotion experience: a theoretical framework”. In: *Psychol Rev*, 109, p.219-259, 2002.

LEIGHTON, S. “A new view of emotion”. In: *Am Philos Q*, 22, p.133-141, 1985.

LYONS, W. *Emotion*. Cambridge: Cambridge University Press, 1980.

MADDELL, G. “Emotion and feeling”. In: *Proc Aristotelian Soc*, 71(suppl): p.147-162, 1997.

MCDOWELL, J. “Values and secondary qualities”. In: Honderich, T. (éd.) *Morality and Objectivity*. London: Routledge; p.110-129, 1985.

MENDELOVICI, A. “Pure intentionalism about moods and emotions”. In: Kriegel, U. (éd.) *Current Controversies in Philosophy of Mind*. New York: Routledge, p.135-157, 2014.

MILL, J. S. *L’utilitarisme*. Paris: Flammarion, 2008.

MULLIGAN, K. “From Appropriate Emotions to Values”. In: Monist, 81, p.161-188, 1998.

NAAR, H. “A dispositional theory of love”. In: *Pacific Philosophical Quarterly*, 94, p.342-357, 2013.

_____; TERONI, F. (éd.) *The Ontology of Emotions*. New York: Cambridge University Press, 2017.

NICHOLS, S. *Sentimental Rules: On the Natural Foundations of Moral Judgement*. New York: Oxford University Press, 2004.

NUSSBAUM, M. C. *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press, 2001.

_____. “Emotions as judgments of value and importance”. In: Solomon, R. C. (éd.) *Thinking About Feeling: Contemporary Philosophers on Emotions*. New York: Oxford University Press, 2004, p.183-199.

PEARS, D. *Questions in the Philosophy of Mind*. London: Duckworth, 1975.

PELSER, A. “Emotion, Evaluative Perception, and Epistemic Justification”. In: Dans S. Roeser & C. Todd (éd.), *Emotion and Value*. Oxford: Oxford University Press, 2014, p.106-122.

PITCHER, G. “Emotions”. In: *Mind*, 74, p.326-346, 1965.

PLAMPER, J.; TRIBE K. *The History of Emotions: An Introduction*. New York: Oxford University Press, 2015.

PLATON. *Protagoras; Euthydème; Gorgias; Ménexène; Ménon; Cratyle* (trad. E. Chambry). Paris: Flammarion, 2016.

PRICE, C. "Affect without object: Moods and objectless emotions". In: *European Journal of Analytic Philosophy*, 2 (1), p.49-68, 2006.

_____. *Emotion*. Malden, MA: Polity Press, 2015.

PRINZ, J. *Gut Reactions: A Perceptual Theory of Emotions*. New York: Oxford University Press, 2004a.

PRINZ, J. "Which emotions are basic?". In: Evans, D. & Cruse, P. (éd.) *Emotion, Evolution and Rationality*. New York: Oxford University Press, 2004b, p.69-87.

_____. *The Emotional Construction of Morals*. New York: Oxford University Press, 2008.

PUGMIRE, D. *Rediscovering Emotions*. Edinburgh: Edinburgh University Press, 1998.

RABINOVICZ, W.; RØNNOW-RASMUSSEN, T. "The strike of the demon: On fitting pro-attitude and value". In: *Ethics*, 114, p.391-423, 2004.

RATCLIFFE, M. "William James on Emotion and Intentionality". In: *International Journal of Philosophical Studies*, 13 (2), p.179-202, 2005.

REISENZEIN, R. "What is an emotion in the belief-desire theory of emotion?". In: Paglieri, F., Falcone, R. & Micel, M. (éd). *The Goals of Cognition: Essays in Honor of Cristiano Castelfranchi*. London: College Publications, 2012.

ROBERTS, R. "Propositions and animal emotion". In: *Philosophy*, 71, p.147-156, 1996.

_____. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press, 2003.

_____. *Emotions in the Moral Life*. New York: Cambridge University Press, 2013.

_____; WOOD, W. J. *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press, 2007.

ROBINSON, J. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art*. New York: Oxford University Press, 2005.

RONNØW-RASMUSSEN, T. *Personal Value*. New York: Oxford University Press, 2011.

ROSENWEIN, B. *Generations of Feeling. A History of Emotions, 600-1700*. Cambridge: Cambridge University Press, 2016.

RUSSELL, J. "Core affect and the construction of emotions". In: *Psychol Rev*, 110, p.145-172, 2003.

RYLE, G. *La notion d'esprit* (trad. S. Stern-Gillet). Paris: Payot, 1949/2005.

SALMELA, M. "Can emotions be modelled on perception?". In: *Dialectica*, 65: p.1-29, 2011.

_____. *True Emotions*. Amsterdam: John Benjamins, 2014.

SCANLON, T. *What We Owe to Each Other*. Cambridge: Harvard University Press, 1998.

SCARANTINO, A. "Basic emotions, psychological construction and the problem of variability". In: Russell, J. & Barrett, L. (éd.) *The Psychological Construction of Emotion*. New York: Guilford Press; 2014, p.334-376.

_____; GRIFFITHS, P. "Don't give up on basic emotions". In: *Emotion Review*, 3: p.444-454, 2011.

SCHACHTER, S.; SINGER, J. "Cognitive, social, and physiological determinants of emotional states". In: *Psychol Rev*, 69, p.379-399, 1962.

SEARLE, J. *L'intentionnalité: Essai de philosophie des états mentaux* (trad. C. Pichevin). Paris: Minuit, 1983/1985.

SMITH, M. *The Moral Problem*. Oxford: Blackwell, 1994.

SOLOMON, R. *The Passions: Emotions and the Meaning of Life*. Indianapolis: Hackett, 1993.

_____. "Back to basics: On the very idea of "basic emotions"". In: *Not Passion's Slave*. New York: Oxford University Press; 2011, p.115-142.

DE SOUSA, R. *The Rationality of Emotion*. Cambridge: MIT Press, 1987.

STEVENSON, C. L. "The emotive meaning of ethical terms". In: *Mind*, 46, p.14-31, 1937.

STRAWSON, P. F. "Freedom and resentment". In: *Proc Br Acad*, 48, 1962, p.1-25.

TAPPOLET, C. *Emotions et valeurs*. Paris: Presses Universitaires de France, 2000.

_____. "Emotions, action, and motivation : the case of fear". In: Goldie, P. (éd.) *Oxford Handbook of Philosophy of Emotion*. Oxford: Oxford University Press, 2010.

_____. "Values and emotions: Tthe prospects of neo-sentimentalism". In: Bagnoli, C. (éd.) *Morality and the Emotions*. Oxford: Oxford University Press, 2011, p.177-134.

_____. "Emotions, perceptions, and emotional illusions". In: Dans C. Calabi (éd.), *The Crooked Oar, the Moon's Size and the Kanizsa Triangle: Essays on Perceptual Illusions*. Houndsmill: Palgrave-Macmillan, 2012, p. 207-224.

_____. "Emotions, reasons and autonomy". In: Dans A. Veltman & M. Piper (éd.), *Autonomy, Oppression and Gender*. New York: Oxford University Press, 2014, p.163-180.

_____. *Emotions, Values and Agency*. New York: Oxford University Press, 2016.

_____. "The metaphysics of moods". In: Dans H. Naar & F. Teroni (éd.), *The Ontology of Emotions*. New York: Cambridge University Press, 2017.

_____; ROSSI, M. "Emotions and wellbeing". In: *Topoi*, 34, p.461-474, 2014.

TAYLOR, G. *Pride, Shame and Guilt*. Oxford: Oxford University Press, 1985.

TERONI, F. “Emotions and formal objects”. In: *Dialectica*, 61, p.395-415, 2007.

_____. “Plus ou moins: Emotions et valence”. In: Dans C. Tappolet, A. Konzelmann et F. Teroni (éd.), *Les ombres de l’âme: Penser les émotions négatives*. Genève: Markus Haller, 2011.

_____. “Emotions et connaissance”. In: Dans J. -M. Chevalier & B. Gaultier (éd.), *Connaître: Questions d’epistémologie contemporaine*. Paris: Ithaque, 2014, p.9-34.

_____. “In pursuit of emotional modes: The philosophy of emotion after James”. In: Dans A. Cohen & B. Stern (éd.), *Thinking about the emotions: A philosophical history*. Oxford: Oxford University Press, 2017.

_____; DEONNA, J. “Emotions as attitudes". In: *Dialectica*, 69, p.293-311, 2015.

TYE, M. “The experience of emotion: an intentionalist theory”. In: *Rev Int Philos*, 243, p.25-50, 2008.

VELLEMAN, J. D. “Love as a moral emotion”. In: *Ethics*, 109, p.338-374, 1999.

WALLACE, R. J. *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press, 1994.

WALTON, K. “Fearing fiction”. In: *J Philos*, 75, p.5-27, 1978.

WHITING, D. “The feeling theory of emotions and object-directed emotions”. In: *Eur J Philos*, 19, p.281-303, 2011.

WOLLHEIM, R. *On the Emotions*. New Haven: Yale University Press, 1999.

WRINGE, B. “The contents of perception and the contents of emotion”. In: *Noûs*, 49, p.275-297, 2014.

ZAGZEBSKI, L. “Emotion and moral judgment”. In: *Philos Phenomenol Res*, 66, p.104-124, 2003.

3

A SOBERANIA DOS AFETOS: UMA CRÍTICA AO UTILITARISMO A PARTIR DA EMPATIA E DA PSICANÁLISE

Érico Andrade*

Introdução

Considerando que para uma teoria utilitarista a felicidade consiste na satisfação dos interesses dos agentes morais ou simplesmente na rejeição do sofrimento por parte desses agentes, meu objetivo neste capítulo consiste em mostrar que o utilitarismo precisa de uma compreensão do bem que seja mais complexa e que não se encerre na procura pela *felicidade.* Defenderei que essa compreensão mais complexa é fundamental para a adoção do cálculo de utilidade. Entre os utilitaristas há uma tendência a considerar que a felicidade não contém em sua extensão os afetos no que concerne, pelo menos, à predicação da moralidade de uma ação. Acredito que essa tendência é a marca da falta de uma compreensão mais ampla das diversas motivações para as ações humanas. Nesse sentido, a minha crítica se concentra na visão austera do utilitarismo quanto ao que pode ser tomado como critério para a avaliação de ações.

Minha hipótese é a de que o utilitarismo parte de uma tese razoável, a saber, a de que os agentes morais têm em princípio o mesmo valor – nas palavras de Bentham: "todos devem contar por um, ninguém por mais de um"[1] – para concluir, de forma equivocada, que essa igualdade de valor se mantém *a posteriori* em qualquer contexto e assume a forma de uma constante para o cálculo racional.

---

*Uma versão ligeiramente modificada deste texto foi publicada na revista Trans/Form/Ação, 36.2 (2013): 105-122. Agradeço a gentileza da revista em permitir a publicação e divulgação do texto.

[1] Bentham *apud*. MILL 1861.

Ou seja, o fato de que todos os agentes morais podem, em princípio, ter um mesmo valor moral não implica que, seja qual for o contexto ou a situação, os agentes morais são todos iguais, pois é perfeitamente possível que um agente moral sacrifique a solução *ótima* (proposta pelo cálculo utilitário e que privilegia a felicidade da maioria) em nome de uma motivação afetiva, a qual potencializa, num determinado contexto, o valor de um agente moral em detrimento de outros.

Diferentemente de várias objeções clássicas ao utilitarismo, mapeadas muito bem por Cláudio Costa (2002), meu ponto é que a compreensão da felicidade só pode ocorrer no interior de um contexto no qual se determina a motivação da ação e os valores sociais (instituídos no horizonte do tempo) que lhe são agregados. Minhas objeções, portanto, não questionam o fato de que a felicidade (bem-estar) pode ser um dos nossos bens morais relevantes, mas meu ponto é que ela não é o único bem, nem se apresenta historicamente de maneira uniforme. Assim, parece-me que o utilitarismo elide uma discussão mais ampla sobre a natureza da felicidade e sobre as razões que movem a ação humana, o que o torna pouco eficiente no que concerne à resolução de alguns dilemas morais.

Para realizar a crítica proposta aqui, estruturarei o capítulo em três etapas. Primeiro, farei breves considerações sobre o que quero designar com o termo 'utilitarismo'. Estou ciente, contudo, de que não há consenso quanto ao emprego desse termo e que certamente as minhas críticas podem não tocar alguns matizes do utilitarismo. Com efeito, as minhas objeções se dirigem mais particularmente à posição utilitarista de Cláudio Costa. O autor supramencionado é um dos raros filósofos no Brasil a apresentar uma defesa clara do utilitarismo e é um potencial interlocutor deste escrito. Se as minhas críticas apresentarem boas questões contra sua posição, julgo que meu trabalho terá obtido êxito em parte considerável de seu intento. Em seguida, tecerei algumas considerações sobre a motivação das ações individuais, no que chamo de discussão antropológica. Essa discussão visa mostrar que a compreensão utilitarista do homem como uma espécie de *átomo de prazer* homogeneíza equivocadamente as motivações humanas à medida que estabelece uma hierarquia – questionável – entre os diferentes sentimentos

humanos e que geralmente destina um papel pouco relevante para a empatia no que diz respeito à predicação da moralidade[2]. Nesse caso, o utilitarismo tem um déficit antropológico por desconsiderar que o altruísmo (por razões neurológicas e afetivas) pode ser tomado como um valor moral que legitima o autossacrifício ou que justifica uma ação dissonante da procura pela própria felicidade. No terceiro e último momento do capítulo, farei uma crítica, chamada por mim de sociológica, por meio da qual defenderei que o utilitarismo tem um importante déficit cultural, visto que, para a resolução de conflitos interculturais, é necessário reconhecer a diversidade de valores culturais. Ela, no entanto, é negligenciada pelo utilitarismo na medida em que este homogeneíza as culturas por meio do apelo à felicidade da maioria. O "truque" do utilitarista é promover a desvinculação entre o indivíduo (cuja racionalidade seria uma marca constituinte independente da história) e a cultura (cuja variação não molda o desejo do indivíduo pela felicidade). Essa desvinculação permite ao utilitarismo tratar agentes morais de diferentes culturas de modo uniforme, como se todos os agentes devessem ter uma mesma compreensão da felicidade, mas o afasta da resolução de conflitos entre culturas, porque subordinada a compreensão da cultura a um agente moral desinteressado e, portanto, não põe em debate a função da cultura na determinação da moralidade. Concluirei que falta ao utilitarismo, por um lado, uma visão menos austera do agente moral e, por outro, um senso sociológico mais acurado.

---

[2] Tenho plena consciência de que Mill tece algumas críticas ao utilitarismo proposto por Bentham que se remetem ao questionamento do aspecto austero da compreensão antropológica do ser humano aventada por aquele autor. Ainda que Mill reconheça o papel das emoções, ele ainda se mantém no registro de que o utilitarismo, em última análise, não introduz nos seus princípios básicos as emoções como critério de avaliação da ação moral. M. C. M. Dias defende, em função de uma análise estrutural e genética, que Mill assume um novo posicionamento quanto ao papel das emoções na moralidade após seu problema mental. Ela se apoia, sobretudo, na Autobiografia de Mill, que efetivamente confessa sua crítica ao utilitarismo que se coloca indiferente às emoções (DIAS, 2011, p.72-73). O ponto central da autora consiste na defesa de que a antropologia de Mill é diferente daquela sustentada por Bentham, o que desembocaria em visões, em certa medida, dissonantes do utilitarismo. Alguns defendem que o utilitarismo de Mill é anômalo.

1. Breves notas sobre o utilitarismo

Eu compreendo por *utilitarismo*, em termos éticos, a teoria moral que defende a otimização da felicidade (bem estar) em detrimento do sofrimento. Não gostaria de sublinhar com essa definição apenas o utilitarismo de matiz hedonista clássico, pois acredito que ela é capaz de abarcar variações mais sofisticadas do hedonismo que ainda trabalham no registro de que o ponto essencial de uma teoria utilitária repousa na promoção *otimizada* da felicidade[3]. Assim, se entre duas alternativas de ação há uma que promove a maior felicidade, esta é a alternativa que o agente moral deve seguir[4]. Nesse sentido, o utilitarismo não é, segundo Mill, "a maior soma da felicidade do próprio agente, mas a maior soma da felicidade conjunta"[5]. O utilitarismo pensa a ação moral em função das consequências dessa ação para a felicidade de um conjunto o maior possível de pessoas[6]. Essa posição subsidia efetivamente um cálculo para aferir em que medida uma ação teve ou não uma consequência boa, pois a variável é a quantidade de pessoas beneficiadas. O mapeamento empírico das possíveis consequências da ação torna exequível o cálculo que determina o maior número de beneficiados. Tenho consciência, contudo, de que não há consenso entre os utilitaristas quanto à natureza da felicidade ou do bem estar, mas várias definições da felicidade tendem a assumir uma das diferentes matizes do hedonismo.

Se a felicidade é o bem moral para o qual tendem as nossas ações individuais, de um ponto de vista social as ações devem estender a felicidade para o maior número possível de pessoas. Isto é, os indivíduos devem procurar a maximização da felicidade e a sociedade deve estender ao máximo os benefícios entre os indivíduos de modo a contemplar um número maior de indivíduos. Assim,

---

[3] Mill é responsável por uma classificação dos prazeres, mas não deixa de assumir o binômio prazer e sofrimento como o principal vetor do utilitarismo (MILL, *Utilitarismo*, Cap. II). Sobre a qualificação do prazer em Mill, ver SIMÕES, 2009, p.48.

[4] Essa definição é próxima da produzida por SINGER 2009, p.11.

[5] 1861, p.196.

[6] Ver, por exemplo, SMART 1956, p.354.

o componente social obriga que as decisões individuais sejam sintonizadas com o desejo da maioria. O utilitarismo defende que se *deve* inserir o interesse dos outros nos meus interesses[7]. Por isso, não me parece muito correto defender que o utilitarismo é uma teoria do egoísmo moral em sentido estrito.

Nessa perspectiva, Cláudio Costa argumenta que uma formulação consistente da máxima utilitária ou do princípio da utilidade tem que levar em consideração a ideia de contrato. Ele retoma uma das formulações da referida máxima: "a ação moralmente correta é a que segue uma regra cuja adoção produz um bem maior para a sociedade que adota o sistema de regras à qual ela pertence"[8]. Esse princípio ou máxima utilitária precisa, seguindo a argumentação de Cláudio Costa, de um adendo para evitar a objeção ao utilitarismo de regras, que preconiza a necessidade de se transgredir algumas normas sociais que oprimem, em vários casos, a maioria. Assim, ele defende uma aplicação geral do princípio supracitado mediante a sua coordenação com o princípio clássico do utilitarismo conforme o qual "uma ação moralmente correta é a que produz maior prazer (bem) e/ou menor sofrimento (mal) para a maioria"[9].

O princípio da utilidade defende que cada ação particular (utilitarismo de atos) ou padrão de ações (utilitarismo de regras) deve agregar e distribuir ao máximo a felicidade ou o bem-estar. As ações são pensadas sempre no horizonte da otimização dos bens que devem, sempre que possível, ser distribuídos de forma equitativa, uma vez que não há no agente moral uma qualidade que possa

---

[7] A ideia de que a ação do indivíduo deve ser compatível com os interesses dos demais torna possível, segundo Hare, uma aproximação parcial com o imperativo categórico de Kant, mais precisamente nas formulações que prescrevem deveres para com os outros (HARE, 1998 p.212). Sobre a crítica da assimilação da filosofia kantiana como um utilitarismo bastante mitigado, por assim dizer, ver Dall'Agnol (2001). Dall'Agnol insiste, entre outras coisas, que a filosofia kantiana tem um princípio essencial que é incompatível com o utilitarismo, qual seja: o ser racional é um fim em si mesmo. Um dos seus argumentos centrais é que a filosofia kantiana prioriza o respeito à pessoa, em detrimento, eventualmente, de sua felicidade. Para o autor, nem Kant pode ser utilitarista, nem Hare pode ser kantiano.

[8] COSTA, 2002, p.169.

[9] COSTA, 2002, p.164.

diferenciá-lo dos demais agentes. Isto é, a homogeneização do agente moral implica a extensão de um benefício a um maior número possível de agentes morais de modo razoavelmente uniforme. As vantagens da posição utilitarista no geral estão circunscritas, pelo menos, a dois pontos: A) ele parece estar de acordo com alguma de nossas intuições básicas quando oferece uma resposta razoável para certos dilemas morais. Por exemplo, considerando que haverá uma enchente após uma abertura imprescindível de uma barragem prestes a se romper, e que algumas pessoas precisam ser desalojadas para que toda a cidade possa sobreviver por meio de uma construção de outra barragem, se pode julgar razoável que algumas pessoas sejam desabrigadas para que as demais sobrevivam. B) O utilitarismo tem um comprometimento empírico interessante, na medida em que estabelece princípios morais que podem ser quantificáveis e testados, quanto à sua validade, na experiência. A atitude utilitária é exequível. Ainda que apresentadas de forma geral, essas razões servem de indicativo da força do utilitarismo que sempre está presente nos debates da ética contemporânea.

A noção de *bem-estar* ou felicidade social é suficientemente ampla para abarcar a otimização do prazer em detrimento da dor, inscrita nos interesses individuais. De qualquer modo, o indivíduo é considerado como o agente moral que age em conformidade com seu próprio bem-estar ou felicidade. Contudo, por que a escolha da otimização da felicidade ou bem-estar deve governar a avaliação de cada ação no que concerne ao uso dos predicados morais?

O epicentro da tese do utilitarismo é a aposta na igualdade do valor dos agentes morais, que compreende cada agente como um átomo que visa invariavelmente à realização dos seus desejos ligados, sobretudo, ao não sofrimento. Como os indivíduos têm um mesmo valor moral, as ações são orquestradas pelo princípio que anula ações que se oponham à maioria, ao passo que "recompensa" as ações que conseguem contemplar a maioria dos agentes morais. O reconhecimento da igualdade de valor dos agentes morais implica a construção de uma moral que avalia as ações em virtude do maior benefício que ela pode trazer para uma quantidade maior de indivíduos. Contudo, permanece a questão referente ao reconhecimento dos indivíduos do igual valor de cada agente

moral. O que levaria os indivíduos a elegerem como principal valor moral a igualdade de valor de cada agente moral independente do contexto? Essa igualdade está presente em qualquer contexto? Essa questão é essencial porque toca na motivação do agente moral para procurar a felicidade ou o bem-estar, visto que é necessário reconhecer moralmente um valor para que ele possa ser o critério de avaliação da ação moral. Defenderei agora que o pressuposto de que cada indivíduo tem um mesmo valor moral, embora sedutor, falha em alguns casos por idealizar o agente moral, desconsiderando a potencial moralidade de ações que envolvem, sobretudo, a empatia pelo outro e os afetos de modo geral.

2. O déficit antropológico do utilitarismo

O utilitarismo pressupõe que os indivíduos são capazes de separar interesses difusos e investirem numa atitude racional que planifica as suas escolhas e hierarquiza os seus prazeres em sintonia com a disponibilidade dos recursos escassos. Por isso, os indivíduos são tomados como átomos racionais que devem agir conforme uma determinação de um cálculo de otimização de recursos. Ainda que os indivíduos possam, em vários contextos, levar em consideração o bem-estar da maioria – enquanto horizonte, porque nem sempre é viável fazer o cálculo do bem-estar, como bem pontua Williams[10] – o ponto que quero ressaltar reside nos problemas de tomar a regra do cálculo utilitário como um critério para a avaliação das ações independente do contexto afetivo no qual elas ocorrem. Assim, se, por um lado, o utilitarismo está de acordo com algumas de nossas intuições, como no exemplo que ofereci da represa, por outro, ele se afasta de algumas outras intuições que me parecem fundamentais. Destaco particularmente a empatia, não apenas como um elemento de motivação moral, mas como um dos critérios de avaliação moral.

Entendo por *empatia* toda disposição para agir em sintonia com um grau de altruísmo que implica algum sacrifício do próprio interesse em nome de outra pessoa. Em situações extremas, é possível que indivíduos façam alguns sacrifícios

[10] 2005, p.151-156.

não propriamente em função da atitude mais racional (no sentido utilitarista), mas em função de um interesse difuso ou que não contribui necessariamente com a maioria. Será que seria possível dizer que *a priori* esse indivíduo não procedeu moralmente de forma correta?

Deixem-me oferecer um exemplo: numa casa que está sendo destruída pelas chamas, um pai tem a opção de fugir quase ileso e garantir certamente o bem estar de três dos seus quatro filhos, que não estão na casa, ou ficar na casa, tentando resgatar o outro filho, mas com uma taxa pequena de chances de sobreviver. Não é necessário traçar um grande levantamento empírico para constatar que boa parte dos pais tomaria a decisão de tentar salvar o filho, ainda que isso implicasse que os outros filhos ficariam, de algum modo, desassistidos. O ponto é que a decisão do pai não é racional (no sentido do utilitarismo), mas é bastante compreensível do ponto de vista do vínculo afetivo que ele guarda com o seu filho. Não é *a priori* contraintuitivo tomar decisões que não se pautam num cálculo racional que otimiza o bem-estar. Elas podem ser bem intuitivas, como acredito que seja o hipotético caso mencionado, e podem eventualmente servir para a avaliação da ação moral. Nesse sentido, não se trata aqui do que Smart (1978) chama de "fraqueza de vontade", pois não seria uma posição meramente particular que mostraria, supostamente, mais uma ação sintonizada com uma dificuldade de agir segundo o cálculo de utilidade do que uma ação recomendável moralmente. O ponto é que a empatia é uma das variáveis para a motivação moral à proporção que também serve como um dos critérios para avaliar a ação. A ação de salvar o filho, conforme narrei aqui, é socialmente louvada e não se restringe a uma "fraqueza de vontade" individual.

Não pretendo argumentar que a empatia é o único critério para a avaliação moral, mas apenas que ela não pode ser completamente excluída como critério moral. Não pretendo entrar na ceara das discussões de metaética sobre a força moral da empatia ou sobre em que medida ela é um sentimento moral imprescindível, mas acredito que é necessário que o utilitarismo justifique o critério da escolha do cálculo racional sem desconsiderar as objeções apresentadas aqui referentes à presença dos sentimentos morais na avaliação da ação. Meu ponto é

que o utilitarismo se compromete com uma rígida noção de indivíduo (átomo racional) que nem sempre está em sintonia com a pulverização de interesses que marca a ação moral e que torna os agentes morais um feixe de múltiplas motivações.

Minha intenção, com o exemplo, consiste em mostrar que a quantidade de beneficiados não pode *a priori* ser condição necessária e suficiente para avaliar moralmente uma ação. O utilitarismo se compromete com uma individuação do agente moral que sedimenta apenas uma das feições da ação humana, inscrita na racionalidade responsável pela maximização do bem-estar ou da felicidade. Ora, se não é natural ou absolutamente espontâneo que nossas ações sejam pautadas apenas na maximização do bem-estar, é necessário avaliar em que medida uma compreensão unívoca do bem-estar ou da felicidade pode ser um critério moral suficientemente abrangente. É necessário, portanto, que se ofereçam outras qualidades ao critério da quantidade de beneficiados para que ele possa ser legitimado como tribunal de julgamento da ação moral. Em suma, é preciso qualificar o agente moral racional para avaliar, não se a sua conduta foi ou não a mais correta quanto à extensão dos beneficiados que ela abarca, mas se moralmente o referido benefício é um bem moral relevante. O critério da quantidade não é suficiente para justificar a adoção do cálculo de utilidade.

A questão que coloco aqui não versa apenas sobre a excessiva idealização do agente moral como um agente pronto para executar um cálculo racional perfeito. A minha questão é se é possível afirmar que o agente moral idealizado se justifica pelas consequências de suas atitudes, cuja sintonia com a felicidade da maioria seria necessária e suficiente para predicar a ação moral, prescindindo de qualquer outro critério que justifique a desigualdade do valor moral do agente em certos contextos. Em geral, os utilitaristas assumem a vida ou o não atentado à vida (o sofrimento) como critério moral último, mas eles homogeneízam a vida a tal ponto que inviabilizam o valor moral das ações que estão em consonância com motivações que atravessam as nossas escolhas de forma profundamente enraizada em nossos sentimentos. O agente moral utilitário deve agir, como costumava defender Mill, como indivíduo desinteressado e estritamente imparcial,

cuja ação deve considerar que a sua vida tem o mesmo valor que a dos demais agentes morais (*Utilitarismo*, cap. II). O utilitarismo elide um fato importante, inscrito no reconhecimento de que os agentes morais podem ser valorados diferentemente em contextos em que o componente da empatia ganha uma particular relevância, como no exemplo que acabo de oferecer.

A confusão que o utilitarismo faz é entre a tese razoável de que nenhum agente moral detém *a priori* o monopólio sobre a escolha do outro ou que um agente não tem o direito de, por decisão do seu arbítrio, diferenciar-se moralmente dos outros agentes, com a obrigação moral de tomar todos os agentes morais como instâncias cujo valor moral é absolutamente igual em qualquer contexto. Ou seja, o fato de que todos os agentes morais podem, em princípio, ter um valor moral igual (ninguém tem o direito de aniquilar arbitrariamente ninguém) não implica que, seja qual for o contexto ou a situação, os agentes morais sejam todos iguais, pois é perfeitamente possível, e foi o que quis mostrar com o meu exemplo, que um agente moral sacrifique a solução *ótima* proposta pelo cálculo racional em nome de uma motivação afetiva que não pode ser reprovada moralmente de forma tão fácil e está longe de ser um "desperdício"[11].

Farei agora algumas considerações críticas ao utilitarismo por meio do recurso à radicalização de algumas de suas premissas. Mostrarei que as premissas do utilitarismo que se apoiam na experiência não são suficientes para a adoção do cálculo de utilidade como critério moral. Alguns estudos empíricos (como o que vou citar em seguida) mostram que a empatia desempenha um papel central na rejeição da atitude utilitária em alguns dilemas morais. Eles mostram uma grande dificuldade de pessoas sacrificarem alguém próximo a elas em nome

[11] Um sacrifício que não beneficia a maioria é um desperdício, segundo Mill. Nesse sentido, ainda que Mill reconheça que o sacrifício é uma atitude possível de ser empreendida pelos seres humanos, ele ressalta que, se ela não estiver em sintonia com o critério da utilidade, é um desperdício. Ele escreve: "A moralidade utilitarista reconhece nos seres humanos o poder de sacrificar o seu bem maior para o próprio bem dos outros. Ele só se recusa a admitir que o sacrifício é em si uma boa. Um sacrifício que não aumenta, ou tende a aumentar, a soma total de felicidade, que ela considera como desperdiçado"(*Utilitarismo,* cap. II).

de uma solução ótima que preserva uma maior quantidade de beneficiados. Dificilmente alguém empurraria, por exemplo, a mãe num trilho de trem para evitar que um trem fique desgovernado e tombe matando cem pessoas. Um recente artigo publicado na *Nature* corrobora a tese da empatia como importante variável que veda a adoção do cálculo de utilidade como único critério para avaliação moral. Esse estudo mostra que apenas pessoas que sofrem danos no córtex pré-frontal tomam, sem nenhuma hesitação, decisões utilitárias que privilegiam a maioria em detrimento de qualquer relação empática que possamos guardar com a alguém (ver Koenigs *et al.* 2007). Ora, essa é justamente a região do cérebro (junto com amígdala) responsável pelas emoções. Ou seja, a motivação para decisões que tomam a atitude utilitária como o único critério para a avaliação moral é seriamente comprometida pelo componente empático que parece mostrar uma indisposição dos homens para as ações em que o cálculo racional é o único critério válido. Uma pessoa não empurra a outra, sobretudo, quando se trata de alguém com algum grau de proximidade, porque a quantidade de beneficiados não a move necessariamente a tomar uma decisão utilitária independente do contexto em que ela esteja. A quantidade de beneficiados também não é suficiente para justificar moralmente a adoção do cálculo racional como critério de avaliação em qualquer contexto. De um ponto de vista da história evolutiva e da estrutura neurológica humana há sérios indícios empíricos de que não podemos homogeneizar o agente moral desconsiderando o feixe de múltiplas emoções ou sentimentos (entre elas destaco a empatia) que compõem a tomada de decisões morais. O estudo mostra que o cálculo utilitário só poderia ser um critério moral absoluto, caso tivéssemos sérios abalos em nosso sistema nervoso, que poderiam nos habilitar a prescindir das emoções ou sentimentos, como a empatia, nos nossos julgamentos morais. Nesse sentido, há uma resistência da experiência quanto à adoção do cálculo de utilidade como único critério moral.

Notem que não tenciono negar que, de um ponto de vista da definição de racionalidade como a otimização de recursos escassos, certas atitudes podem ser mais ou menos racionais. No entanto, isso não significa que elas sejam mais ou menos morais, caso não se leve em consideração outras variáveis na atribuição

dos predicados morais. Ainda que se aceite o critério da quantidade de beneficiados como tribunal para avaliação da ação quanto à sua moralidade, isso não isenta o fato de que deva ser justificado e, preferencialmente, deva estar em sintonia com algumas de nossas intuições, sobretudo, aquelas que estão ligadas às nossas relações empáticas. Com essas considerações acho que é possível mostrar que a compreensão do agente moral do utilitarismo é bastante austera e se distancia de várias motivações de nossas ações que não são completamente redutíveis ao cálculo da utilidade, nem podem ser completamente anuladas na avaliação de nossas ações. O ônus que cabe ao utilitarismo é ter de provar que a ação racional implica ação moral. Isso não é evidente.

Esse problema, na minha exposição, se voltou para uma espécie de déficit antropológico do utilitarismo que iguala os agentes morais de forma tão radical que desconsidera a potencialização do valor de certos agentes em certos contextos. A minha ideia é de que a igualdade inicial do valor moral (ninguém tem *a priori* o direito de destruir ninguém) não pode elidir a potencialização do valor de certos agentes quando se trata de uma decisão de se autossacrificar por uma razão de origem afetiva. O sacrifício em nome de outra pessoa sempre implica o aumento do valor de um agente moral que está sendo salvo por aquele sacrifício. Tentei mostrar que o agente moral é mais complexo, de sorte que a aceitação da tese de que todos os agentes morais podem ter, num primeiro momento, o mesmo peso na decisão quanto a uma determinada atitude, que invariavelmente implica outras pessoas, não significa que eles sejam todos absolutamente iguais *a posteriori*, ao ponto de que as suas ações tenham que visar apenas ao benefício de uma maioria quantitativamente determinada. Particularmente nos contextos em que não se trata de atentar contra a vida de ninguém, mas de escolher salvar alguém, os agentes morais podem agir em função de uma alta dose de envolvimento emocional que está em desacordo com o cálculo de utilidade. Tenciono mostrar agora que os problemas do utilitarismo não repousam apenas na sua austera compreensão do agente moral como um átomo racional, mas que eles residem também numa desconsideração da dimensão social e contextual da ação moral.

3. O déficit sociológico do utilitarismo

Vou enfatizar agora o problema de se eleger a quantidade de beneficiados em detrimento de um critério que qualifique o benefício no interior de um determinado contexto. Minha ideia é que o utilitarismo tem um déficit cultural importante relativo à falta de flexibilidade do critério de bem-estar ou felicidade. Os valores sofrem oscilações que configuram diferentes redes de uso dos predicados morais no transcorrer da história. É evidente que o utilitarista concede variações históricas do que é correto moralmente, mas meu ponto é que essas variações históricas precisam ser pensadas em função de um critério moral que promove essas mudanças em sintonia com o contexto em que elas ocorrem e que respeite as diferenças que se encontram no interior desse contexto. Ou seja, não está *dado* numa sociedade o que ela entende por felicidade.

Quanto ao reconhecimento de que os valores morais variam historicamente, Cláudio Costa oferece um exemplo: a mudança das lutas nas arenas dos circos romanos, que custavam as vidas de alguns para a alegria de vários, para as atuais lutas de boxe. Ele defende, com isso, que a sociedade pode corrigir distorções morais como o circo romano em sintonia com o refinamento do prazer que não precisaria implicar a dor acompanhada de morte[12]. No futuro, sociedades civilizadas não teriam, segundo aquele autor, lutas de boxe. Essas sociedades teriam outras formas de obter prazer que não passariam pelo sofrimento de alguns.

Vou me ater ao próprio exemplo oferecido por Cláudio Costa para mostrar algumas fragilidades do utilitarismo, quando pensamos em contextos culturais distintos. Primeiramente, gostaria de destacar que a uniformização da cultura, tanto romana, quanto atual, implícita no exemplo do referido autor, indica uma negligência recorrente dos utilitaristas quanto à natureza do bem no horizonte de uma cultura. Cláudio Costa aposta num desenvolvimento cultural – sem esclarecer muito bem o que isso significa – que terminaria por condenar moralmente as lutas de boxe e, por fim, extirpá-las, como ocorreu com o circo romano. Entretanto, dificilmente algumas sociedades ocidentais ou parte importante da população

[12] COSTA, 2002, p.169-170.

seriam hoje majoritariamente contra o circo romano, como sugere Cláudio Costa[13], se no lugar de gladiadores (escravos ou prisioneiros de guerra) tivéssemos estupradores e matadores em série de crianças. Meu contraexemplo não é uma forma velada de hierarquizar os agentes morais ou de julgar quem é mais ou menos importante socialmente. Ele consiste numa tentativa de mostrar que, sem uma análise sociológica mais acurada, o utilitarismo se afasta dos valores das culturas vigentes à medida que se compromete com valores morais injustificados socialmente ou para os quais ele não encontra uma justificativa assentida pelos membros da sociedade. Para que o exemplo do circo tenha validade, se deve levar em consideração quem é o "gladiador" submetido ao combate e qual é a sociedade em que o circo está inserido. Não é apenas um compromisso com o não sofrimento que garante que uma sociedade não adote estádios ou circos romanos nos quais as pessoas são dilaceradas por leões.

O problema da argumentação utilitarista, especificamente a de Cláudio Costa, é que ela tem dificuldade em relação às diferentes compreensões culturais do que é moralmente bom. O ponto não repousa na prerrogativa de que as pessoas têm mais ou menos prazer com o sofrimento alheio, mas consiste em investigar a instituição dos valores morais nas diferentes culturas e pensar em que medida é possível ter um corpo homogêneo de valores. Certamente, para se eleger o não sofrimento como um valor moral partilhável interculturalmente, é necessário uma justificativa mais clara, inclusive sobre o que significa sofrer. Novamente, é preciso qualificar mais o agente moral para que os indivíduos de diferentes culturas possam se reconhecer nesse agente.

O argumento naturalista, que é evocado por alguns utilitaristas para lastrear uma posição transcultural, ainda não me parece suficiente para justificar, pelo menos hoje, a adoção do cálculo de utilidade como critério de avaliação da moralidade de uma ação. Mesmo que evitar o sofrimento e promover a otimização do prazer possam ter sido uma boa estratégia evolutiva, isso não implica dizer que esse par conceitual continue sendo um elemento decisivo para pensar a condição

[13] 2002, p.169.

humana. Tudo indica que a conservação foi realizada por meio da planificação de nossas necessidades básicas que passam por evitar privações que muitas vezes geram sofrimento. Contudo, ainda que o utilitarismo possa ter a seu favor a tese de que os homens, de um ponto de vista da sua história evolutiva, e em vários momentos, agiram em sintonia com critérios que evitam o sofrimento e promovem o bem-estar social ou a felicidade – a partir do mecanismo de recompensa para os que agem em sintonia com a maioria e punição para aqueles que fazem o contrário (ver Burnham e Johnson 2005; Fehr e Gächter 2002) – ele não pode estender esse critério para toda a história e determinar de maneira *fixa* a própria noção de felicidade ou bem-estar. A validade do utilitarismo na história evolutiva não pode obliterar seu déficit sociológico face à cultura produzida pelos seres humanos nos últimos milênios.

Por não ser um critério inscrito de maneira inequívoca na condição humana enquanto uma variável fundamental para a sua adaptação ao mundo enquanto espécie, o critério do não sofrimento não está isento de uma justificativa moral para que seja adotado numa hierarquia de valores. Assim, se a história evolutiva pode ratificar alguma perspectiva utilitarista em certos momentos de nossa história e se a ação racional continua presente na estrutura do sistema capitalista, no qual os agentes sociais podem tender a otimizarem os recursos escassos, na história da humanidade é possível recuperar exemplos em que a lógica utilitária não se impõe como a única forma possível de ação moral.

Meu argumento consiste na ideia de que é muito pouco prudente pressupor uma evolução social uniforme que legitima a substituição das arenas dos circos romanos pelos ringues de boxe sem antes discutir no interior de uma cultura, com subculturas distintas, o que é moralmente bom. Claro que a diversidade de valores não implica a impossibilidade de um código moral, por assim dizer, universal ou mesmo que não sejam possíveis juízos objetivos em ética. Não se pode confundir o diagnóstico da diversidade de valores com a legitimação dessa diversidade ou mesmo com a inexorabilidade dessa diversidade. O ponto que tenciono sublinhar consiste no reconhecimento de que a diversidade cultural aponta para o fato de que a quantidade de beneficiados que podem, por assim dizer, evitar o sofrimento

não é necessariamente um critério moral que poderia unificar *per se* a diversidade de opiniões humanas sobre a moralidade em diferentes culturas. É preciso qualificar o benefício e os beneficiados. Isso, por sua vez, só pode ocorrer no contexto cultural. Se, por um lado, o relativismo cultural pode incorrer em contradições relevantes – ver, por exemplo, as críticas de Williams (2005) – o utilitarismo pode oferecer uma concepção moral pouco abrangente por subestimar a diferença cultural entre as sociedades que se dilata ao longo da história.

As relações interculturais não são *a priori* movidas por um único ponto de vista que é capaz de orquestrar interesses diversos em torno de uma mesma sinfonia. Se a busca pelo não sofrimento fosse um bem moral constante e transcultural, não teríamos códigos morais diferentes, em alguns casos antagônicos. A diversidade cultural está presente em cada decisão do agente moral e atravessa a sua pessoa como marca de sua individualidade. Por isso, não se pode elidir tão facilmente os conflitos interculturais em favor da escolha do cálculo de utilidade como critério moral que privilegia o não sofrimento. A aceitação do referido cálculo já implica a compreensão moral de que ele é a melhor forma de resolver conflitos culturais, no entanto, isso só pode ser estabelecido *a posteriori* quando de algum modo os envolvidos no conflito se disponham a recorrer a uma solução que otimize o bem-estar ou felicidade da maioria e que tenha clareza sobre o uso desses conceitos. O ponto é que, para chegar a essa resolução, os indivíduos já devem estar movidos por alguma prerrogativa moral que justifica a disposição para resolver o conflito de acordo com o cálculo de utilidade, visto que o simples cálculo não se configura como uma justificativa, pois é preciso antes decidir pela resolução do conflito e, por conseguinte, pelo recurso ao cálculo.

Sem partir de um acordo entre os membros de diferentes culturas, o utilitarismo tem o ônus de justificar seu critério independente da cultura à qual ele será aplicado. Se o cálculo de utilidade não se apoia no reconhecimento da diversidade cultural, ele termina enredando num círculo vicioso quando visa à fundamentação da moral. Para provar que o recurso ao cálculo de utilidade é um bem moral relevante, independente da cultura, o utilitarismo pressupõe que os agentes morais estão dispostos a serem racionais (imparciais) e realizarem

cálculos de utilidade. Contudo, os utilitaristas precisam provar que os agentes morais estão dispostos a serem racionais.

O raciocínio utilitarista parece se apresentar da seguinte maneira: porque o sujeito é racional (deseja otimizar o bem-estar), ele deve agir racionalmente (calculando a maximização do bem-estar), e deve agir racionalmente (otimizando o bem-estar), porque é racional (deseja maximizar o bem-estar). Para fugir a essa circularidade, o utilitarista precisa justificar moralmente que o cálculo racional é a melhor forma de resolver problemas morais que envolvem diferentes culturas. Para fazer isso, contudo, ele precisa explicar o critério por meio do qual ele elegeu o cálculo racional como norte da ação moral e não pressupor que os agentes morais são racionais e agem racionalmente porque são desterrados e absolutamente indiferentes ao contexto cultural no qual constroem e estipulam os seus valores.

Para os utilitaristas, os indivíduos *devem* agir em conformidade com o bem-estar por meio do recurso ao cálculo de utilidade. Contudo, por que aceitar um prognóstico (a normatividade do cálculo utilitário) cujo diagnóstico (a aposta de que os agentes morais são racionais) não reflete uma análise da diferença cultural que marca a produção de valores morais? Ou seja, por que os agentes morais deveriam abandonar os seus critérios de bem-estar ou felicidade em nome de um critério, materializado num contrato, que preza pela quantidade de beneficiados no que concerne ao não sofrimento? Ora, isso só poderia ser exigido, se os indivíduos se reconhecem independentemente da cultura naquele critério. Contudo, o utilitarismo não pode ser um horizonte para balizar conflitos culturais com alternativas morais relevantes, se ele, em última análise, não reconhece os conflitos culturais como um problema moral. Para isso é necessário antes, por um lado, a procura pela legitimação de critérios de avaliação moral interculturais, que permitam discursar sobre "evolução" social, e, por outro, a caracterização pela contextualização histórica e cultural dos benefícios partilhados intersubjetivamente.

É preciso procurar nas culturas critérios intersubjetivos e não tentar eliminar as diferenças culturais, apelando para um critério supostamente supracultural que termina por oscilar entre um etnocentrismo recalcitrante e um sistema moral inócuo, porque incapaz de dirimir conflitos culturais, visto que não reconhece a

diferença entre as culturas. Se todos estivessem de acordo quanto ao cálculo da utilidade, não precisaríamos de nenhum critério moral, de sorte que o utilitarismo seria um truísmo. Contudo, a divergência sobre como qualificar o bem-estar ou a felicidade é que exige uma discussão da moralidade para que os agentes morais possam estabelecer um critério para a validação de sua ação que não seja apenas restrito ao grupo cultural ao qual pertencem.

O utilitarismo tem o ônus, portanto, de harmonizar seu critério de avaliação dos juízos morais com a motivação humana para a ação moral que nem sempre é a mesma em toda cultura, nem muito menos em todos os indivíduos. Se ficar no patamar da quantificação do benefício pautada na variável não sofrimento, ele será obrigado a esvaziar o agente moral também de seus valores culturais e emotivos, deixando um sujeito tão abstrato quanto desumano.

Referências

BENTHAM. J. *An introduction to the principle of morals and legislation*, 1823. Disponível em: <http://www.econlib.org/library/Bentham/bnthPML0.html>.

COSTA, C. "Razões para o utilitarismo". In: *Ethic@* 1.2, p.155-174, 2002.

DALL'AGNOL, D. "Utilitarismo kantiano?". In: *UFSC* (pré-publicações), Ano IV, Número 49, 2001.

DIAS, M. C. M. *A ampliação do espaço moral no utilitarismo de John Stuart Mill: uma comparação com a moral do utilitarismo de Bentham*. São Paulo: USP (tese), 2011.

KOENIGS, M.; YOUNG, L. *et al*. "Damage to the prefrontal cortex increases utilitarian moral judgments". In: *Nature* 446, p.908-911, 2011.

HARE, R. *Ética: problemas e propostas*. Trad. M. Mascherpe e C. A. Rapucci. São Paulo: Unesp, 1998/1997.

MILL, S. J. *Sobre a liberdade*. Trad. Ari Brito. São Paulo: Hedra, 2008.

_____. *Utilitarianism*, 1863. Disponível em: <http://www.utilitarianism.com/mill1.htm>.

OLIVEIRA, É. A. M. "A crítica de Nietzsche à moral kantiana: por uma moral mínima". In: *Cadernos Nietzsche* 27: p.169-189, 2010.

SIMÕES, M. V. "Rule-utilitarianism". In: *Etic@* 8.3: p.47-62, 2009.

SINGER, P. *Ética Prática*. Trad. J. L. Camargo. São Paulo: Martins Fontes, 2009.

SMART, J. J. C. "Extreme and restricted utilitarianism". In: *The Philosophical Quarterly* 6.25, p.344-354, 1956.

_____. "Utilitarianism and Justice". In: *Journal of Chinese Philosophy* 5.3: p.287-299, 1978.

WILLIAMS, B. *Moral: uma introdução à ética*. Trad. R. Mannarino. São Paulo: Martins Fontes, 2005.

4

A IMPORTÂNCIA DA FENOMENOLOGIA: O QUE A AMABILIDADE E A FRIEZA HUMANA PODEM NOS ENSINAR SOBRE ÉTICA?*

Michael Slote

A fenomenologia referida no título deste capítulo não pertence ao movimento fenomenológico husserliano ou heideggeriano, mas à fenomenologia entendida como descrição do caráter sentido da experiência vivida. É claro que esse movimento filosófico teve coisas importantes a dizer acerca de como devemos fazer a fenomenologia no segundo sentido, o sentido empregado aqui, mas o que eu quero enfatizar é a importância da fenomenologia no segundo sentido para os campos ou disciplinas da ética, psicologia moral, epistemologia e filosofia da mente. E o que tenho a dizer irá chamar ou evocar, pelo menos, ideias do empirismo associacionista inglês, tanto quanto da fenomenologia de Husserl e Heidegger.

Com efeito, considerações fenomenológicas foram abordadas mais notadamente por empiristas como Berkeley e Mill para fins de compreensão da natureza da matemática ou da linguagem em geral, mas me foi ensinado, como estudante de graduação, que tais abordagens eram totalmente equivocadas: que a matemática e os nossos conceitos têm que ser entendidos em termos conceituais, do modo como operamos ou pensamos em termos de conceitos matemáticos ou linguísticos, e não por referência às nossas experiências (sensíveis). E eu não tentarei convencê-los que essas ideias estão equivocadas. O que vou dizer é relevante para a ética, psicologia moral, epistemologia e filosofia da mente, mas isto pode não ter muita relevância para a matemática, linguística e para as

---

*Traduzido por Flavio Williges e Mateus Stein.

disciplinas filosóficas que lidam com tais temas. De modo geral, permita-me acrescentar que não se deve inferir abruptamente do que digo aqui a respeito de vários temas que o que alguém *poderia* pensar deveria ser orientado pelo tipo de fenomenologia que estarei discutindo e descrevendo. No entanto, creio que o que vou dizer tem relevância para o campo da estética, e o que devo dizer na parte final deste capítulo irá ao menos esboçar algumas razões para se pensar assim. E com isso entendido, ou talvez não entendido claramente, deixem-me ser um pouco mais específico acerca do que eu estou falando.

Penso que a fenomenologia, a descrição da experiência, acaba sendo, em aspectos que não foram ainda reconhecidos, essencial para o desenvolvimento da ética, psicologia moral, epistemologia e filosofia da mente, e esta ideia é e será controversa porque vai contra o senso comum de muito do que nós (no Ocidente anglófono) aprendemos e lemos nessas áreas. Por exemplo, lembro-me sempre que, quando estudante de graduação, fui alertado por Philippa Foot para não colocar demasiada ênfase na fenomenologia na tentativa de explicar conceitos fundamentais da ética e da psicologia moral, sendo o orgulho um de seus exemplos favoritos. Os empiristas, dizia ela, muitas vezes identificaram o orgulho com um sentimento (bom) no peito, mas isto, ela alegava, é claramente um equívoco. Se um homem (ela sempre falava de "um homem") anda com um ar alegre e bem humorado ao longo da praia, isto não significa que esteja orgulhoso de algo. Ele não pode orgulhar-se do mar, a menos que tenha certas *crenças;* por exemplo, a crença de que ele ou alguém próximo a ele *criou* o mar ou é *dono* dele. Claro, empiristas como Hume pensaram que até mesmo a crença era, em algum sentido, uma questão de sentimento, e quase todo mundo na filosofia recente concorda que isso representa uma incompreensão total do caráter intelectual ou conceitual da crença. Mas eu vou tentar persuadi-los de que esta concepção herdada e comum é falsa. Não há tal coisa como uma crença puramente intelectual e vou argumentar que (de uma maneira própria e um pouco ingênua) os empiristas farejaram alguma coisa quando eles passaram a enfatizar o caráter sensível, a fenomenologia da crença (no nosso sentido).

Mas deixem-me mencionar também outra razão pela qual especialistas em

ética (e epistemólogos, mas no momento vou simplificar) rejeitam ou minimizam a ideia de uma fenomenologia moral. A maioria dos especialistas em ética são racionalistas e pensam que nossos conceitos morais e nossa aderência a normas morais pode estar fundada na razão. Então, se a emoção e o sentimento entram em cena (de uma forma positiva), entram, eles pensam, como auxiliares na conquista dos objetivos mais fundamentais ou auxiliares da natureza da moralidade. Sentimentos como a solidariedade e a compaixão podem, às vezes, nos ajudar a fazer o que é nosso dever moral fazer, mas o motivo do puro dever, da consciência, é considerado (por muitos kantianos e outros racionalistas), em princípio, suficiente por si só para motivar uma ação ou comportamento e, em qualquer caso, o que *conta* como certo ou errado é considerado uma questão de razão (prática) e não depende constitutivamente de quaisquer emoções ou sentimentos (ordinários), os quais o racionalista considera pelo menos como tipicamente pouco fiáveis quando se trata de fazer as escolhas e ações moralmente certas.

Contudo, não sou um racionalista, mas um sentimentalista moral e considero que os racionalistas estão enganados acerca das questões acima. Tentarei persuadi-los (da forma mais sucinta possível) que isto é assim, e parte do argumento, uma grande parte, na verdade, dependerá da fenomenologia. Eu creio que, se prestarmos atenção em nossa experiência da moralidade, poderíamos obter pistas importantes acerca da natureza ou a constituição da moralidade ou da vida moral, e essas pistas apontam na direção de uma compreensão da moralidade que evidencia elementos emocionais, fenomenologicamente acessíveis ou, ainda, cheios de sentimentos de um modo que os racionalistas não admitem. Uma fenomenologia de certas categorias é essencial não só para a nossa motivação moral, mas até mesmo para os nossos conceitos morais, e eu vou dar as minhas razões mais básicas, a seguir, aqui. (Os argumentos são apresentados mais extensamente e talvez de modo mais persuasivo em meu livro recente *Moral Sentimentalism*[1]). Mas, como sugeri acima, eu também irei defender a importância

[1] New York: Oxford University Press, 2010.

da fenomenologia para a epistemologia e a filosofia da mente: por exemplo, para a compreensão da crença racional ou epistemicamente justificada e para a compreensão do que a crença, em si mesma, realmente é. E vou concluir dizendo algo sobre o modo pelo qual a fenomenologia da percepção estabelece laços com crenças e com a emoção humana. A fenomenologia tem sido bastante negligenciada na ética atual, na epistemologia e na filosofia da mente, porque todo mundo (na filosofia analítica) tem sido um tanto racionalista ou materialista-dualista (mente e corpo), mas o sentimentalismo moral, que esteve durante muito tempo eclipsado, começou a reviver nos últimos anos, e creio que há razões para esperar que outros tipos de sentimentalismos (por exemplo, sobre a natureza da crença e até mesmo da percepção) podem também, eventualmente, ganhar nova vida. Nossas abordagens recentes têm sido muito unilaterais e podemos ser menos unilaterais se apreciarmos a importância da fenomenologia. Não tenho nenhuma pretensão de defender o empirismo como uma abordagem geral a todos os problemas filosóficos; ao contrário de Hume e outros empiristas britânicos, acho que muitos aspectos de nossos conceitos, crenças e justificativas epistêmicas pouco ou nada têm a ver com a fenomenologia. Mas temos de perceber que a fenomenologia também tem um lugar importante dentro de todos estes tópicos.

1

Começarei com a ética e com uma observação fenomenológica sobre um aspecto ou uma parte da ética ou moralidade que pode, à primeira vista, parecer trivial ou sem importância. Nós chamamos as pessoas amáveis ou benevolentes de amáveis (*warm-hearted*) e pensamos que suas ações frequentemente expressam ou exibem tal amabilidade (*warmth*), e eu acho que os racionalistas estão tão conscientes deste fato quanto eu ou vocês estamos. Mas estes fatos não têm o seu significado filosófico assumido, e, por conseguinte, é possível aceitá-los e pensar que eles não são realmente muito importantes para nossa compreensão das fundações moralidade: as fundações da verdade moral e da motivação moral desejável. Isso é precisamente o que os racionalistas fazem (talvez com razão), mas eu acho que os fatos que acabei de mencionar têm um significado mais amplo

do que o racionalismo concede, logo, eu quero ver o quão longe a ideia de amabilidade, por exemplo, pode levar-nos.

Se alguém é amoroso, amável, parece ou age com afetividade para com os outros, isto no mínimo sugere que esse alguém ou suas ações motivadas possuem uma qualidade que podemos chamar apropriadamente de afetividade. Ora, esta ideia de amorosidade pode ter se originado como uma metáfora em conexão com o calor e o frio, físico ou sensorial. No entanto, a ideia de amabilidade pessoal é tomada hoje como uma metáfora frívola. Desculpe-me, mas quando dizemos que alguém se sente amável para com outra pessoa ou é uma pessoa amorosa, o que dizemos pode ser literalmente verdadeiro. E (testem isto em si mesmos!) há um sentimento muito familiar que a maioria de nós tem quando sente pena de alguém ou quando gentilmente nos dispomos a ajudar alguém. E esse sentimento pode ser descrito fenomenologicamente (não somos todos loucos nesta área), como um sentimento amoroso ou um sentimento de afetividade. Mas por que, vocês podem perguntar, isto deve ser especialmente significativo? Como isto esclarece, realmente, o caráter da própria moralidade? Precisamos dizer mais, muito mais, acerca disto.

Talvez o primeiro ponto que deve ser esclarecido aqui é que o tipo de amabilidade que estamos falando é moralmente bom ou louvável. Isto é algo que Kant e muitos kantianos negaram, alegando que a única fonte de bondade moral é uma vontade conscientemente boa. Se alguém age a partir de um senso de dever ou por apreço à moralidade (seja lá o que isso signifique), sua ação pode ser moralmente boa ou louvável, mas de acordo com Kant, se alguém age com base em sentimentos amáveis benevolentes em relação a outra pessoa (ou animal), não há nada moralmente louvável ou bom ou meritório em/ou acerca do que foi feito. Isto pareceu absurdo para muitos filósofos (o dramaturgo Schiller escreveu alguns versos satirizando o ponto de vista de Kant sobre este tema). Mas alguns kantianos recentes pensam que podem lidar com esta dificuldade (de Kant), argumentando que sentimentos (amáveis) de compaixão ou benevolência podem ter uma utilidade auxiliar, de suporte, na vida moral.

No entanto, esta acomodação do sentimento não vai longe o suficiente para

salvar o kantismo da implausibilidade moral. Ela nos diz, provavelmente, que a motivação amável às vezes pode ser louvável, mas ela nunca nega a perspectiva kantiana segundo a qual é sempre moralmente bom e louvável agir a partir de um senso de dever, e que tal motivo nunca é inferior aos motivos mais empíricos ou motivos fenomenologicamente sentidos por simpatia amorosa, entre outras coisas, que os kantianos atuais têm tão frequentemente procurado trazer (mais) para dentro da abordagem ética. Mas é precisamente isto que o famoso exemplo de Bernard Williams sobre a exigência de "pensar demais" ("*one thought too many*") põe em causa. Se um marido que vê a sua esposa e um estranho em algum lugar próximo se afogando tem de consultar a moralidade antes de salvar a sua esposa, se ele tem que se perguntar primeiro se ele é obrigado ou autorizado a dar preferência a sua esposa, então, de acordo com Williams, ele tem que "pensar demais". Do ponto de vista ético, há algo estranho com esse marido; consideraríamos mais apropriado um marido fazer alguma coisa imediatamente para salvar a sua esposa por amor, ao invés de salvar o estranho. Os kantianos não têm um modo de acomodar *esta* conclusão ou *esta* intuição, e essa é uma das razões pelas quais os sentimentalistas morais (e especialmente os adeptos da ética do cuidado), consideram o racionalismo kantiano pouco atraente e inaceitável como uma fundamentação da (boa) motivação moral. A amorosidade (por parte dos agentes morais) desempenha um importante papel moral que os kantianos não conseguem ou não podem admitir, logo, de acordo com o exemplo dado por Williams, a fenomenologia mostra seus primeiros resultados em teoria moral, e precisamos considerar o que ainda pode plausivelmente ser dito em favor da (enfatizando) fenomenologia à luz do que sabemos ou acreditamos sobre a importância, ainda que ocasional, de motivos e sentimentos amáveis. E acontece que, na verdade, muita coisa pode ser dita. (Aliás, os kantianos gostariam de questionar a *confiabilidade* moral de sentimentos amorosos, mas eu não tenho tempo para discutir essa objeção ao sentimentalismo moral aqui. Ela é o foco principal do capítulo 7 de *Moral Sentimentalism*).

Para começar a colocar os pingos nos is, deve-se mencionar que existe insensibilidade, uma (total) falta de sentimentos amorosos em relação aos outros.

(Tal como acontece com o calor físico ou a amabilidade, a amabilidade emocional parece andar lado a lado com a frieza emocional: entendemos por frieza a ausência de sentimentos amáveis). E, neste ponto, nós podemos ver como a frieza e a amabilidade enquanto traços fenomenológicos ou motivacionais de indivíduos (ou algumas de suas ações ou atitudes) podem desempenhar um papel não apenas no sentido de avaliar pessoas ou ações (nós consideramos melhores as pessoas amáveis e suas ações), mas em como ações moralmente aceitáveis ou boas são, de fato, motivadas. Um senso de dever pode muito bem ser capaz de nos levar a fazer alguma coisa (embora isso possa exigir-nos mais e até mesmo uma análise fenomenológica), mas a compaixão amável compreensivelmente também dá origem a várias ações situacionalmente-relevantes, e embora esse seja um fato amplamente aceito e óbvio, ele é um daqueles fatos que o kantismo tendeu a minimizar ou negar, com base nos resultados não intuitivos abordados anteriormente.

Mas, além disso, e isto realmente é um grande passo à frente, os fatos sobre a amabilidade e a fenomenologia que mencionei também podem ser relevantes à nossa compreensão de conceitos morais, e isto, se for verdade, nos levará para uma direção muito diferente de tudo aquilo que nos é familiar na teoria moral racionalista. Isto nos levará, de fato, em direção a um sentimentalismo moral amplo. Percebam que não existe apenas o fenômeno de pessoas agirem de modo amável e de estarmos cientes de tais fatos, mas também de *sermos tocados* pela amabilidade dos outros (outros agentes). Hume foi, talvez, o primeiro filósofo que chamou nossa atenção para esta possibilidade, o fato comum das interações sociais[2]. Mas há também o fato correspondente, algo que Hume *não* mencionou: podemos tornar-nos frios na/ou pela indiferença ou crueldade (novamente a amabilidade entra em cena) de uma determinada pessoa ou agente. Estas reações ocorrem através de processos de empatia, processos nos quais Hume foi a primeira pessoa (ao menos no Ocidente) a descrever (embora ele não tivesse à disposição o termo "empatia"). Logo, creio que preciso dizer algo, neste momento,

[2] Ver a edição de Selby-Bigge do *Tratado da Natureza Humana* de Hume, p. 604ss.

acerca da natureza e do conceito de empatia.

No uso contemporâneo padrão da língua inglesa, a empatia é diferenciada da simpatia. A empatia ocorre quando, como Bill Clinton descreveu, você sente a dor ou a alegria de outra pessoa. Mas simpatia pela pessoa que, digamos, sente dor não obriga-nos a sentir a sua dor e, por outro lado, essa é uma questão de querer a sua condição de melhora (e esperar ou querer fazer algo para conseguir esse resultado). Mas de acordo com grande parte da literatura recente acerca do desenvolvimento moral, a nossa simpática ou benevolente preocupação para com ou sobre os outros *depende* substancialmente de nossa capacidade de empatia. Desse modo, há (pelo menos) dois tipos de empatias: o tipo de "empatia projetiva", que envolve entrar na cabeça de alguém e ver as coisas do seu ponto de vista e, além disso, o tipo de empatia mais emocionalmente receptiva, a "empatia associativa", que é basicamente uma questão de colocar-se no lugar da outra pessoa (isto é, por vezes, descrito como contágio emocional ou osmose). Acredita-se que os psicopatas são capazes de colocarem-se no lugar de outras pessoas (para fins egoístas ou malignos), mas incapazes de sentirem a dor dos outros, incapazes desse mesmo tipo de receptividade emocional ou empática. E, nessa medida, também podemos dizer que os psicopatas são insensíveis (*cold-hearted*). Eles não (receptivamente) empatizam com os sentimentos dos outros, não sendo (portanto) capazes de benevolência afetiva, amabilidade ou compaixão em relação aos outros. Eles são, afinal de contas, psicopatas. Então, simpatia, altruísmo e carinho genuíno em relação aos outros requer o tipo receptivo de empatia que falta aos psicopatas. Porém, isto tem implicações interessantes para o tipo de experiências que Hume estava falando quando ele falou de pessoas sendo aconchegadas pela (visão) da amabilidade que um amigo pode dispensar aos seus ou suas amigas (digamos, ao ajudá-los). A amabilidade exibida por um amigo em relação aos outros envolve ou requer a empatia em relação à condição do amigo, por exemplo, com a sua dor ou seu sofrimento. E assim a amabilidade e a empatia estão interligadas nos agentes. Uma pessoa moralmente decente, ao ajudar um amigo, exibe uma amabilidade que depende da sua capacidade para exibir o tipo de empatia receptiva que um psicopata é incapaz de exibir. Tudo certo até aqui.

Mas se interesses afetivos ou amáveis em relação aos outros requerem uma empatia associativa ou receptiva, pode-se precisar de empatia ao ser afetado pela amabilidade de outra pessoa em relação a uma terceira. Então, quando falei anteriormente do fato de alguém ter agido amavelmente em relação a outra pessoa, eu falei de um fato que até mesmo um psicopata pode (de alguma forma) ser capaz de reconhecer. Eles sabem que deveria haver algo como um calor pessoal mesmo se nunca tivessem sentido isto (os psicopatas são emocionalmente semelhantes a um cego de nascença), assim um psicopata pode, em alguma medida ou grau, reconhecer que alguém agiu amavelmente em relação a um amigo. Entretanto, mesmo assim um psicopata não pode *ser empaticamente afetado* pela amabilidade agencial de alguém em relação a um amigo. Eles não têm a capacidade da empatia receptiva, por osmose empática ou emocional, que é necessária para que alguém seja afetado ou, neste caso, indiferente àquilo que outra pessoa tenha feito (ou queira fazer). Ser afetado pela amabilidade de alguém é uma espécie de amabilidade de segunda ordem, quando comparada com a amabilidade de primeira ordem em que alguém sente amor ou uma ligação afetiva por seus amigos e, com base nisso, quer ajudá-los. Mas os dois tipos de amabilidade (e a indiferença de segunda ordem também) exigem o tipo de empatia receptiva que os psicopatas são incapazes de ter. Assim, parece que a fenomenologia da amabilidade, da frieza ou da indiferença (dizemos que os agentes são frios ou insensíveis, por que falamos deles tornarem-se *frios* pela indiferença), pode desempenhar um papel importante para ajudar-nos a compreender a motivação moral e a sua ausência, e isto já está bem além de qualquer coisa que o racionalismo ou a filosofia moral atual, em geral, parece ter permitido. Mas podemos de fato ir muito, muito além, porque (como eu continuo prometendo), podemos usar as ideias fenomenológicas que acabamos de explorar e aplicá-las para fornecer uma descrição metaética fenomenologicamente rica (e sentimentalista) do que se *entende* por termos como “certo” e “errado”.

Hume pensava que juízos ou “enunciados” que fazem o uso de termos ou conceitos morais repousavam sobre o que ele considerava serem fenômenos psicológicos de aprovação e desaprovação. Ele concebeu a aprovação como uma

espécie de prazer do prazer que alguém dá ou causa a um terceiro, mantendo a desaprovação em paralelo com um tipo de dor ou desconforto com a dor que alguém transfere ou causa na outra pessoa. Deste modo, Hume concebeu a aprovação e a reprovação como tendo uma espécie de fenomenologia (e como sendo fenomenologicamente *opostas* uma a outra), e ele então argumentou (estou falando de um modo muito superficial de toda essa discussão), que expressamos ou descrevemos aprovação ou desaprovação quando dizemos que algo é bom, certo ou errado. Esta ideia foi imediatamente atacada (por Richard Price e Thomas Reid) alegando que a aprovação ou a desaprovação, longe de ser a base para alegações ou enunciados morais *pressupõem* tais alegações ou enunciados. Desaprovar já significa pensar que algo está errado, logo, juízos ou proferimentos que tenham a ver com incorreção (*wrongness*) da ação não podem ser não-circularmente fundados em (em primeiro lugar, conceitualmente) estados de desaprovação. O mesmo é válido para a aprovação e a correção (*rightness*) ou bondade moral. Mas, de certa forma, isso é uma injustiça a Hume. Talvez ele não devesse chamar o prazer e a dor de segunda ordem de aprovação e reprovação, mas esses fenômenos não exigem, em si mesmos, a feitura de afirmações morais (ou a posse de conceitos morais). Assim, não há circularidade na utilização e na referência a eles como uma base para a compreensão de enunciados morais reais (não somente, talvez, em relação à aprovação e a desaprovação real).

No entanto, eu não vou invocar estas idéias particulares de Hume como ponto de apoio para uma defesa adicional, de maior importância, da fenomenologia moral. Por razões que dei na medida certa em *Moral Sentimentalism*, creio que não devemos usar ideias particulares de Hume acerca do que ele chama de aprovação e do que ele chama de desaprovação como uma base para a compreensão de juízos morais ou enunciados – ou conceitos morais. A ideia que precisamos de uma descrição fenomenológica da aprovação e da reprovação ou algo que talvez pudesse ser mais bem descrito como formas primordiais de aprovação e desaprovação, como base para a compreensão da terminologia moral, parece-me estar no caminho certo, mas vou propor uma fenomenologia um pouco diferente para os estados ou sentimentos que, como um sentimentalista metaético, creio

fundamentar e constituir os nossos conceitos e linguagens morais (e permitam-me chamá-los de estados de aprovação e reprovação; no final veremos que há boas razões para usar esses termos para estes estados, mesmo que eles não envolvam ou pressuponham juízos ou conceitos morais).

A minha sugestão alternativa é, na verdade, em grande medida, devida ao próprio Hume. Como mencionei anteriormente, Hume fala de sermos tocados pela amabilidade generosa que alguém mostra em relação aos seus amigos ou amigas, e creio que a amabilidade de segunda ordem é uma espécie de aprovação e também pode ser vista como constitutiva de uma base plausível para termos ou conceitos morais positivos como "certo" ou "moralmente bom". E, então, indo além de Hume da maneira que indiquei anteriormente, creio que ser empaticamente indiferente pela atitude de alguém insensível ou cujo comportamento em relação a terceiros constitui uma espécie de desaprovação (primordial) daquela atitude ou comportamento, pode ser usado como uma base sólida para a compreensão de termos ou conceitos morais negativos. Assim, estou, como Hume, concebendo a aprovação e a reprovação como atitudes de segunda ordem, mas concebo-as como atitudes diferentes daquelas atitudes focadas e mencionadas por Hume. Mas em vez de argumentar contra Hume aqui, deixem-me tentar dizer o que pode positivamente ser dito em favor de pensar-se a aprovação e a reprovação como reações empáticas de segunda ordem de amabilidade e indiferença, respectivamente.3 Deixem-me, então, explicar por que acho que estes estados psicológicos, que não parecem pressupor conceitos ou juízos morais, podem funcionar como uma base para fazermos juízos morais – e, muito mais fortemente, podem fazê-lo de um modo que não irá minar a validade ou objetividade dos juízos morais, ao menos não do modo que os tipos de metaética sentimentalista que tem sido atribuída a Hume claramente fazem. (Lembrem-se: estou argumentando em

[3] Como o próprio Hume aponta no *Tratado*, quando somos afagados pelas ações amáveis de alguém para com um amigo, as lágrimas podem "brotar em nossos olhos". Deste modo, ser empaticamente acolhido (no domínio de segunda ordem) não precisa ser univocamente agradável e minha descrição da aprovação em termos de amabilidade de segunda ordem, portanto, está em desacordo com a ideia básica de Hume de que a aprovação é uma espécie de *prazer* de segunda ordem.

favor de um sentimentalismo moral metaético, mas isso não me compromete, e eu não estou comprometido, com qualquer tipo de empirismo sobre conceitos em geral).

Em primeiro lugar, um ponto sobre a história, a história da filosofia em relação à fenomenologia. O fenômeno de sermos indiferentes (frios) diante da atitude ou das ações de alguém é muito comum na vida cotidiana, e este fenômeno, esta experiência, tem uma fenomenologia definida que a maioria de nós conhece. Não obstante, eticistas nunca falam sobre a experiência de sermos indiferentes (tanto quanto eu sei), e embora a ideia de sermos afetados pela amabilidade de alguém, conforme ela se mostra nas ações, seja mencionada por Hume, ele só faz menção a ela *an passant;* ele não faz uso dela em larga escala para propósitos filosóficos ou metaéticos. E não acho que este fenômeno, esta ideia, tem sido muito utilizado depois de Hume (talvez alguém com um melhor conhecimento da história da ética, do Oriente e Ocidente, possa corrigir-me aqui). Mas se o que estou dizendo é correto, os fenômenos de indiferença e amabilidade de segunda ordem, fenômenos aos quais estamos todos familiarizados enquanto seres humanos ordinários, têm, à nossa revelia e bem debaixo do nosso nariz, funcionado como base para os nossos próprios conceitos morais. Isto é, de certa forma, surpreendente, e alguns podem dizer que toda esta ideia é totalmente implausível: não podemos ser ignorantes dos nossos próprios conceitos. Mas talvez tenhamos sido. Vejamos!

Agora, uma vez que se vê que a empatia pode impulsionar-nos a agir, a ideia de que a amabilidade de segunda ordem equivale a um sentimento positivo de aprovação pode não parecer tão *outré*. Se a empatia receptiva move-nos, ela nos move a ajudar, em vez de nos afastar de alguém que precisa ou sente dor. E alguém que é capaz de ter empatia de segunda ordem vai ser capaz de ter empatia de primeira ordem (embora eu esteja omitindo os argumentos a favor dessa tese). Logo, se sou tocado pela generosidade amável de alguém em relação a um amigo, estarei disposto a fazer essas mesmas coisas amáveis e desprendidas, e isto acontece num sentido em que ser acolhido pela amabilidade de alguém constitui uma atitude positiva em relação a essa pessoa e/ou às suas

ações. Obviamente, a aprovação é também uma atitude positiva, mas esse é justamente o meu ponto. O que estamos dizendo em relação à empatia de primeira e a de segunda ordem e a amabilidade empática torna claro que a amabilidade de segunda ordem é um tipo de atitude positiva que não é de todo implausível considerar como equivalente ou igual a (o que consideramos ser) atitude positiva de aprovação moral. E nossas reações indiferentes de segunda ordem serão direcionadas ao comportamento que, dadas as nossas disposições empáticas ou amáveis de primeira ordem, estaremos inclinados ou dispostos a *não* reproduzir, e, nessa medida, também, portanto, faz sentido identificar a nossa reação de indiferença empática à frieza da atitude ou comportamento de outra pessoa com uma atitude negativa de desaprovação. E, se eu posso pedir-lhe para pensar sobre as suas próprias experiências um pouco, não é sentir-se tocado pela amabilidade de alguém equivalente a aprová-las? E não é ficarmos insensível ou frios perante a atitude ou ações de outra pessoa a mesma coisa que *desaprová-las*? Com certeza parece-me assim e tudo isso favorece a descrição da aprovação e da reprovação que forneci até o presente momento. Essa descrição é apenas parcialmente fenomenológica, no entanto, pois ela equipara, digamos, a aprovação com a amabilidade empaticamente-derivada, e isto carrega uma noção causal ou uma hipótese que não quero equiparar com alguma experiência fenomenologicamente disponível. A empatia é uma noção causal, mas o sentimento amável empaticamente-derivado combina um fator causal (ou origem) com uma experiência fenomenológica (ou resultado). Isto é mais fenomenologia do que aqueles que pensam que a aprovação já exige um juízo moral ou algum conceito consentido; trata-se de uma mistura de fenomenologia com outros fatores, e é interessante notar que, se apenas isso já é suficiente para reconhecer que o que estamos dizendo aqui não é tão extremado como se poderia pensar inicialmente. A aprovação e a desaprovação não são *só* uma questão de fenomenologia (e a teoria de Hume destes estados também mistura fenomenologia com outros fatores causais, embora Hume, diferente de mim, pode ter pretendido entender a causação em si mesma em termos puramente fenomenológicos).

Embora eu acredite que a descrição da aprovação e da desaprovação que

ofereci tão brevemente tem, com efeito, certa plausibilidade fenomenológica, há uma objeção a ela que precisa ser considerada antes que eu passe a usar a descrição da aprovação e da reprovação para explicar o significado de termos morais como "certo" e "errado". E a objeção é, de fato, ou, em certa medida, fenomenológica. Certos filósofos argumentaram que a desaprovação é uma questão de se estar com raiva e dado que a raiva tem uma fenomenologia, nós realmente precisamos considerar se seria melhor (com base em razões fenomenológicas e possivelmente outras razões filosóficas) conceber a desaprovação em termos de raiva e não em termos de insensibilidade empática de segunda ordem. Mas, curiosamente, aqueles que têm procurado identificar a desaprovação com a raiva nem sequer consideraram a possibilidade alternativa de que a desaprovação é ou envolve uma reação empática insensível. E se fizessem, talvez eles pudessem admitir que a reação de insensibilidade também tivesse alguma pretensão de constituir uma forma de desaprovação (embora eu não esteja apostando todas minhas fichas de que eles iriam conceder isto). Poderíamos, então, dizer que a desaprovação (ao contrário da aprovação?) pode ser baseada e identificada com dois estados mentais diferentes (assim como existem dois tipos muito diferentes de jade), e pode acontecer que o estado do qual estive falando, o estado de insensibilidade de segunda ordem, seja mais útil para usar na explicitação (negativa) da linguagem moral do que a ideia ou a experiência de raiva. Eu realmente acho que há razões para preferir a descrição da desaprovação que tenho oferecido acima em vez daquela que equipara a desaprovação a uma espécie de raiva. Com certeza muitas vezes ficamos com raiva de ações que pensamos (tardiamente?) estarem erradas – nós podemos realmente ter tais reações mais frequentemente do que sentirmo-nos indiferentes em relação a alguém insensível. E eu não estou dizendo que a raiva pressupõe conceitos morais e assim não pode constituir uma base sentimentalista adequada para tais conceitos. Mas eu acho que há algumas razões para a fenomenologia e a filosofia pura pensarem na desaprovação como envolvendo insensibilidade, em vez de

raiva. Deixem-me citá-las agora[4].

Primeiro as razões (mais) da filosofia ou teóricas. Nossas opiniões morais comuns distinguem entre nossas obrigações com os outros e com nós mesmos; as obrigações *eu-outro* são, num certo e importante sentido, assimétricas. Por exemplo, prejudicar negligentemente outras pessoas geralmente é concebido como sendo moralmente criticável, mas não estamos inclinados à crítica moral semelhante de alguém que, por negligência, prejudica a si mesmo (a menos que isto tenha muitas implicações causais definitivas no bem-estar de outras pessoas). Em outras palavras, nós desaprovamos moralmente a negligência em relação aos outros de uma forma ou numa amplitude diferente daquela verificada quando desaprovamos moralmente o ato de prejudicar ou ferir a si mesmo negligentemente – embora, é claro, tal auto-negligência ou auto-omissão é muitas vezes pensada como envolvendo estupidez ou falta de discernimento por parte da pessoa que não-intencionalmente machuca a si mesma. (Estas parecem mais críticas racionais do que críticas especificamente morais). Da mesma forma, nos incomodamos com aquilo que algumas pessoas fazem a terceiros, mas não somos sensíveis em relacão àquilo que fazem a si mesmas, e isso acontece porque a empatia faz mais sentido quando focada noutras pessoas, mas não quando o foco é em nós mesmos. Com certeza seria possível sentirmos empatia por uma criança indefesa que sofreu abuso, mas nós não costumamos falar ou pensar em termos de empatia com relação aos acontecimentos de nosso próprio eu, e isso significa que alguém que negligencia a si mesmo não demonstra aquilo que costumeiramente chamamos de falta de empatia. Também nesse caso, e dado o que foi dito anteriormente, não parece haver nada de frio ou insensível na própria pessoa e, assim, não estaremos sendo insensíveis num sentido de segunda ordem em relação àquilo que tal pessoa faz a si mesmo ou a si mesma. Assim, a

[4] Em um artigo sobre o meu livro *Moral Sentimentalism,* que será publicado no *Southern Journal of Philosophy*, Justin D'Arms argumenta a favor do tratamento da desaprovação como uma forma de raiva e não como uma reação fria, de indiferença. As razões para discordar de D'Arms que mencionarei, também foram mencionadas numa resposta ao artigo de D'Arms que será publicada nessa mesma revista.

alegação de que a desaprovação envolve sensibilidade de segunda ordem, juntamente com o que dissemos e estamos dizendo acerca da empatia permite-nos explicar por que não desaprovamos moralmente o ato de ferir-se por negligência e também certas tentativas fracassadas de ajudar a si mesmo (Por exemplo, não desaprovamos moralmente alguém que desnecessariamente rejeita ou recusa vários tipos de comidas e atividades prazerosas).

Mas a raiva é diferente. É muito fácil ficarmos zangados devido às coisas irracionais ou estúpidas que fazemos a nós mesmos, como, por exemplo, quando ferimo-nos por descuido e nos impossibilitamos de disputar, digamos, uma importante corrida ou, para dar outro exemplo vívido, quando nos irritamos com nós mesmos por termos fracassado num primeiro encontro com alguém que gostaríamos de ter conhecido melhor. A raiva não tem a assimetria eu-outro do mesmo modo que as reações frias o têm, e esta é uma boa razão, penso eu, para não identificarmos a raiva com a desaprovação moral (assimétrica). E há ainda um outro ponto.

A ideia de Hume, ainda que tenha outros problemas, de que a aprovação e a reprovação podem ser entendidas, respectivamente, em termos de prazer e dor, preserva uma intuição comum que pretende mostrar que a aprovação e a reprovação são duas coisas opostas. A ideia de que a aprovação e a desaprovação envolvem reações de amabilidade e indiferença ou frieza também em grande parte preserva essa intuição, mas a ideia de que a desaprovação é uma forma de raiva deixa-nos sem qualquer contrapartida óbvia da raiva que pudesse ser considerada equivalente à aprovação[5]. Portanto, há algumas razões teóricas pelas quais (se tivermos que escolher entre elas) podemos considerar a frieza, em vez da raiva, como constituinte do tipo de desaprovação moral que pode ser usada não-circularmente para fundamentar predicados morais ou juízos morais

[5] Jesse Prinz sugeriu-me que a gratidão pode ser o sentimento oposto apropriado. Mas há uma série de razões para questionar essa sugestão: pelo menos o fato de que a gratidão é eu-outro assimétrica (não somos gratos a nós mesmos) de uma maneira que a raiva não é. (Há também o fato de que nós, por vezes, aprovamos afetuosamente as boas ações de pessoas que viveram há muito tempo e não é de todo claro que, em tais casos, tendemos a nos sentir gratos a essas pessoas).

genuínos. Mas há também algumas razões na fenomenologia para preferirmos a ideia de frieza.

Intuitivamente e nos termos do senso-comum há uma certa fenomenologia em torno do que experimentamos ou concebemos como desaprovação moral. A raiva é muitas vezes vista como quente (*hot*), mas a desaprovação, de fato, nunca é. A noção de desaprovação quente parece muito estranha e é totalmente incomum. Então, a desaprovação e a raiva não parecem bem feitas uma para a outra e, de fato, nós normalmente pensamos em desaprovação como envolvendo um sentido ou realidade *distante* do que (ou quem) sofre desaprovação. Novamente, isto é muito diferente de raiva e de calor, mas o sentido ou a ideia de distância cai muito bem com (a noção de) uma reação de frieza (*chill reaction*). E pense também sobre como falamos de desaprovação. É muito idiomático falar de uma desaprovação fria (*cold disapproval*) e há também o fenômeno que designamos pela frase "frieza da desaprovação" ("*the chill of disapproval"*). Esta frase é tipicamente aplicada não a alguém que desaprova, mas a alguém que é desaprovado e para quem sente a desaprovação: como em "eu senti a frieza de sua desaprovação". Mas pense no que esta frase idiomática implica. A fim de sentir o frio da desaprovação de alguém, essa desaprovação tem que ser algo que pode esfriar outra pessoa, e se não há nada de frio ou indiferente na desaprovação, isso não seria possível. Não se pode sentir a frieza da raiva quente de alguém, mas o fato de que se *pode* sentir a frieza da desaprovação de alguém sugere fortemente que a desaprovação é, na verdade, uma forma de sensação de frio ou de gelo. Observe também que, se a reação fria de reprovação é um fenômeno empático de segunda ordem, então é preciso uma reação de terceira ordem para sentir a indiferença dessa desaprovação. Isto torna as coisas complicadas, mas, creio, de uma maneira que faz com que os fenômenos relevantes aqui pareçam bastante ricos, o que permite que sejam descritos – fenomenologicamente – em termos que correspondem muito bem com o modo como efetivamente os experimentamos.

Nada disto pretende sugerir que não podemos ficar bravos com as pessoas que machucam os outros. Mas quer dizer, porém, que se estivermos realmente com raiva deles (e fervendo), então não temos distância suficiente (não estamos

distanciando-nos suficientemente) deles ou da situação, de modo a sentir o que seria apropriado e idiomático chamar de desaprovação. Podemos ter que nos acalmar um pouco e separar-nos fisicamente ou emocionalmente (e praticamente) da situação dada e do que alguém tenha feito nessa situação, a fim de sermos capazes de realmente *desaprovar* o que eles fizeram nessa situação. E isto permitiria que as reações de indiferença e raiva tivessem usos morais. A raiva pode estimular-nos a remediar uma situação de uma maneira que uma reação fria muito possivelmente não o faria. Mas a distância palpável envolvida na frieza pode fazer tal estado mais relevante e apropriado para a realização de juízos morais objetivos (negativos). Portanto, tendo dado as minhas razões para ver a frieza ou insensibilidade, ao contrário da raiva, como constituinte da desaprovação moral, deixem-me dizer algo acerca de como penso que a aprovação ou a desaprovação moral, entendida no sentido que vim sugerindo, podem fundamentar nossas noções morais e nossa capacidade de fazer juízos morais objetivos.

Isto parece à primeira vista impossível ou pelo menos improvável. A descrição sentimentalista de Hume dos juízos morais ou enunciados como sendo baseados em reações empáticas de prazer e dor considerou tais juízos ou enunciados como formas de expressar ou descrever a aprovação e a desaprovação. Deste modo, de acordo com Hume (como frequentemente interpretado), quando dizemos que algo é certo ou errado, nós estamos somente emocionando-nos (ou projetando nossos sentimentos ou emoções para dentro do mundo), ou estamos descrevendo os nossos sentimentos presentes ou estamos fazendo previsões acerca de como os seres humanos em diversas circunstâncias *sentir-se-iam* acerca da ação que estamos dizendo que é certa ou errada. Nenhuma destas sugestões metaéticas faz reivindicações morais totalmente objetivas ou válidas no sentido em que os kantianos geralmente assumem que elas têm. Entretanto, em *Moral Sentimentalism,* sugeri um sentido em que o sentimentalismo metaético pode preservar a intuição pré-filosófica que as alegações morais, por vezes, têm validade real e o conseguem tendo por base o que acabo de dizer acerca da natureza da aprovação e da reprovação.

Como a nossa teoria concebe-os, estes estados são estados de sentimento

com uma certa fenomenologia e uma determinada origem causal, mas, seguindo de certa forma o trabalho de Saul Kripke em *Naming and Necessity*, eu penso que tais estados de sentimento podem ancorar nossos juízos morais, sem os juízos serem *sobre* esses sentimentos. No meu livro eu argumentei que a cordialidade e a indiferença empática de segunda ordem podem corrigir a referência, respectivamente, de "moralmente bom ou certo" e "moralmente errado" a partir de algo parecido com a maneira pela qual (de acordo com Kripke) a experiência de vermelho pode fixar a referência da expressão "objetivamente vermelho". De acordo com Kripke, julgamentos acerca da vermelhidão objetiva são eles mesmos plenamente objetivos, mesmo se eles estão baseados em uma experiência de vermelhidão, a experiência fixa a referência de "objetivamente vermelho" sem que ela própria seja parte da matéria dos juízos que atribuem vermelhidão objetiva. E eu usei um argumento um pouco parecido (embora haja algumas diferenças importantes) para argumentar que (por exemplo) a frieza empática fixa a referência de "moralmente errado", sem ser a matéria de juízos envolvendo esse conceito e, contudo, permitindo que tais juízos tenham real objetividade ou validade. Eu não vou entrar em detalhes aqui, mas se o método de explicitação de termos morais que eu acabei de sugerir, e isso está descrito em detalhes no meu livro, está, enfim, no caminho certo, então há uma defesa a mais para a fenomenologia no contexto ético-teórico. A atenção fenomenológica às nossas reações empáticas de primeira, segunda, e até mesmo de terceira ordem, pode muito bem permitir-nos explicar termos morais com mais sucesso do que outras abordagens (sentimentalistas) permitiram-nos fazer.

2

Mas agora é hora de mudar de direção. Estamos discutindo a fenomenologia em relação à ética, mas, por razões que serão dadas no que segue, eu também creio que a fenomenologia é importante para a nossa compreensão da epistemologia e da filosofia da mente. Eu passarei a argumentar que a fenomenologia é relevante para a nossa compreensão filosófica básica tanto da crença epistemicamente racional ou justificada e da crença em si. E a

fenomenologia da experiência perceptiva também se conecta com a crença perceptiva, emoções ou sentimentos relevantes, em sentidos que eu não acredito que tenham sido previamente descritos.

Eu penso que o sentimentalismo moral é viável em ética, tanto como uma abordagem normativa ou metaética. O significado dos termos morais pode ser dado de uma forma sentimentalista através de uma abordagem semi-kripkeana que corrige a referência de "certo" e "errado" em relação a certas experiências empaticamente-derivadas de afetividade e frieza. E (apesar de eu não ter mencionado isto antes) eu sou, em ética normativa, um representante da ética do cuidado, que considera que as nossas obrigações morais estão baseadas em fatos sobre o que reflete ou não reflete uma completa preocupação empática em relação aos outros. (Em *Moral Sentimentalism* eu também argumento que a ênfase normativa em relação ao cuidado empático pode ser derivada de uma descrição semi-kripkeana do fixador da referência das semânticas morais.) Mas eu não sou – eu repito – eu *não* sou um completo sentimentalista sobre padrões epistêmicos/racionais– mesmo que Hume possa ter sido. Eu penso que existem formas de irracionalidade epistêmica que não têm nada a ver com qualquer falha emocional ou sentimental: por exemplo, a falácia de Monte Carlo, que envolve pensar (por exemplo) que é mais provável que dê vermelho na roleta após a cor preta ter aparecido seis vezes seguidas. Eu, entretanto, penso que a epistemologia e as nossas vidas de seres cognoscentes e sujeitos de crenças contém elementos de emoções ou sentimentos que não foram previamente reconhecidos ou apontados, e que podemos chegar, pelo menos em parte, a algumas dessas novas e, tal como vejo, plausíveis ideias, pela via da fenomenologia. Tanto a crença epistemicamente justificada, quanto a crença em si, têm alguns aspectos sentimentais fenomenologicamente acessíveis, e vamos começar com a justificação epistêmica. (Em nossa seção final, vamos falar sobre como tudo isto se relaciona com a fenomenologia da percepção.)

Para que possamos ter crenças justificadas, não podemos sair por aí cometendo erros como a Falácia de Monte Carlo. Nós acreditamos que a justificação das nossas crenças obriga-nos também a termos uma mente aberta.

Não podemos simplesmente ignorar o que as outras pessoas dizem, mas, se as nossas próprias crenças, argumentos e teorias devem ser plenamente justificadas, temos que ser justos e, em certa medida, também objetivos acerca das crenças, argumentos e teorias das outras pessoas. O que por vezes é chamado de virtudes epistêmicas de "ter uma mente justa, equilibrada" (*fair mindedness*), "ter a mente aberta" (*open-mindedness*) e objetividade (acerca das evidências) são elementos importantes da racionalidade epistêmica ou cognitiva, e eu quero argumentar que estas virtudes, mesmo que não constituam a totalidade da virtude epistêmica, envolvem elementos específicos de sentimento, e isto significa que a virtude epistêmica ou racionalidade como um todo têm importantes aspectos sentimentais. (Claramente, estas virtudes não cobrem toda a virtude epistêmica porque, por exemplo, também é importante não sermos negligentes ou descuidados na pesagem das evidências para nossas crenças.) Vamos nos concentrar em "ter a mente aberta", embora quase tudo que eu diga aplique-se também a "ter uma mente justa ou equilibrada", à objetividade e à imparcialidade como virtudes epistêmicas.

O que está envolvido em ter uma mente aberta em relação aos pontos de vista das outras pessoas? Bem, por um lado, e para inverter os motores por um momento, eu não penso que uma pessoa de mente genuinamente aberta tem de ser aberta ou receptiva a *tudo*. Por razões (algumas delas fenomenológicas) que eu vou citar mais tarde, às vezes é apropriado ou pelo menos não é impróprio para uma pessoa de mente aberta rejeitar inteiramente certas crenças ou visões. Mas onde ter uma mente aberta for importante, essa virtude parece envolver essencialmente uma habilidade e disposição a ver as coisas a partir do ponto de vista das outras pessoas, os pontos de vista de pessoas que, no todo ou em parte, discordam (talvez bastante) sobre um determinado assunto. E isto claramente requer um tipo de empatia.

Anteriormente eu descrevi a empatia acerca dos sentimentos das outras pessoas, digamos, de dor ou alegria, e discuti ou ressaltei que os psicopatas carecem de um tipo de habilidade empática, a capacidade de receber os sentimentos dos outros por contágio ou osmose empática. (Esta incapacidade

relaciona-se à falta de motivação altruísta.) Mas, como Hume apontou no *Tratado da Natureza Humana*, nós não apenas absorvemos os sentimentos dos outros, mas também atitudes e opiniões ou crenças, e assim parece que a empatia associativa ou receptiva pode receber (trocadilho proposital!) não apenas sentimentos, mas também atitudes e opiniões. Logo, é possível empatizar com o ponto de vista, com os argumentos, com as crenças, teorias, etc., das outras pessoas, e isto parece ser necessário, se alguém for descrito como tendo uma mente-aberta frente aos pontos de vista das outras pessoas, etc.

Mas aqui uma objeção muito importante contra a direção que estou seguindo precisa ser considerada. Pode não ser suficiente para uma pessoa de mente aberta estar disposta e estar apta a colocar-se no lugar de outras pessoas, ver as coisas do ponto de vista delas, o que elas podem estar querendo dizer, sem que nada disso envolva ou exija sentimento ou emoção da parte da pessoa que se coloca no lugar da outra pessoa? Em outras palavras, nós anteriormente distinguimos entre o tipo de empatia associativa que envolve involuntariamente ou receptivamente assumir os sentimentos (ou opiniões) das outras pessoas e uma empatia projetiva que coloca a si mesma na situação mental da outra pessoa sem necessariamente ter qualquer sentimento a acompanhando. Se esta última forma de empatia é suficiente para ter uma mente aberta, é difícil ver como essa virtude (ter mente aberta ou justa, etc.) poderia depender de sentimentos ou emoções. O sentimentalismo epistêmico ficaria sem um ponto de apoio ou aquele tipo de ponto de apoio que tem nos aspectos morais das nossas vidas.

No entanto, penso que deveríamos desconfiar aqui. Afinal, a capacidade de colocar-se no lugar das outras pessoas sem sentir o que elas sentem é característica dos psicopatas, e é tão plausível supor que a simples capacidade de ter a mente aberta tem muito em comum com *eles*? Talvez a mente aberta exige que sejamos capazes de *sentir* algo que é relevante para as crenças, argumentos e pontos de vista dos outros. Talvez isto exija um grau de *simpatia intelectual* com o que os outros acreditam, discutem, etc. E se isso acontecer, então a mentalidade aberta envolverá, pelo menos, um elemento sentimental. S*erá* que a mentalidade aberta exige tais reações sentimentais ou a disposição de ter tais reações?

Bem, consideremos uma pessoa que é capaz de colocar-se no lugar de outra pessoa, mas que carece de toda a simpatia pelos seus pontos de vista, uma pessoa, por exemplo, que está simplesmente tentando provar ou explorar a fraqueza das ideias e argumentos da outra pessoa, a fim de ser capaz de dar um melhor argumento para si ou para sua própria posição intelectual. Esta pessoa possui uma mente aberta? Muito claramente, não! E o que parece estar faltando é qualquer tipo de verdadeira abertura à possibilidade de que a outra pessoa está, em alguns aspectos, certa acerca de suas opiniões; falta qualquer tipo de simpatia (mínima), portanto, em relação àquilo que a outra pessoa pensa. Logo, penso que uma mente realmente aberta requer (que se seja capaz de adquirir) um certo grau ou quantidade de simpatia em relação ao que os outros pensam e tal simpatia claramente significa ter uma opinião, ao menos, um pouco favorável em relação ao que os outros pensam. Mas, para favorecer algo ou alguém é preciso ter, pelo menos, uma atitude ligeiramente positiva em relação a elas, sentirmo-nos positivamente, em alguma medida, em relação a elas; e não há razão para pensarmos que isto é menos verdadeiro do que ter uma opinião (ligeiramente) favorável acerca de alguma tese assumida, argumento ou teoria. Então, se uma mente aberta requer reações simpáticas em relação àquilo que se sabe, através da empatia, (e vou dizer mais sobre o que isto envolve mais tarde) acerca de pontos de vista de outras pessoas, isto requer um sentimento ligeiramente positivo e tem, pelo menos, um aspecto decididamente sentimental[6]. No reino da moralidade, podemos distinguir a empatia por alguém que está sofrendo e a simpatia por esse sofrimento, e o quer que digamos acerca da empatia, a simpatia parece envolver um sentimento positivo em relação à *pessoa* que sofre (e um sentimento negativo em relação ao próprio sofrimento, como resultado). Mas não é por acaso que usamos a mesma palavra "simpatia" para crenças, argumentos e teorias. E isso é assim, por que tal simpatia também envolve e requer um sentimento positivo – só

[6] Eu não vou considerar se os psicopatas, que são incapazes de ter simpatia pela dor das outras pessoas e o seu sofrimento, também são incapazes de simpatia intelectual com os pontos de vista das outras pessoas.

que desta vez voltado às coisas mais abstratas, como crenças e argumentos. (Talvez seja uma marca de sermos seres superiores o fato de que podemos ter sentimentos não apenas em relação às pessoas ou as suas ações, mas também sobre seus pontos de vista e opiniões).

Na esfera moral, a empatia de tipo associativo (e do tipo projetivo também) é considerada necessária à simpatia genuína e ao altruísmo, e talvez haja uma relação semelhante entre a empatia e a simpatia na esfera epistêmica. Eu disse anteriormente que podemos basear-nos nas opiniões dos outros através da osmose empática e, assim, torná-las nossas, mas ninguém sustenta que alguém tenha que concordar com as outras pessoas, a fim de ser considerado como alguém que tem uma mente aberta sobre seus pontos de vista. Uma mãe pode simpatizar com um filho que tenha medo de ir ao médico, pode sentir, em certa medida, o seu terror, e pode sentir simpatia pelo que o filho sente, mesmo quando ela insiste em levá-lo ao médico. E algo semelhante pode acontecer com as crenças no contexto da mentalidade aberta. A pessoa de mente aberta não apenas coloca-se no lugar de outra pessoa do mesmo jeito que faz um psicopata. Ela tem empatia pelo ponto de vista do outro, vê as coisas, em certa medida, de uma maneira tão favorável quanto aquela que a outra pessoa vê, e isto também envolve ou evoca o que chamamos de (um certo grau de) simpatia intelectual com o ponto de vista alheio. Mas, como acontece com o filho e o médico, isso não implica que se acabará por acompanhar ou concordar com esse ponto de vista ou as crenças que o compõem.

Em qualquer caso, parece plausível concluir que as virtudes epistêmicas de ter a mente aberta, justeza intelectual, objetividade e imparcialidade (estes diferem entre si de formas que não precisamos mais discutir) exigem certas tendências e a presença de sentimentos favoráveis (e também desfavoráveis), e assim temos de reconhecer que a racionalidade epistêmica completa contém elementos sentimentais ou emocionais anteriormente insuspeitos[7]. Mas, para tornar o

[7] Em "*Intellectual Desire, Emotion, and Action*". In: A. O. Rorty (ed.)*, Explaining Emotions*, Berkeley: University of California Press, 1980*,* p.323-338. Michael Stocker enfatiza os aspectos emocionais do

processo sentimentalista completo lado a lado com a epistemologia e a filosofia da mente, é preciso também considerar se a crença em si possui alguns elementos emocionais. O argumento acima dá vários indícios pontuais de que tem e já é hora de ser mais explícito.

De acordo com o que eu afirmei brevemente acima, a empatia ou a simpatia em relação ao ponto de vista de outra pessoa envolve ver certas coisas da mesma maneira favorável que a outra pessoa as vê. Mas pense no que isto significa: significa que a outra pessoa tem uma visão favorável em relação a um determinado argumento ou teoria, e que a pessoa de mente aberta, ou, como podemos dizer, a pessoa intelectualmente simpática, vê as coisas, pelo menos momentaneamente, na mesma perspectiva favorável em que aquela outra pessoa vê aquele argumento ou teoria. Mas isso significa que uma pessoa que sustenta uma teoria ou que (sinceramente) avança ou advoga um argumento veja aquelas coisas numa perspectiva favorável, favorecendo-as acima de todos os outros (possíveis) argumentos ou teorias; em tal caso, se segue que crer nas teorias ou argumentos como tais envolve uma atitude favorável, algum sentimento positivo ou emoção. E disso resulta que não há tal coisa como uma crença puramente intelectual ou uma opinião. (Mesmo Karl Popper, que disse que não devemos nunca aceitar nenhuma hipótese científica, tinha uma visão favorável e sentia uma emoção positiva em relação à sua própria recomendação ou opinião cética).

Assim, a vida intelectual e a pura aquisição e defesa das crenças na vida cotidiana estão permeadas por elementos emocionais. É preciso emoção para ter opiniões ou crenças e para propor argumentos de uma forma sincera. Precisa-se também de emoção ou de emoções para que se possa tratar com a mente aberta as opiniões e pontos de vista de outras pessoas, e, nessa medida, como já dissemos, a racionalidade intelectual ou epistêmica tem importantes aspectos sentimentais (mesmo se eles não esgotam toda a racionalidade epistêmica). E a relevância da fenomenologia aqui pode agora ter aparecido ou ter se tornado

---

pensamento intelectual, mas não vai tão longe no sentido sentimental como tenho feito aqui – e como fiz em *Moral Sentimentalism*.

transparente. Para que se possa ter uma opinião favorável, para que se possa favorecer uma coisa em detrimento de outra, envolve-se o sentimento tanto quanto quando se favorece um candidato político ou uma criança em relação a outra. Mas quando imaginamos que as crenças podem ser puramente intelectuais ou não-emocionais, nós esquecemos, nós ignoramos, os sentimentos positivos que realmente temos com relação a todas essas coisas. Se olharmos para dentro, se refletirmos um pouco acerca do caráter da nossa experiência comum, poderemos ver, poderemos reconhecer, que isto envolve certos sentimentos. Os sentimentos podem muitas vezes ser tênues o suficiente para que eles não chamem a nossa atenção e, assim, escapem à atenção. Mas uma coisa que eu estou tentando fazer aqui é chamar a atenção para estes aspectos muitas vezes ignorados (inclusive por filósofos) da nossa experiência, para que possamos nos lembrar de uma relação mais clara e atenta com a nossa própria experiência vivida. E eu faço isto, pelo menos em parte, porque eu acho que isto vai ajudar-nos a ver não só a nossa experiência, mas alguns dos nossos mais importantes conceitos filosóficos sob uma nova luz filosoficamente útil. A fenomenologia é importante por que quando prestamos mais atenção à nossa experiência, podemos reconhecer que há elementos sentimentais lá que são inerentemente (mesmo *a priori*) parte da compreensão básica de conceitos epistêmicos, doxásticos ou racionais e isso deve ser dito se quisermos responder (algumas das) perguntas que a epistemologia e/ou a filosofia da mente sempre fizeram e que ainda precisam ser respondidas.

Enfim, tendo chamado a atenção para a questão da (a relevância da) fenomenologia, permita-me mencionar outra área ou aspecto em que a fenomenologia pode esclarecer e iluminar o terreno da epistemologia tradicional. Mencionei anteriormente que, às vezes, é apropriado descartar definitivamente certos pontos de vista, e eu creio que precisamos dizer mais sobre como e quando isto é possível e sobre como tais fatos se relacionam com a ideia de que a receptividade mental é uma virtude epistêmica.

Para colocar as coisas de forma breve, penso que é apropriado rejeitar inteiramente os pontos de vista de outras pessoas, quando essas outras têm uma atitude de mente fechada em relação às opiniões dos outros. Alguém que acredita

que é Napoleão ou que acredita que a Terra é plana, presumivelmente, não vai ouvir às opiniões e argumentos contrários de outras pessoas, e por isso estou dizendo que a virtude de ter a mente aberta tipicamente não obriga-nos a ter a mente aberta face às opiniões e os argumentos daqueles que não possuem uma mente aberta. E há aqui um paralelo com o que se sabe sobre a empatia em contextos morais. Como o psicólogo desenvolvimentista Martin Hoffman chamou a atenção, a nossa preocupação moral empática em relação às outras pessoas diminuirá se conseguirmos vê-las como terceiros, ou seja, de um modo não-empático e despreocupado, machucando, digamos, seus sentimentos gratuitamente[8]. E esta reação não nos ensina a ser pessoas não-empáticas ou despreocupadas, mas é, em certa medida, um sinal de quão empáticos e cuidadosos nós somos. Da mesma forma, o caráter genuíno da receptividade mental empática em relação aos outros não é enfraquecida pelo fato de que alguém é intolerante em relação àqueles mesmos que são intolerantes e carecem de abertura mental. Hoffman diz que uma pessoa cuidadosa terá sua ira despertada quando enxergar outra pessoa machucar um terceiro, mas algo parecido com isto também ocorre em contextos epistêmicos. Nossa intolerância em relação à mentalidade epistêmica fechada ou à intolerância dos outros é uma espécie de *raiva epistêmica*, mas, novamente, este elemento fenomenológico em/ou de nossas vidas epistêmicas nunca foi realmente descrito ou identificado antes. A intolerância em relação às pessoas de cor de pele diferente é uma forma de raiva, mas também e de uma forma parecida, é a intolerância em relação a certas crenças das pessoas. Mas quando essa intolerância é de segunda ordem, é a intolerância do intolerante, então ela pode ser apropriada e compatível com ter a mente aberta – assim como a raiva em relação àquelas pessoas que, com raiva ou de outra forma, machucam os outros, pode ser adequada e consistente com uma atitude geral de preocupação empática em relação aos outros.

Além disso, o próprio fato de que existe tal coisa como uma raiva

---

[8] Martin HOFFMAN, *Empathy and Moral Development: Implications for Caring and Justice*, Cambridge: Cambridge University Press, 2000.

epistêmica apoia a ideia de que o sentimento desempenha um papel importante e inerente à racionalidade epistêmica e na(s) dimensão (ões) cognitiva (s) das nossas vidas. A raiva é um sentimento mais forte do que a atitude levemente favorável que faz parte do processo de, empaticamente, assumir ou compreender a atitude favorável de outra pessoa em relação a uma teoria ou argumento, mas isto é, quando se pensa a respeito, um fato óbvio e atuante em nossas vidas intelectuais ou cognitivas. Muitas vezes somos e estamos justificados em ter, e às vezes somos intolerantes e *não* justificados em ter – raiva ou intolerância diante de certas opiniões. E este fato pode ajudar-nos a reconhecer que as atitudes favoráveis, os sentimentos positivos que estão envolvidos em favorecer certos pontos de vista em relação a outros, também estão amplamente envolvidos na cognição, na aprendizagem e no conhecimento. Além disso, a imagem cumulativa elaborada no decorrer deste capítulo inteiro pode ajudar a tornarmo-nos conscientes de quão importantes são não apenas o sentimento, mas também a consciência fenomenológica e a investigação do sentimento humano, à ética, à epistemologia e à filosofia da mente.

3

Nós ainda não estamos prontos. A crença é uma categoria crucial, um conceito ou elemento em epistemologia e filosofia da mente e argumentamos que a crença, considerada em si mesma, envolve o sentimento em certas formas que não foram previamente percebidas (nem mesmo, eu penso, pelos empiristas ou sentimentalistas britânicos, como Hume). Mas a crença relaciona-se com outros conceitos na filosofia da mente (tradicional) e, se posso colocar as coisas desta forma, funciona como uma espécie de encruzilhada ou centro de intercâmbio entre as funções aferentes e eferentes da nossa psicologia. O que percebemos sobre o mundo dá origem e modela as nossas crenças de um modo que se relaciona e afeta o que fazemos e queremos fazer no mundo. E a nossa discussão anterior neste capítulo efetivamente limitou-se ao lado eferente da crença. Crenças envolvem favorecer certas maneiras de ver ou entender as coisas sobre os outros, e isto é claramente relevante para o modo como vamos (querer) atuar no mundo. E

mais imediatamente ou auto-suficientemente relevante do que se poderia pensar.

Alguém poderia pensar que para a crença influenciar o modo como agimos ou desejamos agir, ela teria que se envolver e atuar com certos desejos especificáveis separadamente – como no modelo do desejo/crença de compreensão da ação humana. Mas este modelo foi recentemente questionado por racionalistas como Thomas Nagel, que pensam que a apreensão de certos fatos, isto é, certas crenças cognitivamente adequadas, pode motivar a ação independentemente de se ter já certos desejos – e às vezes é também sustentado (por exemplo, por T. M. Scanlon, outro racionalista moral) que a apreensão das nossas razões aparentes (e das crenças sobre) para agir podem motivar-nos de algum modo, independentemente do desejo[9]. E o modelo (simples) desejo/crença da ação humana também pode ser questionado por alguém que, de uma maneira não-racionalista, enfatiza a fenomenologia da crença de um modo semelhante àquele sugerido anteriormente. Se (um estado de) crença envolve uma atitude favorável, pelo menos um mínimo de emoção positiva, então a crença, *qualquer* crença, será sem dúvida menos inerte do que a maneira como a crença, em si mesma, tem sido tradicionalmente vista. Se a crença é considerada como um estado puramente intelectual ou estritamente cognitivo, então ela mesma pode parecer não estar relacionada com a motivação e ação humanas. Mas independentemente de todos os argumentos racionalistas contra esta hipótese, o fato de que a crença envolve emoção pode fazê-la parecer relevante para a ação e a motivação humanas, de uma forma que, de outro modo, poderia não ser. Afinal, mesmo que a razão tenha sido durante muito tempo vista por muitos filósofos como sendo motivacionalmente inerte, ninguém ou quase ninguém pensou desta forma em relação à emoção e ao sentimento (eu não estou falando sobre um sentimento do tipo envolvido na percepção da rugosidade de uma superfície, mas do sentimento no sentido em que este se relaciona essencialmente com a emoção). Sentimentos (por exemplo) de ciúme, inveja ou gratidão, fazem-nos

[9] Thomas NAGEL, *The Possibility of Altruism*, Oxford: Oxford University Press, 1970. T. M. SCANLON, *What We Owe to Each Other*, Cambridge: Harvard University Press, 1998.

querer fazer certas coisas, e por isto não são apenas desejos, mas também sentimentos que mais comumente entendemos como imediatamente ou diretamente relevantes e ativos em nossas motivações e ações. Deste modo, o fato de que o sentimento está envolvido com a crença pode tornar mais fácil a compreensão de como a crença pode motivar ações e fazer isto, em certa medida, independentemente das considerações racionais que filósofos como Nagel ou Scanlon invocam para conectar as nossas crenças (ou apreensões racionais) com a motivação relevante e (às vezes, mas nem sempre) à ação.

Deste modo a estranheza do que J. L. Mackie chamou de "prescritividade objetiva", isto é, a estranheza aparente de assumir que a apreciação racional de uma crença em certos fatos pode *ipso facto* constituir uma (mais ou menos eficaz) motivação ou razão para a ação, pode parecer menos estranha se conseguirmos mostrar – como temos procurado demonstrar aqui – que a crença em si tem o tipo de carga emocional que consideramos mais fácil de conectar com a motivação e a ação motivada. (Sou grato neste ponto a Elijah Chudnoff). Deste modo, a ênfase na fenomenologia aqui pode realmente tornar mais fácil para o racionalista tornar o processo de conectar a crença e a ação motivada, embora eu mesmo quereria rejeitar o racionalismo e argumentar, em contrapartida, que a crença pode relacionar-se ou ser relevante à motivação e ação simplesmente através da sua conexão com a emoção e pelo fato de que podemos facilmente entender como os estados emocionais podem motivar.

Muito mais precisaria ser dito acerca do que acabei de tratar, mas tenho outro assunto para abordar antes de terminar o capítulo, e eu penso que é melhor seguir em frente. O outro tópico refere-se ao lado aferente da crença, a crença, como em tantos casos importantes, resultante da percepção ou da experiência perceptiva. A fim de completar nosso panorama das muitas maneiras em que a fenomenologia pode ser importante e útil para a filosofia, eu gostaria de permanecer, tão brevemente quanto o assunto permite fazê-lo com propriedade, na fenomenologia da percepção e como ela relaciona-se com a crença perceptiva. Como sugeri, de um modo um tanto obscuro anteriormente, epistemólogos e filósofos da mente têm (o que parecerá ser às pessoas comuns e ao seu senso

comum) uma atitude um tanto deflacionária ou desdenhosa em relação à fenomenologia da percepção. Os materialistas ou fisicalistas como D. M. Armstrong e George Pitcher trataram a percepção como sendo idêntica a um certo modo de adquirir ou reter crenças e os funcionalistas, como Sydney Shoemaker, parecem considerar a percepção como um conceito definido em parte pela tendência de suas instâncias a dar origem a certos tipos de crenças[10]. E estas abordagens tiveram (pelo menos) duas características interessantes. Elas capturam e até mesmo enfatizam o fato de que a percepção (adulta) normalmente dá origem, causa, certas crenças perceptivas: ver a diferença de tamanho entre dois objetos normalmente faz-nos pensar (ou ver) que eles diferem em tamanho. Mas elas também subestimam o papel da fenomenologia, tanto na percepção em si mesma e, de acordo com o que eu quero dizer, quanto em (descrever) a conexão entre a experiência perceptiva e a crença relevante. A intenção é, antes, entender a percepção dentro de um ponto de vista metafisicamente materialista ou fisicalista da mente e dado que qualquer ênfase na qualidade sentida da nossa experiência perceptiva, na fenomenologia de tal experiência, torna mais difícil evitar o dualismo e terminar com uma visão fisicalista, há um motivo filosófico para economizar ou minimizar a fenomenologia sensorial/perceptiva (por exemplo, tratá-la como redutível ou compreensível em termos de distinções sutis entre crenças perceptivas metafisicamente não-problemáticas e crenças perceptivas menos problemáticas).

Bem, eu mesmo não quero ser um dualista e creio que se pode evitar esse perigo ainda enfatizando a fenomenologia. Então, deixem-me continuar como se isso fosse, de fato, possível. A primeira coisa que eu quero dizer é que a fenomenologia da experiência perceptiva adulta pode realmente ser rica o suficiente para *incluir* ou *envolver* a crença perceptiva. Basta pensar o que significa para nós ver as coisas ao redor do nosso mundo. As coisas não têm, a maioria das

[10] D. M. ARMSTRONG, *A Materialist Theory of the Mind*, London: Routledge and Kegan Paul, 1968; George PITCHER, *A Theory of Perception*, Princeton: Princeton University Press, 1971; e Sydney SHOEMAKER, *Physical Realization*, Nova York: Oxford, 2007.

coisas ao nosso redor, uma imediaticidade ou vivacidade que *ipso facto* envolva a consideração da sua realidade? A maioria das coisas que eu vejo (e vamos nos concentrar na visão) parecem-me extrair sua realidade do modo como as percebo, e isto significa, em particular, que a tendência a crer não é algo adicional ou existente ao lado de certas experiências perceptivas ordinárias, mas está contido nelas mesmas[11].

Se isto é assim, então se a crença também implica uma certa quantidade de emoção positiva, a relação entre a experiência perceptiva ou a percepção e a crença será análoga à ligação que os racionalistas esboçaram entre a apreensão racional de certos fatos e nossa motivação ou razões para agir (uma conexão que parece mais provável de ser válida se alguém reconhece fenomenologicamente a conexão entre a crença e a emoção, positiva e, às vezes, suave). E isso pode, dessa forma, revelar, muito surpreendentemente, que a percepção envolve o afeto positivo e a motivação de uma forma intrínseca.

No entanto (como Brian McLaughlin lembrou-me), nem toda percepção vívida faz-nos acreditar de forma adequada. Há a ilusão de Mueller-Lyer, que persiste mesmo quando estamos convencidos que as duas linhas representadas são de igual comprimento, e há vários casos em que sabemos de antemão que estamos sendo submetidos a algum tipo de alucinação (por exemplo, hologramas). Mas mesmo nestes casos, a nossa experiência perceptiva, sua vivacidade, sua

---

[11] Depois que escrevi o primeiro esboço deste capítulo, encontrei um artigo de A. D. SMITH ("*Perception and Belief*", *Philosophy and Research* 62, p.283-309, 2001), que converge com muitos dos mesmos pontos que acabo de demonstrar acerca da relação entre a experiência perceptiva e a crença perceptiva. No entanto, Smith incorpora suas ideias dentro da premissa husserliana segundo a qual toda percepção de (a realidade de) objetos ocorre contra um pano de fundo da mundanidade assumida, e não penso que alguém precisa de tais pressupostos amplos, a fim de estabelecer a conexão entre a percepção e a crença. (Na verdade, penso que essas pressuposições nos distraem e desviam dos pontos mais mundanos que tanto Smith, quanto eu gostaria de assumir). Smith também não sugere a ligação entre a crença e sentimento positivo/emoção que nos permite fortalecer o caso geral para (a não-estranheza da) prescritividade objetiva; e ele nem invoca o tipo de analogia entre a percepção estética e a percepção sensível, que, conforme argumentarei, ajuda a tornar o processo de uma conexão íntima entre a percepção e a crença perceptiva.

amplitude, nos desperta uma tendência a crer que só pode ser combatido ou minado por nossas outras crenças. E se é preciso de tais outras crenças para derrubar ou neutralizar o nosso sentido de ou nossa crença na realidade do que percebemos, então, dada a fenomenologia da nossa percepção do mundo que nos rodeia, podemos dizer que há uma ligação essencial entre tal experiência e a tendência a crer na realidade do que ela é "de" (seu objeto) e entre a experiência e as crenças reais relevantes para essa experiência (onde estas incluem algumas crenças que não estão relacionadas diretamente à fenomenologia). Estas conclusões parecem-me ser um pouco surpreendentes; mas, novamente, penso que é a fenomenologia que nos levou a elas e fez com que parecessem plausíveis, e, nessa medida, eu gostaria, muito brevemente, de tratar de um último ponto que se refere a como as coisas que falei relacionam-se com a estética.

A fenomenologia tem sido mais relevante para a estética e para a apreciação, digamos, de pinturas belas, excelentes ou boas. Mas (de certa forma inspirado pelo que Platão diz acerca da experiência que uma pessoa racional pode ter da Forma do Bem) eu penso que a fenomenologia da apreciação da arte é também relevante para o que acabamos de dizer acerca da percepção e da crença. Às vezes, é sustentado que para que possamos perceber (inteiramente), reconhecer ou apreciar a beleza de uma obra de arte, por exemplo, uma pintura, é necessário estarmos automaticamente atraídos a contemplar esteticamente ou envolver-nos com ela e estarmos relutantes a abandonar essa experiência muito rapidamente. No entanto, esta ligação parece mais um exemplo de prescritividade objetiva, de como a dimensão cognitiva pode essencialmente (e talvez inextricavelmente) misturar-se com e/ou envolver elementos motivacionais ou emocionais. Eu penso que um exemplo tão aparentemente plausível ou intuitivo do campo da estética ajuda no processo da prescritividade objetiva no domínio da moralidade. (Se a prescritividade objetiva pode ocorrer em uma área do pensamento ou da vida, então quão facilmente pode ocorrer uma objeção de princípio à sua ocorrência em outra área do pensamento ou da vida?). Mas a plausibilidade do processo estético também dá suporte ao que eu estava dizendo anteriormente sobre a conexão entre a percepção e a crença. Afinal, se o caráter

fenomenológico da percepção ou da apreciação da beleza ajuda a dar sentido a uma conexão entre a percepção e o sentimento/motivação no reino estético, por que não poderia haver uma ligação entre algo análogo e a pura/mera percepção sensorial e o sentimento/motivação, algo mediado pelo papel da crença perceptiva em nosso engajamento com o mundo, porém, em grande medida, devido a (e compreensível) em termos do caráter fenomenológico da percepção dos sentidos. Nesse caso, espero tê-los convencido que a fenomenologia pode ajudar-nos não apenas na esfera da moralidade, mas também nas grandes áreas da epistemologia e da filosofia da mente.

5

EM DEFESA DA EMPATIA: UMA RESPOSTA A PRINZ*

Cláudia Passos-Ferreira

1. Empatia e moralidade

A visão dominante em psicologia moral sustenta que as habilidades cognitivas da empatia e da tomada da perspectiva afetiva do outro [*affective perspective-taking*] desempenham um papel crucial na moralidade e em comportamentos pró-sociais e altruístas. De acordo com diversos psicólogos (Eisenberg & Strayer, 1987; Batson *et al*., 1981; Batson & Shaw, 1991; Zhan-Waxler & Radke-Yarrow, 1990; Hoffman, 2000; Vaish et al., 2009, 2011; Decety, 2011) e filósofos (Hume, 1739/1978; Smith, 1759/2009; Slote, 2010; Goldman, 2006; Darwall, 1998; de Vignemont & Frith, 2008), a empatia é fundamental à moralidade.

Recentemente, Jesse Prinz (2005, 2011a, 2011b) desafiou essa tendência em psicologia moral. Prinz defende duas teses centrais. Primeiro, ele argumenta que a empatia não é uma precondição necessária para a aprovação ou desaprovação moral. Segundo, argumenta que a empatia tende a levar a distorções no juízo moral que a tornam potencialmente nociva, e, frequentemente, produz resultados moralmente indesejáveis. As conclusões contraintuitivas de Prinz são que a empatia não desempenha um papel essencial na moralidade, e que ela interfere negativamente com os fins da moralidade, e que, portanto, não deve ser cultivada.

Argumentarei contra ambas as teses. Primeiro, argumentarei que a empatia desempenha um papel necessário no desenvolvimento moral humano. Defendo

---

* A versão em inglês deste artigo foi primeiro publicada na Revista *Abstracta* Vol. 8, n.2, p.31–51, 2015. Agradeço ao editor Alex Tillas a autorização para publicação da versão em português

que a empatia – entendida ou como compartilhamento vicário de emoções ou como tomada da perspectiva afetiva por meio de simulação e reconstrução imaginativa – é fundamental ao desenvolvimento da agência moral. A ausência ou deficiência desses processos leva à ausência ou deficiência de um elemento crucial de nossa moralidade. Em segundo lugar, argumentarei que há um benefício moral associado aos sentimentos empáticos. Argumento também que há certas situações moralmente exigentes nas quais a empatia é nosso melhor guia para o julgamento moral.

O capítulo está dividido em quatro seções. Na primeira seção, explico a visão negativa de Prinz de uma moralidade baseada na empatia, e esclareço minha tese sobre a necessidade da empatia no desenvolvimento moral, sugerindo que essa tese deve ser entendida como uma afirmação sobre o desenvolvimento moral em seres humanos.

Na segunda seção, apresento o argumento de Prinz contra a tese da empatia como necessária ao desenvolvimento moral. A discussão de Prinz sobre o desenvolvimento moral se centra em três questões: os déficits morais em psicopatas, os déficits morais em pessoas autistas e as teorias de desenvolvimento moral. Ele argumenta contra a ideia de que os prejuízos das habilidades empáticas em psicopatas e autistas seriam responsáveis pelos déficits na competência moral dos mesmos. Prinz propõe uma abordagem do desenvolvimento moral baseada na imitação para explicar como o desenvolvimento da competência moral envolve a aquisição de capacidades emocionais via aprendizagem imitativa. Examino as pesquisas empíricas sobre psicopatas e autistas e proponho uma explicação alternativa para o déficit moral dos psicopatas e autistas, que favorece a abordagem do desenvolvimento moral baseada na empatia que proponho.

Na terceira seção, defendo uma conceptualização mais ampla da empatia que nos ajude a entender a competência moral. Então, distingo duas fontes da empatia – empatia perceptiva e empatia imaginativa – baseadas em mecanismos subjacentes distintos.

Na quarta seção, discuto o argumento de Prinz contra a tese normativa, e

argumento que, apesar do potencial de distorção da empatia, existem maneiras de resistir à conclusão de que a empatia não deve ser cultivada. Argumento que há contextos morais nos quais a empatia é um bom guia moral.

2. A empatia desempenha um papel necessário na moralidade?

Começo com a questão de se a empatia é necessária à moralidade. Prinz não foi o primeiro a desafiar a concepção segundo a qual a empatia é necessária. Jannette Kennett, em "Autism, empathy and moral agency" (2002), Victoria McGeer, em "Varieties of moral agency: lessons from autism (and psychopathy)" (2008), Heidi Mailbom, em "Feeling for others: empathy, sympathy, and morality" (2009), e Peter Goldie, em "Anti-empathy" (2011) também adotaram essa perspectiva. Muitos dos que argumentam contra a empatia como uma precondição para a moralidade adotam uma concepção racional kantiana da moralidade. Kennett, por exemplo, argumenta que um exame do pensamento moral em pessoas autistas mostra que a agência moral pode ser desenvolvida na ausência de empatia, e essa evidência pode apoiar uma perspectiva kantiana da agência moral.

Diferente dessas críticas racionalistas[1], Prinz (2007) defende uma concepção sentimentalista da moralidade. Como sentimentalista, Prinz compartilha a intuição humeana de que emoções são essenciais para o juízo moral e para a motivação moral, e que os juízos morais envolvem aprovação e desaprovação. De acordo com Prinz (2011a), o sentimentalismo de Hume pode ser formulado como

[1] McGeer (2008) também endossa uma concepção humeana da moralidade. Ela argumenta que os juízos baseados na razão desempenham um papel instrumental na moralidade, e que apenas as emoções têm a força motivacional necessária que acompanha as atitudes morais, e que, consequentemente, todos os tipos de agência moral humana estão enraizados no afeto. No entanto, ela nega que as habilidades de empatia e da tomada de perspectiva sejam a base da moralidade. Ela sugere que as pessoas com autismo desafiam esta concepção. No autismo, o déficit de empatia e das habilidades de tomada de perspectiva não levam a um déficit na moralidade. McGeer conclui que a empatia não deve ser considerada a única emoção capaz de fornecer motivação moral; outros tipos de disposições afetivas (que estão disponíveis para pessoas com autismo, como a preocupação afetiva) também desempenham um papel na moralidade.

uma afirmação constitucional: "acreditar que algo é moralmente certo ou errado consiste em aprová-lo ou desaprová-lo". Prinz endossa a concepção humeana da moralidade, mas descarta a empatia como a base para a aprovação ou desaprovação moral. Ele argumenta em favor do que chama de *sentimentalismo anti-empático*.

Em "Against empathy", Prinz questiona a tese humeana de que a empatia é uma precondição para a moralidade. Ele analisa dois pontos centrais do projeto de Hume: 1) a *tese definicional* de que a empatia consiste em sentir uma emoção que acreditamos que outra pessoa também sente; 2) *a tese precondicional* de que a empatia é uma precondição constitutiva da aprovação ou desaprovação moral. Quando a tese precondicional é combinada com o sentimentalismo de Hume, segue-se que a empatia é uma precondição para o juízo moral. Prinz endossa o sentimentalismo de Hume, mas rejeita sua tese precondicional.

Mais precisamente, Prinz nega que "a aprovação moral envolve qualquer tipo de congruência entre as emoções daquele que aprova e daqueles em quaisquer lados da ação a ser aprovada"[2]. Há muitas maneiras de conectar empatia e juízo moral. Prinz rejeita todas as teorias que conectam necessariamente empatia e moralidade. Ele argumenta contra seis versões da *tese precondicional*, correspondentes aos seguintes seis tipos de precondição: 1) constitutiva (a empatia é um elemento necessário do juízo moral); 2) causal (a empatia é necessária para causar o juízo moral); 3) desenvolvimental (a empatia é necessária para o desenvolvimento da agência moral); 4) epistêmica (a empatia desempenha um papel epistêmico necessário ao juízo moral); 5) normativa (a empatia é necessária para justificar o juízo moral); e 6) motivacional (a empatia é necessária para a motivação moral).

A forma geral dos argumentos de Prinz contra essas teses pode ser reconstruída do seguinte modo:

P1. Juízos morais são constituídos por sentimentos de aprovação ou desaprovação.

[2] 2011a, p.218.

P2. A empatia desempenha um papel contingente nos sentimentos morais de aprovação ou desaprovação.

P3. Se a empatia desempenha um papel contingente no juízo moral, a empatia não é necessária para a moralidade.

C1. Portanto, a empatia não é necessária para a moralidade.

A premissa P1 expressa o emocionismo[3] de Prinz, e não será alvo da minha discussão. A premissa P2 é uma tese modal, que Prinz sustenta com contraexemplos e estudos psicológicos. A premissa condicional P3 abarca todos os tipos de conexão necessária entre a empatia e a moralidade negadas por Prinz.

Com efeito, há seis versões diferentes do argumento, correspondentes aos seis tipos diferentes de precondições. Como vimos, as seis versões da *tese precondicional* fazem seis tipos diferentes de afirmações sobre o papel necessário da empatia na moralidade. Por exemplo, a *tese precondicional desenvolvimental* afirma que a empatia é necessária para o desenvolvimento moral. No argumento de Prinz contra essa tese, a premissa P2 diz que a empatia desempenha um papel contingente no desenvolvimento moral, e a conclusão diz que a empatia não é necessária para o desenvolvimento moral. Essa é a versão em que me concentrarei.

Negarei a premissa P2, argumentando que a empatia desempenha um papel necessário no desenvolvimento moral dos seres humanos. Argumento que há aspectos de nossa moralidade para os quais emoções empáticas são necessárias. Onde Prinz usa resultados empíricos para argumentar contra a tese precondicional desenvolvimental, sugiro interpretações diferentes dos resultados. Trabalhos experimentais recentes mostram que a empatia desempenha um papel

[3] O emocionismo – tal como definido por Prinz (2007) – consiste em qualquer teoria que afirme que emoções são essenciais à moralidade. Prinz distingue o termo 'emocionismo' – que ele define como um rótulo genérico para qualquer concepção que afirme que sentimentos são essenciais à moralidade – do termo 'emotivismo', que é uma versão específica do emocionismo. Prinz argumenta em favor de uma versão forte da concepção emocionista, defendendo o que chama de 'sentimentalismo construtivo', a concepção de que "sentimentos literalmente criam a moralidade, e os sistemas morais podem ser criados de modos diferentes" (PRINZ, 2007, p.9).

fundamental no surgimento da agência moral em seres humanos. Também endosso uma abordagem ontogenética do desenvolvimento da agência moral baseada na empatia. Uso essas considerações para rejeitar a conclusão C1, que afirma que a empatia não é necessária para a moralidade, ao menos quando essa afirmação está restrita a seres humanos.

A ideia de que a empatia não é necessária para o juízo moral pode ser entendida como uma afirmação sobre todos os sistemas morais ou como uma afirmação sobre seres humanos. Na primeira maneira de entendê-la, afirma-se que existem sistemas morais possíveis em sociedades possíveis nos quais há juízos morais sem empatia. Concordo que a empatia não é necessária ao juízo moral nesse sentido. Podemos conceber um sistema primitivo, por exemplo, como uma savana mítica (Prinz 2007) habitada por humanos primitivos em estado natural, governados por valores pré-morais, em que as pessoas não precisam de empatia para seguir regras e respeitar autoridade. Nessa situação primitiva, a empatia pode não estar envolvida em atitudes de aprovação e desaprovação. Em um estado natural, um agente primitivo pode usar outros elementos para pensar moralmente [*moralizing*]. Ele pode, por exemplo, usar reações de aversão ou a raiva como guias para a ação moral (embora uma questão importante seja a de saber se esse tipo de comportamento pode ser considerado como pensamento moral, ou como uma mera regulação do comportamento).

Para os propósitos presentes, a tese da necessidade deve ser entendida como uma afirmação sobre seres humanos. Ou seja, nos seres humanos, tal como estão constituídos psicologicamente, a empatia é necessária ao seu desenvolvimento moral. Essa é a afirmação que defenderei. É verdade que até mesmo os seres humanos podem desaprovar moralmente uma ação sem diretamente passar pela empatia ou pelo processo de tomada da perspectiva afetiva. Estudos psicológicos recentes (Haidt 2012) mostram que juízos morais podem ser resultado de reações afetivas automáticas. Podemos formar um juízo moral com base em emoções viscerais [*gut feelings*]. Entretanto, na concepção que advogo, emoções viscerais em seres humanos só contam como pensar moralmente [*moralizing*] quando expressam atitudes de desaprovação de um

agente moral; e os traços característicos de um agente moral dependem de ele ser capaz de formar atitudes morais como resultado de um processo de simulação empática e tomada da perspectiva afetiva.

3. Autismo, psicopatia e desenvolvimento moral

Muitos estudos contemporâneos de psicologia moral enfatizam o papel constitutivo da empatia nos primeiros estágios do desenvolvimento moral (ver Batson 1981, 1991, 2011; Hoffman 2000, 2011; Eisenberg 2000; Tomasello & Vaish 2013). A tese central defendida por vários psicólogos morais é que processos empáticos são mecanismos psicológicos subjacentes à agência moral. A moralidade envolve processos de regulação comportamental relativamente aos outros. Bem cedo na ontogenia, bebês começam a usar processos empáticos para regular seu comportamento em relação aos outros. Quando uma criança não mostra empatia em relação aos outros, podemos prever que ela não adquirirá a capacidade de se preocupar com os outros ou de apreciar como suas ações afetam os outros. Consequentemente, a falta de empatia afetará sua capacidade de avaliar suas próprias ações e as ações de outros como sendo corretas ou erradas, e de reagir a essas ações expressando aprovação ou desaprovação. Podemos prever que as pessoas que exibem dificuldades nas habilidades empáticas apresentarão déficits morais.

A ideia de que a empatia desempenha um papel central no desenvolvimento moral é apoiada por diversos estudos que investigam a correlação entre empatia, comportamentos prossociais e cooperação (Batson 1981, 1991, 2011; Hoffman 2000, 2011; Eisenberg & Strayer 1987; Eisenberg 2000; Tomasello & Vaish 2013). No início da ontogenia, as crianças começam a se comportar prossocialmente e a se engajar em ações cooperativas e altruístas. A questão aqui é quais são os mecanismos psicológicos que tornam as crianças aptas a se comportarem prossocialmente. O que esses estudos mostram é que processos empáticos fornecem às crianças habilidades afetivas, cognitivas e motivacionais para se comportarem prossocialmente. Processos empáticos motivam comportamentos altruístas, tais como o de ajudar, cuidar e confortar os outros, e se relacionam

negativamente com comportamentos antissociais e agressivos (Batson 1981, 1991, 2011; Eisenberg 2000; Hoffman 2000). Dados empíricos mostram que a capacidade para o raciocínio moral e para a agência moral é fortemente dependente de capacidades de resposta empática aos estados afetivos de outros e de tomar a perspectiva do outro em consideração.

Prinz desafia essa interpretação e argumenta que não há dados empíricos que evidenciem a conclusão forte de que a empatia está na base do desenvolvimento moral. Ele não nega a correlação positiva entre empatia e juízo moral. O que ele nega é que essa correlação seja evidência para a tese precondicional desenvolvimental. Prinz (2011a) aborda três fontes potenciais de evidência para essa tese: evidências sobre psicopatas, evidências sobre pessoas autistas, e as teorias de desenvolvimento moral.

As duas populações patológicas – psicopatas e autistas – têm sido de especial interesse na psicologia moral. Ambas as populações apresentam deficiências na compreensão social, na responsividade social e competência moral. Ambos os psicopatas e autistas também apresentam deficiências nas habilidades empáticas. Uma visão amplamente aceita é que essas deficiências empáticas explicam suas deficiências morais (Nichols 2004). As pesquisas psicológicas sobre essas populações parecem apoiar a visão de que a falta de empatia afeta a competência moral (Blair 2005), sugerindo que a empatia desempenha um papel fundamental no desenvolvimento moral.

Prinz (2005, 2011a, 2011b) rejeita esta explicação. Ele argumenta que essas evidências sobre autistas e psicopatas não apoiam a tese precondicional desenvolvimental. Considerarei seus argumentos em ambos os casos, focalizando especialmente as evidências sobre psicopatas.

Psicopatia. A psicopatia é um distúrbio associado a traços de insensibilidade e ausência de emoção (falta de medo, culpa, remorso e afeto raso) e comportamento antissocial e agressivo (Blair et al., 2005). É amplamente sustentado que a deficiência emocional dos psicopatas explica sua falta de empatia, e sua empatia prejudicada (seu déficit empático) explica sua falta de competência moral. Os psicopatas não possuem emoções que facilitem a

educação moral; eles não possuem as respostas emocionais que são constitutivas dos juízos morais.

Segundo Prinz (2005), os déficits morais dos psicopatas podem ser explicados sem recorrer ao déficit de empatia. É a falta de emoções básicas, como o medo e a tristeza, e não a falta de empatia que explica a deficiência no raciocínio moral detectado em pessoas com psicopatia. A explicação que Prinz endossa se baseia na hipótese do medo disfuncional[4]. Segundo essa hipótese, os psicopatas apresentam deficiência nos sistemas que modulam o comportamento do medo (Fowles, 1988). Os psicopatas apresentam uma diminuição do condicionamento aversivo, e diminuição de respostas emocionais em antecipação do castigo e na imaginação de eventos ameaçadores. O déficit de medo dos psicopatas os impede de serem socializados e de desenvolverem competência moral. A socialização moral é, em geral, obtida por meio do uso da punição. A punição agressiva instila o medo, e o medo da punição é frequentemente usado no aprendizado moral. Uma criança assustada pelo castigo associará esse medo à ação que resultou na punição e desenvolverá respostas aversivas condicionadas a ameaças antecipadas. Uma criança que não tem medo de punição não aprenderá boa conduta se for ameaçada com punição.

Prinz argumenta que a deficiência moral dos psicopatas pode ser explicada por um déficit no comportamento inibitório e emoções inibitórias. Ele afirma que o mesmo sistema disfuncional que danifica o medo em psicopatas também pode danificar a tristeza, outras emoções negativas, e outras reações negativas. O

[4] Grosso modo, há dois modelos cognitivos na literatura empírica para explicar o déficit moral na psicopatia. Um deles é o sistema comportamental inibitório proposto por Fowles (1988) e outros, que afirma que os psicopatas têm um déficit básico em seu sistema de comportamento rudimentar que está subjacente a muitos aspectos das emoções e provoca comprometimento no comportamento aversivo e no medo. O outro é o modelo inibitório de violência precoce (VIM), e sua versão atualizada para os sistemas de emoção integrados desenvolvidos por Blair (Blair et al., 2005), que explica a natureza do comprometimento emocional em indivíduos com psicopatia como resultado de deficiências em diferentes sistemas, como empatia disfuncional, o medo disfuncional, e o VIM disfuncional.

sentimento de tristeza, afirma Prinz, é crucial para a moralidade porque está envolvido no reconhecimento e na resposta à tristeza dos outros, e é um elemento básico que pode ser usado para criar emoções morais (como vergonha e culpa).

As emoções morais são emoções complexas que surgem em contextos que envolvem conformação ou violação de uma regra moral. Prinz (2005, 2007) considera que as emoções morais (como a vergonha, a culpa, o arrependimento e a indignação) são geradas por uma mistura de emoções básicas (por exemplo, medo, tristeza e raiva) combinadas com um processo de calibração [*calibration process*]. No processo de calibração, como proposto por Prinz, uma emoção básica que inicialmente tinha um conjunto de condições desencadeadoras pode ser atribuída a um novo conjunto de condições desencadeadoras que foram elaborados através da experiência para formar um mecanismo de desencadeamento independente. Por exemplo, Prinz afirma que "a culpa é apenas tristeza que foi calibrada para situações em que alguém causou um dano a alguém com quem se preocupa"[5]. Ou seja, a mistura emocional pode ser associada a situações em que a criança "capta" os estados emocionais de outra pessoa (sofrimento, reação negativa, desaprovação, etc.) copiando os estados emocionais de outra pessoa por meio de imitação ou contágio emocional.

Em "*Imitation and Moral Development*" (2005), Prinz dá à aprendizagem imitativa um lugar fundamental na explicação do desenvolvimento moral. Para desenvolver a competência moral, a criança tem que ser capaz de reagir com emoções negativas na presença da desaprovação dos cuidadores ou na presença de angústia de outro. Ser capaz de reagir com sentimentos negativos nesses contextos requer não apenas a disposição básica dos sentimentos de medo e tristeza, mas também a capacidade de "captar" as emoções dos outros e "captar" a angústia dos outros. Segundo Prinz, essas disposições emocionais são estabelecidas pela imitação. Em nosso processo de socialização, imitamos as emoções percebidas (expressões faciais e vocalizações), e acabamos por copiar (por meio da imitação) os estados internos dos outros, como a vergonha, a culpa e

[5] 2005, p.273.

a angústia dos outros. Ver/Observar a angústia dos outros desencadeia angústia vicária e, posteriormente, no desenvolvimento, desencadeia respostas de consolação. Assim, Prinz conclui que os psicopatas não são bons pensadores morais porque não conseguem aprender a reação emocional apropriada à sua conduta no contexto da angústia de suas vítimas, nem através de aprendizado imitativo nem por contágio emocional.

No entanto, nas seções seguintes, argumentarei que a aprendizagem imitativa não é suficiente para explicar plenamente o desenvolvimento moral. A aprendizagem imitativa pode explicar o reconhecimento de emoções básicas, mas não pode explicar o desenvolvimento de emoções morais como culpa, vergonha, arrependimento, admiração e preocupação empática. Como argumentarei mais adiante, não é possível captar essas reações emocionais complexas "copiando o estado afetivo de outrem" através da aprendizagem imitativa, mesmo no amplo sentido de imitação adotado por Prinz (2005). Os elementos intencionais e motivacionais desses estados afetivos não estão disponíveis para percepção direta e aprendizagem associativa. Uma explicação completa para o desenvolvimento de emoções morais deve envolver empatia. Se isso é correto, a hipótese de Prinz sobre psicopatas é, na melhor das hipóteses, incompleta.

Isso não significa que devamos abandonar a hipótese do medo disfuncional. Em vez disso, podemos usar essa hipótese, mas acrescentar um papel para a empatia. Pode ser que a deficiência no medo[6] cause prejuízos no medo compartilhado e na empatia, e essas deficiências, por sua vez, causam o déficit moral. É claro que ser capaz de sentir emoções (como medo, raiva, tristeza, alegria, repulsa e surpresa) é um pré-requisito para compartilhar essas emoções e para a empatia (emocional). Portanto, a incapacidade de sentir as emoções (como afirmado na hipótese do medo disfuncional), necessariamente, causará um déficit nas emoções compartilhadas e na empatia. É natural supor que esse déficit é o que leva a déficits morais em psicopatas (pelo menos em uma abordagem

[6] Os sentimentos de tristeza não serão abordados, pois se sabe muito pouco sobre como a tristeza é afetada na psicopatia.

sentimentalista da moralidade). De fato, a hipótese é que a empatia comprometida (prejudicada) medeia a conexão entre emoções e moral prejudicada. Isso fornece uma alternativa à hipótese de Prinz de que a aprendizagem imitativa prejudicada medeia a conexão.

É também plausível que a hipótese de Prinz não possa funcionar a menos que seja dado a empatia um papel-chave. Acredita-se amplamente que as deficiências morais dos psicopatas estão especialmente ligadas a deficiências no reconhecimento e resposta à angústia de suas vítimas. Estas deficiências são explicadas naturalmente em termos de deficiências no compartilhamento da angústia das vítimas, o que pode ser visto como uma forma de preocupação empática. A hipótese de Prinz exige que a aprendizagem imitativa sozinha possa explicar o reconhecimento da angústia dos outros. Ele argumenta que a preocupação com o sofrimento da vítima é uma habilidade metacognitiva que emerge tarde no desenvolvimento e deriva do sofrimento vicário inicial, que é uma habilidade mais básica usada para capturar o sofrimento dos outros através do contágio emocional. No entanto, nas seções a seguir, vou argumentar que a angústia vicária requer habilidades empáticas. Ela não pode ser explicada simplesmente através do contágio emocional.

É verdade que há uma forma simples de aflição vicária por contágio emocional em bebês que não requer habilidades empáticas. No entanto, esse tipo de angústia vicária inicial por contágio emocional acontece antes do desenvolvimento da plena auto-diferenciação. No sofrimento vicário inicial, a criança não está experimentando ou reconhecendo a angústia dos outros; a criança está experimentando sua própria angústia, o que leva à angústia pessoal e não à preocupação empática. Esse fenômeno rudimentar não pode explicar os elementos de desenvolvimento moral que envolvem o reconhecimento da angústia dos outros. Para explicar isso, é preciso uma forma mais complexa de aflição vicária envolvendo empatia. Por minha conta, a angústia vicária precoce no contágio emocional evolui primeiro para a angústia vicária empática e, finalmente, para a preocupação empática. A empatia nos ajuda a obter informações sobre a maneira pela qual um evento ou uma ação pode afetar emocionalmente uma

pessoa e causar a angústia dos outros.

Minha explicação do papel da empatia na conexão de emoções e desenvolvimento moral em psicopatas se encaixa bem com pesquisas recentes sobre psicopatas por Blair e outros. Os psicopatas mostram comprometimento no reconhecimento de expressões de medo (face, corpo e voz), redução da experiência de medo (Marsh et al., 2011), comprometimento da resposta ao medo em outros (Marsh, Cardinale 2012), e comprometimento da capacidade de identificar comportamentos que causam medo e julgar a aceitabilidade moral desses comportamentos (Marsh, Cardinale, 2012). Além disso, a psicopatia afeta os julgamentos de transgressões associadas a dano. Os psicopatas tendem a errar no tratamento de violações convencionais, como violações morais, e são menos propensos a justificar seus julgamentos referindo-se ao bem-estar da vítima (Blair, 1995; 2005). Sua propensão a infligir danos a outros indica uma profunda perturbação na sua resposta empática ao sofrimento dos outros (Blair 2005). A capacidade de reconhecer a angústia dos outros é crucial para a experiência de preocupação empática (Nichols 2001). Qualquer prejuízo na capacidade de reconhecimento emocional precoce ou uma incapacidade inata na capacidade de perceber e responder às expressões afetivas dos outros levará a uma empatia emocional disfuncional. Como Blair sugere (2005), um indivíduo que mostra comprometimento na empatia emocional é difícil de socializar através da indução empática, uma prática que envolve o socializador focando a atenção do transgressor sobre o sofrimento da vítima. Tudo isso é mais uma evidência de um papel de empatia na explicação do prejuízo moral dos psicopatas.

Autismo. Prinz sugere que os estudos experimentais (Blair 1996, 2005) mostram que os autistas, diferentemente dos psicopatas, parecem adquirir uma compreensão das regras morais e exibem um déficit de empatia. Ele conclui que, se esta interpretação é correta, "a aquisição de competência moral *pode* não depender de uma capacidade robusta de empatia"[7]. Kennett (2002), Nichols (2004) e McGeer (2008) também argumentam que no autismo o déficit de empatia não

[7] 2011b, p.222.

leva inevitavelmente a um déficit de moralidade. Pessoas com autismo exibem falta de empatia, mas ainda assim têm/apresentam um senso de moralidade. De acordo com Nichols (2004), a habilidade preservada dos autistas de fazer juízos moral, apesar de sua incapacidade de simular a perspectiva de outra pessoa, revela que os relatos de perspectiva da moral devem ser empiricamente errados. A partir desta evidência, Prinz (e também McGeer e Kennett) conclui que a empatia não é necessária para o desenvolvimento da agência moral; se a empatia desempenha algum papel no desenvolvimento moral, desempenha um papel instrumental, portanto, um papel contingente.

Uma maneira de resistir a esta conclusão tem sido mostrar que, enquanto as pessoas com autismo são prejudicadas em empatia cognitiva e habilidades de leitura da mente, eles são capazes de experimentar emoções, mostrar empatia afetiva e emocionalmente responder a angústia dos outros. Isso sugere que sua competência moral pode derivar de suas habilidades emocionais empáticas (Blair 1996; Nichols 2004).

A ideia de que pessoas autistas exibem moralidade sem empatia também foi desafiada por de Vignemont e Frith em "Autismo, Morality, and Empathy" (2008). Elas questionam ambas as ideias: que as pessoas autistas apresentam uma falta de empatia e que eles possuem um senso moral. De Vignemont e Frith argumentam que as pessoas autistas têm algum grau de empatia emocional automática: elas mostram reconhecimento emocional e respostas autonômicas às expressões de tristeza e medo de outros. Estudos psicológicos (Blair 2005) evidenciam que as pessoas autistas podem ter os componentes emocionais dos comportamentos empáticos. Eles são capazes de reagir com empatia ao sofrimento dos outros. Consequentemente, enquanto o autismo pode envolver comprometimento da empatia cognitiva (a capacidade de saber o que outra pessoa pensa), parece que algum nível de empatia emocional permanece intacta nas pessoas com autismo. A falta de comportamento empático no autismo tem sido atribuída a déficits nos processos de mentalização (Batson et al., 1987, Blair, 2005). Apesar de exibirem empatia emocional preservada e capacidade de reconhecimento emocional preservada, estudos baseados em relatos dos pais

sugerem que pessoas com autismo apresentam prejuízos específicos em seu relacionamento afetivo com outras pessoas (Hobson et al., 2006). Eles manifestam claramente sinais de felicidade, angústia, raiva e medo como respostas emocionais ao estado de espírito dos outros; porém, apresentam limitações em experimentar e expressar sentimentos centrados na outra pessoa [*other-person-centered feelings*], tais como simpatia e preocupação; também, raramente expressam sentimentos por e em relação a outras pessoas (Hobson et al., 2009). De acordo com os relatos parentais, eles mostram ciúme em relação aos outros e são afetados pelo estado de espirito dos outros, mas mostram pouca preocupação, culpa ou tristeza empática. As pessoas com autismo são mais propensas a descrever situações em termos de quebrar as regras do que em termos de causar danos físicos ou emocionais a outros (Hobson et al., 2009).

De Vignemont e Frith (2008) sugerem que a presença do componente emocional em pessoas com autismo pode explicar por que a competência moral parece estar aparentemente preservada nos mesmos. As pessoas com autismo são capazes de detectar a transgressão das regras normativas e detectar a angústia dos outros. No entanto, eles não parecem capazes de detectar violações morais. Essa detecção requer a correlação de dois fatos: uma transgressão moral e o sofrimento de alguém sem justificativa moral. As pessoas com autismo parecem não conseguir correlacionar esses dois fatos. De Vignemont e Frith (2007) também sugerem que o problema com as pessoas autistas na detecção de violações morais pode estar relacionado com a forma como fazem a distinção entre representação alocêntrica e egocêntrica. As pessoas com autismo exibem um egocentrismo extremo desconectado do alocentrismo, significando que o seu mundo social é autocentrado. Elas carecem de intuições sociais, fazem análises abstratas do seu entorno, e "estão mais interessados em regras normativas do que em emoções devido a um alocentrismo abstrato desconectado de suas interações egocêntricas com os outros"[8]. Sua conclusão é que não podemos excluir a possibilidade de que as regras seguidas pelos autistas sejam meramente

[8] DE VIGNEMONT & FRITH 2008, p.280.

percebidas por eles como regras convencionais e que sua aparente capacidade de julgamento moral seja o resultado da aplicação dessas regras convencionais.

Embora esta conclusão não seja suficiente para derrotar o argumento de Prinz que as pessoas autistas são capazes de fazer julgamentos morais, ela pode oferecer uma interpretação alternativa a este fenômeno. Primeiro, as habilidades empáticas emocionais parecem estar preservadas em pessoas autistas de alto funcionamento, e essa habilidade preservada pode explicar sua capacidade de fazer julgamentos morais, apesar de suas limitações em experimentar e manifestar preocupação empática, e oferecer conforto em situações de angústia dos outros. Em segundo lugar, não há evidência forte de que sua aparente capacidade de fazer juízos morais seja o resultado da aplicação de regras morais ou demonstração de preocupação moral.

Teorias do Desenvolvimento Moral. O terceiro argumento de Prinz é contra as teorias desenvolvimentistas que enfatizam o papel da empatia no desenvolvimento moral. A psicologia moral do desenvolvimento descreve como evoluímos para nos tornarmos agentes morais, como chegamos a distinguir entre o certo e o errado, e como aprendemos a distinção entre regras convencionais e regras morais. A história desenvolvimentista de Prinz (2005) enfatiza o papel central da imitação no aprendizado de como responder emocionalmente aos juízos morais. Ele sugere que a aprendizagem moral requer um tipo diferente de imitação. As crianças podem "copiar os estados internos dos outros" e não apenas "seus comportamentos dirigidos a metas [*goal-directed behaviors*]". Seu principal argumento é que a imitação nos ajuda a adquirir formas de compreensão moral. Nosso entendimento moral envolve uma gama de capacidades emocionais que dependem do aprendizado imitativo a ser adquirido. Prinz descreve cinco estágios do desenvolvimento moral normal. Na primeira fase, os bebês experimentam as emoções dos outros através da mímica facial; a responsividade moral começa com contágio emocional em recém-nascidos. Esta etapa contribui para o surgimento da preocupação e empatia. No segundo estágio, as crianças pequenas se envolvem em comportamento pró-social e de consolação precoce. Na terceira etapa, as crianças se tornam responsivas às regras morais. Na etapa seguinte, as crianças

começam a se envolver em comportamento reparador e condenação moral. Na última etapa, as crianças distinguem diferentes classes de normas por meio da obtenção de emoções morais (culpa, vergonha, indignação). De acordo com Prinz (2005), a aprendizagem imitativa contribui para todos esses estágios, e nenhuma dessas formas de aprendizagem moral requer empatia. Prinz conclui que a aquisição de competência moral não depende de uma capacidade robusta de empatia.

Há pelo menos dois pontos obscuros na história do desenvolvimento moral de Prinz. Primeiro, ele negligencia explicitamente processos empáticos bem conhecidos que surgem no desenvolvimento humano e são fundamentais para o desenvolvimento da sociabilidade e da competência moral (2005). É amplamente aceito que esses processos contribuem para a aquisição da competência moral. Como discuti em um trabalho anterior (Rochat & Passos-Ferreira 2008), os processos de imitação e espelhamento [*mirroring*] são condições necessárias, mas não suficientes, para que as crianças desenvolvam a moralidade. A imitação fornece um senso básico de conexão social, incluindo o reconhecimento mútuo de existir com outros que são "como eu" [*like-me*] (Meltzoff 2007). No entanto, para que a moral se desenvolva, os processos de imitação e espelhamento precisam ser complementados por um sistema aberto de reciprocidade e representações compartilhadas (de emoções e outros estados mentais). A imitação e o contágio emocional diminuem à medida que o indivíduo desenvolve outras capacidades cognitivas. As pesquisas em psicologia do desenvolvimento mostram que, a partir do segundo mês de idade, a mímica, a imitação e outras respostas de contágio emocional são superadas. A imitação dá lugar a sinais de reciprocidade [*reciprocation*] e regulação emocional mútua. À medida que a atenção conjunta aos objetos se desenvolve, também surgem as representações afetivas compartilhadas com outras pessoas. Com o tempo, um senso moral explícito se desenvolve, acompanhando o aparecimento da leitura da mente [*mind-reading*] e da imaginação aos 4 anos de idade. Por volta dos 5 anos de idade, as crianças mostram uma compreensão explícita dos estados mentais que impulsionam os outros em seus comportamentos e crenças, permitindo que as crianças

compreendam os aspectos motivacionais que desencadeiam as atitudes morais (Rochat & Passos-Ferreira, 2008).

Em segundo lugar, a história desenvolvimentista baseada na imitação de Prinz é apenas uma história parcial sobre a compreensão das mentes dos outros. Imitação e contágio emocional são apenas o ponto de apoio para a compreensão dos outros como portadores de estados mentais. Novos desenvolvimentos são necessários para a compreensão de outros tipos de estados mentais, que estão longe de se resumirem a ações intencionais, desejos, percepção visual e emoções básicas. Para esses estados mentais, há um acoplamento relativamente próximo entre os estados mentais subjacentes e sua expressão na ação corporal. Podemos inferir esses estados através da percepção e imitação, como argumentou Prinz. Nesse sentido, a empatia não é a única maneira de entender a mente dos outros. No entanto, é psicologicamente nosso método mais penetrante para identificar estados mentais em outros. Ele nos permite inferir outros estados mentais de uma maneira mais rápida e precisa. A empatia nos permite fazer previsões mais rápidas e precisas sobre as necessidades das outras pessoas, suas emoções e as intenções de suas ações. Além disso, a empatia é o único mecanismo confiável para a compreensão dos estados mentais dos outros aos quais não temos acesso direto pela percepção e cujos pensamentos não são expressos abertamente em suas ações. A empatia é especialmente relevante na apreensão de crenças falsas, crenças divergentes, perspectivas afetivas e cognitivas divergentes e emoções morais secundárias.

Moralizar – isto é, pensar moralmente – depende de compartilhar os estados afetivos dos outros e levar em conta a perspectiva afetiva dos outros. Compartilhar, simular e imaginar os estados emocionais dos outros é necessário para desenvolver emoções secundárias, tais como sentimentos de preocupação empática, vergonha, culpa, pesar, ressentimento, indignação e admiração. Argumento que um mecanismo empático básico é necessário para aquisição de emoções morais secundárias. A mera capacidade de imitar emoções básicas (medo, raiva, repulsa, felicidade ou tristeza) ou copiar os estados internos dos outros não são suficientes para a compreensão e internalização das emoções

secundárias, as quais são componentes fundamentais de nossa competência moral. Para desenvolver a agência moral, precisamos de um complexo sistema regulatório emocional que seja mais sofisticado do que mera imitação e processos de contágio emocional. Os sistemas empáticos desempenham esse papel (Rochat & Passos-Ferreira, 2008). Na próxima seção, vou sugerir uma outra história de desenvolvimento que vai da imitação à empatia perceptual e a empatia imaginativa.

4. Da imitação a empatia perceptiva e empatia imaginativa

A empatia tem sido definida de várias maneiras (Eisenberg & Strayer, 1987, Eisenberg, 2000, Batson, 1998, Hoffman, 2000, de Vignemont & Singer, 2006, Decety e Jackson, 2006). O termo "empatia" ("sentir como o outro sente") e o termo associado "simpatia" ("preocupar-se com o outro") têm sido usados para se referir a uma ampla família de processos psicológicos. Para definir a empatia, é importante distingui-la de uma variedade de outros fenômenos, como contágio emocional, simpatia, projeção mental e preocupação empática. O contágio emocional é um fenômeno pelo qual uma emoção é automaticamente propagada de um indivíduo para outro, e é caracterizado pela não-diferenciação entre self e outro (por exemplo, um bebê que começa a bocejar quando vê um outro bebê bocejar). Em contraste, a empatia implica uma diferenciação self-outro. A projeção mental é um processo mental no qual nos colocamos na posição do outro para compreendê-lo, por meio da simulação (Goldman 2006; Decety 2004), da leitura da mente [*mind-reading*] (Gopnik & Meltzoff 1997) ou da tomada de perspectiva do outro [*perspective-taking*]. A simpatia é caracterizada pela participação em uma emoção experimentada pelo outro. Trata-se de sentir preocupação, de compartilhar o sofrimento com os outros e buscar seu bem-estar. A preocupação empática, tal como definida por Batson, é "*uma emoção orientada para o outro, suscitada por e congruente com o bem-estar percebido de alguém em necessidade*"[9]; e inclui emoções empáticas, tais como sentimentos de simpatia, compaixão, tristeza,

[9] 2011, p.11.

angústia e preocupação.

Adotarei aqui a concepção amplamente aceita de empatia de Nancy Eisenberg. De acordo com Eisenberg & Strayer (1987), a empatia envolve o compartilhamento da emoção percebida de outro; é uma reação afetiva vicária que "pode ocorrer como uma resposta a pistas abertamente perceptivas indicativas do estado afetivo de outra pessoa (por exemplo, a expressão facial de uma pessoa) ou como consequência de inferir o estado de outrem com base em pistas indiretas (por exemplo, a natureza da situação do outro)"[10].

Tradicionalmente, os psicólogos distinguem dois processos psicológicos envolvidos na empatia: *empatia emocional* (compartilhamento vicário da emoção) e *empatia cognitiva* (tomada da perspectiva mental do outro) (Smith, 2006; Davis, 1983; Hoffman, 1977). A empatia cognitiva envolve tomada da perspectiva cognitiva dos pensamentos e crenças dos outros. A empatia emocional envolve o compartilhamento de estados afetivos com outra pessoa. As diferentes maneiras de conceituar a empatia enfatizam um ou outro desses dois componentes. Alguns pesquisadores se concentram nos aspectos emocionais da empatia, enquanto outros se concentram no processo cognitivo de inferir os estados mentais dos outros. Psicólogos e filósofos distinguem esses processos usando definições estreitas e amplas de empatia. A definição estreita tenta capturar a empatia em sua forma mais básica, identificando-a com contágio emocional, como um processo automático de ressonância afetiva. A definição ampla descreve a empatia como um fenômeno multidimensional que combina os processos (afetivos e cognitivos) envolvidos – ou um conjunto de processos como proposto por Davis (1983) – à medida que aparecem no inicio do desenvolvimento[11].

---

[10] EISENBERG & STRAYER, 1987, p.5.

[11] A abordagem multidimensional da empatia tem sido sugerida por diferentes estudos. Hogan (1969) e Davis (1983) sugeriram uma escala de empatia (cognitiva ou emocional) em que a empatia é considerada como um conjunto de construções que dizem respeito a responsividade e sensibilidade a outros. Para alguns psicólogos, a empatia é um processo unitário que inclui uma classe de fenômenos – como o contágio emocional, a simpatia, a angústia pessoal e a tomada da perspectiva cognitiva – que compartilham o mesmo mecanismo. Hoffman (1977, 2000) sugere uma abordagem

Prinz define a empatia estreitamente[12]. Ele define a empatia como uma emoção vicária que envolve "sentir o que alguém acha que a outra pessoa está sentindo"[13]. De acordo com sua abordagem, a empatia é "uma questão de sentir uma emoção que consideramos que outra pessoa como tendo" como uma resposta a um contágio automático ou como resultado de um exercício da imaginação[14]. Em sua abordagem sentimentalista, Prinz enfatiza os aspectos perceptivos e emocionais da empatia e minimiza as noções racionalistas e intelectualistas de empatia que enfatizam o papel da imaginação, da simulação e da leitura da mente [*minding-reading*]. No entanto, Prinz afirma que a empatia nem sempre é um processo automático na maneira que o contágio emocional é; "às vezes a imaginação é necessária e às vezes experimentamos emoções que achamos que alguém estaria experimentando, mesmo que não tenhamos visto evidências diretas de que a emoção está sendo, de fato, vivida"[15]. Não obstante, Prinz afirma que a imaginação é "excessivamente intelectual"; é "um ato mental que requer esforço por parte do imaginador"[16].

Como observa Prinz, sua definição de "empatia" é semelhante à definição de "simpatia" usada na tradição da filosofia moral, incluindo David Hume e Adam Smith. Contudo, a empatia emocional sozinha não pode desempenhar o papel

---

unitária, na qual o desenvolvimento ontogenético da empatia começa desde o nascimento com a empatia global (contágio emocional), levando ao aparecimento de angústia empática egocêntrica aos 14 meses e ao surgimento da empatia propriamente dita na segunda metade do segundo ano de vida, quando as crianças diferenciam totalmente entre eu e outro. Inversamente, Blair (2005) afirma que o termo "empatia" inclui uma variedade de processos neurocognitivos diferentes e dissociáveis, variando da empatia emocional e empatia perceptiva à empatia cognitiva.

[12] De Vignemont e Singer (2006) também propõem estreitar o conceito de empatia. No entanto, defendem a exclusão do componente automático como parte de sua definição. Elas definem a empatia como um estado afetivo consciente, isomorfo ao estado de outra pessoa, que é suscitado pela observação ou imaginação do estado afetivo da outra pessoa.

[13] 2011b, p.215.

[14] PRINZ, 2011b, p.215.

[15] 2011a, p.212.

[16] 2011a, p.212.

crucial requerido para uma abordagem sentimentalista da moral. Mesmo para sentimentalistas como Hume e Smith, a aprovação moral e desaprovação envolvem colocar-se imparcialmente na perspectiva da pessoa afetada e compartilhar suas emoções e reações. Para compreender o papel que a empatia desempenha na moralidade – pelo menos de acordo com os sentimentalistas – devemos adotar uma concepção mais ampla de empatia que inclua empatia emocional (compartilhamento de emoções) e empatia cognitiva (tomada da perspectiva afetiva).

A definição ampla (incluindo ambas a empatia cognitiva e afetiva) especifica o conteúdo da empatia como uma reação às experiências observadas do outro que é compartilhada (compartilhando estados cognitivos e estados emocionais). Entretanto, a distinção entre empatia cognitiva e empatia emocional não captura todos os processos envolvidos no compartilhamento de emoções. Sugiro uma distinção adicional que não se concentra nas reações empáticas (cognitivas ou afetivas), mas nos mecanismos psicológicos subjacentes necessários para acessar os estados afetivos dos outros. Distinguirei a empatia como uma resposta a uma percepção direta das emoções dos outros – o que eu chamo de empatia perceptiva –; e a empatia como resposta à simulação imaginativa ou projetiva da perspectiva afetiva dos outros – o que chamo de empatia imaginativa. Esses mecanismos ajudam a explicar como a empatia emocional se desenvolve em um estado emocional mais sofisticado que nos permite perceber ou imaginar diretamente ou simular os estados emocionais dos outros. Estes dois processos são parte de um contínuo de processos empáticos que emergem no inicio do desenvolvimento humano da capacidade de compreender e identificar o estado emocional de outra pessoa. À medida que as habilidades cognitivas se desenvolvem, há uma cadeia ontogenética de processos que vai da imitação e do contágio emocional à empatia, simpatia, compaixão e tomada da perspectiva. Essa distinção nos ajuda a entender a base desenvolvimental da conexão entre empatia e moralidade.

No início do desenvolvimento, distinguimos diferentes níveis de empatia. Este processo começa desde o nascimento através de imitação neonatal e contágio emocional, e leva à capacidade de imitar e reverberar os estados

emocionais dos outros. Mais tarde, desenvolve-se uma compreensão dos estados emocionais e intenções dos outros, juntamente com a tomada da perspectiva afetiva, através da atenção compartilhada [*joint-attention*], simulação e imaginação. A habilidade de compreender e responder ao estado emocional de outra pessoa aparece bem no início do desenvolvimento do bebê e aumenta para níveis complexos de empatia ao longo do tempo.

Elisabeth Pacherie (2004) sugere três níveis de empatia e de seus respectivos mecanismos psicológicos, num contínuo que vai do contágio da imitação e da emoção aos processos empáticos perceptivos e imaginativos, cobrindo diferentes estágios do desenvolvimento infantil. Em cada estágio da ontogênese, as crianças desenvolvem habilidades empáticas correspondentes à compreensão de três aspectos dos estados mentais dos outros: 1) o tipo de emoção experimentada pelos outros; 2) a situação que está causando a emoção específica vivida pelos outros; e 3) os fatores motivacionais desencadeados pela emoção. Os três níveis de empatia são a habilidade de identificar uma emoção, a habilidade de compreender o objeto intencional da emoção, e a habilidade de entender a conexão entre o tipo de emoção, seu objeto intencional e os fatores motivacionais desencadeados pela emoção. Nesse sentido, minha proposta desenvolvimentista pode ser vista como uma elaboração da abordagem de Pacherie.

O primeiro nível é o do reconhecimento da emoção, que é a capacidade *de identificar o tipo de emoção* experimentada pelos outros. Como surge nossa capacidade de usar pistas perceptivas para entender a emoção experimentada pelos outros? Este nível começa com a imitação precoce de expressões faciais e vocais em recém-nascidos. De acordo com Andrew Meltzoff (1977), os recém-nascidos são capazes de equiparar seus próprios comportamentos não vistos com gestos e expressões faciais que eles vêem outros executarem. A imitação facial sugere um mapeamento inato entre a observação da expressão do outro e a execução de uma ação motora. Na imitação, há uma correspondência automática entre a informação visual da expressão facial observada e a informação proprioceptiva da representação motora. Quando um bebê imita uma expressão

facial, sua imitação é baseada em uma representação motora formada quando ele está observando a expressão de outra pessoa. Na imitação precoce há uma correspondência entre observar uma expressão facial, adotar a expressão facial ou postura corporal observada, e sentir a emoção correspondente (Meltzoff, 1977). A imitação facial dos recém-nascidos leva ao contágio emocional por meio do feedback facial e vocal. Aos dois meses, os bebês se envolvem em proto-conversações face-a-face, em interações reciprocidade [*reciprocating*] com os outros, num processo de mútua regulação emocional e sintonia afetiva (Rochat & Passos-Ferreira 2008). A imitação e o contágio emocional se baseiam em dois processos distintos: uma conexão direta entre percepção e ação e uma conexão direta entre percepção proprioceptiva e expressão facial (Pacherie 2004). A imitação neonatal e o contágio emocional sempre envolvem propriocepção – uma consciência dos movimentos e posições do nosso corpo – mas não envolvem uma explícita distinção eu-outro.

Ao contrário da imitação e do contágio emocional, a empatia emerge quando o bebê se torna consciente da distinção eu-outro. No desenvolvimento precoce, surgem funções cognitivas específicas que permitem aos bebês distinguir o contágio emocional – que não envolve a consciência da distinção eu-outro – e a empatia – que envolve a consciência da distinção eu-outro. Como observa Pacherie (2004), o primeiro nível de empatia envolve o surgimento de uma conexão direta entre a representação motora evocada e a experiência emocional *sem ter que passar necessariamente pelo estágio proprioceptivo*, isto é, sem a correspondente imitação da expressão dos outros. Na forma inicial da empatia perceptiva, os bebês têm acesso perceptivo ao estado emocional do outro através de gestos faciais e expressões vocais sem necessariamente formar uma representação motora através da propriocepção. Isso permite que os bebês distingam entre sentir suas próprias emoções, observar as mesmas emoções nos outros e compartilhar os estados afetivos dos outros.

O segundo nível de empatia é a capacidade de *compreender o objeto da emo*ção. Neste nível, o sujeito identifica a relação que liga a emoção do outro a uma dada situação. Essa capacidade surge com o desenvolvimento de processos

de atenção conjunta, referências sociais e comunicação intencional. Com o surgimento, aos nove meses, do impulso de vivenciar conjuntamente com os outros eventos e objetos no ambiente, os bebês começam a aprender e a desenvolver significados compartilhados sobre eventos e objetos e compreender as intenções dos comportamentos dos outros. O significado de um evento percebido (por exemplo, se algo é perigoso ou ameaçador, ou é desaprovado) é agora referido através das respostas emocionais dos outros, e, em certa medida, é avaliado em relação aos outros (Rochat & Passos-Ferreira 2008). Essa capacidade cognitiva permite que o sujeito compreenda os estados afetivos dos outros. Na atenção conjunta e nos processos de referência social, quando a criança observa um objeto ou um evento que é o foco da atenção de sua mãe, a criança trata as emoções da mãe e sua expressão facial e vocal como um comentário sobre o objeto ou evento. Interpretamos as emoções de outrem como um comentário e uma avaliação de situações e eventos, que nos dão informações sobre o nosso ambiente. Tais processos permitem que a criança compreenda o papel causal das emoções e compreenda as motivações das reações afetivas dos outros. Como observa Pacherie (2004), ao se tornarem referenciais, as crianças desenvolvem acesso às motivações dos agentes e desenvolvem a habilidade de identificar as intenções imediatas do agente observando o modo como o mesmo se comporta. Nossas intenções estão refletidas em nossos movimentos corporais e a simples observação de uma ação realizada por outros nos permite detectar as intenções e os estados motivacionais dos outros.

Nos dois primeiros níveis de empatia, há uma conexão direta entre percepção e ação, que permite ao sujeito identificar o tipo de emoção e compreender o objeto intencional associado à emoção observada (Pacherie 2004). Eu chamo esta forma de empatia, que envolve uma *percepção direta* dos estados afetivos do outro, de *empatia perceptiva*. A empatia perceptiva desempenha um papel crucial em situações em que o sujeito tem pistas perceptivas que permitem o acesso direto ao tipo de emoção e seu objeto intencional por meio de mecanismos perceptivos. Permite ao sujeito compreender os estados mentais dos outros, como metas, atitudes, motivações e estados afetivos, e identificar a situação que está

causando as emoções das outras pessoas.

O terceiro nível de empatia envolve a capacidade de *compreender a correlação entre o tipo de emoção, seu objeto intencional e seus fatores motivacionais*. Esta forma de empatia depende de simulação e capacidades imaginativas. Aos dois anos de idade, as crianças começam a se envolver em jogos elaborados de imaginação e de fazer de conta simbólicos em que objetos e ações no mundo real são levados a assumir as posições de objetos e ações no reino do faz-de-conta. Eles começam a imaginar situações hipotéticas e a criar personagens imaginativos. Esta capacidade se amplia uma vez que surgem as habilidades cognitivas adicionais durante o desenvolvimento da criança. As crianças adquirem progressivamente flexibilidade imaginativa e capacidade de simular a perspectiva cognitiva e afetiva de outros. Esta forma mais elaborada de empatia – que chamo de *empatia imaginativa* – é necessária quando a situação não oferece ao observador um acesso transparente ou direto aos estados mentais dos outros. Em contextos opacos, as emoções não são expressas abertamente, e os aspectos motivacionais das ações dos outros podem diferir de nossas próprias motivações em contextos semelhantes. Nos estágios iniciais da empatia, os processos de imitação e contágio emocional envolvem principalmente emoções básicas (por exemplo, felicidade, medo, tristeza, raiva, surpresa e nojo), que são caracterizadas por expressões faciais universais das quais o sujeito pode ter percepção direta e acesso transparente. Na empatia imaginativa, a imaginação e simulação mental são mecanismos fundamentais que permitem ao sujeito compreender emoções secundárias (sociais e morais) e inferir seu potencial motivacional. Segundo Pacherie (2004), em *contextos transparentes*, ambas as formas de empatia perceptiva – identificação do tipo de emoção e compreensão da conexão entre emoção e seu objeto intencional – podem emergir de mecanismos perceptivos que estabelecem uma conexão direta entre percepção e ação. Em contextos transparentes, podemos perceber claramente os sinais que indicam o estado afetivo de outra pessoa (por exemplo, expressões faciais de uma pessoa ou gestos corporais). No entanto, em *contextos opacos*, nos quais as pistas perceptíveis estão ausentes, devemos confiar na empatia imaginativa para

apreender a conexão ternária entre o tipo de emoção, seu objeto intencional e os fatores motivacionais desencadeados pela emoção.

Durante a maior parte de nossas vidas, estamos envolvidos em contextos opacos onde precisamos de empatia imaginativa e simulação mental para entender e inferir os estados emocionais dos outros. A empatia, definida como esta capacidade de compreender, via percepção ou imaginação, o tipo de emoção e a conexão entre emoção, aspecto motivacional e objeto intencional, é essencial para o desenvolvimento moral. A capacidade de expressar atitudes morais envolve a capacidade de compreender e identificar reações emocionais secundárias como culpa, vergonha, desprezo, arrependimento, admiração, raiva e preocupação. A empatia imaginativa desempenha um papel central na compreensão dessas reações afetivas e nos permite internalizar essas reações emocionais conforme as imaginamos ou simulamos com base em outras. Podemos experimentar, por exemplo, sentimentos de vergonha, culpa, arrependimento, admiração ou indignação em certas circunstâncias, porque podemos nos colocar na pele dos afetados pela ação e compartilhar suas atitudes reativas. É assim que as crianças compreendem e internalizam regras morais e atitudes morais.

De acordo com essa concepção, a empatia envolve a simulação mental e a imaginação dos sentimentos dos outros, a imaginação de como os outros percebem nossas ações e a imaginação de se os outros aprovam ou não nossas ações. A internalização dos sentimentos imaginados e a simulação da perspectiva afetiva dos outros é crucial para o desenvolvimento de um agente moral capaz de seguir regras morais e se comportar moralmente. As características típicas de um agente moral dependem do mesmo ser capaz de expressar as atitudes morais como resultado de um processo de simulação empática e tomada da perspectiva afetiva. Como foi discutido, a imitação e o contágio emocional são apenas o primeiro passo deste processo. O surgimento da empatia perceptual e imaginativa é necessária para o desenvolvimento da capacidade de pensar moralmente.

5. A empatia é benéfica para moralidade?

A empatia deve desempenhar um papel na moralidade? Devemos cultivar

empatia na moralidade? De acordo com Prinz (2011a, 2011b), uma moralidade baseada na empatia é prejudicial à sociedade. Ele argumenta que emoções empáticas podem levar a imprecisões em nossos juízos morais e não contribuem para boas práticas morais. Uma moralidade baseada na empatia tem muitas limitações como um guia para a motivação moral. A empatia pode não ter a força motivacional necessária para a condução de comportamentos pró-sociais e ações altruístas. A empatia é suscetível a distorções e tende a ser altamente seletiva. Ela também pode levar ao tratamento preferencial e a crimes de omissão. De acordo com Prinz, se a empatia produz distorções no juízo moral e interfere negativamente com a moralidade, então, ela deve ser evitada como um guia para a moralidade. Prinz conclui que a empatia deve ser desencorajada como o componente motivacional central de um sistema moral.

O argumento prescritivo de Prinz pode ser reconstruído assim:

P4. A empatia produz vieses [*biases*] no juízo moral e interfere negativamente na moralidade.

P5. Se a empatia produz distorções no juízo moral e interfere negativamente na moralidade, ela deve ser evitada como um guia para moralidade.

C1. Portanto, a empatia deve ser evitada como um guia para a moralidade.

A premissa P4 baseia-se em pesquisas experimentais que sugerem que a empatia é prejudicial e produz vieses no juízo moral. Esses resultados levam Prinz à conclusão de que a empatia não deve ser cultivada. Eu argumento contra esta conclusão prescritiva ao rejeitar a segunda premissa P5.

O argumento deriva uma conclusão prescritiva (a empatia deve ser evitada) a partir de uma premissa epistêmica (a empatia é prejudicial ao juízo moral). Argumentos semelhantes foram defendidos por Holton e Langton (1999) e Struchiner (2011). Eles enfatizam as limitações e distorções que a empatia pode trazer à moralidade. Holton e Langton (1999) tem o uso da identificação imaginativa como ferramenta epistêmica para a moralidade. Uma das suas

principais preocupações é que a empatia leva ao paroquialismo[17]. A empatia é propensa ao paroquialismo, porque ocorre mais facilmente diante indivíduos que são salientes, aqueles percebidos no momento presente e espacialmente mais próximos de nós, ou que se parecem com nós mesmos (Goldman, 2006). Como observa Hoffman (2000), embora tenhamos empatia por quase qualquer um em situação de perigo, é mais fácil termos empatia com aqueles semelhantes a nós mesmos.

Uma objeção que os defensores da empatia poderiam levantar ao argumento de Prinz é que a empatia poderia ser melhorada combinando-a com ferramentas epistêmicas adicionais e dispositivos úteis. Goldman (2006) mostra que a empatia, como processo envolvendo imaginação e simulação, pode ser aprimorada por informações derivadas perceptivamente para gerar representações mais precisas de um indivíduo anônimo e distante e para transcender o paroquialismo de uma perspectiva afetiva autocentrada. Prinz sugere que melhorias deste tipo podem levar a empatia a desempenhar um papel causal inerte neste processo e que ele poderia facilmente ser substituído por outras emoções, como raiva e ultraje.

Noel Struchiner (2011) também argumenta que a empatia não é necessária para juízos morais, desenvolvimento moral ou motivação moral. Em acordo com Prinz, Struchiner afirma que a empatia é uma emoção "perigosa" que leva a atos de crueldade e injustiça, e deve ser evitada, ou mesmo eliminada, nos sistemas legais. Ele argumenta que, nos sistemas jurídicos, um modelo de tomada de decisão baseado na empatia pode resultar em erros, distorções e abusos. Struchiner argumenta que a empatia é tão potencialmente prejudicial às decisões legais quanto à moralidade. Nessa base, ele conclui que um modelo de tomada de decisão baseado em regras capta a essência da lei, e um bom modelo de tomada de decisão para decisões legais deve invalidar qualquer componente empático[18].

---

[17] N.T. O termo paroquialismo é aqui usado no sentido de uma visão circunscrita a um grupo local, ignorando perspectivas, atividades ou interesses mais vastos e abrangentes.

[18] O modelo de decisão baseado em regras proposto por Struchiner (2011) baseia-se na concepção de regras de Frederick Schauer, que é em si uma resposta aos argumentos de Ronald Dworkin contra

Struchiner propõe que os sistemas legais devem ser guiados pelo raciocínio baseado em uma perspectiva autista. Ele defende o que chama de "a moralidade contingente da tomada de decisões autista baseada em regras". A ideia essencial aqui é que as pessoas com autismo possuem as virtudes certas para um bom modelo jurídico – as virtudes das regras. Pessoas autistas adoram sistematizar; seguem regras rígidas, e levam a sério a literalidade na qual as regras são formuladas. Struchiner sugere que o sistema legal deve abraçar essas virtudes. O pensamento autista, do qual a empatia está ausente, produziria menos viés nos juízos morais e menos distorção nas decisões legais.

Claramente, há um problema com esta caracterização da mente autista. Como discuti anteriormente, a capacidade de seguir regras e detectar transgressões normativas que caracterizam o raciocínio autista não resulta em uma capacidade de detectar transgressões morais. Consequentemente, a presença dessas habilidades não resulta em competência moral em pessoas com autismo. Além disso, mesmo admitindo que possamos encontrar mentes autistas como Struchiner as descreve – nas quais a competência moral se baseia em seguir regras normativas – isso não vindicará o sistema legal baseado em regras. Mesmo que tal sistema funcione melhor em certas circunstâncias, ainda temos de considerar situações nas quais os sistemas legais não podem tomar decisões baseadas apenas em seguir regras. Existem situações que necessariamente requerem o nosso poder imaginativo e nossa capacidade de nos colocarmos na perspectiva afetiva daqueles afetados pela ação – por exemplo, as situações de conflito moral são uma dessas circunstâncias. Nessas situações, os sistemas jurídicos devem ser capazes de transcender as distorções das visões de uma perspectiva egocêntrica. Os juízes devem raciocinar como participantes desinteressados que podem tomar a perspectiva dos afetados e tomar decisões com base na flexibilidade imaginativa.

Existem duas ideias implícitas na abordagem de Prinz. A primeira idéia é que se a empatia não é necessária para todos os tipos de juízo moral, então a

---

o positivismo jurídico em "The model of rules" (1967).

empatia não é necessária para a moralidade; e, consequentemente, poderia ser sistematicamente substituída por outras emoções. A segunda ideia é que a moralidade é, no entanto, baseada em um único tipo de emoção. Prinz sugere que *o ultraje (a indignação)* poderia ser o tipo de emoção que desempenharia o papel central na moralidade, porque tem mais poder de motivação do que a empatia, e é menos susceptível a distorções. De acordo com Prinz (2011a, 2011b), devemos cultivar uma moral baseada no ultraje (na indignação)[19].

Segundo a visão que defendo, a empatia é uma característica crucial do desenvolvimento moral. Importantes aspectos de nossa moralidade estão relacionados a sentimentos empáticos. Além disso, a empatia deve ser cultivada. O fato de que a empatia produz distorção e preconceito não implica que ela não seja benéfica para a moral em geral. Emoções empáticas podem distorcer a nossa percepção e são propensas aos vieses de uma perspectiva egocêntrica; porém, isso é verdadeiro para outras emoções também, como medo, vergonha, arrependimento, admiração, culpa, ultraje, raiva e ciúme. O ultraje, por exemplo, pode levar ao linchamento, uma violência coletiva em que um grupo pune um indivíduo que transgrediu as regras morais.

Nossas reações emocionais podem distorcer nossa visão e interferir negativamente com nosso julgamento moral, levando a juízos incorretos e ações moralmente erradas. Podemos ser mal orientados por nossas reações afetivas (Goldie 2002). Reações emocionais podem levar a equívocos no julgamento moral. Isso mostra que não devemos confiar em nossas reações afetivas como a única

[19] Por exemplo, Prinz (2011a) afirma: "Nós devemos protestar contra o errado. [...] De uma perspectiva prática, talvez seja melhor tentar cultivar um sentimento de indignação pela injustiça onde quer que ela ocorra, e um sentimento de alegria em ajudar os necessitados onde quer que estejam. A suposição de que a empatia é essencial para esses fins pode ser equivocada e os esforços para expandir nossos horizontes morais por indução empática podem nos tornar mais vulneráveis aos erros de alocação". E em uma crítica à ética feminista (2011b), ele afirma "uma moralidade feminista inclinada à libertação não deve ser uma moralidade baseada na empatia se esse rótulo é destinado a descrever uma moralidade que faz da empatia o seu recurso emocional primário. Uma moralidade baseada na indignação pode ser mais eficaz".

fonte de nossa atitude moral. Pensar moralmente [*moralizing*] envolve transcender nossa perspectiva afetiva egocêntrica (Rochat & Passos-Ferreira 2008) e levar em conta a perspectiva afetiva dos afetados pela ação.

Se excluímos todas as emoções que poderiam levar a distorções, limitações ou preconceitos no julgamento moral, os sentimentalistas morais serão deixados com muito pouco para contar como um guia positivo para aprovação moral e desaprovação moral. Os sentimentalistas precisam de emoções para distinguir o certo do errado. Para os não-sentimentalistas, a exclusão dessas emoções será menos problemática. Por exemplo, os kantianos e outros racionalistas morais podem contar com as habilidades de raciocínio para desenvolver a competência moral (como sugere Kennett (2002)). No entanto, sentimentalistas morais (como Prinz e eu) não têm outra alternativa senão contar com as disposições afetivas, incluindo algumas emoções tendenciosas.

Na visão sentimentalista que advogo, a empatia é um elemento crucial da moralidade, e com força moral motivacional. Em algumas circunstâncias específicas, é o nosso melhor guia para a moralidade. A explicação para isso é que a empatia nos permite transcender nossa perspectiva afetiva egocêntrica e simular as perspectivas daqueles afetados por uma ação. Transcender nossa própria perspectiva e tomar a perspectiva afetiva daqueles afetados por uma ação são requeridos para o juízo moral; e a empatia, na maioria das vezes, fornece a melhor maneira de acessar as perspectivas dos outros. Este acesso pode ser imperfeito, mas, no entanto, é altamente benéfico em geral.

Uma abordagem semelhante tem sido defendida por alguns proponentes da ética do cuidado. Esta tradição tem ajudado a enfatizar o papel das emoções relacionadas à empatia. Como observa Virginia Held (2006), a ética do cuidado valoriza emoções, como empatia e simpatia, como uma ferramenta epistêmica para determinar o que a moral recomenda. A ética do cuidado também rejeita a idéia de que você deve favorecer o raciocínio abstrato e a imparcialidade para evitar distorções e arbitrariedades (Held 2006). A empatia desempenha um papel fundamental nos contextos de ajuda e cuidado aos indivíduos que não podem expressar suas emoções, desejos e crenças.

Além disso, nesses casos de cuidado, a empatia imaginativa é muitas vezes benéfica. Mark Coeckelbergh (2007) analisa o processo de tomada de decisão em uma unidade de terapia intensiva para bebês e mostra que a empatia imaginativa desempenha um papel central nas decisões sobre a vida daqueles que não podem se expressar de uma forma transparente. As unidades de saúde infantil são contextos opacos onde não podemos confiar exclusivamente nas pistas perceptivas e na empatia perceptiva para decidir o que é moralmente certo ou errado. Temos que usar nossa imaginação e simular a perspectiva afetiva interna dos bebês para tomar decisões. Devemos exercer a empatia com pessoas que sofrem e não são capazes de expressar seus desejos e preocupações. Eles apelam para nossos poderes imaginativos e eles podem querer que compartilhemos de sua vulnerabilidade e sofrimento como companheiros humanos. Como afirma Coeckelbergh (2007): eles apelam a "um esforço imaginativo por parte do cuidador para imaginar *como é ser a outra pessoa* [*what it is like to be the other person*] ao tomar a perspectiva interna (imaginar como seria ser aquele que sofre) e pela comunicação desse esforço imaginativo"[20].

Em nossa sociedade, cuidar daqueles que carecem de autonomia, incluindo aqueles que nunca desenvolveram autonomia e aqueles que estão temporariamente incapazes de agir como agentes morais autônomos, é um valor central. Como parte de nossa moralidade, aprovamos a prática de cuidar daqueles que, por várias razões, não são capazes de expressar suas preocupações e desejos. Este grupo inclui bebês, pessoas com deficiência e pacientes em estados terminais que perderam a consciência ou têm outras doenças mentais. Nessas situações, as habilidades empáticas são essenciais para evitar as distorções de tomar nossa própria perspectiva ou aplicar regras rígidas que possam potencialmente produzir erros e injustiças. Nos casos em que devemos prestar ajuda e cuidados àqueles cuja experiência afetiva é opaca a nós, a falta de empatia imaginativa levará a distorções no juízo moral e também nas decisões legais. Nesses casos, a empatia é moralmente benéfica.

[20] 2007, p.69.

Conclusão

Argumentamos que a empatia é um elemento crucial na moralidade e que em algumas circunstâncias específicas é o nosso melhor guia para a moralidade. Eu argumentei contra duas teses do sentimentalismo anti-empático de Prinz. Argumentei contra sua tese desenvolvimental, que diz que a empatia não é necessária para o desenvolvimento moral. Eu também argumentei contra a tese normativa de Prinz, que diz que a empatia deve ser evitada como um guia para a moralidade.

Para pensar moralmente, precisamos transcender nossa perspectiva afetiva egocêntrica para corrigir as limitações e distorções desta perspectiva. Podemos fazer isso compartilhando estados afetivos e imaginando as reações daqueles afetados por nossas ações. Desta forma, a empatia serve como guia positivo para o juízo moral.

Referências

BATSON, Daniel C.; et al. "Is empathic emotion a source of altruistic motivation?" In: *Journal of Personality and Social Psychology* 40, p.290-302, 1981.

_____., FULTZ, J., & Schoenrade, P.A. "Adults emotional reactions to the distress of others". In: N. Eisenberg & J. Strayer (Eds.), *Empathy and its development*. Cambridge, England: Cambridge University Press, p.163-184, 1987.

_____.; SHAW, Laura L. "Evidence for altruism: Toward a pluralism of prosocial motives". In: *Psychological Inquiry* 2, p.107-122, 1991.

_____; SAGER, K., et al. "Is empathy-induced helping due to self-other merging?" In: *Journal of Personality and Social Psychology* 73, p.495-509, 1997.

_____. *Altruism in humans*. New York: Oxford University Press, 2011.

BLAIR, R. James R. "A cognitive developmental approach to morality: investigating the psychopath". In: *Cognition* 57, p.1-29, 1995.

_____. “Brief report: Morality in the autistic child”. In: *Journal of Autism and Developmental Disorders* 26.5, p.571-579, 1995.

_____. “Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations”. In: *Consciousness and Cognition* 14, p.698-718, 2005.

_____. et al. *The psychopath: emotion and the brain*, Blackwell Publishing, Oxford, 2005.

COECKELBERGH, Mark. “Who needs empathy? A response to Goldie’s arguments against empathy and suggestions for an account of mutual perspective-shifting in contexts of help and care”. In: *Ethics and Education* 2.1, p.61-72, 2007.

COPP, David. “Jesse Prinz, The Emotional Construction of Morals (Oxford: Oxford University Press, 2007): Prinz’s Subjectivist Moral Realism”. In: *Noûs* 45.3, p.577-594, 2011.

DARWALL, Stephen. “Empathy, sympathy, care”. In: *Philosophical Studies* 89.2/3, p.281-282, 1998.

DAVIS, Mark. “Measuring individual differences in empathy: evidence for a multidimensional approach”. In: *Journal of Personality and Social Psychology*, 44.1, p.113-126, 1983.

DECETY, Jean. “L’empathie est-elle une simulation mentale de la subjectivite d’autrui?” In: A. Berthoz & G. Jorland (org.), *L’Empathie*. Paris: Odile Jacob, p.53-88, 2004.

_____; JACKSON, P. L. “A social-neuroscience perspective on empathy”. In: *Current Directions in Psychological Science* 15.2, p.54-58, 2006.

DE VIGNEMONT, Frederique; SINGER, Tania. “The empathic brain: how, when and why?”In: *Trends in Cognitive Sciences* 10.10, p.435-441, 2006.

_____; FRITH, Uta. “Autism, morality and empathy”. In: W. Sinnott-Armstrong (Ed.), *The neuroscience of morality: emotion, brain disorders, and development*. Cambridge: MIT Press, p.273-280, 2008.

DWORKIN, Richard. “The model of the rules I”, In: R. Dworkin. *Talking rights seriously*. Cambridge, MA: Harvard University Press,1977.

EISENBERG, Nancy; STRAYER J. “Critical issues in the study of empathy”. In: N. Eisenberg & J. Strayer (Eds.), *Empathy and its development*. Cambridge: Cambridge University Press, 1987, p.3-13.

_____; “Emotion, regulation, and moral development”. In: *Annual Review of Psychology* 51, p.665-697, 2000.

FOWLES, D. C. “Psychophysiology and psychopathy: a motivational approach”. In: *Psychophysiology* 25: p.173-391, 1988.

GOLDIE, Peter. “Can we trust our emotions?” In: *Richmond Journal of Philosophy* 1.1, p.27-35, 2002.

_____. “Anti-empathy: against empathy as perspectiveshifting”. In: P. Goldie & A. Coplan (Eds.), *Empathy: Philosophical and psychological perspectives*. New York: Oxford University Press, 2006.

GOLDMAN, Alvin. *Simulating minds*. New York: Oxford University Press, 2006.

GOPNICK, Alison; MELTZOFF, Andrew. *Words, thoughts, and theories*. Cambridge, Mass.: MIT Press, 1997.

HAIDT, Jonathan. *The righteous mind*. New York: Pantheon, 2012.

HELD, Virginia. “The ethics of care”. In D. Copp (Ed.) *The Oxford Handbook of ethical theory*, New York: Oxford University Press, 2006, p.537-566.

HOBSON, R.P.; et al. “Foundations for self-awareness: an exploration through autism”. In: *Monographs of the Society fro Research in Child Development Serial* 71.284, 2006.

HOBSON, J. A., et al. “Anticipatory concern: A study in autism”. In: *Developmental Science* 12.2, p.249-263, 2009.

HOFFMAN, Martin. “Empathy: its development and prosocial implications”. In: C. B. Keasey (Ed.), *Nebraska Symposium on Motivation*, vol. 25. Lincoln: University of Nebraska Press, p.167-218, 1977.

_____. *Empathy and moral development: the implications for caring and justice*. Cambridge, UK: Cambridge University Press, 2000.

_____. “Empathy, Justice and the Law”. In: A. Coplan and P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives*. Oxford: Oxford University Press, 2011, p. 230-254.

HOGAN, R. “Development of an empathy scale”. In: *Journal of Consulting and Clinical Psychology* 33, p.307-316, 1969.

HOLTON, R.; LANGTON R. “Empathy and animal ethics”. In: D. Jameson (ed.), *Singer and his critics*. Oxford: Blackwell, 1999, p.209-232.

HUME, David. *A treatise of human nature*. P. H. Nidditch (Ed.). Oxford: Oxford University Press, 1739/1978.

KENNETT, Jennett. “Autism, empathy and moral agency”. In: *The Philosophical Quarterly* 52, p.340-357, 2002.

MAIBOM, Heidi. “Feeling for others: Empathy, sympathy, and morality”. *Inquiry* 52: p.483-499, 2009.

MARSH, A. A.; BLAIR, R. J. “Deficits in facial affect recognition among antisocial populations: A meta-analysis”. In: *Neuroscience and Biobehavioral Review* 32, p.454-465, 2008.

_____. et al. “Adolescents with psychopathic traits report reductions in physiological responses to fear”. In: *J. Child Psychol. Psychiatry* 52, p.834-841, 2011.

_____.; CARDINALE, E. M. “Psychopathy and fear: specific impairments in judging behaviors that frighten others”. In: *Emotion* 12, p.892-898, 2012.

MCGEER, Victoria. “Varieties of moral agency: lessons from autism (and psychopathy)”. In: W. Sinnott-Armstrong (Ed.), *The neuroscience of morality: Emotion, brain disorders, and development*, Cambridge: MIT Press, 2008 p.227-258.

MELTZOFF, A. N.; MOORE, M. K. “Imitation of facial and manual gestures by human neonates”. In: *Science* 198: p.75-78, 1977.

_____. “The ‘like me’ framework for recognizing and becoming an intentional agent”. In: *Acta Psychologica* 124.1, p.26-43, 2007.

NICHOLS, Shaun. “Mindreading and the cognitive architecture underlying altruistic motivation”. In: *Mind and Language* 16(4): p.425-455, 2001.

_____. “Sentimental rules: On the natural foundation of moral judgment”. Oxford: Oxford University Press, 2004.

PACHERIE, Elisabeth. “L’empathie et ses degrés”. In: A. Berthoz & G. Jorland (org.), *L’Empathie*. Paris: Odile Jacob, 2004.

PATRICK, C. *Handbook of psychopathy*. The Guilford Press, New York, 2005.

PRINZ, Jesse. “Imitation and moral development”. In: S. Hurley and N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science* (Vol 2: Imitation, human development, and culture). Cambridge, Mass: MIT Press, 2005, p.267-283.

_____. *The emotional construction of morals*. New York: Oxford University Press, 2007.

_____. “Is empathy necessary for morality?” In: P. Goldie & A. Coplan (Eds.), *Empathy: Philosophical and psychological perspectives*. New York: Oxford University Press, 2011a , p.211-239.

_____. “Against empathy”. In: *Southern Journal of Philosophy* 49, p.214-233, 2011b.

ROCHAT, Philippe; PASSOS-FERREIRA, Claudia. “From imitation to reciprocation and mutual recognition”. In: J. Pineda (Ed.), *Mirror neurons systems. The role of mirroring processes in social cognition.* New York: Humana Press, 2008, p.191-212.

SLOTE, Michael. *Moral sentimentalism*. Oxford: Oxford University Press, 2010.

SMITH, Adam. *The theory of moral sentiment*. New York: Penguin Books, 1759/2009.

_____. “Cognitive empathy and emotional empathy in human behavior and evolution”. In:*The Psychological Record* 56, p.3-21, 2006.

STRUCHINER, Noel. “No empathy towards empathy: Making the case for autistic decision-making”, based on the so-named paper presented at McMaster University Philosophy of Law Conference “The Nature of Law: Contemporary Perspectives”, May 13-15, 2011, in Hamilton, Ontario (Canada), 2011.

TOMASELLO, Michael; VAISH, Amrish. “Origins of human cooperation and morality”. In: *Annual Review of Psychology* 64, p.231-255, 2013.

VAISH, Amrish, et. al. “Sympathy through affective perspective taking and its relation to prosocial behavior in toddlers”. In: *Developmental Psychology* 45.2, p.534-543, 2013.

_____; WARNEKEN, Felix. “Social-cognitive contributors to young children’s empathic and prosocial behavior”. In: J. Decety (Ed.), *Empathy: From bench to bedside*, Cambridge: MIT Press, 2011.

ZAHN-WAXLER, Carolyn; RADKE-YARROW, M. “The origins of empathic concern”. *Motivation and Emotion* 14.2, p.107-130, 1990.

_____; ROBINSON, J. “Empathy and guilt: early origins of feelings of responsibility”. In: J. P. Tangney & K. W. Fischer (Ed.), *Self-Conscious Emotions.* New York: Guilford, 1995, p.143-173.

6

O PAPEL DA INDIGNAÇÃO DE UMA PERSPECTIVA PERFECCIONISTA*

Jônadas Techio

Introdução

A motivação inicial para a redação este capítulo foi o choque causado por um comentário de Stanley Cavell a uma cena do filme *Mr. Deeds Goes to Town*[1]. O protagonista desse filme é Longfellow Deeds (Gary Cooper), um sujeito simples do interior que ganhava a vida escrevendo versos para cartões, até que inesperadamente herdou uma fortuna de um tio milionário e mudou para sua mansão em Nova York – uma história que causou bastante interesse na imprensa local, que chegou a apelidá-lo de "Cinderella Man". Na cena que me interessa Deeds sai para jantar em um restaurante cuja propaganda afirma que lá você pode "comer com os literatos". Numa grande mesa encontram-se poetas e escritores que, ao serem informados da presença de Deeds, convidam-no a juntar-se a eles. Sabendo que Deeds escreve poemas, os intelectuais começam a fazer perguntas, em tom condescendente e debochado, sobre seus métodos de escrita. Passado algum tempo Deeds se dá conta de que está servindo apenas como objeto de piada, e ao levantar para sair afirma: "Creio que nem todas as pessoas famosas são grandes pessoas". Dois dos escritores protestam desfaçadamente e tentam impedir sua saída, ao que Deeds reage derrubando ambos a socos.

---

* Uma versão deste texto em língua inglesa foi publicada anteriormente na revista Ethic@ - Florianópolis, Santa Catarina, Brasil, v. 15, n. 2, p. 260-278. Nov. 2016. Agradeço a autorização dos editores para a publicação da versão em português.

[1] Frank Capra, USA/Columbia, 1936.

Como vocês avaliariam moralmente essa situação? Não há dúvidas de que os literatos agiram de forma errada, mas e quanto à reação de Deeds – será que foi legítima ou adequada? Numa análise abstrata e descontextualizada do episódio, é provável que muitos de vocês sintam-se inclinados a recriminar a atitude de Deeds; no entanto, para um espectador que está a par dos detalhes da cena é quase inevitável concluir que *no mínimo* Deeds estava legitimado a expressar sua indignação, *de uma forma ou de outra*. (Descrevo aqui uma impressão compartilhada por vários outros espectadores que viram o filme comigo em aula.) É claro que socar alguém é uma forma um tanto radical de fazer isso, mas é preciso recordar que essa não foi a *primeira* reação de Deeds, tendo sido suscitada apenas depois do fracasso da tentativa de expressar *verbalmente* sua indignação. Deixando de lado, por ora, a questão sobre graus aceitáveis de reação, o ponto de princípio que me interessa explorar é este: será que não existem ocasiões em que a atitude moral e racionalmente adequada é justamente mostrar indignação, por oposição a continuar o diálogo?

Uma das razões pelas quais essa questão me interessa é que ela pode nos ajudar a testar certas intuições acerca da natureza e da finalidade de uma discussão moral, e, portanto, acerca da própria natureza da racionalidade prática. Muitos filósofos parecem comprometidos com uma concepção do raciocínio moral que toma como seu fim o *acordo racional* entre os agentes. Tal concepção, se tomada ao pé da letra, exclui (quase que por definição) o tipo de desfecho apresentado acima do âmbito propriamente moral; desse ponto de vista, expressar indignação seria apenas uma maneira *irracional* de livrar-se das pressões exercidas sobre o sujeito no contexto de uma discussão moral.

Numa passagem que faz alusão à cena que descrevi acima, Cavell assume uma posição diametralmente oposta, afirmando que a constituição de uma comunidade moral por vezes poderá exigir que "alguns sejam socados [*punched out*], silenciados [*made speechless*] em seus esforços de usurpar ou desvalorizar a fala de outros"[2]. Na sequência gostaria de explorar alguns aspectos importantes da

[2] *Cities of Words: Pedagogical Letters on a Register of the Moral Life*. Cambridge (MA) & London:

posição cavelliana acerca da moralidade, tentando esclarecer e defender, ainda de que forma limitada, essa chocante alegação.

1. Qual é a finalidade de uma discussão moral?

Na parte III de *The Claim of Reason*[3] Cavell cita o seguinte trecho do diálogo Eutífron:

> Sócrates: Mas que tipo de desacordo, meu caro, produz inimizade e fúria? Consideremos desta maneira. Se discordássemos, eu e tu, a respeito de qual dentre dois números é o maior, será que o desacordo a respeito disso nos tornaria inimigos e faria com que nos enfurecêssemos um com o outro? Não deveríamos resolver essa disputa de uma vez por todas recorrendo ao cálculo?
> Eutífron: Sem dúvida alguma.
> Socr: E então, se discordássemos quanto ao tamanho relativo de duas coisas, não deveríamos pôr fim ao desacordo de uma só vez, medindo-as?
> Euth: É isso mesmo. […]
> Socr: Mas sobre o que, então, discordando e não sendo capazes de chegar a decisão alguma nos tornaríamos de fato inimigos e enfureceríamo-nos um com o outro? [...] Não seria a questão do justo e injusto, do belo e do vergonhoso, do bem e do mal? Não são perguntas sobre estes assuntos que fazem com que você e eu e todos os demais entremos em disputas, quando de fato disputamos, se divergimos sobre elas e não somos capazes de chegar a nenhum acordo satisfatório? (*Apud*. CR 253)

Neste trecho do diálogo Sócrates aponta para uma diferença entre discussões matemáticas ou científicas, por um lado, e discussões morais, por outro. O fulcro da diferença tem a ver com as condições para se obter acordo em cada caso. Assim, se eu e meu interlocutor estivermos em meio a uma disputa sobre determinada magnitude, por exemplo, basta que sejamos ambos

---

Harvard University Press, 2004, p. 207. Doravante "CW".

[3] *The Claim of Reason: Wittgenstein, Skepticism, Morality and Tragedy*. Oxford: Oxford U. P., 1979. Doravante "CR".

competentes nas práticas de contagem e medição para que possamos dirimir esse desacordo. Entretanto, pelo menos em grande parte dos casos, o desacordo acerca do melhor curso de ação a ser tomado não parece tão simples de resolver. Qual é a razão dessa dificuldade? Será que bastaria sermos ambos (eu e meu interlocutor) moralmente competentes para que pudéssemos chegar a um tal acordo? Mas que tipo de competência exatamente seria essa? Seria ela análoga à competência de um matemático ou de um cientista – uma questão de *conhecimento* acerca de princípios, regras ou fatos morais, ou ainda acerca do significado de noções morais?

Essas questões tem dividido os filósofos ao longo da história. O próprio Sócrates teria defendido que o conhecimento é a base da virtude: assim, *em uma situação ideal* em que ambos os interlocutores estivessem num mesmo patamar cognitivo no tocante às condições para bem agir, supostamente não seria possível desacordo racional. Mais próximo de nossos dias, Moore pode ser citado como um exemplo adicional de cognitivista moral: de acordo com sua posição "intuicionista", juízos morais acerca do "bem" (ou qualquer outro termo moral) deveriam concordar tanto quanto juízos empíricos acerca do tamanho relativo de dois objetos[4].

Do outro lado dessa disputa temos as posições não-cognitivistas, tais como aquelas defendidas pelos principais expoentes do positivismo lógico no início do séc. XX. A. J. Ayer pode ser tomado como um representante. Em *Language, Truth and Logic*, ele defendeu que:

> […] a filosofia ética consiste simplesmente em dizer que os conceitos éticos são pseudo-conceitos e, portanto, inanalisáveis. [...] Não pode haver tal coisa como uma ciência ética, se por uma ciência ética queremos dar a entender a elaboração de um sistema 'verdadeiro' de moral. Pois, como vimos, dado que juízos éticos são meras expressões de sentimento, não pode haver nenhuma forma de determinar a validade de qualquer sistema ético. E, de fato, não faz sentido perguntar se qualquer sistema dessa

[4] Ver MOORE, G.E. *Principia Ethica*. Cambridge: Cambridge University Press, 1960.

espécie é verdadeiro[5].

A disputa entre cognitivismo e não-cognitivismo em ética e metaética leva ao que Stanley Bates caracterizou como uma antinomia: o cognitivismo implica uma exigência de acordo entre os agentes que parece absolutamente irrealista; já o não-cognitivismo parece completamente inconsistente com nossas práticas morais. Como resume Bates: "se a teoria emotiva estivesse correta, então o uso da linguagem ética por parte de alguém seria ou bem um ato de má-fé ou bem de alienação, conforme essa pessoa acreditasse ou não na teoria"[6].

Ora, a "solução" padrão para antinomias consiste em mostrar que há um pressuposto problemático compartilhado por ambos os lados da disputa, e essa será justamente a estratégia adotada por Cavell. O pressuposto em questão, nem sempre formulado explicitamente, é que "a lógica e, mais particularmente, a ciência fornece os modelos para a racionalidade de um argumento" (CR 260-261); ou, em outras palavras, que a argumentação moral só poderia ser considerada racional caso tivesse uma estrutura similar ou análoga à do raciocínio dedutivo, levando "de premissas que todos aceitam, por meio de passos que todos podem acompanhar, até uma conclusão que todos devem aceitar" (CR 254). Cognitivistas como Moore (e, antes dele, Platão, ou, pelo menos, o Sócrates de Platão) aceitam esse pressuposto e procuram desesperadamente mostrar que o raciocínio moral pode satisfazer essas exigências, ainda que o custo seja a postulação de uma faculdade intuitiva especial; não-cognitivistas como Ayer também aceitam tacitamente esse pressuposto, mas por perceberem que as discussões morais ficam aquém do padrão de racionalidade empregado no âmbito das disputas científicas acabam excluindo a moralidade da esfera da avaliação racional, relegando-a à "expressão de sentimentos".

Tal como Ayer e outros não-cognitivistas, Cavell considera irrealista esperar que argumentos morais levem a conclusões que todos deveriam aceitar; contudo, argumenta ele, assim como a falha em alegações epistêmicas cotidianas não

[5] AYER, A. J. 1936, p.168.

[6] "Stanley Cavell and Ethics". In: Eldridge, R. (Ed.) (2003), *Stanley Cavell*, p. 22-3.

mostra que o conhecimento é impossível, nem que a empreitada epistemológica é inútil (um ponto estabelecido nas partes I e II de CR), tampouco o desacordo em relação às conclusões das discussões morais precisa ser tomado como índice de uma falha *geral* de sua racionalidade. Ainda que a *esperança* de obter um acordo seja uma motivação essencial dessas discussões – caso contrário não estaríamos *levando a sério* o ponto de vista de nosso interlocutor – Cavell defende a legitimidade e o valor do *desacordo racional* acerca de um curso de ação. Isso é possível por que, a seu ver, o que distingue a racionalidade da irracionalidade em qualquer domínio não é a adesão a um *conjunto específico* de procedimentos de justificação (por exemplo, dedução + indução, ou sejam quais forem os procedimentos considerados corretos na lógica, na matemática e nas demais ciências), mas sim o comprometimento em seguir os padrões considerados adequados *em seus respectivos domínios*, procurando fornecer suporte e justificação ao que dizemos ou ao modo como agimos com base nesse comprometimento.

Nesse sentido, um dos traços distintivos da racionalidade própria da lógica e das ciências é justamente o fato de que se espera que os usuários competentes desses padrões de argumentação *concordem sobre suas conclusões*; em outras palavras, o próprio acordo é, nesses casos, uma marca de competência, e, portanto, de racionalidade. Mas existem outros tipos de racionalidade, tais como aqueles expressos em discussões estéticas e morais. De modo geral, pode-se dizer que uma discussão é racional à medida que os juízos apresentados pelos interlocutores envolvidos na mesma forem *suportados por razões*; mas nada, exceto um compromisso tácito com uma concepção intelectualizada e cientificista da racionalidade, nos obriga a pensar que a habilidade de fornecer razões deva ser identificada com a habilidade de aplicar princípios gerais (*a priori*?) a casos particulares, ou com a capacidade de extrair regras gerais a partir da experiência de várias instâncias. O próprio Kant, normalmente considerado um expoente paradigmático dessa concepção intelectualista da racionalidade prática, já notara a peculiaridade dos juízos estéticos, cujo funcionamento não segue a lógica dedutiva nem indutiva, mas nem por isso os excluiu do âmbito da racionalidade, nem

colocou em questão sua pretensão de objetividade ou universalidade[7]. O tipo de competência que importa no caso dos juízos estéticos é uma capacidade altamente desenvolvida de um sujeito – digamos, um crítico – para detectar o que poderíamos chamar de saliências estéticas, isto é, aspectos ou características objetivas dos fenômenos que se encontram na origem da experiência estética, ancorando seu juízo em tais aspectos. Cavell resume esse ponto ao dizer que "o problema do crítico, assim como o do artista, não é excluir sua subjetividade, mas incluí-la; não é superá-la em busca de acordo, mas dominá-la de maneiras exemplares"[8]. Assim, um juízo estético pode ser visto como um convite do crítico para que outros compartilhem de sua maneira de ver, ou, de modo mais geral, de experienciar uma obra, e é por essa razão que Cavell defende que:

> É essencial para fazer um juízo estético que estejamos, em algum momento, preparados para dizer em seu suporte: mas você não vê, não ouve, não saca [*dig*]? O melhor crítico conhecerá os melhores momentos. Porque se você não vê *algo*, sem explicação, então não há mais nada a discutir (MWM 93).

Essa última frase é absolutamente fundamental para meus propósitos; como afirmou Wittgenstein em um contexto diferente "as explicações devem acabar em algum lugar"[9], e saber *quando* e *como* parar em um contexto específico também constitui um indício importante da competência de um sujeito em determinada esfera de discussão[10]. Adiantando alguns resultados, gostaria de defender que algo análogo deve valer também para as discussões morais, razão

---

[7] Ver KANT, *Critique of Judgement*. Walker N. E Meredith, C. (Tr.). Oxford: Oxford University Press, 2007; especialmente §§7-8 e §19.

[8] *Must We Mean What We Say?*. Cambridge: Cambridge University Press, 1976, p. 94. Doravante "MWM".

[9] WITTGENSTEIN, L. *Philosophical Investigations*. 3rd Ed. G.E.M Anscombe (Ed. & Tr.). Oxford: Blackwell Publishing, §1, 2001.

[10] Talvez não seja necessário lembrar, mas é claro que a competência nos procedimentos argumentativos da matemática e das demais ciências comporta igualmente esse traço. Não precisamos estar dispostos a discutir *indefinidamente* com *qualquer* interlocutor.

pela qual, justamente, elas poderão terminar abruptamente, sem que isso seja um indício de irracionalidade. Mas antes disso quero enfatizar esta importante diferença entre os tipos de acordo esperados nos campos respectivos da estética e das ciências. Ecoando as palavras de Cavell citadas acima, pode-se dizer que no caso das ciências o acordo é garantido justamente através da *exclusão da subjetividade*; já no caso das discussões estéticas o acordo *depende essencialmente* de um uso controlado ou exemplar da subjetividade. É justamente por causa dessa característica que as discussões estéticas – do mesmo modo que as discussões morais, mas diferentemente das discussões científicas – permitem que os participantes desvelem, para si mesmos e para os demais, aspectos íntimos de si mesmos, articulando e tornando inteligíveis as posições que estão adotando, e pelas quais assumem responsabilidade. Nisto reside o interesse, mas também o risco peculiar das discussões estéticas e morais: elas fornecerão aos participantes uma oportunidade importante para desenvolverem suas individualidades, potencializando um aumento no nível de autoconhecimento, bem como a construção ou descoberta de uma comunidade; mas há sempre a possibilidade de falha, a qual pode acarretar na descoberta, por parte do sujeito, de sua própria confusão e opacidade, podendo levar à humilhação, à rejeição e finalmente ao isolamento.

Deixando de lado o paralelo com a estética, cabe elucidar melhor que tipo de razões podem ser *moralmente* relevantes, ou seja, que tipos de considerações são legítimas do ponto de vista dos procedimentos de uma discussão moral. Cavell dá uma pista para responder a essa pergunta em uma passagem na qual critica a posição de Charles Stevenson, justamente por não fornecer um bom critério dessa espécie, considerando legítimo o uso de *qualquer enunciado* que possa vir a *alterar as atitudes* de um interlocutor numa discussão moral. O principal problema dessa posição, de acordo com Cavell, é que sua adoção implica tratar o interlocutor como um mero objeto a ser manipulado, ao invés de uma *pessoa*, "uma criatura com compromissos e preocupações [*commitments and cares*]" (CR 283). Como elucida Mulhall:

> para Cavell, os compromissos de uma pessoa não são mais ou menos

> externos aos seus desejos, posições ou formas de conduta, mas sim implicações do que ela faz e de quem ela é. Se, por exemplo, alguém faz uma promessa, então essa pessoa se compromete a realizar um curso de ação; se ela falhar ao executar tal ação, então, a fim de manter sua credibilidade como agente moral, ela deverá explicar por que as circunstâncias em que se encontrava justificam sua incapacidade de honrar aquele compromisso, por que ela não poderia ter dado um aviso prévio àqueles que dependiam de sua promessa, e assim por diante[11].

Assim como um agente não pode *simplesmente* deixar de levar a sério seus próprios compromissos anteriores, sob pena de ser exilado do reino moral, tampouco se pode considerar moralmente competente alguém que procure criticar o comportamento de outro agente sem levar em conta (ou ao menos sem dispender um esforço para *tentar compreender*) os compromissos e preocupações desse agente. Em outras palavras: avaliações morais (competentes) não devem concentrar-se excessiva ou mesmo exclusivamente no conjunto das ações e escolhas de um sujeito em um determinado momento, mas devem ser feitas contra o pano de fundo constituído por seus compromissos e preocupações anteriores. O erro a ser evitado é tomar a identidade de um ser moral como uma mera soma de ações ou escolhas "discretas" (certas ou erradas), e a alternativa é focar na sua identidade narrativa, a qual, embora esteja sujeita a uma revisão contínua, normalmente está longe de ser completamente instável[12].

---

[11] MULHALL, S. *Stanley Cavell: Philosophy's Recounting of the Ordinary*. Oxford: Oxford University Press, 1994, p.37.

[12] Agradeço a Flávio Williges por sugerir esta formulação. Aproveito para registrar uma conexão entre essa discussão e o tratamento do papel da estabilidade em nossas práticas morais fornecido por Peg O'Connor no livro *Morality and Our Complicated Form of Life: Feminist Wittgensteinian Metaethics* (Pennsylvania State University Press, 2008). O ponto que me interessa particularmente nessa obra é a crítica de certa "imagem filosófica" segundo a qual apenas *fundamentos* poderiam fornecer *estabilidade* às nossas práticas morais. A alternativa que ela apresenta é tão simples na sua formulação quanto fecunda em consequências, a saber, tentar mudar a metáfora dominante para lidar com a normatividade em metaética, abandonando a busca por *fundamentos* ou *alicerces*, em prol da compreensão das condições que permitem a obtenção e manutenção de *estabilidade* entre os vários

Se a descrição da lógica do raciocínio moral apresentada até aqui estiver correta, então podemos concluir que discussões morais serão competentes à medida que exibirem o apelo a pelo menos dois tipos de razões, as quais Cavell denominará, respectivamente, de "bases para preocupação" ("*basis of care*") e "fundamentos do compromisso" (*"grounds of commitment*"); a primeira "fornece qualquer sentido que possa haver em sua atitude de confrontar alguém com o que ela 'deveria' fazer" *[ought];* a segunda "fundamenta aquilo que você diz que alguém 'está obrigado' a fazer *[must]* dados seus compromissos, incluindo tanto suas obrigações explícitas quanto as implicações do que ele faz e de onde ele se encontra, pelas quais ele é responsável" (CR 325). Uma implicação importante desta análise, a qual gostaria de enfatizar, é que se um interlocutor tiver desafiado um determinado comportamento meu apelando ao tipo de razões recém descritos – respeitando minhas "preocupações e compromissos" – então não poderei simplesmente ignorar tal desafio, sob pena de manifestar incompetência e, nesta medida, irracionalidade em um contexto moral. Mas isso não significa que um desafio competente exija sempre *aceitação* de minha parte; é possível que eu *reconheça a relevância* da "base para preocupação" e do "fundamento do compromisso" apresentados por meu interlocutor sem concordar com o *peso* ou a *importância* que ele dá aos mesmos.

Eis mais uma diferença importante entre discussões morais e disputas científicas: nestas procura-se determinar se devemos ou não *aceitar* determinada alegação cognitiva de um interlocutor com base nas *evidências* que ele nos

---

aspectos de nossas práticas. A estabilidade, como O'Connor a define, "é uma questão de relações equilibradas entre um conjunto de fatores, e depende de um constante reconhecimento das limitações e da localização" (ibid., p.14). Essa noção origina-se no contexto da arquitetura, onde um dos objetivos centrais é combinar elementos heterogêneos, obtendo um equilíbrio entre a imobilidade e a flexibilidade: "O concreto só pode curvar até certo ponto, o aço só pode sustentar um certo peso, o vidro só pode aguentar uma certa pressão" (ibid.); ao combinar esses materiais e suas propriedades, um arquiteto pode criar uma estrutura que se sustenta devido ao equilíbrio e à tensão: "apenas considere a importância do movimento em um edifício alto ou em uma ponte" (ibid.). E, de acordo com O'Connor, algo análogo vale para o caso da normatividade em geral, seja ela ética ou linguística.

apresenta e em sua *competência* geral nesse assunto; no caso da discussão moral o que realmente importa determinar é se podemos *compreender* e *respeitar* (mas não necessariamente *concordar com*) posições ou atitudes assumidas por outrem. Em vez da adequação de uma alegação a certos princípios universais e impessoalmente estabelecidos, o raciocínio moral é construído em termos de responsividade entre agentes, e coloca à prova a qualidade e, no limite, a possibilidade mesma de se criar ou manter uma relação baseada em compromissos e preocupações compartilhados. Essa diferença é apresentada por Cavell nos seguintes termos:

> Questionar uma alegação cognitiva toma a forma das perguntas "Como você sabe?" ou "Por que você acredita nisso?", e poderíamos dizer que avaliar tal alegação é uma questão de avaliar se a sua posição [...] [é] adequada a ela. Questionar uma alegação de correção moral [...] toma a forma das perguntas "Por que você está fazendo isso?", "Como você pode fazer isso?", "O que você *está* fazendo?", "Você pensou realmente no que está dizendo?", "Você sabe o que isso significa?"; e avaliar essa alegação é [...] determinar *qual é* sua posição, e desafiar a posição ela mesma [...]. O propósito da avaliação não é determinar *se* ela é adequada [...] [mas] determinar [...] *qual é a posição pela qual você está assumindo responsabilidade* – e se essa é uma posição que eu posso respeitar. O que está em jogo nessas discussões não é, ou não é exatamente, se você conhece o nosso mundo, mas se, ou até que ponto, podemos viver no mesmo universo moral. O que está em jogo [...] não é a validade da moralidade como um todo, mas a natureza ou a qualidade do nosso relacionamento mútuo (CR 268).

A lição que me interessa enfatizar é que a possibilidade de manter um relacionamento moral depende essencialmente dos compromissos e preocupações que estão em jogo; assim, estaremos dispostos a exigir ou a tolerar mais ou menos de nosso interlocutor conforme o peso dado por cada um de nós a esses fatores. Vejamos como isso funciona analisando um caso concreto de discussão moral, apresentado por Cavell:

> A: Decidi não oferecer o emprego a ele.
> B: Mas ele está contando com isso. Você prometeu isso explicitamente a ele.
> A: Eu sei, mas de repente se tornou muito inconveniente ter ele por perto, e, de todo modo, há alguém realmente melhor qualificado.
> B: Se você fizer isso com ele, nunca mais falo com você.
> A: Não faça tempestade em copo d'água. Vou me assegurar de que ele consiga um emprego, e vou dar-lhe algum dinheiro para dar uma força.
> B: Adeus.

Neste pequeno diálogo B critica a intenção expressa por A de quebrar uma promessa feita a um terceiro. Vale notar que B não está avaliando se a atitude de A é correta ou errada, boa ou má *enquanto tal* (não se trata de determinar, por exemplo, se ela expressa um "princípio universalmente válido"); o que está em questão é se *A* está realmente em posição de assumir essa atitude de maneira responsável, dados seus próprios compromissos anteriores. Em resumo, B está acusando A de ser um hipócrita, e, dado que em sua réplica A não faz mais que confirmar essa acusação, B acaba por concluir que vive em um universo moral distinto, e que talvez tenha se iludido a respeito de A até aqui. É nesse sentido que a discussão moral serve como uma "arena para a revelação mútua dos sujeitos"[13].

2. Perfeccionismo e os limites da moralidade

No caso examinado acima temos apenas um desacordo momentâneo entre sujeitos que poderiam, pelo menos em princípio, encontrar novas bases para levar adiante a discussão (não se deve desconsiderar a importância do tempo e da paciência em assuntos sublunares). Mas há casos em que o conflito poderia colocar a moralidade como um todo em questão, o que, de acordo com Cavell, simplesmente indica que a moralidade deve ser vista como limitada no tocante a suas potencialidades, deixando espaço para ideias como a da "salvação do sujeito [*self*] através do repúdio da moralidade" (CR 269), o que, como veremos, já aponta

[13] MULHALL, 1994, p.41.

para o tema do "perfeccionismo moral", que será explorado de maneira mais sistemática nas obras mais recentes desse autor. Eis a passagem crucial de *The Claim of Reason* para esse assunto:

> A moralidade deve deixar-se aberta ao repúdio; ela fornece *uma* possibilidade de resolução de conflito, uma forma de contê-lo de modo a permitir a continuidade das relações pessoais contra o fato duro e aparentemente inevitável dos mal-entendidos, dos desejos, compromissos, lealdades, interesses e necessidades mutuamente incompatíveis [...]. Outras formas de resolução ou contenção de conflitos são fornecidas pela política, pela religião, pelo amor e pelo perdão, pela rebelião e pelo afastamento. A moralidade é uma forma valiosa porque as demais são muito frequentemente inacessíveis ou brutais; mas ela não é tudo; ela fornece uma porta através da qual uma pessoa, alienada ou em perigo de alienar-se de outra através de sua ação, pode retornar com o oferecimento e a aceitação de explicações, desculpas e justificativas, ou pelo respeito que um ser humano mostrará a um outro que vê e pode aceitar a responsabilidade por uma posição que ele próprio não adotaria. Não temos que concordar uns com os outros a fim de viver no mesmo mundo moral, mas temos de conhecer e respeitar nossas diferenças mútuas (CR 269).

Lembremos que, de acordo com nossa análise anterior, a marca distintiva do raciocínio moral é que ele deve obedecer à exigência da inteligibilidade[14]. Mas, como Cavell indica em um outro contexto, isso ainda deixa muitas questões em aberto, tais como, por exemplo:

> se há limites para a obrigação de sermos inteligíveis, se todos nós não temos direito a uma certa obscuridade ou sensação de confusão, e por vezes alguns mais do que outros. Talvez nem sempre haja algo a ser dito; e existe a questão de saber o que se deve fazer em relação à discordância persistente, o quão longe se deve ir na tentativa de resolvê-la [...] (CW 25).

[14] Um indício da centralidade dessa característica é o fato de que ela perpassa as mais diferentes concepções tradicionais da moralidade – por exemplo, o "cálculo de consequências" dos utilitaristas ou a "interpretação de motivos e princípios" do kantismo (ver CW p. 25).

Neste momento me vejo obrigado a exortar: sejamos honestos – o quão longe estamos dispostos a ir em uma discussão moral marcada por tais "discordâncias persistentes"? Indefinidamente? Não creio. Mas o que isso mostra? Que não somos perfeitos, e, particularmente, que não somos completamente racionais? Mas por que exatamente seria sempre irracional desistir de uma discussão, se estivermos convencidos (racionalmente, vale acrescentar) de que ela não levará (mais) a lugar algum, que a possibilidade de fornecer razões adicionais esgotou-se (pelo menos momentaneamente)? – Voltemos ao caso de *Mr. Deeds*: será que ele teria sido mais sensato caso tentasse *convencer racionalmente* aqueles literatos de que, digamos, eles não deveriam sair por ai debochando das pessoas? Ou isso seria simplesmente patético?[15]

Ao comentar esses assuntos em termos mais gerais, Cavell afirma o seguinte:

> […] o ódio e a raiva não são essencialmente irracionais, mas podem claramente ser requeridos. Viver uma vida moral não deve exigir que nos tornemos Socrateses, budas ou cristos, virtualmente imperturbáveis. Mas somos convidados a tornar até mesmo a raiva justificada e o ódio compreensível, a nos responsabilizarmos pela expressão [desses sentimentos] em nossas vidas e, às vezes, mas não sempre e nem em toda parte, a colocá-los de lado (CW 25-6).

Quero enfatizar três pontos dessa passagem: primeiro, Cavell propõe que é possível distinguir raiva e ódio *justificados* e *injustificados*, e, segundo, ele também afirma que, mesmo no caso em que esses sentimentos (e suas consequências) estiverem justificados, ainda seremos instados a torná-los inteligíveis e (nessa medida) assumir responsabilidade sobre eles. Claramente não estamos pensando nos casos em que nos arrependemos de um acometimento de raiva imediatamente depois do acontecido, quando estamos "de cabeça fria"; o caso interessante é aquele em que, considerados todos os fatores, continuamos convictos, pelo menos

[15] Aqui cabe outra injunção apresentada por Wittgenstein para fins distintos: "não pense, mas olhe!" (op. cit. §66).

parcialmente, de que aquela atitude era adequada à situação (pense em B depois de ter dado adeus a seu amigo A). Digo "parcialmente" por que, na vida real (mas também na boa literatura e nos bons filmes!) as coisas não são tão simples, e nem sempre temos clareza sobre nossas próprias motivações. É por isso – e aqui está o terceiro ponto que quero enfatizar – que Cavell encerra a passagem chamando atenção para o fato de que, pelo menos em algumas ocasiões, a demanda por inteligibilidade nos fará reconsiderar a situação, "deixando de lado" nossa reação inicial.

Pensemos novamente na conclusão abrupta da discussão entre Deeds e os literatos, devido ao sentimento de indignação do primeiro. Nesse caso, a quem exatamente Deeds deveria justificações e explicações? Não me parece plausível pensar que ele as deva aos próprios literatos que ocasionaram nele tal sentimento – pelo menos não imediatamente. Aquela discussão está momentaneamente encerrada, e o melhor que se pode esperar é que, *depois* de ambas as partes reconsiderarem suas atitudes e motivações, tomando como pano de fundo suas preocupações e compromissos, um pedido de desculpas possa ser apresentado, permitindo que retomem o relacionamento[16]. Mas quais são as condições para uma tal reconsideração em "primeira pessoa"? Se o agente estiver confuso, será mesmo possível encontrar inteligibilidade por conta própria, fornecendo justificações e explicações *a si mesmo*? É por causa dessa dificuldade que muitos filósofos têm salientado o papel do amigo na busca por inteligibilidade e educação moral. Mas em nenhuma teoria moral esse papel recebe tanta atenção quanto no perfeccionismo.

Cavell descreve o perfeccionismo como "uma dimensão ou tradição da vida moral que se estende por todo o percurso do pensamento ocidental, e que se ocupa com o que costumava ser chamado o estado de nossa alma, uma dimensão

[16] Registro, para possível tratamento futuro, que a dinâmica das tentativas de reconciliação e sua importância da manutenção de uma comunidade moral são temas caros a Cavell. A noção-chave empregada na análise dessa dinâmica, herdada de J. L. Austin, é a de "elaborativos" – "aquelas desculpas, explicações, justificativas [...] que constituem o grosso da defesa moral [*the bulk of moral defense*]" (CR 296; ver também MWM 26-30 e CR 324-5).

que coloca um peso enorme nas relações pessoais e na possibilidade ou necessidade de transformação, tanto do próprio sujeito quanto de sua sociedade"[17]. Essa dimensão estaria expressa em um conjunto bastante diversificado de textos, incluindo cânones da tradição filosófica – autores como Platão, Aristóteles, Locke, Kant, Mill, Nietzsche, Rawls – da literatura – Shakespeare, Ibsen, George Bernard Shaw – e também "textos de escritores que não são normalmente considerados pensadores morais pelos filósofos profissionais" – particularmente Emerson e Freud (CW ix)[18]. Tais textos são vistos como variações sobre o tema da natureza humana como dividida ou duplicada, pendendo entre a aceitação do estado atual do mundo enquanto palco das nossas atividades e perspectivas, e o desejo por uma reforma ou transfiguração do mesmo (ver CW 1-2). Mas Cavell apressa-se em contrastar o tipo de busca perfeccionista que lhe interessa daquilo que denomina "perfeccionismo religioso", comprometido com a ideia de uma perfeição humana *final*. Em minha experiência ao tentar introduzir a ideia de perfeccionismo tem sido comum perceber uma aversão imediata causada pelo uso do próprio termo, justamente devido a essa conotação de "perfeição última" que ele, de maneira mais ou menos inevitável, carrega. Some-se a isso uma certa leitura superficial de Nietzsche ou mesmo de Emerson como filósofos "elitistas" e "anti-democráticos", e está montado o cenário para que essa posição sequer seja levada a sério por alguns interlocutores. Não é minha pretensão aqui tentar desfazer esses mal-entendidos[19]. O que me interessa é deixar claro que uma leitura alternativa é possível. O próprio Cavell denuncia o

---

[17] CAVELL, S. *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism*. Chicago: Univ. of Chicago Press, 1990, p.2.

[18] Sem falar no "cinema americano da assim chamada Era Dourada do talkie hollywoodiano" (ibid.). Mais sobre a relação entre cinema e perfeccionismo no Epílogo.

[19] James Conant procurou fazer isso com relação a Nietzsche no texto "Nietzsche's Perfectionism: A Reading of Schopenhauer as Educator" (in Schacht, R. (ed.), *Nietzsche's Postmoralism*, CUP, 2000, p.181-257), e Cavell tem procurado fazer o mesmo tanto em relação a Nietzsche quanto a Emerson em várias obras recentes, sendo a principal delas *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism* (Chicago: Univ. of Chicago Press, 1990).

perfeccionismo elitista, preocupado apenas com o "cultivo individual" como "aviltado" [debased] (CW 18). Em seu lugar ele propõe uma busca coletiva, democrática e contínua por aquilo que Emerson descreve como "um eu inatingido mas atingível" ["*an unattained but attainable self*"][20]: "um eu que é sempre e nunca nosso, um passo que nos leva não do mal para o bem, ou do errado para o certo, mas da confusão e da constrição para o auto-conhecimento e a sociabilidade" (CW 13). Ao tratar desses assuntos Cavell propõe um contraste bastante elucidativo com Platão:

> A ideia de Platão de um caminho na direção de um objetivo (aquele procurado pelo sábio) não se ajusta exatamente à ideia de Emerson sobre como viver. Em ambos, a ideia da filosofia como um modo de vida desempenha um papel na avaliação da sua vida agora, mas Emerson está menos interessado em apresentar a vida do sábio como um modelo para a nossa do que em nos lembrar que o poder de questionar nossas vidas, por exemplo, quando julgamos o que chamamos de suas necessidades, direitos e bens, está ao alcance de todos os seres humanos (ou, de todo modo, dos que são livres para falar sobre suas vidas e modificá-las) (CW 13).

O que Emerson propõe, então, no lugar do sábio de Platão? A resposta a essa questão nos devolve ao ponto anteriormente levantado – trata-se justamente da função do amigo, daquele exemplar ou modelo de um "eu inatingido mas atingível". Se estou confuso e tenho dificuldade em tornar minhas próprias ações e atitudes transparentes e inteligíveis, fica difícil tomar o próximo passo na direção daquele ideal, o "*further self*", de modo que necessito de algum tipo de atração externa. O amigo pode servir como tal atração, na medida em que pode nos indicar falhas – fazer com que confrontemos nossas confusões – de um modo que não gerará tanta resistência, dada a sua postura moral específica em relação a nossas "preocupações e compromissos". O amigo não me confronta apresentando razões

[20] A frase encontra-se no ensaio "History" (Essays, First Series [1841]), disponível online aqui (acesso em 10 de maio de 2012):
<http://www.vcu.edu/engweb/transcendentalism/authors/emerson/essays/history.html>.

impessoais, mas sim a partir de uma posição que *ele* ocupa em relação àqueles compromissos. Ora, se alguma disposição para compreender e ser compreendido é necessária para a argumentação moral, então é fácil perceber como um contexto de amizade e respeito mútuo torna-se particularmente adequado para esse fim. É sobretudo nesse tipo de contexto que poderemos avançar em nossa educação moral – uma educação que, de acordo com Cavell, não visa primeiramente "proporcionar um aumento de aprendizagem, mas uma transformação da existência" (CF 325)[21].

Mas dizer que um contexto de amizade é *particularmente adequado* para o avanço na educação moral não é o mesmo que dizer que *só é possível* avançar em tais contextos. Como um amigo alertou, "inimigos e estranhos também podem nos ensinar algo sobre a moralidade de nossa conduta. Aprendemos lições definitivas ouvindo o que não queremos ouvir, daqueles que mal conhecemos"[22]. Certamente; entretanto, parece-me que isso só será possível naqueles casos em que mesmo nossos inimigos ou estranhos compartilharem de *pelo menos alguns* dos nossos próprios compromissos e preocupações[23]. O único cenário que está sendo excluído como propício à discussão moral é o de sujeitos que, nos termos de Cavell, vivem em mundos morais completamente distintos. Pensemos uma última vez no que aconteceu entre Deeds e os literatos: estes não demonstraram *nenhum* interesse genuíno nas preocupações e compromissos de Deeds, tratando-os desde o início como sem sentido ou risíveis (lembremos do quanto riem do fato de ele ser um poeta de cartões[24]). Já Deeds partiu de uma perspectiva oposta, de admiração, e por isso mesmo sua frustração foi tão grande. Ele foi humilhado e

[21] A passagem continua tratando da importância do "casamento": "aqueles que não podem se inspirar mutuamente a uma educação desse tipo não estão casados; eles não têm o tipo certo de interesse um pelo outro". Infelizmente esse é um tema que não pude tratar no espaço da presente comunicação.

[22] Devo essa consideração a Eduardo Vicentini de Medeiros, em uma troca de e-mails.

[23] Lembremos do grande (embora limitado) respeito mútuo que arqui-inimigos invariavelmente demonstram na ficção.

[24] Agradeço a Nykolas Friedrich Correia Motta por chamar atenção a esse ponto.

descobriu-se isolado, e por isso decidiu fazer o mesmo com os literatos, restabelecendo os termos de seu relacionamento. Ao fazer isso, creio que propiciou uma ótima oportunidade para os literatos repensarem suas atitudes e aprenderem com seus erros. Ele próprio aprendeu uma lição muito importante: "nem todas as pessoas famosas são grandes pessoas". Momentos de confronto como esse são de fato cruciais em nosso aprendizado moral. Mas, como espero ter indicado, essa constatação não me parece contradizer a visão cavelliana sobre a importância da conversa para a revelação mútua dos agentes – pelo contrário, creio que ela sai reforçada.

Epílogo: cinema e perfeccionismo

Ao longo deste texto retornei continuamente a uma única cena de uma comédia hollywoodiana dos anos 1930. Esse procedimento é familiar aos leitores de Cavell, sobretudo daquelas obras explicitamente voltadas ao tema do perfeccionismo. A convicção por trás desse procedimento é que os temas, motivos e preocupações expressos em alguns filmes – particularmente aqueles que Cavell agrupa sob os gêneros da "comédia do recasamento" e do "melodrama da mulher desconhecida" – justificam sua inserção naquela mesma tradição do pensamento ocidental que "preocupa-se com o que costumava ser chamado de estado das nossas almas"[25]. Por serem, eles mesmos, "estudos perfeccionistas" em que os protagonistas se envolvem em uma jornada na direção de um "eu inatingido" [*further self*], marcada por incontáveis conversas entre amigos[26] – figuras que "podem ocorrer como o objetivo da jornada, mas também como sua instigação e acompanhamento"[27] – esses filmes servem como:

> um pequeno laboratório para estudar a conversação moral, não como a

[25] Passagem citada integralmente acima.

[26] Cavell os descreve como "filmes cujas conversas estão entre as glórias do cinema mundial" (Cavell, S. "The Good of Film". In: Rothman, W. (ed.), *Cavel on Film*. Albany: SUNY Press, 2005, p.338).

[27] CW 27. Ainda sobre isso: "A presença da amizade nos filmes que vamos considerar (incluindo a falta por vezes drástica dessa relação no melodramas) é da maior importância específica para estabelecê-los como narrativas perfeccionistas." (id. ibid.)

> tentativa de convencer alguém a uma linha de ação, nem como a avaliação de uma instituição social, mas sim como algo que às vezes penso que é prévio e reparatório para esses objetivos familiares do raciocínio moral, outras vezes que é subsequente e complementar, a saber, a *capacidade de uma alma de responder a e de ser examinada por outra*. Isso é prévio, pois fornece-nos os estudos da postura que um agente moral reivindica ao confrontar outro com seu julgamento; isso é subsequente, pois fornece o espaço para avaliar o enquadramento moral dentro do qual você está raciocinando. […] Pode-se dizer que o perfeccionismo concentra-se na demanda para tornarmo-nos inteligíveis a nós mesmos e mutuamente. E suponho que nenhuma perspectiva seria considerada moral a menos que desse lugar a tal demanda[28].

Os filmes que interessam a Cavell retratam o esforço (mas não necessariamente o sucesso) dos protagonistas para tornarem-se pessoas melhores, escolhendo um modo de vida melhor[29]. Ao insistir na pertinência de se incluir esses filmes no conjunto de textos que exploram temas perfeccionistas, Cavell não quer dar a impressão de que "a filosofia, restrita a seus próprios recursos, exige uma compensação por meio de revelações próprias do meio do cinema" (CW 5), mas sim, pelo contrário, indicar que esses filmes podem ser pensados como

> caminhos intelectuais e emocionais diferentemente configurados que a filosofia já está explorando, mas que às vezes ela talvez tenha motivos para abandonar prematuramente, particularmente nas formas que [a filosofia] assumiu depois de sua profissionalização, ou academização [...] A alegação implícita é que o cinema, a última das grandes artes, mostra que filosofia é

[28]"The Good of Film" (op. cit.) p.339. Itálicos adicionados.

[29] No caso específico "comédias do recasamento" essa busca é apresentada com alternativa à ameaça do "cinismo moral": "a tentação de desistir de uma vida mais coerente e admirável do que aquela que nos parece acessível depois que os compromissos da vida adulta passam a obscurecer a promessa e os sonhos da juventude. O fato de que o par principal dessas comédias é um pouco mais velho do que os jovens pares da comédia clássica fornece um contexto em que determinadas formas de realizar antigos sonhos entraram em colapso, e um novo regime deve ser instituído tal que nele o consentimento possa, sob reflexão, ser obtido, ou tal que nele possamos apostar" (CW 23-24).

> o acompanhamento muitas vezes invisível das vidas comuns, as quais o cinema é tão competente em captar (até mesmo, e talvez sobretudo, quando as vidas retratadas são históricas, elevadas ou cômicas, perseguidas ou assombradas) (CW 5-6).

Com estas duas ideias – a saber, que deveríamos lutar contra a tentação de abandonar prematuramente as complexidades de nossas vidas comuns, e que os filmes são particularmente aptos a capturar essas complexidades – creio ter alcançado um ponto adequado para concluir esta exposição, dando continuidade à nossa conversa sobre a natureza da moralidade.

7

CETICISMO DARWINIANO SOBRE O REALISMO MORAL*

David Copp

Introdução

Parece haver explicações evolutivas darwinianas plausíveis para uma gama de fenômenos psicológicos com importância moral, fenômenos que são tanto de natureza conativa como cognitiva. Chamemo-los de *fenômenos da psicologia moral*. Esses fenômenos incluem o fato de que os humanos têm uma tendência a desenvolver e a tornarem-se dispostos a expressar uma variedade de estados conativos, incluindo desejos, sentimentos e emoções moralmente significativos. Há, por exemplo, uma tendência a desenvolver atitudes morais que favoreçam a cooperação, uma tendência dos pais a desejar cuidar de sua prole e uma tendência daqueles que receberam bens de outrem a desejar retribuir. Os fenômenos que tenho em mente incluem capacidades cognitivas variadas, incluindo aquelas que são necessárias para se formar crenças morais, e uma tendência a formar essas crenças. E incluem também a capacidade para o que Allan Gibbard chama de "governo normativo"[1], e que possivelmente depende de uma capacidade para sentir vergonha e culpa, entender normas sociais e raciocínio prático complexo. Direi mais sobre esses fenômenos no que segue. Para simplificar minha tarefa, suporei que a natureza desses fenômenos não está em questão.

Com base no trabalho de biólogos evolutivos, vários filósofos defenderam

---

* Traduzido por Marcelo Fischborn, com autorização, a partir do original "Darwinian skepticism about moral realism", publicado em *Philosophical Issues* 18, p.186-206, 2008.

[1] 1990, p.61-80.

que as forças darwinianas influenciaram fortemente o fenômeno da psicologia moral, tendo isso como resultado que a natureza do fenômeno, incluindo o conteúdo de nossas crenças morais, pode ser explicado ao menos em parte a partir da teoria evolutiva[2]. Chamemos isso de *a hipótese darwiniana*. Suponho, a título de argumento, que ela está correta. A hipótese darwiniana é compatível com propostas específicas variadas sobre a maneira exata pela qual as pressões evolutivas influenciaram o fenômeno. E é compatível com concepções diferentes sobre quão fortemente as forças darwinianas influenciaram o fenômeno.

A questão na qual estou interessado é se há um argumento forte, mesmo que não seja um argumento definitivo, contra o realismo moral que possa ser formulado com base na *suposição* de que a hipótese darwiniana esteja correta. Concentro-me em um argumento que creio ser especialmente útil para evidenciar a natureza do desafio que o darwinismo põe ao realismo moral. O argumento é de Sharon Street (2006). Elementos dele aparecem no trabalho de Richard Royce (2006), Philip Kitcher (2006) e outros.

Street defende que a verdade da hipótese darwiniana acarretaria um dilema para o realismo moral e também para o realismo sobre o valor em geral[3]. Pois, se as forças evolutivas "desempenharam um papel forte na formação do conteúdo das atitudes [morais] humanas", os realistas têm de explicar a relação entre essas forças e os fatos morais. A questão crucial é se as forças evolutivas fizeram ou não com que nossa psicologia moral seja tal que nossas crenças morais tendem a acompanhar os fatos. Se os realistas respondem a essa questão negativamente, defende ela, defrontam-se com o resultado cético de que é improvável que nossas crenças morais sejam verdadeiras. Mas, se respondem afirmativamente, estão comprometidos com a hipótese darwiniana específica de que "a seleção natural favoreceu os ancestrais que eram capazes de apreender essas verdades". Essa hipótese, defende Street, é implausível. Uma abordagem mais plausível da

---

[2] Ver, por exemplo, Blackburn (2000), Gibbard (1982 e 1990), Joyce (2006), Kitcher (1993 e 2006), Ruse (1986), Sober e Wilson (1998), Street (2006).

[3] O alvo de Street é o realismo valorativo, mas foco no realismo moral. Essa mudança não afeta a força de seu argumento.

evolução do fenômeno da psicologia moral, *a abordagem da ligação adaptativa*, não postula a existência de verdades morais. Por isso, defende Street, a hipótese darwiniana força os realistas a escolherem entre um resultado cético implausível e uma explicação cientificamente implausível do fenômeno da psicologia moral. Para evitar ambas as alternativas do dilema, tem-se de negar o realismo moral[4].

Inicio explicando a ideia básica por trás da hipótese darwiniana. Então, explico o dilema darwiniano em maiores detalhes e defendo que os realistas podem escapar-lhe. Os realistas não são obrigados a escolher entre o resultado cético e a hipótese implausível. O desafio remanescente para os realistas que aceitam a hipótese darwiniana é mostrar a plausibilidade de uma concepção segundo a qual os seres humanos evoluíram e desenvolveram-se de uma maneira tal que suas crenças morais sejam reações a fatos que *são* fatos morais. Ao final do artigo, ilustro como essa concepção pode ser defendida. A conclusão é que um argumento bem sucedido contra o realismo moral não pode basear-se no dilema darwiniano, mas tem de, em vez disso, valer-se de objeções filosóficas de um tipo mais familiar.

1. Evolução e crenças morais

Street, Joyce e Kitcher oferecem elaborações similares da hipótese darwiniana. A ideia básica é que tendências para certos tipos de respostas valorativas, ao invés de outras, promoveram o sucesso reprodutivo entre nossos ancestrais remotos, porque "forjaram ligações adaptativas entre as circunstâncias de nossos ancestrais e suas respostas a essas circunstâncias, levando-os a agir, sentir e crer de maneiras que se revelaram evolutivamente vantajosas"[5]. Essa é a abordagem da ligação adaptativa. Parece claro, por exemplo, que uma tendência a cooperar com aqueles que têm uma tendência a retribuir tenderia a promover o sucesso reprodutivo, e também uma tendência a buscar o bem dos próprios

[4] STREET, 2006, p. 109, e em todo o artigo. Para as observações citadas, ver p.109.
[5] STREET, 2006, p.127; JOYCE 2006, p.117; ver KITCHER 2006, p.175-181.

familiares e a evitar aqueles que nos prejudicariam[6]. De acordo com a abordagem da ligação adaptativa, uma tendência para certos tipos de respostas valorativas, em vez de outras, contribuiu para o sucesso reprodutivo no ambiente ancestral e resultou no fato de que os seres humanos tendem a manifestar certas emoções e outras disposições conativas moralmente significativas[7].

A hipótese darwiniana também afirma, é importante dizer, que as forças darwinianas afetaram o conteúdo de nossas crenças morais. Isso pode parecer enigmático. Pois, primeiro, como Street nota, nossos ancestrais remotos provavelmente não tinham a capacidade cognitiva de formar crenças morais[8]. E, além disso, é difícil ver como a atuação da seleção natural sobre as características desses ancestrais distantes poderia ter afetado o conteúdo das crenças morais que temos atualmente. A ideia de Street, no entanto, é que o conteúdo de nossas crenças morais poderia ter sido influenciado indiretamente pelas forças darwinianas a partir de sua influência mais direta sobre os estados mentais pré-conceituais primitivos de nossos ancestrais remotos[9].

Street parece ter em mente um processo que podemos decompor em aproximadamente três estágios. Primeiro, a seleção natural levou ao desenvolvimento de algumas capacidades e disposições psicológicas básicas em nossos ancestrais remotos que os levaram a desenvolver uma tendência a ter disposições para certas atitudes valorativas. Temos razões para acreditar que alguns tipos de comportamento teriam sido adaptativos para nossos ancestrais. Neste primeiro estágio, os humanos evoluíram de modo a estarem dispostos a formar atitudes valorativas que tendessem a conduzi-los a esses tipos de comportamento[10]. Em segundo lugar, as capacidades e disposições que se desenvolveram no primeiro estágio afetaram em alguma medida o desenvolvimento do repertório conceitual dos ancestrais menos remotos, de modo

---

[6] STREET, 2006, p.129; KITCHER 2006, p.166-174

[7] JOYCE, 2006, p.13-142.

[8] 2006, p.118-119.

[9] 2006, p.118-120.

[10] STREET 2006, p.129; JOYCE 2006, p.13-50; KITCHER 2006, p.166-174.

que afetaram o conteúdo dos juízos valorativos que passaram a estar dispostos a fazer. Street afirma que a abordagem da ligação adaptativa pode explicar dessa maneira por que os seres humanos evoluíram de modo a estarem dispostos a fazer juízos favorecendo aqueles tipos de comportamento. Por exemplo, ela afirma que pode explicar por que os seres humanos evoluíram de modo a tenderem a julgar que o fato de uma pessoa ter ajudado alguém é uma razão para lhe retribuir a ajuda[11]. Em terceiro lugar, como resultado de tudo isso, temos uma tendência não-reflexiva a formar certas crenças morais básicas.

Kitcher acrescenta a essa imagem uma abordagem sobre a função evolutiva de uma capacidade para governo normativo juntamente com uma abordagem sobre o impacto da "evolução cultural". Ele sugere que a função evolutiva de uma capacidade para governo normativo foi "promover a coesão social" por meio da oposição às pressões dos desejos egoístas e antissociais com "pressões que reforçariam o tecido social"[12]. Ele defende que "os hominídios com a tendência a agir com base em disposições altruístas [e que induzem a cooperação] teriam se saído melhor que aqueles sem <essa tendência>", uma vez que isso lhes teria dado um acesso mais fácil a "alianças vantajosas"[13]. Ele sugere, além do mais, que uma capacidade para governo normativo resultante da evolução teria permitido aos nossos ancestrais desenvolver uma "proto-moralidade", a qual reforçaria essas disposições e seria transmitida ao longo das gerações. Um processo de evolução cultural poderia então desenvolver-se, no qual certos sistemas normativos seriam mais bem-sucedidos que outros para sobreviver nas gerações seguintes. Esperar-se-ia que certos tipos de normas sobrevivessem e tornassem-se mais amplamente aceitas. Estas incluiriam, por exemplo, normas projetadas para promover a estabilidade social, a paz e a cooperação[14]. A abordagem da ligação adaptativa, portanto, parece capaz de explicar por que os seres humanos teriam desenvolvido uma capacidade para governo normativo,

[11] STREET 2006, p.127; JOYCE 2006, p.50-57.
[12] 2006, p.172.
[13] 2006, p.172; ver também Gibbard 1990.
[14] 2006, p.173.

envolvendo uma tendência a aceitar normas em vigência na cultura local e a governar seu comportamento de acordo com essas normas. Ela também parece capaz de explicar por que haveria uma tendência robusta para as normas em vigência em uma cultura favorecerem comportamentos 'pró-sociais', incluindo o comportamento que favorece a própria prole e parentes, que retribui benefícios recebidos e favorece a cooperação.

Na proposta de Kitcher, então, houve um quarto estágio no processo. Nesse estágio, devido a um processo de evolução cultural, normas que promovem a estabilidade social, a paz e a cooperação tenderam a tornar-se mais amplamente aceitas, tendendo, assim, a levar os humanos com capacidade para governo normativo a formar crenças morais que favoreçam essas formas de comportamento. Por conveniência, entenderei a abordagem do processo de quatro estágios acima como sendo parte da abordagem da ligação adaptativa.

A observação crucial é que crenças morais, juntamente com outros fenômenos da psicologia moral, tais como nossas emoções e desejos, tendem a motivar o comportamento. É por isso que o fenômeno da psicologia moral poderia ter evoluído de modo a fornecer motivação para o comportamento adaptativo. Nada aqui exige ou implica que a capacidade de formar crenças morais evoluiu para permitir-nos detectar os fatos morais[15].

Dado isso, e dada uma abordagem funcionalista da distinção entre crenças e desejos, pode parecer que a abordagem da ligação adaptativa ameaçaria o realismo moral mediante um apoio ao não-cognitivismo, a concepção segundo a qual o estado mental de uma pessoa que aceita uma alegação moral (pura e básica) não é uma crença, mas, em vez disso, um tipo de estado conativo. Kitcher defende isso[16]. Sua ideia parece ser que, na abordagem da ligação adaptativa, a função evolutiva do fenômeno da psicologia moral é motivar o comportamento adaptativo. Por isso, mesmo que seja plausível que evoluímos de modo a tender a ter um repertório adaptativo de disposições e atitudes emocionais e

[15] JOYCE 2006, p.131; KETCHER 2006, p.176.
[16] 2006, p.175-176.

comportamentais, há pouca plausibilidade na ideia de que teríamos evoluído de modo a formar crenças valorativas propriamente ditas, em acréscimo a esse repertório de estados não-cognitivos. Não posso analisar esse argumento aqui. Parece-me que, mesmo se a função evolutiva do fenômeno da psicologia moral for motivar o comportamento adaptativo, é defensável que essa motivação seja produzida de uma maneira usual através da combinação de crença e conação. A abordagem da ligação adaptativa não nos força ao não-cognitivismo[17].

A abordagem da ligação adaptativa é obviamente especulativa. Também é especulativo afirmar, como Street faz, que as forças darwinianas "afetaram fortemente" a nossa psicologia moral, incluindo o conteúdo de nossas crenças morais[18]. Poderíamos concordar que essas forças tiveram *algum* efeito, sem concordar que o efeito foi especialmente forte. Não obstante, a título de argumento, seguirei Street na suposição de que a abordagem da ligação adaptativa é ao menos aproximadamente verdadeira e que as forças darwinianas tiveram um impacto significativo na psicologia moral humana. Suporei que algo como o processo de quatro estágios ocorreu e que explica em alguma medida significativa o conteúdo de nossas crenças morais. Isso certamente não significa negar que nossas crenças morais foram afetadas por uma variedade de fatores, além das pressões darwinianas, incluindo a cultura bem como a deliberação e a reflexão[19].

2. O dilema darwiniano

Pode-se salientar as peculiaridades do argumento de Street contrastando-o com um argumento darwiniano mais direto contra o realismo moral. Este argumento baseia-se na premissa de que a explicação darwiniana mais plausível do fenômeno da psicologia moral não exige a suposição de que quaisquer crenças

---

[17] Gibbard também pensa que as considerações evolutivas apoiam o não-cognitivismo (1982, p.43). Para uma discussão, ver STURGEON, 1992.

[18] 2006, p.109; ver também p.113, p.121.

[19] Para uma discussão útil, ver SOBER, 1994, p.95-99. Ver também JOYCE, 2006, p.33-40, p.140, p.180-181.

morais sejam verdadeiras. O argumento alega, com base nisso, que considerações sobre parcimônia ontológica nos dão razões para supor que não há verdades morais[20].

Esse argumento, no entanto, não é bem-sucedido em sua ameaça ao realismo moral por pelo menos duas razões. Primeiro, as explicações darwinianas mais plausíveis de nossas capacidades de ter crenças sobre o "big bang" e sobre a teoria das cordas também não postulam a existência de verdades nessas áreas, mas isso não nos dá razões para concluir que não há tais verdades. De fato, temos razões independentes para supor que há tais verdades. Por tudo o que a história darwiniana mostra, também poderíamos ter razões independentes para supor que há fatos morais[21]. Em segundo lugar, como Joyce aponta[22], se os fatos morais são "redutíveis" a fatos não-morais do tipo que é mencionado na hipótese darwiniana, não há perda de parcimônia em supor-se que há fatos morais. Nesse caso, a Navalha de Occam não teria nada a fazer[23].

A própria Street aceita uma versão do reducionismo naturalista, uma versão que descreve como "construtivista" e "antirrealista", para contrastá-la com as posições "realistas" que seu argumento pretende atacar[24]. Poderíamos, em vez disso, categorizar seu construtivismo como um tipo de realismo, mas nada depende de nossa escolha terminológica. O fato importante é que Street não pretende argumentar contra as teorias construtivistas, de acordo com as quais os fatos morais são redutíveis a, ou são "uma função de", fatos sobre nossas "atitudes valorativas" – sendo que nossas "atitudes valorativas" incluem todos os fenômenos da psicologia moral, tanto conativos quanto cognitivos, incluindo "juízos valorativos

[20] Ver RUSE, 1986, p.253-254. Para críticas, ver CAMPBELL (1996).
[21] Ver JOYCE 2006, p.182-184; KITCHER 2006, p.176.
[22] 2006, p.188-189.
[23] Joyce defende que nenhuma redução de fatos morais a fatos naturais pode dar certo, porque a moralidade tem uma "autoridade prática inescapável" que nenhuma versão do naturalismo moral pode explicar (2006, p.199-209). Respondi a esse argumento em outro lugar (Copp 2007b).
[24] Ver STREET, 2006, n.37; também STREET, 2008.

sustentados consciente e inconscientemente"[25].

A afirmação que define o realismo valorativo, diz ela, "é que há pelo menos alguns fatos ou verdades valorativas que são independentes de todas as nossas atitudes valorativas", incluindo atitudes valorativas que *poderíamos* ter, "mediante reflexão", bem como atitudes que de fato temos[26]. Um realista *moral* defende, então, que há algumas verdades *morais* que são independentes de nossas atitudes valorativas. O tipo relevante de 'independência', explica Street, é o que Russ Shafer-Landau chamou de "independência de perspectiva"[27]. Como Shafer-Landau explica, os fatos ou verdades morais são "independentes de perspectiva" se, e somente se, não é o caso que são "constituídos" por fatos sobre as "perspectivas" das pessoas. Para Street, esses fatos dizem respeito aos fenômenos da psicologia moral.

Focarei agora no argumento do dilema darwiniano de Street. A estrutura básica do argumento é apreendida por Joyce, que sugere que, dada a hipótese darwiniana, a menos que "seja provável que a seleção natural tenha produzido crenças [morais] verdadeiras", há razão para pensar que nossas crenças morais positivas são injustificadas[28]. A alegação central de Street é, então, que o realista moral, diferentemente do construtivista, não pode aceitar a proposição de que é provável que a seleção natural tenha produzido crenças morais verdadeiras – exceto adotando uma abordagem implausível da maneira pela qual a seleção natural afetou nossa psicologia moral, uma abordagem que conflitua com a abordagem da ligação adaptativa.

Esse argumento é muito mais complexo que o argumento direto a partir da parcimônia ontológica com que começamos. É útil formulá-lo como apresentando um dilema. Os realistas têm de ou afirmar ou negar uma tese que chamarei de *tese do acompanhamento*, a qual diz que a seleção natural afetou nossa psicologia

[25] 2006, p.111.
[26] 2006, p.110-111.
[27] 2006, p.111; SHAFER-LANDAU 2003, p.14.
[28] 2006, p.179-182.

de modo tal que nossas crenças morais tendem a acompanhar os fatos morais[29]. Em uma das alternativas do dilema, a "alternativa que afirma o acompanhamento", os realistas afirmam a tese do acompanhamento. Mas, contesta Street, a tese do acompanhamento é insustentável, porque afirmá-la compromete os realistas com negar a abordagem da ligação adaptativa. Na outra alternativa do dilema, a "alternativa que nega o acompanhamento", os realistas negam a tese do acompanhamento. Mas, contesta Street, dada a hipótese darwiniana, se os realistas negam a tese do acompanhamento, estão comprometidos com um resultado cético implausível. Se Street está correta, os realistas têm dificuldades quer afirmem, quer rejeitem a tese do acompanhamento.

É interessante considerar por que o construtivismo poderia ser imune ao argumento. O construtivismo sustenta que os fatos morais são redutíveis ao fenômeno da psicologia moral, o qual, de acordo com a hipótese darwiniana, pode ser explicado, em parte, como resultado das influências darwinianas. Precisamos perguntar se o construtivismo pode lidar bem com a proposição de que é provável que as forças darwinianas produziram crenças morais verdadeiras e se pode fazê-lo sem adotar uma versão da hipótese darwiniana que conflitue com a abordagem da ligação adaptativa. A resposta parece ser que pode. Os fatos morais são fatos de um tipo que tornaria nossas crenças morais verdadeiras. De acordo com o construtivismo, esses fatos são constituídos pelo fenômeno da psicologia moral.

---

[29] Street diz que os realistas têm de escolher entre supor que não há *nenhuma relação* entre os fatos valorativos e a seleção natural e supor que *há* tal relação (2006, p.125). Isso não é o que ela deveria dizer. Ela defende que se, no primeiro lado do dilema, não há relação entre os fatos valorativos e a seleção natural, então "a seleção natural tem de ser vista como uma influência puramente distorcedora" sobre nossas crenças valorativas (2006, p.121). Mas, se a seleção natural teve uma influência distorcedora sobre nossas crenças, ela mantém com os fatos morais a relação de *fazer nossas crenças valorativas representá-los mal*. Por isso, se a seleção natural não está em *nenhuma* relação com os fatos valorativos, então ela *não* distorce nossas crenças morais. Felizmente esse problema é superficial. A questão central é se a seleção natural esteve em uma "relação de acompanhamento" epistemologicamente significativa com os fatos morais (2006, p.125). Eu reformulo o dilema, portanto, como oferecendo aos realistas uma escolha entre afirmar e negar a tese do acompanhamento.

Portanto, de acordo com o construtivismo, as forças darwinianas que influenciaram o fenômeno da psicologia moral também influenciaram, por essa mesma razão, quais são os fatos morais e se os fatos morais tornam ou não nossas crenças morais verdadeiras. Dado isso, parece que teria de haver uma maneira de formular uma teoria construtivista em que os fatos morais que postula, dada a natureza do fenômeno da psicologia moral previsto pela ligação adaptativa, são tais que tornam verdadeiras as crenças morais que a abordagem da ligação adaptativa prevê que tenderemos a ter[30]. Deixo de lado a questão de quão plausível seria uma teoria dessas.

Se esse argumento sobre o construtivismo estiver correto, então uma teoria construtivista formulada adequadamente pode combinar sua abordagem das condições de verdade das crenças morais com a abordagem da ligação adaptativa para explicar por que a tese do acompanhamento é verdadeira. São os realistas morais que, defende Street, não podem usar a abordagem da ligação adaptativa para explicar por que a tese do acompanhamento se sustenta. Defenderei, no entanto, que isso é um erro. Dada uma abordagem adequada das condições de verdade das crenças morais, uma teoria realista pode aceitar a abordagem da ligação adaptativa e usá-la para apoiar a tese do acompanhamento.

3. A alternativa do dilema que nega o acompanhamento

Na alternativa do dilema que nega o acompanhamento, os realistas negam que a seleção natural tenha afetado nossa psicologia moral de modo a fazer com que nossas crenças morais tendam a acompanhar os fatos morais. Street contesta que, nessa posição, "as forças da seleção natural têm de ser vistas como uma influência puramente distorcedora sobre nossos juízos valorativos, tendo impelido-nos em direções valorativas que não têm absolutamente nada a ver com a verdade valorativa"[31]. Nós "evoluímos para" ter crenças morais com qualquer conteúdo que tendesse a "promover o sucesso reprodutivo". Fomos "guiados pelo tipo errado de

[30] Street oferece um argumento diferente (2006, p.153-154).
[31] 2006, p.121.

influência". Mas, então, salvo "um acaso da sorte", é muito provável "que muitas ou a maioria dos nossos juízos valorativos estejam fora do trilho". Por isso, um realista que aceite a hipótese darwiniana e que adote a alternativa do dilema que nega o acompanhamento compromete-se com "o resultado cético duvidoso" de que seria "puramente uma questão de chance" se qualquer de nossas crenças morais fosse verdadeira[32]. Para evitar esse resultado cético, defende Street, os realistas que aceitam a hipótese darwiniana têm de aceitar também a tese do acompanhamento[33].

Joyce oferece um argumento similar, mas chega a uma conclusão diferente. Street pensa que os realistas que aceitam a hipótese darwiniana, mas negam a tese do acompanhamento, estão comprometidos com a posição de que é improvável que nossas crenças morais sejam verdadeiras. Joyce defende que esses realistas estão comprometidos com a posição de que nossas crenças morais positivas podem ser injustificadas. Ele aponta que, se sabemos que nossas crenças sobre um dado tópico foram influenciadas por fatores que não tendam a acompanhar a verdade, então é um dever racional que sejamos menos confiantes a respeito dessas crenças. Por exemplo, se sei que minhas crenças sobre as qualificações profissionais de alguém foram influenciadas por minha raiva dessa pessoa, então é um dever racional que eu seja menos confiante a esse respeito do que seria de outro modo. Quanto mais fortes penso que essas influências foram, menos confiante devo ser[34]. Por isso, os realistas que aceitam que as forças darwinianas tiveram um efeito significativo sobre o conteúdo de nossas crenças morais, e que negam a tese do acompanhamento, devem ter um nível de confiança significativamente reduzido em suas crenças morais. Em sua posição, há razões para pensar que nossas crenças morais positivas são injustificadas.

Alguém poderia objetar que a reflexão racional pode corrigir a influência potencialmente distorcedora da seleção natural sobre nossas crenças morais.

---

[32] 2006, p.121-122.
[33] 2006, p.134-135.
[34] 2006, p.179-182; ver STREET 2006, p.155; SOBER 1993, p.107.

Street responde que, a menos que as forças evolutivas que afetaram o conteúdo de nossas crenças morais acompanhassem de algum modo a verdade, as concepções morais com que a reflexão começaria provavelmente apontariam na direção errada. Por isso, se a reflexão racional sobre questões morais consiste em buscar um "equilíbrio reflexivo" para nossas concepções morais, é improvável que a reflexão racional aproxime-nos da verdade[35]. Muitos realistas objetariam que a reflexão racional pode conduzir a uma mudança significativa em nossas concepções morais, e levar-nos para além da mera busca de um equilíbrio reflexivo para concepções previamente dadas. Não obstante, seria difícil negar que os resultados da reflexão racional seriam fortemente influenciados por nossos pontos de partida, bem como por nossas capacidades intelectuais e conceituais evoluídas. Se rejeitamos a tese do acompanhamento, não é claro que razão temos para pensar que o exercício dessas capacidades na reflexão sobre nossas crenças morais iniciais permitiria aproximarmo-nos da verdade.

Poder-se-ia responder que, mesmo que as forças darwinianas tenham influenciado o conteúdo de nossas crenças morais, houve várias *outras* influências sobre seu conteúdo, incluindo influências culturais, e essas outras influências poderiam ter tendido a impelir nossas crenças morais na direção da verdade[36]. Por exemplo, mesmo que a seleção natural tivesse feito com que os homens tendessem a pensar irrefletidamente que não há nada errado com o estupro, a cultura poderia ter levado muitos deles a pensar diferente. A empatia pela vítima também poderia tê-los levado a pensar diferente. O problema com essa resposta é que, na abordagem da ligação adaptativa, a natureza de nossas emoções morais também foi afetada pela seleção natural, e nossas culturas também foram afetadas pelas forças darwinianas. Assim, se rejeitamos a tese do acompanhamento, não é claro que razão temos para pensar que nossas emoções ou nossas culturas tenderão a impelir nossas crenças morais na direção da verdade.

A objeção mais importante ao argumento de Street é que os realistas que

---

[35] 2006, p.124.

[36] Ver SINNOTT-ARMSTRONG 2006, p.40-45.

aceitem a hipótese darwiniana não precisam aceitar a tese do acompanhamento para evitar comprometer-se com o resultado cético de Street ou Joyce. Os realistas não precisam aceitar que as forças darwinianas tenderam a fazer com que nossas crenças morais acompanhassem a verdade. Seria suficiente a seguinte relação complexa existisse: o efeito indireto das forças darwinianas sobre o conteúdo de nossas crenças morais foi tal que, dado o efeito cumulativo de todas as influências sobre seus conteúdos, nossas crenças morais tendem a acompanhar *suficientemente bem* a verdade moral, de tal modo que a reflexão racional pode em princípio corrigir suficientemente qualquer influência distorcedora e enfraquecer a preocupação cética. Isto é, seria suficiente se as forças darwinianas tivessem influenciado o conteúdo de nossas crenças morais de um modo tal que, dadas todas as influências sobre o seu conteúdo, elas acompanham a verdade em um 'grau epistemicamente suficiente'[37]. Por brevidade, falarei em "acompanhamento parcial" quando tiver em mente o acompanhamento em um grau epistemicamente suficiente.

A tese do companhamento parcial diz, basicamente, que as forças darwinianas afetaram nossa psicologia de tal modo que nossas crenças morais tendem a *acompanhar parcialmente* os fatos morais. Defendi que os realistas que aceitam a hipótese darwiniana não precisam aceitar a tese do acompanhamento para evitar o resultado cético. Deve-se notar, no entanto, que esses realistas têm de aceitar a tese do acompanhamento parcial para evitar um resultado cético. Podemos, portanto, reformular o dilema de modo que ofereça aos realistas uma escolha entre aceitar e negar a tese do acompanhamento parcial. Estou disposto a conceder que os realistas que aceitem a hipótese darwiniana têm de aceitar a tese do acompanhamento parcial.

4. A alternativa do dilema que afirma o acompanhamento

Na alternativa do dilema que afirma o acompanhamento, como originalmente formulei o argumento, os realistas aceitam a tese do

[37] Para uma observação similar, ver BOYD, 1988, p.190, p.208-209.

acompanhamento. Street defende que os realistas que aceitam a tese do acompanhamento têm de explicar por que essa tese é verdadeira. Para explicar isso, defende ela, os realistas têm de aceitar *a abordagem do acompanhamento*, que é a hipótese de que nossas crenças morais tendem a acompanhar a verdade porque a capacidade para detectar verdades morais promoveu o sucesso reprodutivo entre nossos ancestrais[38]. "Não há outro caminho", diz ela, para que o realismo evite a alternativa do dilema que nega o acompanhamento[39]. Ela contesta, no entanto, que a abordagem do acompanhamento é "inaceitável", porque não é compatível com a abordagem empiricamente mais plausível da ligação adaptativa[40].

Há um problema aqui. A *tese* do acompanhamento diz, basicamente, que as forças darwinianas fizeram com que nossas crenças morais acompanhassem os fatos morais. A *abordagem* do acompanhamento é a hipótese de que, de fato, as forças darwinianas fizeram com que nossas crenças morais acompanhassem os fatos morais *porque* a capacidade para detectar verdades morais promoveu o sucesso reprodutivo. Parece que se pode rejeitar essa explicação sem negar a tese do acompanhamento. Em vez disso, parece, os realistas deveriam poder invocar a abordagem da ligação adaptativa para explicar por que chegamos a ter a capacidade de formar crenças morais que acompanham os fatos. A própria Street pensa que o construtivismo pode empregar a abordagem da ligação adaptativa para explicar por que chegamos a ter essa tendência.

Nós reformulamos o dilema, no entanto. Na alternativa do dilema reformulado que afirma o acompanhamento, os realistas aceitam a tese do *acompanhamento parcial*. Novamente, o aspecto importante é que a *tese* do acompanhamento parcial diz que as forças darwinianas fizeram com que nossas crenças morais acompanhassem parcialmente os fatos morais. *A abordagem do acompanhamento parcial* é a hipótese segundo a qual as forças darwinianas

---

[38] 2006, p.125-126.
[39] 2006, p.135.
[40] 2006, p.128-135.

fizeram com que nossas crenças morais acompanhassem parcialmente os fatos morais *porque* a capacidade de detectar verdades morais promoveu o sucesso reprodutivo entre nossos ancestrais. Parece que os realistas podem rejeitar essa explicação sem negar a tese do acompanhamento parcial. Como defenderei, eles podem recorrer à abordagem da ligação adaptativa para explicar por que a tese do acompanhamento parcial é verdadeira.

Parece, ademais, que a abordagem do acompanhamento, ou ao menos um parente próximo dessa abordagem, é de fato compatível com a abordagem da ligação adaptativa. Os realistas, portanto, podem combinar essas abordagens de modo a explicar por que a tese do acompanhamento parcial é verdadeira. Uma analogia pode ajudar. Presumivelmente, evoluímos de modo a sermos capazes de acompanhar fatos sobre predadores. Pode-se explicar isso com base na hipótese de que a capacidade de detectar predadores foi parte de um sistema que nos permitiu fugir de predadores, e certamente a habilidade de fugir de predadores foi adaptativa[41]. Os realistas podem adotar uma posição análoga a respeito da capacidade de detectar verdades morais. Podem dizer que nossas crenças morais tendem a acompanhar parcialmente os fatos morais porque a capacidade para detectar verdades morais foi parte de um sistema de respostas valorativas que foi adaptativo entre nossos ancestrais. Esse é o parente próximo da abordagem do acompanhamento parcial que mencionei. Os realistas podem então usar a abordagem da ligação adaptativa para explicar que o sistema de respostas valorativas em questão foi adaptativo porque motivou o comportamento que aumentou o sucesso reprodutivo.

A ideia básica, então, é que a capacidade de acompanhar parcialmente a verdade moral foi parte de um sistema de respostas valorativas que foi adaptativo porque motivou o comportamento que aumentava o sucesso reprodutivo, como explicado pela abordagem da ligação adaptativa. Como a entendo, a abordagem da ligação adaptativa postula um processo de quatro estágios por meio do qual as forças darwinianas afetaram o conteúdo de nossas crenças morais. Em algum

[41] STREET 2006, p.135-136.

estágio desse processo, nossos ancestrais adquiriram a disposição para fazer juízos morais e ter crenças morais com um conteúdo que tendia a motivá-los a responder às suas circunstâncias de maneiras que fossem reprodutivamente vantajosas. Essa história é compatível com a afirmação realista de que há fatos morais. E é compatível com a afirmação de que o processo de quatro estágios levou-nos a adquirir uma tendência a ter crenças morais que acompanhassem parcialmente a verdade. Os realistas, portanto, podem acrescentar essas afirmações à história. Podem defender que a disposição de ter respostas valorativas que tendeu a levar nossos ancestrais a responder às suas circunstâncias de maneiras que fossem adaptativas conduziu, por sua vez, através do processo de quatro estágios, a uma tendência em nossos ancestrais mais recentes a ter crenças morais que se aproximassem da verdade. E podem defender que essa tendência foi parte de um sistema de respostas que foi adaptativo.

Se tudo isso é correto, os realistas podem combinar a abordagem da ligação adaptativa com um parente próximo da abordagem do acompanhamento parcial para explicar por que a tese do acompanhamento parcial é verdadeira. Isso significa que, dado tudo o que Street mostrou, os realistas podem seguramente escolher a alternativa do Dilema Darwiniano que afirma o acompanhamento.

5. O desafio remanescente

Mesmo se os realistas puderem usar a abordagem da ligação adaptativa, da maneira que sugeri, para explicar por que a tese do acompanhamento parcial é verdadeira, permanece a questão sobre se a explicação é plausível. De fato, parece haver uma lacuna na explicação. Pois não é evidente o que, a respeito da *verdade* moral, faz com que, se a abordagem da ligação adaptativa estiver correta, a capacidade para acompanhar parcialmente a *verdade* moral seja parte de um sistema de respostas valorativas que foi adaptativo. Fica a questão-chave, então, de se os realistas que aceitem a hipótese darwiniana podem garantir a plausibilidade da tese do acompanhamento parcial. Se não podem, então, como Street argumentaria, se supomos a verdade da hipótese darwiniana, o realismo

defronta-se com a preocupação cética que ameaçava na alternativa do dilema que nega o acompanhamento.

Isso deveria ser fácil de ver, dado o modo como defini o acompanhamento parcial. A tese do acompanhamento parcial diz, basicamente, que o efeito de todas as influências sobre o conteúdo de nossas crenças morais, incluindo as influências darwinianas, foi tal que a reflexão racional pode em princípio corrigir suficientemente quaisquer influências distorcedoras, de modo a enfraquecer a preocupação cética. Se os realistas são incapazes de defender essa tese, então são incapazes de defender a afirmação de que o realismo moral pode evitar a preocupação cética.

A preocupação cética em questão é a preocupação que Street levantou em sua discussão da alternativa do dilema que nega o acompanhamento. Poderia ser útil retomar essa discussão. Lembremos que supondo, nesse contexto, que a hipótese darwiniana e a abordagem da ligação adaptativa são ao menos em grande medida verdadeiras. Com base nessas suposições, o conteúdo de nossas crenças morais foi fortemente influenciado, primeiro pelo fato de que certo conjunto complexo de atitudes e disposições morais foi adaptativo no ambiente ancestral, e, segundo, pelo fato de que certos tipos de sistemas normativos, uma vez que se tenham tornado vigentes em uma sociedade, foram mais suscetíveis do que outros a serem transmitidos de modo bem-sucedido de uma geração à próxima. Se supomos que alguma forma de realismo moral é verdadeira, influências desse tipo certamente parecerão "independentes" dos fatos morais[42]. Por isso, ao que parece, o realista está diante de uma conclusão cética. Com certeza, não é evidente como formular melhor essa conclusão. Street sugere que os realistas podem ser forçados a concluir que é improvável que nossas crenças morais sejam verdadeiras[43]. Joyce sugere, mais cautelosamente, que os realistas podem ser forçados a concluir que suas crenças morais são injustificadas, ou a reduzir

[42] SOBER 1993, p.107.
[43] 2006, p.121-122.

significativamente seu grau de confiança nessas crenças[44]. O desafio para o realista é evitar ou tornar inócua qualquer dessas conclusões.

Em resposta, os realistas poderiam defender que as influências darwinianas não são independentes, em um sentido relevante, dos fatos morais. Chamemos essa opção de "estratégia agressiva". Essa abordagem buscaria mostrar que, com efeito, as forças darwinianas tendem a acompanhar os fatos morais. Ou os realistas poderiam rejeitar o argumento que parte da afirmação de independência para qualquer conclusão cética interessante. Chamemos essa estratégia de "estratégia defensiva". Essa abordagem menos ambiciosa poderia buscar mostrar que o argumento cético não é correto ou persuasivo. Dado o modo como defini o acompanhamento parcial, ambas as estratégias podem ser descritas convenientemente como visando apoiar a tese do acompanhamento parcial. Uma abordagem completamente adequada combinaria, talvez, ambas as estratégias, mas focarei na estratégia agressiva, uma vez que me parece ser potencialmente a mais satisfatória das duas.

A estratégia defensiva é certamente familiar. Os realistas precisam responder a vários desafios céticos porque nossas crenças morais estão sujeitas a várias influências potencialmente distorcedoras, incluindo o autointeresse, a ideologia, a superstição e a cultura[45]. Pode-se pensar que a influência das forças darwinianas é simplesmente um fator potencialmente distorcedor adicional, e que nada no desafio darwiniano justifica adotar a estratégia agressiva.

Esse pensamento omite o fato de que, na hipótese darwiniana e na abordagem da ligação adaptativa, as forças darwinianas tiveram um efeito forte e penetrante sobre o fenômeno da psicologia moral como um todo. Incluídos entre esses fenômenos não estão apenas as nossas emoções morais e disposições reativas, mas também o conteúdo de nossas crenças morais e a natureza das capacidades e processos psicológicos, incluindo processos de raciocínio, que estão envolvidos na formação e avaliação das crenças morais. Se supomos a

[44] 2006, p.179-182.
[45] STURGEON 1992.

verdade desse panorama, como estou fazendo, então, parece-me, os realistas precisam mostrar que as influências darwinianas podem entrar em uma explicação de nossa habilidade de discernir a verdade moral. A estratégia agressiva visa mostrar isso.

Ao levar a cabo a estratégia agressiva, não é suficiente argumentar que as crenças morais básicas que a seleção natural nos preparou para ter, de acordo com a abordagem da ligação adaptativa, parecem, intuitivamente, estar pelo menos próximas da verdade. Não há qualquer interesse nessa afirmação, se a abordagem da ligação adaptativa é verdadeira. Pois, se a abordagem é verdadeira, as proposições morais nas quais prevê que tenderemos a acreditar têm certamente de ser proposições que tendemos a considerar plausíveis.

Para levar a cabo a estratégia agressiva, os realistas precisam explicar por que, se a abordagem da ligação adaptativa é correta, a habilidade de acompanhar parcialmente a verdade teria sido parte de um sistema de respostas que foi adaptativo. Ademais, a explicação precisa ser contrafactualmente robusta. Isto é, deve apoiar a ideia de que, mesmo se a seleção natural tivesse levado nossa psicologia moral a ser um tanto diferente do que é, e mesmo se tivéssemos tido, como resultado, crenças morais um tanto diferentes, é provável que nossas crenças morais ainda tenderiam a aproximar-se da verdade[46]. De outro modo, haveria um sentido em que seria simplesmente um feliz acaso que haja uma tendência em nossas crenças morais a aproximarem-se da verdade.

O desafio básico, no entanto, é explicar o que, a respeito da verdade moral, é tal que, se a abordagem da ligação adaptativa é correta, é provável que nossas crenças morais tendam a aproximar-se da verdade. Para enfrentar esse desafio de maneira completamente satisfatória, penso que um realista precisaria propor uma teoria das condições de verdade dos juízos morais. O realista poderia, então, usar essa teoria para explicar por que as crenças influenciadas pelas forças darwinianas, de acordo com a abordagem da ligação adaptativa, tenderiam a acompanhar parcialmente fatos que, sob a luz da teoria, são fatos morais. Algumas

[46] Agradeço a Mark Schroeder e Jon Tresan pela ajuda na formulação dessa ideia.

teorias realistas podem estar em condições de enfrentar o desafio, como tentarei mostrar agora.

6. A teoria moral centrada na sociedade

Para realizar a estratégia agressiva, os realistas que aceitem a hipótese darwiniana precisam fazer três coisas. Primeiro, precisam aceitar a tese do acompanhamento parcial. Segundo, suponho, precisam aceitar a abordagem da ligação adaptativa. Terceiro, precisam explicar por que, se nossa psicologia moral formou-se de acordo com a abordagem da ligação adaptativa, é provável que nossas crenças morais tendam a acompanhar parcialmente a verdade. É preciso explicar por que uma tendência a formar crenças morais que é explicada com base na abordagem da ligação adaptativa seria uma tendência a formar crenças morais que são em geral aproximadamente verdadeiras.

Quero agora esboçar uma versão do realismo moral naturalista, que defendi em outros lugares, e defender que pode enfrentar o desafio darwiniano. Chamo-lhe de teoria centrada na sociedade[47]. Está fora do escopo deste artigo considerar se outras formas de realismo moral podem realizar a estratégia agressiva[48].

A teoria centrada na sociedade é um tipo de funcionalismo moral. Vivemos em sociedades e precisamos viver em sociedades, mas, porque temos talentos diferentes e interesses um tanto conflitantes, nossas sociedades precisam motivar-nos a viver juntos pacificamente, cooperativamente e produtivamente. Para ser bem-sucedida nisso, uma sociedade precisa ser governada por normas ou padrões compartilhados, os quais podemos pensar como constituindo um código moral social. Nesse sentido, a 'moralidade' tem a função de permitir a uma sociedade atender suas necessidades. Ela o faz fornecendo regras para o governo de nossas vidas que, quando internalizadas de modo suficientemente amplo em uma sociedade, motivariam a cooperação e o comportamento pacífico e produtivo que

[47] COPP, 1995 e 2007a. Apresentei versões um tanto diferentes da teoria em lugares diferentes. Para mais sobre isso, ver a introdução a Copp (2007a).
[48] Para observações sugestivas, ver BOYD, 1988, p.208-209.

são necessários para seus membros. Diferentes códigos morais difeririam em quão bem sua vigência em uma sociedade serviria a essa função. Dado isso, como agora explicarei, a teoria diz que uma proposição moral básica, tal como a proposição de que a tortura é errada, seria verdadeira somente se o código moral que servisse melhor à função de permitir à sociedade satisfazer suas necessidade incluísse ou implicasse uma norma correspondente em um sentido relevante, tal como a proibição da tortura.

A teoria tem duas partes. A primeira é uma abordagem das condições de verdade de proposições normativas, a qual chamo de abordagem "baseada em padrões". A segunda é a abordagem centrada na sociedade sobre o modo 'embasador da verdade' dos 'padrões' morais.

A abordagem baseada em padrões baseia-se em uma distinção entre *proposições* morais, tal como a proposição segundo a qual a tortura é errada, e *padrões* morais, tal como o padrão proibindo a tortura que poderia ser expresso pelo imperativo "Não torture ninguém!". A abordagem propõe um esquema que pode ser usado para explicitar as condições de verdade de proposições morais em termos do modo [*status*] dos padrões correspondentes relevantes. O esquema diz que uma proposição moral (pura e básica) é verdadeira se, e somente se, um padrão moral correspondente tem o modo embasador da verdade relevante[49]. Por exemplo, é verdade que a tortura é errada se, e somente se, um padrão correspondente – presumivelmente o padrão que proíbe a tortura – tem o modo embasador da verdade relevante.

A ideia é que, se há quaisquer verdades morais, há *algum* modo que um padrão pode ter, tal que, se um padrão desfruta desse modo, uma proposição moral (pura e básica) correspondente é verdadeira. Tem de haver *alguma* propriedade que distinga o padrão que proíbe a tortura de padrões que não tenham qualquer caráter moral, tal como o padrão que proíbe homens de usar

[49] Uma proposição moral 'pura' não tem nenhuma implicação ou pressuposição não-moral (outras que aquelas dadas pela própria teoria baseada no padrão). Uma proposição moral 'básica' atribui uma propriedade moral a algo.

chapéu dentro de casa. Podemos chamar o modo embasador da verdade dos padrões morais de seu modo de serem moralmente "autorizadas". A abordagem baseada em padrões, então, diz que, se a tortura é errada, ela é proibida por um padrão moralmente autorizado.

Agora considere a intuição de que uma sociedade precisa de um 'código moral social'; isto é, uma sociedade precisa dele para ser o caso de que há, entre seus membros, um sistema de padrões morais que é geralmente subscrito e que se faz cumprir socialmente e que é culturalmente transmitido e que pede por comportamentos pró-sociais de vários tipos. Outras coisas sendo iguais, uma sociedade com tal código moral experienciaria menos conflito entre seus membros, e menos conflito danoso, que experienciaria se carecesse desse código. Outras coisas sendo iguais, haveria mais cooperação entre seus membros do que haveria se carecesse de tal código, e seus membros seriam mais bem-sucedidos em satisfazer suas próprias necessidades e seguir seus valores. Uma sociedade com tal código é melhor do que seria de outro modo, se tudo o mais é igual, em satisfazer sua necessidade de que haja cooperação entre seus membros e sua necessidade de evitar o conflito interno danoso.

Obviamente, alguns códigos morais sociais seriam melhores para possibilitar a uma sociedade satisfazer as suas necessidades básicas do que outros. De acordo com a teoria centrada na sociedade, o código que serviria melhor às necessidades básicas de uma sociedade, se fosse servir de código moral social daquela sociedade, é o código que é moralmente autorizado com respeito àquela sociedade. Ele tem o modo embasador da verdade relevante. Padrões que estão incluídos ou são implicados pelo código subscrevem à verdade das proposições morais básicas correspondentes. Nesse sentido, diz a teoria, uma proposição moral básica é verdadeira somente se um padrão moral correspondente está incluído no, ou é implicado pelo código moral cuja vigência na sociedade relevante permitiria à sociedade servir melhor às suas necessidades básicas do que serviria a vigência de outros conjuntos de normas, e melhor do que

seria o caso se nenhum conjunto de normas vigorasse na sociedade[50]. Nessa proposta, resulta que ser errado é relacional. A teoria, com efeito, diz que o termo "errado" expressa uma relação, ser errado em relação à sociedade S, onde, grosso modo, essa é a propriedade de ser proibido pelo código moral cuja vigência em S seria 'melhor' para a sociedade S, e onde a identidade da sociedade S é fixada pelo contexto.

Essa abordagem básica precisa ser aprimorada de várias maneiras[51]. Por exemplo, precisa ser aprimorada para permitir ligações entre códigos morais distintos, cuja vigência em uma sociedade serviria igualmente bem às necessidades da sociedade. No que segue, ignoro essa e outras complicações.

Quais são as necessidades básicas de uma sociedade? Essa é uma questão importante e difícil, mas certas coisas parecem óbvias. Uma sociedade precisa garantir que sua população continue a existir. Precisa garantir que haja e continue havendo um sistema de cooperação entre seus membros. Precisa garantir a harmonia social interna. Precisa de relações pacíficas e cooperativas com sociedades vizinhas. A questão sobre como promover essas necessidades é empírica, e as implicações morais da teoria são contingentes e um tanto especulativas. No entanto, penso que é provável que a teoria favoreça um código moral deontológico de um tipo familiar[52]. Uma vez que as sociedades têm as mesmas necessidades básicas, os códigos morais que são autorizados relativamente a diferentes sociedades tenderão a ser similares em conteúdo. Ainda, as sociedades podem estar em circunstâncias diferentes e, então, é improvável que os 'melhores' códigos morais para diferentes sociedades sejam exatamente os mesmos.

Essa abordagem levanta um número de questões. O que são sociedades? Qual é a sociedade 'relevante'? Pode-se defender a teoria contra as objeções familiares ao relativismo moral? A teoria fornece uma abordagem plausível sobre a

---

[50] Ver COPP, 1996, p.252-253.
[51] COPP, 1995, p.198-200; COPP, 1996, p.257-258. Ver também a introdução a COPP, 2007a.
[52] COPP, 1995, p.201-209.

normatividade das proposições morais? Não posso tratar dessas questões aqui[53].

A questão importante para os propósitos presentes é se a teoria pode ilustrar uma estratégia que permitiria aos realistas morais escapar ao desafio darwiniano. Como vimos, a chave para escapar o desafio é explicar por que, concedendo a hipótese darwiniana e supondo a abordagem da ligação adaptativa, nossas crenças morais poderiam tender a acompanhar parcialmente a verdade moral. A questão é se a teoria centrada na sociedade pode explicar isso.

Argumentarei que, se uma população começa tendo crenças morais com um conteúdo mais ou menos previsto pela abordagem da ligação adaptativa, suas crenças iniciais aproximam-se suficientemente da verdade moral, sob a luz da teoria centrada na sociedade, e que, dadas a deliberação e a reflexão adequadas, sendo o restante igual, é provável que suas crenças aproximem-se da verdade. Isto é, argumentarei que as proposições morais que a teoria centrada na sociedade afirma serem verdadeiras "acabam por ser [de modo suficientemente próximo] os mesmos juízos que forjam as ligações adaptativas entre as circunstâncias e a resposta"[54]. Dado isso, a teoria centrada na sociedade pode evitar a conclusão cética de Street.

7. A teoria centrada na sociedade e o desafio darwiniano

Para construir o argumento, preciso conectar a explicação do desenvolvimento de nossa psicologia moral fornecida pela abordagem da ligação adaptativa com a abordagem da verdade moral fornecida pela teoria centrada na sociedade.

De acordo com a abordagem da ligação adaptativa, as forças darwinianas afetaram o conteúdo de nossas crenças morais através de um processo de quatro estágios do tipo que esbocei anteriormente. Em um estágio inicial, nossos ancestrais remotos desenvolveram uma tendência a ter disposições altruístas e de indução da cooperação, porque essas disposições foram adaptativas no ambiente

[53] COPP, 1995, COPP, 2007a.
[54] STREET, 2006, p.132.

ancestral. Uma capacidade para governo normativo também teria sido adaptativa. Em um estágio subsequente do processo, os ancestrais com essa capacidade teriam passado a compartilhar um sistema de normas que reforçaria disposições pró-sociais. Em um estágio posterior, como resultado de um processo de evolução cultural, certos tipos de normas teriam se tornado mais amplamente aceitas, incluindo, especialmente, as normas projetadas para promover a estabilidade social, a paz e a cooperação[55]. Finalmente, como resultado de tudo isso, os humanos desenvolveram uma tendência a formar crenças morais que favoreçam tipos importantes de comportamento pró-social, incluindo comportamentos que promovea a estabilidade social, a paz e a cooperação.

A abordagem não depende da verdade dos juízos morais que prevê que os seres humanos estariam dispostos a aceitar por causa das forças darwinianas. Todavia, é compatível com sua verdade ou verdade aproximada. E, como explicarei, a teoria centrada na sociedade implica que é provável que esses juízos seriam aproximadamente verdadeiros.

O aspecto importante é que, nessa concepção do processo de quatro estágios, provavelmente o processo da evolução cultural teria levado à vigência de códigos morais favorecendo o comportamento que aumentaria a capacidade das sociedades de atenderem a suas necessidades. Sociedades precisam que haja estabilidade social, cooperação entre seus membros e relações pacíficas com seus vizinhos. A vigência de normas que promovam a estabilidade social, a paz e a cooperação promoveria o comportamento que tende a atender a essas necessidades. Ademais, em nossa imagem do processo de quatro estágios, a evolução cultural tenderia a levar à vigência de códigos morais que incluam normas desse tipo, normas que exijam tipos de comportamento que promovam a estabilidade social, a paz e a cooperação. Uma pessoa que subscrevesse a tal código tenderia a formar as crenças 'correspondentes'[56]. Por exemplo, se o código

[55] KITCHER, 2006, p.173.

[56] Uma crença de que *p* "corresponde" a um código moral *C* apenas quando (1) *C* inclui ou implica um padrão moral *s* e (2), de acordo com a teoria centrada na sociedade, *p* é verdadeira apenas no caso em que *s* está incluído em ou é implicado pelo 'melhor' código moral para a sociedade relevante.

pede que seja cooperativa, tenderia a formar a crença de que esse comportamento é certo ou bom. E tenderia a agir de acordo, dado que tem uma capacidade para governo normativo. Seu comportamento tenderia, então, a aumentar a habilidade da sociedade de satisfazer as necessidades que esbocei anteriormente. Então, na abordagem da ligação adaptativa e na abordagem do processo de quatro estágios, a evolução cultural plausivelmente favoreceria os códigos morais cuja vigência em uma sociedade aumentaria a habilidade da sociedade de satisfazer às suas necessidades[57].

De acordo com a teoria centrada na sociedade, dada a abordagem da ligação adaptativa, na medida em que o processo de evolução social prossegue, outras coisas sendo iguais, as crenças morais das pessoas tenderiam mais e mais a aproximarem-se da verdade moral. Isto é, as crenças que correspondem ao código moral social que emergiu dessa maneira de um processo de evolução cultural tenderia crescentemente a aproximar-se da verdade. Pois a teoria implica que as crenças morais que correspondem a um código moral, cuja vigência em uma sociedade tende a promover a habilidade da sociedade de satisfazer às suas necessidades, são em geral aproximadamente verdadeiras. E é razoável pensar que a deliberação e a reflexão adequadas poderiam aproximar ainda mais essas crenças da verdade.

Nesse sentido, afirmo que a teoria centrada na sociedade pode explicar por que, outras coisas sendo iguais, e dada a abordagem da ligação adaptativa e a abordagem do processo de quatro estágios, é provável que nossas crenças morais tendam a acompanhar parcialmente a verdade moral. A teoria pode explicar por que uma tendência a formar crenças morais com um conteúdo que foi influenciado pelo processo de quatro estágios seria uma tendência a formar crenças que seriam suficientemente próximas da verdade moral e que, dadas a reflexão e deliberação apropriadas, nossas crenças morais tenderiam a chegar ainda mais perto da verdade.

A abordagem implica que se pode esperar, *outras coisas sendo iguais*, que

[57] COPP, 2007a, p. 85.

nossas crenças morais estejam suficientemente próximas da verdade da qual a deliberação e a reflexão tenderiam a nos aproximar, mas não há *garantia* de que seja assim[58]. Muitas contingências poderiam ter interferido, incluindo superstição e ideologia. Ademais, se nossas crenças morais 'iniciais' estivessem tão fora do trilho, a deliberação moral poderia simplesmente ter nos levado mais longe no erro. Richard Boyd apontou, similarmente, que a confiabilidade da metodologia científica depende da medida em que "as teorias atualmente aceitas são aproximadamente verdadeiras, de modo relevante"[59]. Parece provável, no entanto, que as crenças morais que tenderíamos a ter (dada a influência do processo de quatro estágios) tenderiam a estar próximas o suficiente da verdade (de acordo com a teoria centrada na sociedade), de modo tal que haja uma chance realista de que os seres humanos que começaram com essas crenças, e que refletiram sobre elas em um esforço para alcançar equilíbrio reflexivo, tenderiam, ao longo do tempo, a chegar mais perto da verdade, sendo outras coisas iguais – dado que os seres humanos têm o tipo de psicologia moral que a abordagem da ligação adaptativa sugere que têm. Creio que o argumento mostrou, ao menos, que, se nossas crenças morais são verdadeiras ou aproximadamente verdadeiras, isso não é uma questão de chance. Na teoria centrada na sociedade, isso deve-se à natureza da moralidade e das condições de verdade das proposições morais.

8. Objeção: um dilema darwiniano de segunda ordem

A teoria centrada na sociedade é uma teoria realista de acordo com a qual os fatos morais são idênticos a certos fatos naturais ordinários que têm a ver com as necessidades das sociedades. Street sustenta que qualquer versão genuinamente realista do naturalismo moral confronta um dilema darwiniano de segunda ordem[60]. Isso porque, pensa ela, um naturalista tem de valer-se amplamente de intuições morais para elaborar uma teoria sobre quais fatos

[58] Para uma discussão útil, ver SOBER,1994, p.95-99.
[59] BOYD, 1988, p.190. A passagem está em itálico no original.
[60] 2006, p.139-141.

naturais são idênticos a fatos morais e, na hipótese darwiniana, o conteúdo dessas intuições foi fortemente influenciado pela seleção natural. O naturalista, então, enfrenta a escolha entre sustentar ou negar, grosso modo, que a seleção natural fez com que nossas intuições morais acompanhassem parcialmente os fatos morais[61].

Pelas razões que forneci, o naturalista claramente deveria adotar a alternativa que afirma o acompanhamento. Como defendi, o naturalista pode combinar a afirmação de que as forças darwinianas fizeram com que nossas intuições acompanhassem parcialmente os fatos morais com a abordagem da ligação adaptativa. Pode-se escapar do dilema de segunda ordem da mesma maneira que se pode escapar do dilema de primeira ordem.

Meu argumento a favor da teoria centrada na sociedade, ademais, vale-se em grande medida de intuições filosóficas de segunda ordem, incluindo a ideia de que a moralidade tem a função de tornar a sociedade possível, e da abordagem baseada em padrões da semântica dos juízos normativos. Certamente, não consideraríamos a concepção centrada na sociedade plausível se suas implicações conflituassem drasticamente com nossas concepções morais, especialmente se pensássemos que haveria ainda conflito depois da deliberação adequada. Todavia, meu argumento não depende de nossas intuições morais. De fato, penso que podemos imaginar um filósofo marciano, com intuições morais substantivas diferentes das que temos, que, não obstante, pudesse aceitar a teoria centrada na sociedade, pelas mesmas razões que me conduziram à concepção centrada na sociedade[62]. Parece, portanto, que meu argumento para a concepção centrada na sociedade não é vulnerável a um desafio darwiniano do tipo que Street tem em mente.

Talvez, no entanto, ele seja vulnerável de uma maneira diferente[63]. Talvez as intuições de segunda ordem em que me apoiei tenham sido influenciadas pelas

---

[61] Presumivelmente, um construtivista também defrontaria esse dilema de segunda ordem. Mas veja STREET, 2006, n.57.

[62] Ver STREET, 1996, n. 57.

[63] Jon Tresan pressionou-me sobre esse tópico.

forças evolutivas de uma maneira que torna meu argumento aberto a um novo dilema darwiniano. Talvez eu me defronte com uma escolha entre um resultado cético, se *nego* que as forças darwinianas fizeram com que minhas intuições de segunda ordem acompanhem os fatos, tais como fatos sobre a função da moralidade, e uma hipótese darwiniana implausível, se afirmo que as forças darwinianas *fizeram* com que as intuições de segunda ordem acompanhassem esses fatos.

É implausível, no entanto, que o conteúdo das intuições de segunda ordem que invoquei tenha sido fortemente influenciado pelas forças evolutivas. É plausível que o fenômeno da psicologia moral tenha sido influenciado por pressões evolutivas, porque nossa psicologia moral incita comportamento que pode ser ou adaptativo ou desadaptativo. Mas as concepções de segunda ordem sobre a moralidade e sobre as condições de verdade dos juízos normativos não têm a conexão direta com o comportamento que as crenças morais de primeira ordem têm. Elas são concepções filosóficas, antes que concepções morais normativas que dizem como comportar-se. Por isso, embora de fato tenhamos evoluído para sermos capazes de ter pensamentos filosóficos, é implausível que o *conteúdo* desses pensamentos tenha sido moldado pelas forças evolutivas. Não há, portanto, nenhuma plausibilidade na ideia de que uma pessoa que pense que haja fatos sobre questões filosóficas defronte-se com um dilema darwiniano.

Conclusão

Na hipótese darwiniana, os fenômenos da psicologia moral, incluindo o conteúdo de nossas crenças morais, foram fortemente afetados pelas forças darwinianas. Street argumenta que, nessa abordagem, os realistas morais defrontam-se com um dilema. Na alternativa que aceita o acompanhamento, estão comprometidos em negar a abordagem da ligação adaptativa, a abordagem mais plausível da atuação das forças darwinianas sobre nossa psicologia moral. Na alternativa que nega o acompanhamento, precisam conceder que é “pura chance” se nossas crenças morais tendem a ser verdadeiras. Em resposta, argumentei que os realistas podem adotar a alternativa do dilema que afirma o acompanhamento e

que podem usar a abordagem da ligação adaptativa para explicar por que a tese do acompanhamento parcial se sustenta. Apresentei a teoria moral centrada na sociedade, que é um tipo de naturalismo moral. Defendi que ela pode explicar por que nossas crenças morais tenderiam a acompanhar parcialmente a verdade moral, se seu conteúdo foi afetado pelas forças darwinianas, tal como sugere a abordagem da ligação adaptativa. Na teoria centrada na sociedade, outras coisas sendo iguais, uma tendência a formar crenças morais com um conteúdo que foi influenciado da maneira sugerida pela abordagem da ligação adaptativa seria uma tendência a formar crenças que se aproximam da verdade. A teoria centrada na sociedade mostra, então, que o desafio darwiniano não representa nenhum risco para o realismo moral. O argumento contra o realismo moral tem de valer-se de objeções filosóficas mais familiares à teoria centrada na sociedade e a outras formas de realismo e não no dilema darwiniano.

Referências

BLACKBURN, Simon. *Ruling Passions*. New York: Oxford University Press, 2000.

BOYD, Richard. "How to be a Moral Realist". In: Geoffrey Sayre-McCord (ed.), *Essays on Moral Realism*. Ithaca, New York: Cornell University Press, 1988, p.181-228.

CAMPBELL, Richmond. "Can Biology Make Ethics Objective?". In: *Biology and Philosophy* 11, p.21-31, 1996.

COPP, David. *Morality, Normativity, and Society*. New York: Oxford University Press, 1995.

_____. "Moral Knowledge in Society-Centered Moral Theory". In: Walter Sinnott-Armstrong e Mark Timmons (eds.). *Moral Knowledge? New Readings in Moral Epistemology*. New York: Oxford University Press, 1996 p.243-266.

_____. *Morality in a Natural World*. Cambridge: Cambridge University Press, 2007a.

_____. “Moral Naturalism and Three Grades of Normativity”. In: *Morality in a Natural World*. Cambridge: Cambridge University Press, 2007b.

GIBBARD, Allan. “Human Evolution and the Sense of Justice”. In: *Midwest Studies in Philosophy* 7: p.31-46, 1982.

_____. *Wise Choices, Apt Feelings*. Cambridge, Mass.: Harvard University Press, 1990.

JOYCE, Richard. *The Evolution of Morality*. Cambridge, Mass.: MIT Press, 2006.

KITCHER, Philip. “The Evolution of Human Altruism”. In: *Journal of Philosophy* 90: p.497-519, 1993.

_____. “Biology and Ethics”. In: David Copp (ed.), *The Oxford Handbook of Ethical Theory*. New York: Oxford University Press, 2006, p.163-185.

RUSE, Michael. *Taking Darwin Seriously*. Oxford: Basil Blackwell, 1986.

SINNOTT-ARMSTRONG, Walter. *Moral Skepticisms*. New York: Oxford University Press, 2006.

SHAFFER-LANDAU, Russ. *Moral Realism: A Defense*. Oxford: Clarendon Press, 2003.

SOBER, Elliott. “Prospects for an Evolutionary Ethics.” In: Elliot Sober, *From a Biological Point of View*. Cambridge: Cambridge University Press, 1994, p.93-113.

_____; WILSON, David Sloan.*Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, Mass.: Harvard University Press, 1998.

STREET, Sharon. “A Darwinian Dilemma for Realist Theories of Value”. In: *Philosophical Studies* 127, p.109-166, 2006.

_____. “Constructivism About Reasons”. In: Russ Shafer-Landau (ed.). *Oxford Studies in Metaethics*, vol. 3. Oxford: Clarendon Press, 2008.

STURGEON, Nicholas L. “Nonmoral Explanations”. In: *Philosophical Perspectives* 6, “Ethics”, p.97-117, 1992.

8
MORALIDADE, MOTIVAÇÃO E OBJETIVIDADE

Alcino Eduardo Bonella

Introdução

O expressivismo é criticado normalmente porque não descreveria adequadamente a natureza dos juízos morais, e excluiria a possibilidade da objetividade moral. Tentarei mostrar neste trabalho que a crítica está errada, porque o expressivismo é totalmente compatível com as demandas da objetividade moral, e mesmo com a aceitação, no devido sentido, da verdade moral *independente da opinião* (individual ou coletiva): ele descreve adequadamente o juízo moral, sendo mesmo uma teoria mais coerente com a fenomenologia comum da moralidade (por exemplo, à luz da função principal do juízo moral e das alternativas descritivistas corriqueiras – intuicionismo e naturalismo); e, bem compreendido, não é relativista. Ele não é relativista porque não só admite, mas impõe, aos juízos morais, a possibilidade lógica da autocontradição, que torna tais juízos incompatíveis entre si, e, logo, também torna impossível que ambos sejam verdadeiros. Estes pontos são a parte principal de um argumento em prol do expressivismo em que, ao mesmo tempo, se tenta esclarecer que a objetividade moral não se reduz à objetividade moral fatual (de tipo descritivista).

1. Moralidade e a crítica ao expressivismo

Segundo Shafer-Landau, o expressivismo: não descreveria bem a natureza dos juízos morais, excluiria a possibilidade do argumento racional e da objetividade moral, e seria refutado pelo fenômeno da akrasia (fraqueza da vontade)[1]. Como a maioria dos escritores em metaética, Shafer-Landau define o expressivismo como

[1] 2012, p.312-319.

a posição não-realista e não-cognitivista acerca dos juízos morais. Um julgamento como "a tortura é errada" não relataria ou descreveria um fato natural ou moral do mundo (a posição realista) e, por isso, não estaria apto à verdade ou falsidade, nem seria uma peça de nosso conhecimento sobre o mundo (a posição cognitivista). Tal julgamento serviria, antes, para manifestar nossos sentimentos de desaprovação da tortura, para dar ordens ou comandos para que não se torture, e para expressar nossos compromissos com a civilidade (nossas atitudes contra a tortura policial, por exemplo). Assim, um expressivista discordaria do realista cognitivista (em sentido estrito), mas também discordaria do ficcionalismo e da teoria do erro, ou seja, do subjetivismo ético, pois a prática moral estaria em ordem, seria ainda um tipo de atividade racional, exatamente porque é simples e essencialmente a maneira de expressarmos nossas escolhas, aprovações, atitudes[2].

Agora, se perguntássemos às pessoas em geral se é isso que elas fazem quando dizem que a tortura é errada, Shafer-Landau acredita (quem sabe ele fez a pergunta e obteve tal resposta!) que dirão: 'não é nada disso, *não estamos expressando nossa condenação da tortura*, estamos descrevendo que ela tem essa característica, de ser errada, ou imoral, acreditando que isso é algo que não depende de nossa vontade ou da vontade de nossa sociedade, e querendo que outros acreditem que isso é realmente verdade'. Shafer-Landau devia ter em mente algo como 'quando você diz que a tortura é errada você acredita que ela realmente é errada ou você acha que são apenas suas emoções sendo manifestadas?' Mas isso seria nos predispor contra o expressivismo: eu mesmo, se fosse um expressivista, diria que a tortura é realmente errada. Parece que há um empobrecimento que distorce a posição filosófica do expressivismo, mas é nesse sentido empobrecido que Shafer-Landau acredita que ele falha em captar o essencial, ou algo essencial, do juízo moral. Shafer-Landau desafia o expressivista a traduzir frases como "se a guerra é imoral, então generais militares *são menos*

[2] Estranhamente, para Shafer-Landau, o expressivismo é uma posição *nihilista*, que rejeita a objetividade dos juízos morais.

virtuosos do que parecem", "algumas ações, apesar de *preencherem os requisitos dos deveres morais,* ainda *carecem de valor moral*"[3]. Mas isso não parece de modo algum ser o caso. Podemos dizer: "se a guerra deve ser proibida, então generais militares não devem chamados de virtuosos"; e, "algumas ações são exigidas, mas não é bom que sejam feitas". (Essa última frase é estranha nas duas formas! Que ação é um dever e ao mesmo tempo carece de valor?).

Ligado a isso está a segunda crítica, essa também muito comum em vários filósofos. A objeção é mais ou menos a seguinte: se juízos morais não descrevem fatos, mas, antes, expressam atitudes ou comandos, então eles não têm lógica e não podem ser objetivos. Por objetivos entenda-se: que valem para todos (universais) e excluem juízos em contradição com eles (são necessários), ou, no devido sentido, que não são arbitrários e relativos ao que pensa o indivíduo ou sociedade que expresse as atitudes morais. Shafer-Landau escreve que "expressivistas querem um modo de ter confiança na moralidade ao mesmo tempo em que *rejeitam a objetividade ética*"[4], ou que expressivistas "concordam com teóricos do erro *em negar que a ética é objetiva,* e negar que deveres morais nos dão razões categóricas"[5]. Como em todos os discursos, a objetividade de algum modo se vincula ao argumento lógico, e como a interpretação expressivista dos juízos morais excluiria tal coisa, então o expressivismo e a objetividade moral não poderiam se casar, de jeito nenhum.

3. A função essencial do juízo moral e a adequação do expressivismo

Shafer-Landau parece pensar que as expressões de desaprovação ou ordens não podem ser universais, mas sempre particulares, e que quando transformamos tais sentenças expressivas em proferimentos universais, estamos a modificar sua gramática. Mas isso só seria confortante se a universalidade deste tipo de sentença não fosse compatível com o expressivismo por alguma razão

[3] 2012, p.316.
[4] 2012, p. 313.
[5] 2012, p. 317.

inerente a ele, algo do tipo, é necessário vincular arbitrariamente à ideia de expressar uma atitude, a ideia de tal coisa como universal, pois a mera desaprovação é sempre particular, e, além disso, se tal expressão universal fosse uma forma disfarçada de realismo ou descritivismo, algo do tipo, quando alguém diz uma ordem universal ele está de fato enunciando ou relatando um fato do tipo ("todas as ações que desumanizam pessoas são *realmente* erradas"), que, tomadas pelo valor de face, descreveriam uma propriedade das ações deste tipo. Mas não descrevem, ao menos não literalmente! Basta explicitar de modo inequívoco para que usamos frases e argumentos deste tipo. Suponha que a pessoa que nos ouve dizer "todas as ações que desumanizam pessoas, incluindo tortura, que é uma ação deste tipo, são erradas" diga que acredita em nós, que concorda conosco. O que esperamos desta concordância? Que ela acredite nisso, conheça isso, no sentido de conhecer que a proposição é verdadeira? Ou que ela não torture? O que pensaríamos se ela dissesse que concorda com o enunciado e fosse torturar alguém logo depois? À luz da finalidade principal do argumento moral, é a frase "todas as ações que desumanizam pessoas são imorais" que tem de ser traduzida ou parafraseada por "Não se deve desumanizar pessoas", e, lógica formal à parte, com certeza, por "Não desumanize pessoas!" (ou "Ações desumanizantes – arrhhh! Não faça isso!).

Se o ouvinte diz que concorda e depois tem a possibilidade de desumanizar (torturar) pessoas, então ele, se estava sendo sincero, não irá fazer isso. Se fizer, ou é um mentiroso, ou concordou com alguma outra coisa (não entendeu), ou sofre de um problema psicológico chamado fraqueza da vontade (akrasia). Mas, fora o problema psicológico da akrasia, o argumento moral existe exatamente para eliminar as duas primeiras possibilidades. E na verdade, fazendo isso, ele existe exatamente como meio para produzir objetividade moral, ou seja, eliminar o juízo "a tortura é moral" e levar à manutenção do juízo "a tortura é imoral", e, então, permitir que se expresse a contradição entre (algumas) diferenças morais, e que se resolva a contradição, mas em termos práticos (que são os que importam). Hare

estava dizendo a mesma coisa[6] quando escreveu:

> Isso torna bastante implausível o argumento usado por alguns descritivistas de que, uma vez que "o ato dele foi errado" parece uma sentença sujeito-predicado ordinária, sua gramática superficial apoia a visão de que o erro é uma propriedade no sentido ordinário, e assim apoia o realismo ético. O fato de que "o ato dele foi errado" significa quase o mesmo que "Ele não devia ter feito o que fez", que tem uma estrutura superficial totalmente diferente, deveria ao menos fazer com que nos perguntássemos se a estrutura superficial da primeira sentença é um bom guia para seu significado e sua lógica[7].

Ou seja: quando alguém diz na forma indicativa que uma ação é errada, como a tortura, ele está de fato expressando desaprovação, ou seja, a sentença não deve ser tomada pelo valor de face, não é apenas uma forma de descrever uma propriedade das ações deste tipo, mas uma forma de prescrever, e é com tal instrumento que expressivistas fundam a objetividade, ao menos a objetividade que realmente importa. Esse é nosso último e mais difícil passo argumentativo.

4. Relativismo e objetividade moral

Frankena, antes de Shafer-Landau, também sugeriu que o expressivismo continha um tipo de relativismo. Ele escreveu:

> Em minha opinião, mesmo as formas menos extremadas dentre as teorias não-descritivistas [ele se referia ao prescritivismo de Hare] recentes carecem de profundidade. Apressaram-se a admitir uma espécie de relativismo básico. […] admitem ou mesmo insistem em que a validade das razões é, afinal, relativa, considerado o indivíduo ou a cultura a que ele pertence e, assim, juízos básicos conflitantes podem ser ambos justificados ou justificáveis […] esquecem por completo […] a exigência de serem

[6] 2000, p.136.
[7] Tradução em HARE, 2003, p.185-186.

> justificados ou considerados objetivamente e racionalmente[8].

David Lyons chamou de *relativismo individualista* a teoria de Hare, exposta em *A Linguagem da Moral,* e sobre isso escreveu: "Porque os princípios não descrevem o mundo, eles não podem nem corresponder aos fatos objetivos, e portanto não podem ser considerados como objetivamente obrigatórios"[9]. Habermas[10] parece que pensou que o não-cognitivismo presente nas teorias imperativista, emotivista e prescritivista, leva ao ceticismo, porque não se poderia associar uma pretensão de verdade ou qualquer pretensão de validez universal, destinada à argumentação, e tal noção seria, para tais abordagens, uma ilusão do modo cotidiano de compreender. Por fim, consideremos a citação de Philippa Foot:

> O emotivista ou prescritivista está comprometido com alguma forma de relativismo, mesmo que ele não goste da pecha. Stevenson, que reivindicou ter refutado o relativismo moral, não deixa de tornar-se ele próprio um tipo de relativista moral[11].

A principal razão deste compromisso, para Foot, é que, tanto para emotivistas quanto para prescritivistas, que são teóricos para quem predicados morais são algo mais do que *descritores* de uma dada realidade, a verdade moral seria, por isso, relativa aos padrões do indivíduo, que endossa tal padrão, ao expressar sua aprovação ou desaprovação. Isso é assim porque:

> […] essas teorias negam a presença de critérios objetivos, ou qualquer método objetivo pelo qual diferenças entre indivíduos com princípios morais básicos radicalmente diferentes pudessem ser resolvidas. Se tais teorias são corretas, qualquer um que interroga a verdade de um juízo moral, e ainda possui o recurso de testá-la pelos seus princípios morais mais básicos, usa "verdadeiro" substancialmente, mas não pode ir além deste ponto. Segue-se que o emotivista ou prescritivista está comprometido com

[8] FRANKENA, 1981, p.129.
[9] LYONS, 1992, p.34, 205.
[10] 1989, p.76.
[11] Citado por HORGAN & TIMMONS 2006, p.73.

uma forma de relativismo[12].

Do outro lado do debate, Stevenson sustentou que o emotivismo (que é uma forma de expressivismo), não só não defendia o relativismo, como era uma resposta a ele, e devia nos fazer compreender por que os erros do relativismo eram tentadores[13]. Hare sustentou que todo descritivismo, reduzindo a correção moral às condições de verdade dos enunciados, e fazendo isso exatamente para obter a objetividade, cai vítima do relativismo, pois condições descritivas de verdade podem variar entre culturas ou indivíduos, enquanto o não-descritivismo (nome que ele dá para o que aqui chamamos de expressivismo), que não reduz assim o significado completo dos termos morais, mas mantém que há algo extra, *além* das condições de verdade, não cai, pois o aspecto prescritivo não varia e não torna possível que, de duas proposições morais que se opõem, ambas sejam verdadeiras: ou seja, para escaparmos do relativismo tínhamos de ser expressivistas![14] Blackburn sempre se opôs ao relativismo e à visão de que não há nenhuma verdade, valendo, em moralidade, tudo[15]. Ele defendeu que o expressivismo, que pensa a partir de dentro ou do que projetamos, é, do modo como disseram Stevenson e Hare, uma solução ao relativismo[16]. Stevenson, Blackburn e Hare, nunca negaram que haja *também* significado descritivo e espaço para a verdade moral e conhecimento moral, e ao menos Blackburn e Hare rejeitam a expressão *não-cogitivismo* para suas teorias[17]. Gibbard, como Hare, argumentou que é um erro confundir expressivismo com subjetivismo e que a objetividade é bem explicada pelo seu normo-expressivismo[18].

Todos os filósofos citados acima no primeiro grupo, e Shafer-Landau, incorrem nos mesmos defeitos: não descrevem apropriadamente o expressivismo,

---

[12] Citado em HORGAN & TIMMONS, 2006: p.189.
[13] Cf. HORGAN & TIMMONS, 2006, p.73.
[14] Ver HARE, 1999; 2000.
[15] Cf. 2001a, p.19
[16] Cf. BLACKBURN 2001a, p. 133-135; 2001b, p.237-238.
[17] Ver HORGAN & TIMMONS 2006, p.74-75, n. 1; HARE, 2000, p.56-57.
[18] Cf. 1992, p.153-154; ver HARE, 1999, p.21-22.

tornando-o uma caricatura fácil de rejeitar, ao negar que também haja significado descritivo nos termos morais, e, também, pretensão de verdade nos argumentos morais. Eles também confundem expressivismo com subjetivismo, que é a interpretação de que, ao enunciarmos proposições morais, estamos relatando o que ocorre em nossa mente; e, finalmente, confundem objetividade com objetividade fatual (a correspondência do enunciado moral aos fatos, naturais ou morais). Parece que essa concepção fatualista de verdade moral é parte da crítica. Mas segundo Boisvert, já Stenvenson incluiu em sua descrição do emotivismo a ideia de que atitudes e interesses, que se caracterizam como estados *de ser a favor* ou *de ser contra,* podiam ser governados racionalmente, e que sentenças morais tinham ambos, significado emotivo *e descritivo.* Eles podiam ser racionalmente governados porque atitudes eram atos volitivos que afirmavam ou negavam uma intenção de ação, e com isso tornavam logicamente incompatíveis duas sentenças contraditórias entre si, e argumentos podem ser elaborados a partir dessas afirmações/negações, conjugados com fatos acerca das ações. Ao tomar partido, um arguidor não está somente expressando uma atitude particular, sua ou de sua sociedade, mas também endossando a aplicação dos predicados morais a certas ações de forma objetiva. Pois, dado que a atitude em favor de X ("tortura é imoral") elimina a possibilidade de não X ("tortura não é imoral"), a atitude, na sua função prática, também é um endosso de que o significado descritivo presente na sentença (o padrão para aplicação do predicado moral, no caso, que todos os atos que tenham as características de tortura ou de desumanização são errados) torna verdadeira tal predicação, torna-a "o caso".

Mas aqui, expressar uma atitude não deve ser entendido como relatar ou descrever o próprio estado psicológico, ou constatar o que se aprova ou reprova de fato. E, principalmente, não pode ser confundido com a crença de que, ao afirmar que a tortura é errada, dado que 'errada' é uma expressão de desaprovação, a tortura se torna errada ou é tida como errada simplesmente porque é desaprovada subjetivamente. A tortura é errada porque tem as características que tem, e porque é realmente o caso (ou, é verdade) que não se deve praticar atos que tenham tais características. Usamos a linguagem moral

para desafiar qualquer um ou nós próprios a pensar e agir assim, ou seja, pensando tal prescrição de não torturar como objetiva. Isso não é o mesmo que tornar o dever de não torturar dependente do que nós queremos ou da nossa vontade subjetiva: expressões de atitude moral

> não são *enunciados que* o falante de fato quer alguma coisa. [Prescitivistas] não estão então expostos à crítica de fazer nossos deveres *dependentes do que* nós queremos, como se pudéssemos tornar algo nosso dever querendo isso. Se este tipo de subjetivismo fosse correto, então, se alguém dissesse que algo era seu dever, ele estaria a dizer que, como matéria de fato, ele queria aquilo; e se realmente ele quisesse, seria impossível discordar dele. Antes, prescritivistas sustentam que, ao fazermos um enunciado moral, estamos *expressando* nossa vontade racional; e se alguém mais quer algo diferente, o desacordo tem de ser resolvido pela razão […][19].

O subjetivismo elimina a possibilidade de haver contradição entre os arguidores, pois, se cada um apenas relata seu estado psicológico (sua desaprovação pessoal ou social) vigente, e se esse é realmente o caso *psicológico* (é verdadeiro *em termos de psicologia daquele que expressa a sentença*), então ambas as sentenças morais que se opõem ("torturar é errado", equivalendo a pensar ou sentir subjetivamente que é errado, e "torturar não é errado", equivalendo a pensar ou sentir subjetivamente que não é) serão verdadeiras: é verdade que um não gosta e outro gosta da tortura, e só isso. Mas aqui haverá dois significados diferentes para a palavra 'errado', poderíamos grafar como '$errado_1$' e '$errado_2$', por exemplo. Para que as sentenças se contradigam, o significado de 'errado' tem de ser o mesmo para ambas as frases e ambos os arguidores, e esse é o caso se a explicação expressivista do que é dizer tais frases, for correta. Mas não é o caso se a explicação descritivista for correta. Por isso é que Hare concluiu que o descritivismo é que conduz ao relativismo, enquanto que o expressivismo é um meio de evitá-lo.

A crítica ao expressivismo, como em Shafer-Landau e em Foot, combinando

[19] HARE, 1999, p.21-22.

a premissa de que o expressivismo, por não interpretar o juízo moral como que se referindo a fatos independentes, eliminaria a possibilidade de tais juízos serem verdadeiros ou objetivos, com a premissa de que *prescrições*, *imperativos* ou *convites/pedidos* de que se formem determinadas reações práticas perante os fatos, são sempre subjetivos, concluem que o expressivismo tornaria qualquer juízo moral, um juízo arbitrário. E, como tal, não seria somente incapaz de lidar com a contradição entre juízos morais divergentes, mas seria a própria encarnação deste tipo de arbítrio. O tipo de subjetivismo ou relativismo supostamente implicado seria a negação da racionalidade, pois um juízo moral como a condenação da tortura seria, para esta interpretação ruim do expressivismo, apenas o relato da desaprovação subjetiva, feita por alguém, ou por uma cultura, da tortura (a tortura ser errada seria algo como "eu não aprovo a tortura"; não ser errada seria algo como "eles aprovam a tortura"), e seria algo eminentemente particular, no sentido de que a condenação da tortura englobaria apenas o sujeito que realmente tem o sentimento de desaprovação disto e apenas casos particulares da ação de torturar ("a tortura é errada" equivaleria somente a "tortura: arrggghhh!", que não poderia ser premissa de um argumento). Mas ambas as premissas principais são equivocadas.

A segunda, por exemplo, implica em entender o aspecto emotivo ou prescritivo de maneira errada, de uma maneira que neutraliza o potencial prático do que seja expressar uma posição pró ou contra uma ação, ou seja, *tomar uma posição*, e não relatá-la, nem interpretá-la como necessariamente pessoal e particular. E se isso procede, não se pode concluir que os juízos morais passariam a ser arbitrários e meros meios de se expressar opiniões subjetivas, e, como tais, não poderiam mais valer em sentido forte, valer objetivamente, mas apenas em sentido fraco (seriam sempre subjetivos nos sentidos de que cada um *tem o seu,* de que eles podem variar temporalmente *apenas porque as pessoas podem ter opiniões diversas*, e de que a validade é dada pela concordância individual, pelo endosso individual). Para o expressivista, juízos morais expressam o que deve ser aprovado ou desaprovado, nos dispõem e dispõem o interlocutor a agir de um determinado modo, e implicam logicamente a eliminação da proposta que

contradiz tal prescrição. Exatamente por serem essencialmente emotivos (ou prescritivos, ou expressivos) eles são formas de dizer que se deva agir de um modo *ao invés de outro*, que se deva formar ou ter certas disposições subjetivas ao invés de outras, que se deva aceitar tais normas ao invés de outras. Isso supõe que o juízo seja seguido e tornado o caso, o que é o mesmo que ser válido para todos, ou, universal: se a tortura não deve ser feita (é errada), então torturar deve estar fora de cogitação, e não esta tortura neste momento, ou mesmo, *para mim,* e sim, que qualquer ação com as mesmas características, e *para todos,* é errada. Mas esses dois ingredientes, a possibilidade de contradição entre enunciados ou proposições, e a sua validez objetiva ou para todos, são os ingredientes normalmente pensados como necessários para que o argumento moral seja racional, e seja verdadeiro, ou, objetivamente válido.

O expressivismo, como proposta de racionalidade moral não-descritivista, também tem então um discurso sobre a verdade, almeja a objetividade, e se diferencia do subjetivismo. Para o subjetivismo, "X é bom" equivaleria a "aprovamos X", e "aprovamos" equivaleria a ter esse sentimento e relatá-lo descritivamente. Mas para o expressivismo, "X é errado" equivaleria a "não se deve aprovar X", "não faça X!", com a consequente oposição (e é para isso que há tais juízos) a juízos que contradizem isso, como "X não é errado". Como X (não torturar, por exemplo), possui características descritivas a que se aplica o termo valorativo e tais características exprimem um padrão de escolha presente no juízo, o juízo tem significado descritivo. O expressivismo não é exatamente a negação de que há tal significado, mas a interpretação de que há algo mais do que descrever tal padrão ou exprimi-lo: queremos que os juízos e argumentos morais orientem a conduta, motivem a ação no devido sentido prático. Queremos, por exemplo, que policiais passem a agir sem torturar de fato, e não que simplesmente digam que aceitam ou entendem que a tortura é errada, e este ingrediente a mais, que envolve não a crença, mas a atitude, é o significado valorativo ou expressivo das sentenças. O expressivismo dá proeminência a este significado valorativo formal, que estaria presente em todo uso de expressões valorativas ou normativas (seria constante e universal), e seria independente do significado descritivo: o significado

principal seria o valorativo e ele dirige e controla o descritivo (por exemplo, ele pode ser usado para modificar o significado descritivo das expressões). O problema todo está em interpretar adequadamente esta tese, ou ao menos problematizar o modo como os críticos como Shafer-Landau interpretam isso como a negação da objetividade moral.

Referências

BLACKBURN, S. *Essays in quasi-realism*. Oxford: Oxford University Press, 1993.

_____.*Being good.* Oxford: Oxford University Press, 2001a.

_____. *Pense.* Lisboa: Gradiva, 2001b.

BONELLA, A. E. "Prescritivismo universal e utilitarismo". In: M. C. Carvalho; D. Dallagnol. *Utilitarismo em Foco.* Florianópolils: UFSC, 2007.

_____. "Racionalidade prescritiva". In: R. Bueno. *Racionalidade, justiça, direito.* Uberlândia: Edufu, 2013.

COPP, D. "Realist expressivism – A neglected option for moral realism". In:*Social Philosophy and Policy* 18, p.1-43, 2001.

FRANKENA, W. *Ética.* Rio de Janeiro: Zahar, 1981.

GIBBARD, A. *Wise Choices, Apt Feelings.* Harvard University Press, 1992.

LYONS, D. *As regras morais e a ética.* Campinas: Papirus, 1992.

HABERMAS, J. *Consciência moral e agir comunicativo.* Rio de Janeiro: Tempo Brasileiro, 1989.

_____. *Verdade e justificação.* São Paulo: Loyola, 2006.

_____. *A ética da discussão e a questão da verdade.* São Paulo: Martins Fontes, 2009.

HARE, R. M. *The language of morals.* Oxford: Clarendon Press, 1952.

_____. *Moral thinking.* Oxford: Clarendon Press, 1981.

_____. *Essays in ethical theory.* Oxford: Clarendon Press, 1993.

_____. *Objective prescriptions.* Oxford: Clarendon Press, 1999.

_____. *Sorting out ethics.* Oxford: Clarendon Press, 2000.

_____. *Ética, problemas e propostas.* São Paulo: Unesp, 2003.

HORGAN, T.; TIMMONS, M. "Expressivism, yes! Relativism, no!" In: R. Shafer-Landau, *Oxford studies in metaethics*, 1. Oxford: Clarendom Press, 2006.

MILLER, A. *An introduction to contemporary metaethics.* Cambridge: Polity, 2003.

SHAFER-LANDAU, R. "Moral nihilism". In: *The fundamentals of ethics.* Oxford: Oxford University Press, 2010.

_____. *What happened to good and evil.* Oxford: Oxford University Press, 2004.

9

ADEQUAÇÃO SENTIMENTAL, UNIVOCIDADE E RECALCITRÂNCIA

Leonardo de Mello Ribeiro

Introdução

> Na adequação ou inadequação, na proporção ou desproporção que o afeto parece ter com a causa ou objeto que o excita consiste a propriedade ou impropriedade, a probidade ou improbidade da ação consequente.
>
> Adam Smith 1790, I.I.iii.

Neste capítulo argumentamos que uma tese neossentimentalista sobre valor (defendida por Blackburn, Gibbard, D'Arms & Jacobson) é mais bem-sucedida do que uma tese judicativa (defendida por Nussbaum, Foot, Solomon, Greenspan) e a tese da "não-prioridade" (defendida por McDowell e Wiggins) em uma série de questões que qualquer teoria sobre a relação entre valor e sentimento deve acomodar. As questões nas quais nos concentraremos podem ser classificadas como recaindo sob duas categorias gerais: o que chamaremos de 'problema da univocidade' e 'problema da recalcitrância'. É um pressuposto do capítulo que qualquer teoria sobre a relação entre valor e sentimento deve ser capaz de (1) tornar possível o desacordo acerca dos padrões para julgar algo valoroso e (2) explicar como o sentimento ou emoção pode estar em erro à luz de juízos contrários. Como tentaremos mostrar, uma proposta neossentimentalista deve ser preferida a uma proposta judicativa e à tese da "não-prioridade" no que diz respeito a acomodar aquelas duas questões gerais. Na parte final do texto,

levantaremos algumas objeções relevantes à tese neossentimentalista que estão relacionadas às questões discutidas ao longo de nossa discussão das teses judicativa e da "não-prioridade". Esboçaremos, então, uma estratégia geral de resposta, em nome do neossentimentalista, às objeções levantadas.

1. As teses judicativa, da "não-prioridade" e neossentimentalista

Em psicologia, a tese judicativa foi, por décadas, a abordagem dominante a sentimentos ou emoções[1]. Após o declínio do behaviorismo, a tese judicativa tornou-se o padrão substituto em psicologia cognitiva. Porém, este cenário passou por mudanças significativas desde pelo menos a década de oitenta. Em oposição à tese judicativa, a tese sentimentalista surgiu como uma alternativa e deu início ao que é geralmente chamado de 'revolução afetiva' em psicologia[2]. O trabalho seminal e as descobertas empíricas do psicólogo Zajonc, assim como seus desdobramentos posteriores, consolidaram-se como um desafio genuíno à tese judicativa (se não a toda a teoria, ao menos a alguns de seus mais importantes compromissos).

Esta tendência em psicologia, coincidentemente ou não, também se refletiu na filosofia (do valor em geral). A tese judicativa foi a tradição dominante na literatura filosófica ao longo dos séculos, encontrando suporte em Platão, Aristóteles e nos Estóicos[3]. Mas, assim como na pesquisa em psicologia, o debate contemporâneo filosófico sofreu significativas mudanças ao longo das últimas décadas. A tese sentimentalista tem se consolidado como uma real alternativa à tese judicativa. Em particular, desde pelo menos os esforços de Blackburn e

---

[1] Iremos usar 'sentimentos' e 'emoções' como sinônimos. Para uma distinção entre estes termos, ver Broad (1954).

[2] As teses judicativa e sentimentalista são comumente chamadas 'cognitivismo' e 'afetivismo' (ou 'não-cognitivismo', respectivamente, na literatura psicológica sobre sentimentos ou emoções. Mas, a fim de evitar uma duplicação desnecessária de terminologia, manteremos teses 'judicativa' e 'sentimentalista', que são os termos mais comuns na literatura filosófica.

[3] Autores como Foot, Solomon, Nussbaum e Greenspan (para mencionar apenas alguns) estão entre os contemporâneos que produziram extensos trabalhos que pertencem à tradição judicativa.

Gibbard em fornecer os detalhes de uma versão da tese sentimentalista que possa responder a objeções sérias a versões anteriores da mesma (defendida provavelmente por Hume, Adam Smith, Ayer e Stevenson), o programa sentimentalista em filosofia, mesmo que não possuindo um impacto tão revolucionário quanto o seu desenvolvimento análogo em psicologia, tem desfrutado de uma imagem de respeitabilidade perante seus adversários e atraído mais simpatizantes aos seus compromissos[4].

Por que mencionar essa comparação? A razão é simplesmente que é uma hipótese substantiva deste capítulo que podemos compreender melhor questões sobre filosofia do valor a partir de contribuições da pesquisa em psicologia e da chamada "revolução afetiva". Como? Se a pesquisa em psicologia puder nos ajudar a compreender melhor o que um sentimento ou emoção é, ela poderá também contribuir para a explicação do discurso e pensamento valorativos que possuem claramente um toque sentimental ou emocional. Em especial, se o programa afetivo em psicologia for bem sucedido, poderíamos identificar e explicar sentimentos ou emoções sem recorrer a conceitos e juízos valorativos, de modo que se tornaria agora possível compreender os últimos a partir de informações sobre os primeiros. E isto é exatamente o que a abordagem sentimentalista em filosofia do valor tenta fazer.

Antes, porém, de continuarmos e expormos os detalhes de nossa principal hipótese, devemos esclarecer três pontos que *não* são parte de nossas tarefas aqui. Em primeiro lugar, não é parte de nosso objetivo discutir a relevância de descobertas científicas na pesquisa filosófica e argumentar a favor do programa afetivista ou sentimentalista em *psicologia*. Ao contrário, isto funcionará como um ponto de partida condicional em nossa hipótese. *Se* o programa afetivista em psicologia for bem sucedido, uma proposta sentimentalista em *filosofia do valor* poderá se beneficiar disto, recorrendo à literatura em psicologia como um aliado. Assim, quaisquer que sejam os nossos resultados aqui, eles serão, no máximo,

[4] Ver Griffiths (1997) para um panorama geral do desenvolvimento destas questões nas literaturas psicológica e filosófica.

condicionais.

Em segundo lugar, nossa hipótese não depende de um argumento a favor da tese de que *todos* os nossos juízos e conceitos valorativos são associados a um sentimento. É suficiente para os nossos propósitos pressupor que ao menos *parte* de nosso discurso valorativo pode ser entendido em tais termos: isto é, envolvendo uma relação entre valor e sentimento.

Em terceiro lugar, uma vez que tentaremos esclarecer algo sobre o discurso e o pensamento valorativo através de uma discussão sobre a natureza de sentimentos e emoções, nosso debate não seguirá as tendências mais comuns nas investigações meta-éticas e meta-valorativas. Estas envolvem normalmente questões sobre a semântica, a ontologia e a epistemologia do discurso valorativo. Nosso foco será, em vez disto, em algumas características comuns do fenômeno do discurso valorativo como um todo, que contém certas características que não estão em disputa e são endossadas por todas as partes do debate. Perguntaremos, então, qual das três abordagens consideradas aqui (sobre a relação entre sentimento e valor) sai-se melhor para acomodar tais características comuns do fenômeno do discurso valorativo.

Tendo esclarecido este ponto, podemos agora fornecer os detalhes de nossa principal tarefa. Qual é exatamente esta? Em termos muito gerais, o que está em jogo no debate que nos interessa aqui (tanto em psicologia quanto em filosofia) é como explicar a relação entre (a aplicação de) certo tipo de conceito valorativo (ou a identificação da propriedade associada), a formulação do respectivo juízo valorativo, e o sentimento associado. Por exemplo, a relação entre o que é errado, perigoso, vergonhoso ou engraçado e os sentimentos de culpa, medo, vergonha e de achar graça[5]. Existem basicamente três principais abordagens neste tema, a saber: as teses judicativa, da "não-prioridade" e sentimentalista. No que se segue, veremos cada uma delas separadamente.

A tese judicativa defende a prioridade de juízos valorativos (que possuem

[5] Usaremos aqui a expressão 'achar graça' para representar o sentimento associado a situações engraçadas ou cômicas. Funcionará, assim, como o equivalente de 'amusement' em inglês.

conteúdo cognitivo, de acordo com esta tese) sobre sentimentos no seguinte sentido: sentimentos devem ser compreendidos como envolvendo (como parte de seu conteúdo) o juízo de que algo é valoroso. Assim, segundo esta tese, é parte do conteúdo do sentimento ou emoção de medo, por exemplo, o juízo de que algo é considerado perigoso. Dessa forma, não podemos compreender se um sentimento é de medo sem julgar algo perigoso.

Os psicólogos Lazarus & Lazarus colocam o ponto do seguinte modo:

> Uma emoção é gerada quando uma pessoa com certos objetivos e crenças sobre si mesma e sobre o mundo avaliam que aquilo que ocorre é danoso, ameaçador ou benéfico. Esta avaliação depende da razão, ainda que o raciocínio possa estar longe de ser preciso[6].

E Martha Nussbaum torna claro que o conteúdo cognitivo relevante em questão é valorativo:

> […] As crenças nas quais emoções estão baseadas incluem proeminentemente nossas crenças valorativas, nossas crenças sobre o que é bom ou mau, valoroso ou sem valor, útil ou danoso[7].

A tradição judicativa diverge em si mesma sobre o papel e o status da crença ou do juízo na explicação de tal relação. Algumas versões sustentam que o conteúdo (ou juízo) cognitivo (valorativo) da emoção é necessário e suficiente para a emoção (de forma que um pode ser identificado com o outro), enquanto que outras versões sustentam que o juízo é apenas necessário, mas não suficiente, para se possuir uma determinada emoção (ou seja, a emoção deve envolver mais do que um juízo valorativo). Ainda, outras versões mantêm que o juízo é apenas causalmente necessário para se possuir uma determinada emoção[8].

---

[6] LAZARUS & LAZARUS 1994, p.199; ver também LAZARUS 1982.

[7] NUSSBAUM, 1994, p. 370.

[8] Ver Griffiths (1997), Nussbaum (1994), Helm (2001), e Roberts (2003), para discussões detalhadas dessas posições. Algumas versões, como uma defendida por alguns estóicos, tal como descrita por Nussbaum, defendem que, para que um agente seja inteiramente racional, ele precisa promover uma extirpação total das emoções. De acordo com Nussbaum, para Crisipo, as emoções são sempre

Uma versão radical da tese judicativa é defendida por Martha Nussbaum, que sustenta que:

> [Emoções] possuem, elas mesmas, um conteúdo cognitivo; elas estão intimamente relacionadas a crenças ou juízos sobre o mundo de forma tal que a remoção da crença relevante removerá não apenas a razão para a emoção, mas também a própria emoção. A crença é a base necessária e "fundamento" da emoção. Ela pode até mesmo ser considerada uma parte constituinte da própria emoção. Raiva, por exemplo, é definida por Aristóteles, o primeiro grande proponente desta ideia, como um composto de sentimento doloroso com a crença de que me fizeram algum mal. Isto implica (como parece correto) que, se eu descobrir que a minha crença é falsa – que o mal aparente não ocorreu de fato – irei, descartando a minha crença falsa, deixar de ficar com raiva. Se algum resíduo de sentimento doloroso permanece, ele não será mais considerado raiva mas, em vez disto, como um resíduo de irritação ou excitação irracional[9].

Nussbaum parece identificar emoção com juízo ao ponto de que, segundo esta versão da tese judicativa, o abandono do juízo – como resultado, por exemplo, da descoberta de sua falsidade – implica que o sentimento ou emoção não está mais presente. Esta é a consequência natural a se esperar uma vez que juízo é identificado com emoção[10]. Mas, mesmo que a relação aqui não seja tão estrita quanto a de identidade, é importante notar que Nussbaum enfatiza que, assim que o juízo é abandonado, o que permanece é um "resíduo de irritação ou excitação irracional", mas não um sentimento ou emoção genuíno. Este é um ponto importante ao qual retornaremos posteriormente.

Para muitos, esta é uma tese controversa. Seria talvez mais plausível se o judicativista sustentasse uma relação mais fraca entre juízo e sentimento, de tal

---

falsos juízos. Shaffer (1983) também defende a dispensabilidade das emoções. Para uma tese contrária, sobre a utilidade das emoções, ver Nesse (1990).

[9] Ver Solomon (2007) para uma tese similar, especialmente o seu ensaio 'On Emotions as Judgments'.

[10] Isto parece ser confirmado pela defesa de Nussbaum (1994) da tese de Crisipo.

forma que pudesse acomodar a ideia de que podemos em princípio julgar algo valoroso e não sentir a emoção ou sentimento correspondente (ou sentir algo distinto em seu lugar), de modo que julgar X valoroso seria julgar que é *adequado* ou *apropriado* ter um sentimento ou emoção correspondente como resposta a X, embora possa não ter esta resposta (ou sentimento)[11]. Como veremos, é com base em casos como este que outras versões da tese judicativa propõem uma importante alteração em formulações mais radicais da tese, como a de Nussbaum.

Mas, seja como for, há claramente algo atraente na tese judicativa. Dado que é muito plausível sustentar que exista uma relação estreita entre juízo valorativo e emoção, este ponto é muito facilmente assimilável à tese judicativa. Se nos concentrarmos na *articulação* de sentimentos ou emoções – isto é, no uso de nossa capacidade linguística e conceitual para explicá-los e descrevê-los – como poderíamos avaliar que estamos respondendo adequadamente a uma situação de perigo se não a descrevermos, de algum modo, como envolvendo o juízo de que isto é o caso? Parece que medo é o sentimento adequado a se possuir quando quer que julgamos que algo é perigoso. E se medo é a resposta adequada, isto deve ser assim porque a situação pode ser caracterizada (valorativamente) como perigosa. Este é certamente um tipo de juízo que comumente fazemos.

Ademais, a tese judicativa também parece ganhar alguma plausibilidade de um ponto de vista fenomenológico. Como alguns sugerem, possuir uma emoção parece similar a ter uma experiência valorativa que pode ser interpretada (ao menos de um ponto de vista fenomenológico) como envolvendo uma forma de *percepção* (ou juízo perceptual). Soa defensável a ideia de que parecemos nos deparar com *o engraçado*, *o errado*, *o vergonhoso* quando normalmente experienciamos achar graça, culpa ou vergonha.

É neste sentido que Philippa Foot critica Hume (a quem ela atribui uma forma de tese sentimentalista *naïve* sobre a relação entre sentimento e valor) quando diz que:

---

[11] Os termos 'adequado' e 'apropriado' são extraídos de Broad (1954), Wiggins (2002), e do trabalho conjunto de D'Arms & Jacobson.

> Hume, penso, estava cometendo um erro quando ele tentou explicar o que significava dizer que uma ação ou qualidade é virtuosa em termos de um sentimento especial; pois a explicação do pensamento aparece na descrição do sentimento, não ao contrário[12].
> Parece estranho dizer que ele não precisa trazer à tona nenhum *fato* especial *sobre a ação* para embasar o que ele diz. Assim, obviamente, o que se quer saber é […] o que o levou a dizer que a virtuosidade de uma ação não podia ser uma simples questão de fato provável. Às vezes, de fato, sugere que não se pode simplesmente encontrar tal fato. Mas, uma vez que ele mesmo diz que todas as coisas chamadas virtudes são qualidades agradáveis ou úteis à humanidade, por que não deveria dizer que é nisto que *consiste* a sua virtuosidade? Por que tem que insistir em um sentimento de aprovação e, ao tornar isto a parte essencial do juízo moral, ancorar enunciados sobre virtude nos sentimentos do observador e não nos fatos?[13].

Foot está dizendo que simplesmente soa implausível tentar caracterizar emoções meramente como algum tipo de estado fenomenológico que possui um "sentir" peculiar. A maneira como individuamos uma emoção, como orgulho, por exemplo, se dá, de acordo com ela, ao identificarmos certo tipo de conteúdo cognitivo, nomeadamente, que o objeto ao qual a atitude é direcionada é julgado bom e este pertence a mim (algo que deve talvez também ser julgado bom)[14]. Dessa forma, emoções parecem não apenas *intencionais* e *cognitivas* (diferentemente de meras sensações, como coceiras, dores, surpresa, assim como estados apetitivos como fome e sede, que parecem não possuir "foco")[15], mas também formas de valoração. Elas são direcionadas a objetos que são articulados em termos valorativos.

---

[12] FOOT, 1978, p.76.
[13] 1978, p.77-78.
[14] Ver '*Moral Beliefs'* em FOOT, 1978.
[15] Deixaremos de lado a objeção de Griffiths à tese judicativa segundo a qual existem realmente emoções que não possuem "foco" e são similares a coceiras e dores, que a tradição judicativa tem dificuldades para explicar.

Não obstante toda essa representação atraente do que seja sentimento ou emoção, a tese judicativa deixa alguns autores insatisfeitos não apenas porque ela parece requerer (ou estar facilmente associada a) uma metafísica e epistemologia que é independente das preocupações e dos afetos humanos, mas também por seus compromissos *semânticos* ou *conceituais*[16]. De acordo com o judicativista, se algo é perigoso ou não deve ser tomado como um fato independente de nossas reações ou respostas ao mundo. Poderíamos especificar e definir os critérios de periculosidade sem precisar recorrer a qualquer aspecto da dimensão afetiva humana. Isto parece problemático não apenas no que diz respeito à caracterização de uma metafísica e epistemologia valorativas apropriadas (problemas que levaram muitos a evitar o tipo de teorias valorativas Mooreanas), mas também no que diz respeito ao fato de que, com relação a certos conceitos valorativos (como vergonhoso, perigoso, engraçado, feliz, nojento e, possivelmente, mesmo conceitos estritamente morais como certo e errado), parecemos pensar que nossas respostas humanas são decisivas e informativas (isto é, não-triviais) acerca de seu *significado* e *racionalidade*. Parece não fazer muito sentido entender aquilo que é engraçado ou perigoso sem recorrer a reações especificamente humanas de sentimentos de achar graça e medo como respostas a certas situações.

Assim, a tese judicativa na sua versão mais forte não é uma explicação de sentimentos ou emoções dependente das respostas (*response-dependent*) subjetivas humanas. Isto soa implausível a muitos autores. Alguns tentam evitar o problema caracterizando o discurso valorativo como envolvendo conceitos dependentes de respostas subjetivas (*response-dependent*), mas ao mesmo tempo preservando certos aspectos da tese judicativa. Segundo tal proposta, a melhor maneira de compreender e explicar a natureza do valor se dá, de alguma forma, por referência às distintivas respostas subjetivas humanas a certas

[16] Estamos aqui deixando de lado também outras objeções tradicionais à tese judicativa, tais como encontrar espaço para emoções em crianças e animais, já que estes (muito provavelmente) carecem de capacidades conceituais para articular juízos sobre as suas situações. Veja Deigh (1994) para uma discussão deste ponto e objeções à tradição judicativa sobre como esta tenta lidar com emoções em crianças e animais.

situações. Mas isto, sustenta a proposta, não implica que precisemos abandonar a relevância do conteúdo cognitivo e do juízo valorativo totalmente – ou, ao menos, é o que a proposta tenta legitimar.

McDowell e Wiggins tentam fazer isto ao proporem compreender conceitos valorativos em termos de qualidades secundárias. Julgar algo perigoso é pensar que é apropriado temê-la. De acordo com aqueles filósofos, não podemos compreender a periculosidade sem compreendermos nossos sentimentos de medo, isto é, nossas respostas ao perigo. Algo é perigoso em função do sentimento de medo que ele desperta em nós. Ainda assim, ao mesmo tempo não há outra forma de caracterizar o medo, segundo aqueles autores, sem apelar à periculosidade. Como diz Wiggins[17], "este é um subjetivismo de sujeitos e propriedades *mutuamente* ajustados". Wiggins explica o ponto detalhadamente nos seguintes termos:

> Suponha que os objetos que regularmente nos agradam ou nos são úteis ou nos divertem […] ou nos prejudicam ou nos desagradam ou nos importunam […] de várias maneiras passem a ser agrupados por nós sob várias categorias ou classificações às quais damos vários nomes abertamente antropocêntricos; e suponha que passem a ser agrupados assim precisamente *porque* eles são tais que nos agradam, nos são úteis, nos divertem, […] ou nos prejudicam ou nos desagradam ou nos importunam […] nas suas várias maneiras. Desse modo, frequentemente não haverá como falar das propriedades que esses nomes representam independentemente das reações que elas provocam. (O ponto de chamar esta posição de subjetivismo é que as propriedades em questão são explicadas por referência às reações dos sujeitos humanos.) Mas igualmente – ao menos quando o sistema de propriedades e reações se diversifica, se torna complexo e enriquece – não haverá como falar exatamente sobre *qual* reação uma coisa com a propriedade associada provocará sem direta ou indireta alusão à própria propriedade. Achar graça, por exemplo, é uma reação que precisamos caracterizar por referência ao

[17] 2002, p.199.

> seu próprio objeto, através de algo percebido como engraçado (ou incomum ou cômico ou o que quer que seja)[18].

McDowell e Wiggins tentam tornar a sua abordagem ainda mais clara traçando uma analogia com cores. Da mesma forma que no caso, por exemplo, de vermelho, que, segundo a teoria das qualidades secundárias, é aquela propriedade responsável por nossas experiências (subjetivas) de vermelhidão, periculosidade será a propriedade que é responsável por nossas experiências (subjetivas) de medo. Assim, não há como compreender medo sem recorrer a periculosidade. Mas, também, *ex hypothesi*, não há como compreender periculosidade sem referência ao nosso aparato de respostas subjetivas. Como Wiggins diz:

> Em questões desse tipo, uma analogia com cores é sugestiva. 'x é vermelho se e somente se x é tal que fornece, sob certas condições especificáveis como normais, uma certa impressão visual' naturalmente suscita a questão 'que impressão visual?' E essa questão atrai a resposta 'uma impressão de ver algo vermelho', que reintroduz *vermelho*[19].

E McDowell, ao explicar o mesmo ponto, afirma que

> A ideia da experiência de valor envolve a atitude de admiração para representar o seu objeto como possuindo uma propriedade que (embora presente no objeto) é essencialmente subjetiva, de forma muito parecida com a propriedade que um objeto é representado como possuindo por uma experiência de vermelhidão – isto é, compreendida adequadamente apenas em termos da modificação apropriada da sensibilidade humana (ou similar)[20].

Não obstante, McDowell e Wiggins argumentam, abandonando a analogia com uma explicação puramente *disposicional* de qualidades secundárias como cores (de acordo com a qual x é vermelho se e somente se x é *causalmente*

---

[18] 2002, p.195.
[19] 2002, p.189.
[20] McDOWELL, 1998b, p.143.

responsável por gerar a experiência de vermelhidão em um observador adequadamente situado), que o tipo de relação "conceito/propriedade-resposta" no discurso e pensamento valorativo – isto é, relação entre os cenários aos quais respondemos e as nossas respostas – é ineliminavelmente *normativa* ou *valorativa*, e não causal. Esta é a razão por que McDowell e Wiggins sustentam que medo é a resposta digna (*merited*) a periculosidade.

> A ausência de analogia, nesse caso, é que uma virtude, por exemplo, não é concebida meramente como tal que gera uma "atitude" apropriada (como uma cor é meramente tal que causa as experiências apropriadas), mas, em vez disto, é tal que é digna [de ter tal atitude como resposta]. E isto lança dúvidas sobre se explicações meramente causais em experiências de valor são relevantes para o teste explicativo […][21].

E Wiggins torna o ponto ainda mais explícito quando explica a natureza do 'porque' que fundamenta a relação entre um conceito/propriedade/juízo e a resposta:

> Explicaremos isto insistindo que coisas que são *genuinamente* [divertidas, lamentáveis, revoltantes, reconfortantes, tranquilizadoras, nojentas, prazerosas, agradáveis etc.] são coisas que não apenas [divertem, provocam lamento, revoltam, reconfortam, tranquilizam, enojam, dão prazer, nos agradam etc.] mas têm também estes efeitos precisamente porque elas são [divertidas, lamentáveis, revoltantes, reconfortantes, tranquilizadoras, nojentas, prazerosas, agradáveis etc.] – ao mesmo tempo insistindo que este 'porque' introduz uma explicação que tanto explica quanto justifica[22].

Como deve estar claro, a abordagem de McDowell e Wiggins é claramente *circular*, uma vez que apela à dignidade ou mérito de uma resposta (que é um conceito valorativo) como forma de explicar um sentimento, e apela ao mesmo tempo a um conceito ou juízo valorativo como constitutivo do sentimento. Mas os seus defensores sentem-se à vontade com esse tipo de circularidade. McDowell e

[21] MCDOWELL, 1998b, p.143.
[22] WIGGINS 2002, p.199-200.

Wiggins afirmam que, embora a explicação de valor seja circular, ela é informativa por meio do pano de fundo sentimental que ela introduz.

> Circularidade em si mesma não é uma objeção se a formulação incômoda for também *verdadeira*. Mas (perguntarão) que uso tem uma tal formulação circular? Minha resposta é que, ao traçar tal círculo, o subjetivista espera elucidar o conceito de valor exibindo-o em seu real envolvimento com os sentimentos. De acordo com ele [isto é, com o subjetivista], não se teria suficientemente elucidado o que é valor *sem* aquele percurso indireto [detour] [23].

Na verdade, eles chegam ao ponto de afirmar que a melhor maneira de caracterizar as suas abordagens é compreendê-las como *não* pleiteando qualquer *prioridade* de conceitos ou juízos valorativos sobre sentimentos, e vice-versa. Como vimos, para McDowell e Wiggins, é impossível compreender conceitos (e juízos) valorativos sem recorrer às respostas apropriadas que os caracterizam; mas é igualmente impossível compreender um sentimento sem recorrer aos conceitos e juízos valorativos; de tal forma que:

> Ao negar que as características extras têm prioridade sobre os sentimentos relevantes, essa abordagem se distancia da ideia de que elas [as características] pertencem, misteriosamente, a uma realidade que é totalmente independente e superior a nossa subjetividade. Não se segue disto [porém] que os sentimentos tenham uma prioridade. Se não há como compreender os sentimentos corretos independentemente dos conceitos das características extras relevantes, uma tese de não-prioridade é certamente indicada[24].

Esta abordagem tenta tornar a explicação de valores dependente de sentimentos humanos mas, ao mesmo tempo, tenta garantir racionalidade e

---

[23] WIGGINS 2002, p.189. Esta é a razão por que McDowell e Wiggins explicitamente rejeitam chamar a abordagem deles de uma *análise* de valor. Se não é possível ter uma análise que não seja circular, eles se dão por satisfeitos em chamar a sua explicação de valor de um mero "comentário".
[24] MCDOWELL, 1998a, p.159-160.

objetividade em um sentido forte – dado que as nossas respostas apropriadas ao mundo na forma de sentimentos seriam parcialmente cognitivas – um ponto que é compartilhado com a tradição judicativa[25]. Porém, mesmo que suponhamos que as abordagens de McDowell e Wiggins sejam bem sucedidas acerca de uma série de questões problemáticas[26], a circularidade da exposição, tal como surge na tese em questão, exige um fardo excessivamente pesado a alguns filósofos que concordam com a tese da não-prioridade no que diz respeito à ideia de que o discurso valorativo deve ser compreendido por referência a conceitos dependentes de respostas subjetivas (*response-dependent concepts*). Alguns destes tentam resolver a questão garantindo uma proeminência, por assim dizer, à face humana, maior do que na tese da não-prioridade, na explicação de valor. Isto é, em linhas gerais, o que a tese sentimentalista tenta fazer, à qual nos voltamos agora.

Sentimentalismo, tal como entenderemos aqui, é a tese de acordo com a qual há uma prioridade de sentimentos ou emoções sobre conceitos e juízos valorativos. Assim, para o sentimentalista, compreendemos pensamento e discurso valorativo descobrindo, antes, algo sobre nossos sentimentos ou emoções. Sentimentalismo, neste sentido, é uma tese que foi inaugurada (de modo sistemático) por David Hume e Adam Smith. No século XX, ela foi revisitada e defendida por autores como Ayer, Carnap e Stevenson[27]. Da forma que aparece

---

[25] Ver especialmente a defesa deste ponto em McDowell (1998a).

[26] Estas incluem não apenas a necessidade de explicar como circularidade pode ser informativa e não-viciosa, mas também carregam o fardo de propor uma filosofia da mente alternativa, que abandona o modelo tradicional, segundo o qual operações da razão são distintas de operações de sentimentos e possuem diferentes funções em nossa economia mental – o abandono deste modelo parece ser um traço distintivo das propostas de McDowell e Wiggins.

[27] Muitos incluiriam Shaftesbury e Hutcheson entre os fundadores dessa tradição. Não é, porém, inteiramente claro que eles seriam considerados sentimentalistas em nosso sentido aqui. Por vezes, Shaftesbury e Hutcheson falam de emoções ou sentimentos como *percepções* de valor. Se percepções envolvem uma forma de cognição, eles não seriam sentimentalistas em nosso sentido, pois os sentimentos seriam veículos de informação valorativa sobre o mundo. Neste sentido, embora sentimentos pudessem ser indispensáveis para se conhecer valor, eles não desfrutariam do tipo de prioridade que a tese sentimentalista, em nosso sentido, atribui. Assim, isto tornaria Shaftesbury e

nas obras destes filósofos, a tese sentimentalista enfrenta muitos problemas. Estes giram sobretudo em torno de questões relativas à verdade, objetividade e racionalidade do discurso valorativo[28]. Mas não se restringem a estes. O tipo de dificuldade com aquelas versões antigas de sentimentalismo que mais nos interessa aqui diz respeito à análise de valor que elas fornecem. Aquelas versões de sentimentalismo defendiam a ideia de que sentimentos ou emoções desfrutam de uma prioridade sobre juízos valorativos, mas elas não eram claras sobre que ocorrências de sentimentos ou emoções são realmente representativas de valor. Afinal, não devemos supor que qualquer resposta ao mundo ou qualquer ocorrência de sentimento ou emoção seja *adequada* ou *apropriada*. Portanto, a tese sentimentalista precisa de um critério que confira sentido às respostas adequadas sem, ao mesmo tempo, abandonar seu compromisso básico segundo o qual compreendemos conceitos e juízos valorativos compreendendo, antes, nossas respostas sentimentais ao mundo.

Todos esses problemas que pareciam afetar aquelas versões determinaram a agenda de trabalho para novas e mais sofisticadas versões da tese sentimentalista. Dessa forma, autores como Blackbrun e Gibbard têm tentado fornecer respostas àquelas questões e problemas enquanto permanecem fiéis aos compromissos de um projeto sentimentalista em termos gerais. O *Neossentimentalismo*, como chamaremos esta tese, e no sentido que nos interessa aqui, defende, da mesma forma que aquelas versões antigas, a prioridade de sentimentos sobre conceitos e juízos valorativos. Porém, contrariamente a versões antigas de sentimentalismo, a tese neossentimentalista defende (de maneira similar aos defensores da tese da não-prioridade) que pensar

---

Hutcheson mais próximos da tese da não-prioridade, como vimos defendida por McDowell e Wiggins.

[28] Por exemplo, como acomodar a gramática do discurso valorativo de senso comum que parece estar comprometida com o descritivismo, o cognitivismo (e o proposicionalismo) e o realismo? Se o discurso valorativo não incorpora estes compromissos, como podemos explicar um argumento ou apresentar razões para que se adote um ponto de vista no discurso valorativo? Considerações desta natureza tornaram-se famosas sob o nome de 'Frege-Geach problem'. Este é, (in)felizmente, um tópico para uma outra oportunidade.

que algo é, por exemplo, engraçado é pensar que é adequado ou apropriado achar graça em relação a este mesmo algo. Antigas versões da tese sentimentalista pareciam ser menos sofisticadas do que isto na medida em que não pareciam conceder espaço para aquele tipo de atitude ou juízo de ordem superior (que deve, por sua vez, ser também a expressão de um sentimento ou emoção para que a tese neossentimentalista seja inteiramente coerente) que avalia uma reposta (de primeira-ordem) como adequada ou apropriada. Antigas versões de sentimentalismo pareciam entender o tipo de prioridade de sentimentos sobre conceitos e juízos valorativos de uma forma mais direta.

Um ônus deste tipo de neossentimentalismo é que, como a tese da não-prioridade, ele incorpora circularidade em sua caracterização de valor em um nível superior de pensamento, no qual a avaliação de sentimentos e emoções ocorre. Julgar uma resposta *adequada* é fazer uso de um conceito valorativo ou normativo. Há, porém, uma diferença crucial para a tese da não-prioridade. É o fato de que a tese neossentimentalista rejeita a ideia de que sentimentos ou emoções devem apelar a conceitos ou juízos valorativos a fim de serem identificados ou individuados. A tese neossentimentalista tenta fornecer uma caracterização independente de sentimentos, de forma que puderíamos identificar um sentimento de medo, por exemplo, sem precisar recorrer a um conceito valorativo, como 'periculosidade', ou ao juízo de que algo é perigoso. Se o neossentimentalista for capaz de fazer isto, ele pode evitar o problema da circularidade que é uma ameaça à tese da não-prioridade (que a contamina não apenas no nível superior de pensamento, mas também no nível inferior), enquanto ao mesmo tempo evita as implicações problemáticas da tese judicativa. Como o neossentimentalista pode fazer isto?

Gibbard (1990) e o trabalho conjunto recente de D'Arms & Jacobson (1994, 2000a, 2000b, 2006a, 2006b) parecem fornecer a melhor caracterização disto[29]. Recorrendo ao trabalho de psicólogos e à literatura evolucionista recente em

[29] Blackburn (1998) é menos claro acerca deste ponto, embora certamente esteja em larga escala em acordo com aqueles autores (ver cap. 5, especialmente p.126-134).

filosofia da emoção[30], eles tentam fornecer uma explicação independente de sentimentos e emoções, argumentando que estes desempenham um papel adaptativo essencial.

De acordo com Gibbard, é plausível entender sentimentos como *síndromes* que envolvem uma série de respostas complexas, coordenadas ao meio ambiente e automatizadas, cujo *output* é primariamente motivacional[31]. Segundo Gibbard,

> A combinação destes constitui uma síndrome, e a emoção é qualquer estado do organismo que suporta a síndrome. É qualquer estado que possua essas causas típicas, seja expresso naquelas formas, e tenda para tais tipos de ações[32].

Além disso, devemos enfatizar que há fortes evidências a favor de expressões comuns entre culturas (*cross-cultural*) de sentimentos e emoções[33], assim como padrões compartilhados de respostas emocionais[34]. Se isto se confirma, podemos especular como explicação para tal a postulação de um

---

[30] Aqui a obra de Paul Griffiths (1997) ocupa um papel central.

[31] Griffiths diz: a complexidade de um tal sistema "normalmente inclui (a) alterações de expressões faciais, (b) respostas músculo-esqueléticas, tais como esquivar-se e orientar-se, (c) alterações de expressões vocais, (d) alterações no sistema endócrino e consequentes alterações no nível de hormônios, e (e) alterações no sistema nervoso autônomo" (1997, p.77).

[32] 1990, p.132.

[33] A obra de Ekman & Friesen (1971) sobre reconhecimento facial através de conceitos emocionais é uma pesquisa fundamental neste campo. Eles apresentaram evidências empíricas a favor de aspectos comuns entre culturas nas manifestações faciais das seguintes emoções: felicidade, tristeza, raiva, surpresa, nojo e medo.

[34] A acalorada discussão sobre emoções básicas toca neste ponto. Esta é a lista de Ekman (1994; 1999) de emoções básicas: achar graça, raiva, desprezo, contentamento, nojo, embaraço, exaltação, medo, culpa, orgulho em obter algo, alívio, tristeza/angústia, satisfação, prazer sensório e vergonha. E as características seguintes, segundo Ekman, distinguem emoções básicas umas das outras: sinais universais distintivos, fisiologia particular, apreciações (*appraisal*) automáticas associadas a: universais típicos em eventos antecedentes e característico desenvolvimento, presença em outros primatas, rápida irrupção, duração breve, ocorrência espontânea, pensamentos, memórias e imagens característicos, experiência subjetiva peculiar. Para uma discussão e crítica das teses de Ekman, veja os ensaios de Averill, Scherer, e Shweder em Ekman & Davidson (1994).

sistema ou mecanismo que opera por propósitos adaptativos evolucionistas. De fato, Paul Ekman, por exemplo, diz que "emoções evoluíram por seu valor adaptativo ao lidar com tarefas vitais fundamentais"[35]:

> Há várias maneiras de descrever estas tarefas vitais fundamentais. Johnson-Laird & Oatley (1992) dizem que existem obstáculos humanos universais, que envolvem realizações, perdas, frustrações etc. Cada emoção, assim, nos impele em uma direção que, no curso da evolução, foi mais bem sucedida do que outras soluções em circunstâncias recorrentes que são relevantes a objetivos[36].

De modo similar, Tooby & Cosmides afirmam que tarefas vitais envolvem "situações adaptativas, disputas, amar, escapar de predadores, confrontar infidelidade sexual, e assim por diante, cada uma das quais foi recorrente inúmeras vezes na história da evolução"[37]. E Nesse diz que

> […] emoções são sistemas coordenados de respostas que foram moldadas pela seleção natural porque aumentaram a adaptação em certas situações[38].
>
> As emoções são modos especializados de operação moldados pela seleção natural para ajustar os parâmetros fisiológicos, psicológicos e comportamentais do organismo de forma a aumentar a sua capacidade e tendências de responder adaptativamente às ameaças e oportunidades características de tipos específicos de situações[39].

Ekman resume a questão nos seguintes termos:

> A função primária da emoção é mobilizar o organismo para lidar prontamente com importantes encontros interpessoais, preparado para agir assim pelos muitos tipos de atividades que foram adaptativas no passado.

---

[35] 1994 p.15.
[36] EKMAN 1999, p.46.
[37] 1990, p.407-408.
[38] 1990, p.264.
[39] 1990, p.268.

> O passado se refere em parte àquilo que foi adaptativo no passado histórico de nossa espécie, e o passado se refere também àquilo que foi adaptativo em nossa própria história individual de vida[40].

Esta explicação das situações recorrentes precipitadoras de muitos de nossos sentimentos e emoções agora sugere a seguinte descrição de Zajonc. A ideia seria, então, a de que em um típico cenário de medo:

> Um coelho confrontado por uma serpente não tem tempo de considerar todos os atributos perceptíveis da serpente na esperança de que ele possa ser capaz de inferir deles a probabilidade do ataque da serpente, o tempo do ataque ou a sua direção. O coelho não pode parar para contemplar o tamanho das presas da serpente ou a geometria das suas marcas. Se o coelho for capaz de escapar, a ação deve ser realizada muito antes de completar um simples processo cognitivo […]. A decisão de fugir dever ser realizada sob engajamento cognitivo mínimo[41].

Dessa forma, de acordo com Zajonc, se esta é uma representação plausível da operação dos afetos em suas manifestações básicas:

> É uma possibilidade adicional que possamos temer ou gostar de algo antes de sabermos precisamente o que é e talvez mesmo *sem* saber o que é. […] É certo que a reação afetiva inicial é bruta e vaga. Não obstante, é capaz de influenciar o processo cognitivo posterior em um nível significativo[42].

A teoria evolucionista explica o desenvolvimento de todo o mundo natural vivo. Dessa forma, supõe-se que a história que acabamos de ouvir aplique-se ao menos a todos os seres naturais capazes de afeto e cognição. Mas também aprendemos que a evolução atuou nos seres naturais de maneiras diferentes, por razões aleatórias. Em particular, sabemos que as capacidades humanas evoluíram de tal modo que nos tornamos seres altamente complexos, possuindo uma série

---

[40] 1999, p.46.
[41] ZAJONC 1980, p.156
[42] 1980, p. 154. Ver também Kunst-Wilson & Zajonc (1980); Logue, Ophir & Strauss (1981); e Wilson (1979).

de recursos intelectuais e emocionais dos quais outros seres naturais carecem. É, assim, muito provável que existam sentimentos ou emoções que sejam tipicamente humanos. E talvez a melhor explicação para tal seja a alta complexidade que nossas vidas sociais adquiriram. Como diz Gibbard, "em seres humanos as possibilidades são mais amplas"[43], e é provável que "algumas destas possibilidades originais constituam adaptações biológicas específicas às necessidades da vida social"[44]. Ainda mais importante para os nossos propósitos aqui, vale dizer que é provável também que, quando se trata de seres humanos,

> Muito frequentemente essas experiências ocorrem quer se queira ou não. Pode-se ser capaz de controlar a expressão de uma emoção, mas não a própria experiência da mesma[45].

Assim, podemos frequentemente experienciar um sentimento ou emoção indesejados. Mas não apenas isto. Aparentemente, podemos julgar algo *inadequado* – isto é, como uma resposta inadequada ao mundo. Parece ser uma característica comum do fenômeno do valor em seres humanos que podemos, por exemplo, experienciar medo (e continuar experienciando-o irresistível e persistentemente) ainda que se julgue a situação ausente de perigo. Para explicar este fenômeno, parecemos precisar ser capazes de caracterizar sentimentos ou emoções independentemente de juízos de ordem superior sobre a sua adequação.

---

[43] 1990, p.146.

[44] Gibbard parece pensar que sentimentos (provavelmente) tipicamente humanos, como culpa, são o resultado da vida social altamente complexa que seres humanos experienciaram ao longo do curso de sua evolução, dando origem ao discurso moral. Mas Gibbard também aponta que é plausível interpretar culpa como a contraparte de primeira pessoa de raiva (1990, p.146, p.272), e raiva é possivelmente compartilhada com outros animais (sendo direcionada ao comportamento de outros quando quer que alguém seja afetado de forma negativa). A hipótese de Gibbard é que nossas capacidades sentimentais e cognitivas evoluíram – como resultado de (ou mutuamente reforçada por) nossa necessidade de sistemas cooperativos altamente complexos – de tal forma que desenvolvemos um senso de culpa, direcionado a nós mesmos, quando quer que raiva seja adequada em outros. Nesse (1990, p.278) diz algo similar.

[45] ZAJONC 1980, p.156.

Alguns filósofos e psicólogos tentaram explicar este fenômeno dando um passo além no cenário que apresentamos acima, propondo que podem existir dois sistemas operacionais: um de atitudes complexas de ordem superior (envolvendo tanto afeto quanto cognição) e outro de atitudes de ordem inferior, que envolvem respostas mais imediatas ao mundo. Por exemplo, Paul Griffiths e alguns filósofos do valor (simpáticos à tese sentimentalista em nosso sentido) parecem propor que este mecanismo deve operar (ao menos parcialmente) de modo independente das operações de ordem superior de nosso pensamento (isto é, do outro sistema operacional)[46].

Em psicologia especulativa, uma hipótese similar tornou-se famosa ao ser proposta por Jerry Fodor em *The Modularity of Mind* (1983). Segundo Fodor, muitos de nossos sistemas de reações automáticas (*reflex*) possuem informação encapsulada (*informationally encapsulated*). Sobre nosso sistema motor visual, Fodor diz que:

> […] existe a amplamente notória persistência de muitas ilusões perceptuais (por exemplo, [os casos da] sala de *Ames*, do fenômeno *Phi*, da ilusão visual Müller-Lyer; os efeitos na fala da *phoneme restoration* e *click displacement*) mesmo contrariamente ao explícito conhecimento do sujeito de que a percepção é ilusória. O mesmo indivíduo que é capaz de dizer que as flechas da Müller-Lyer são idênticas em comprimento, que até mesmo as mediu, continua vendo uma maior do que a outra. Nesses casos, é difícil encontrar uma alternativa melhor à tese de que ao menos alguma informação do pano de fundo disponível ao sujeito seja inacessível ao menos a alguns de seus mecanismos perceptuais [47].

[46] Ver D'Arms & Jacobson (2003, 2006a, 2006b) e D'Arms (2005) para uma sugestão de endosso dessa proposta.

[47] 1983, p.66. Considere também o exemplo em que Fodor nos convida a supor que "você sabe perfeitamente bem que sob nenhuma condição eu enfiaria meu dedo em seu olho. Suponha que esta crença sua é tanto explícita quanto profundamente arraigada. […] Ainda assim, se eu empurrar meu dedo bastante próximo aos seus olhos, e rápido o bastante, você irá piscar. Dizer, como dissemos acima, que o reflexo do piscar é inevitável é dizer, *inter alia*, que ele não tem acesso àquilo que você

O ponto agora é sugerir que, da mesma forma que Fodor propõe que parte de nosso sistema visual é encapsulado, automatizado e opera independentemente de atividades de ordem superior de nossa mente (como pensamento, reflexão e uso de linguagem), poderíamos propor que sentimentos e emoções em suas manifestações mais básicas poderiam ser caracterizados de forma apropriada em termos similares. Assim, poderíamos dizer que respostas sentimentais ou emocionais ao meio ambiente são operadas por um mecanismo específico (adaptativo do ponto de vista evolutivo) que atuam independentemente de atividades de ordem superior e são encapsuladas e automatizadas no sentido de que suas respostas são rápidas e fazem uso de informação que não está (necessariamente) disponível a outras atividades cognitivas e afetivas[48].

Se esta proposta evolucionista e em psicologia especulativa funcionar para sentimentos e emoções, então poderemos ser capazes de compreendê-los independentemente de juízos valorativos de ordem superior (que pertenceriam mais propriamente a nossas atividades mentais de ordem superior). Com isto em mãos, o neossentimentalista aparentemente possui agora a teoria psicológica necessária para dar sentido à nossa proposta de análise inicial. Ele pode afirmar que juízos valorativos de ordem superior acerca da adequação de respostas sentimentais ou emocionais são possíveis na medida em que podemos caracterizar independentemente sentimentos e emoções.

Um último ponto deve ser mencionado sobre tais juízos de ordem superior

---

sabe sobre o meu caráter ou, no que toca a questão, a qualquer de suas crenças, utilidades e expectativas. Por esta razão o reflexo do piscar é frequentemente produzido mesmo quando reflexão sóbria o consideraria desnecessário…" (FODOR, 1983, p.71).

[48] Se levarmos esta proposta ao extremo, poderíamos formular a hipótese, nos termos da psicologia especulativa, de atribuir um módulo para respostas emocionais (ver FODOR, 1983, p.71). Mas, como veremos posteriormente, temos razões para resistir a essa analogia se esta propuser traçar uma estrita distinção entre emoções como respostas encapsuladas e atividades de ordem superior como não-emocionais ou exclusivamente racionais. Como Blackburn sugere, não há necessidade de compreender aquela distinção em termos de uma forma de dualismo. Podemos explicar as atividades de ordem superior em questão como uma forma de "ascensão emocional" (ver BLACKBURN, 1998, p.129). Retornaremos a este ponto na seção 5.

no seio de um programa neossentimentalista. A forma que Gibbard propõe para explicar tais juízos é chamá-los de juízos sobre a *racionalidade* de um sentimento ou emoção, com o que Gibbard quer dizer que *faz sentido* sentir de uma certa maneira em resposta a uma dada situação (ver 1990, especialmente partes I e II). Tal juízo de ordem superior é, nos termos de Gibbard, *expressivista*. Ele é a expressão de um estado mental. Mais especificamente, ele é a expressão de uma rica e complexa atitude não-cognitiva direcionada a uma certa situação. (Voltaremos a este ponto abaixo.) Desse modo, ele não é descritivo de coisa alguma. Segundo Gibbard, fornecemos o significado de um termo valorativo explicando o que fazemos quando usamos tal termo: expressamos uma atitude direcionada a um sentimento ou emoção relacionados a uma certa situação no mundo.

Esta tese é uma forma de *projetivismo*. Ela faz eco a Hume quando notoriamente disse que "pintamos os objetos naturais com cores extraídas de um sentimento interno" e que, dessa forma, "produzimos uma nova criação"[49]. Assim, nos limites de nosso escopo aqui, basta dizer que, para o neossentimentalista, ao menos alguns de nossos juízos valorativos são expressões de uma atitude não-cognitiva através da qual projetamos valor sobre um mundo desprovido de valor. É vital para esta forma de projetivismo neossentimentalista que emoções e sentimentos *não* sejam cognições. Se essa forma de projetivismo pretende explicar valor por referência a sentimentos, é preciso que não explique valor anteriormente à sua explicação de sentimento ou emoção.

2. Sentimento e valor: duas exigências gerais

Toda tese que pretende fornecer uma explicação da relação entre valor e sentimento precisa ser capaz de satisfazer ao menos duas exigências gerais. Ela deve ser capaz (1) de fornecer uma indicação de que duas partes disputantes em um debate sejam capazes de falar sobre a mesma coisa de tal forma que seja possível discordar sobre os padrões para julgar algo valoroso; e (2) de explicar

[49] 1751, Ap. I.

como o sentimento ou emoção (especialmente recalcitrante ou irresistível) pode estar em erro (em nossa terminologia, ser inadequada) mesmo diante de um juízo contrário[50]. Chamaremos a primeira exigência de 'problema da univocidade' (ver Wiggins 2002; e D'Arms & Jacobson 2003) e a segunda de 'problema da recalcitrância' (ver D'Arms & Jacobson 2003).

Mais precisamente, poderíamos dizer, sobre o problema da univocidade, que qualquer tese que se proponha a explicar a relação entre sentimento e valor deve ser capaz de explicar como indivíduos podem discordar sobre o que julgam, por exemplo, vergonhoso. Se eu julgo X vergonhoso e penso que vergonha é a resposta adequada a X, enquanto que você julga que X não é vergonhoso e diz que vergonha não é a resposta adequada a X, nós estamos discordando sobre X ser vergonhoso. Ora, para explicar este fenômeno de discordância, parecemos precisar ser capazes de especificar algo ou sobre o conceito (e juízo) valorativo ou sobre o sentimento acerca do qual concordamos e compartilhamos (conceitualmente) como uma característica comum da disputa. Devemos ser capazes ou de identificar independentemente o sentimento de vergonha e avaliar se a situação com a qual nos deparamos é tal que vergonha é uma resposta adequada, ou identificar o que é vergonhoso apontando para características objetivas (naturais ou não, dependendo da análise proposta de 'vergonhoso') sobre as quais devemos concordar que são constitutivas de vergonhoso – estas sendo possivelmente independentes de sentimentos. Parece que é apenas nestes sentidos que podemos garantir que estamos falando sobre a mesma coisa quando nos perguntamos se uma situação é vergonhosa ou não. Precisaríamos concordar ou sobre o que é objetivamente vergonhoso ou sobre o que é sentir vergonha.

Com relação ao problema da recalcitrância, poderíamos dizer, em primeiro lugar, que um dos pontos centrais do discurso valorativo parece se fundar na possibilidade de que sejamos capazes de avaliar nossas respostas psicológicas à luz do que é julgado adequado (com base nos relevantes conceitos valorativos) em

[50] Na verdade, uma tese relativista sobre juízos de valor não precisaria responder a (1). Porém, estamos pressupondo aqui que todas as teses discutidas não são relativistas.

uma dada situação. Assim, se julgamos que uma determinada situação é vergonhosa, estamos julgando ao menos que, se há algum tipo de reação que é adequada, esta é sentir vergonha; nenhuma outra resposta serviria. Desse modo, a fim de tornar possível explicar o discurso valorativo sobre a aplicabilidade de tais conceitos em juízos, precisamos ser capazes de distinguir casos em que se responde adequadamente aos juízos de casos em que isto não ocorre.

Parece ser um fato empírico trivial sobre o agir humano e sua motivação que, às vezes, sofremos de certas respostas recalcitrantes ou irresistíveis a uma dada situação enquanto a julgamos *inadequada*. Estamos cientes da resposta inadequada e, ainda assim, somos, por vezes, incapazes de nos livrar dela. Estes são casos em que somos capazes de identificar um tipo de "erro" (ou inadequação) em nossa economia mental diante de nossas valorações. Assim, parece que qualquer tese que pretenda explicar a relação entre sentimento e valor, se for bem sucedida como uma explicação do discurso valorativo, deve ser capaz de acomodar tais "erros" como possíveis. Eles são, por assim dizer, a expressão de ausência de integração entre juízos e uma reposta (motivacional) emocional.

Devemos agora tomar essas duas exigências como guia para a nossa subsequente discussão sobre a relação entre sentimento e valor. Como tentaremos argumentar, a tese neossentimentalista parece estar mais apta a acomodar aquelas exigências do que as teses judicativa e da não-prioridade.

3. Problemas com as teses judicativa e da não-prioridade

Comecemos com a tese da não-prioridade pois há algumas complicações e adições a serem feitas à tese judicativa que irão exigir mais tempo e detalhes em nossa discussão.

O que, então, a tese da não-prioridade tem a dizer sobre o problema da univocidade? Basicamente, uma vez que a tese da não-prioridade defende que há uma interdependência entre sentimentos e conceitos (ou propriedades) valorativos, de forma que não poderíamos compreender o primeiro sem recorrer ao segundo, ela enfrenta dificuldades em explicar a discordância, porque não parece ser capaz de identificar um ponto comum ou compartilhado de comparação entre partes

disputantes em um debate valorativo. Dado que não podemos compreender sentimento sem recorrer a conceitos valorativos, e não podemos explicar conceitos valorativos sem compreender sentimentos, não possuímos um critério independente para fixar ou sentimentos ou as propriedades objetivas (que justificam a aplicabilidade dos conceitos) como um ponto de vista comum sobre o qual uma disputa poderia estar fundada.

Se a tese da não-prioridade dispusesse de um critério independente de sentimento, ela poderia explicar a discordância alegando que as partes disputantes estão aplicando diferentes padrões de avaliação com respeito à adequação de certa resposta sentimental a uma situação no mundo. Mas, dado que a tese da não-prioridade parece tornar isto impossível, quando quer que haja uma disputa entre partes, esta será provavelmente apenas aparente e elas não estarão genuinamente discordando. Não há garantia de que haja uma característica comum sendo compartilhada na disputa. Para a tese da não-prioridade, quando apelo a um sentimento eu necessariamente também apelo a certas características no mundo que são valorativas, de modo que, se tivermos padrões distintos de aplicação de nossos conceitos, parece que não nos restaria outra opção senão estarmos fadados a alguma forma de solipsismo valorativo. Não parece ser possível que eu discorde de sua resposta sentimental, pois ter tal sentimento poderia envolver padrões (valorativos ou normativos) completamente diferentes dos meus.

D'Arms coloca o problema nos seguintes termos:

> Wiggins insiste, e McDowell parece concordar, que as respostas às quais a tese sentimentalista apela não podem ser compreendidas como conceitualmente primárias em relação aos conceitos ou propriedades valorativos que tais respostas são invocadas para explicar. Propriedade e resposta são parceiros comuns. Assim, quando eles levantam a questão sobre quais são os sentimentos em questão, cada um daqueles autores sugere que qualquer resposta apropriada deve apelar às propriedades às quais os sentimentos são respostas. […] Isto levanta questões sobre circularidade, obviamente. Mas, mesmo que se conceda que elucidações

> circulares são, às vezes, informativas, identificar os sentimentos apelando às propriedades às quais são respostas apropriadas abandonaria quaisquer vantagens que prometessem assegurar univocidade juntamente com contestabilidade genuína. Se as partes em uma dada disputa possuem diferentes pontos de vista sobre a extensão da propriedade e diferentes opiniões sobre que características fazem algo possuir tal propriedade, então a preocupação inicial sobre univocidade é apenas uma preocupação sobre se estão falando sobre a mesma propriedade. Assim, se o sentimento que cada um sente pode ser individuado apenas apelando-se à propriedade à qual cada indivíduo está respondendo, então a alegação de que estejam falando sobre a adequação do mesmo sentimento compartilhado não é mais confiável do que era a alegação de que estejam falando sobre a mesma propriedade (ou empregando o mesmo conceito) [51].

Agora, por outro lado, pode a tese da não-prioridade acomodar a exigência relacionada ao problema da recalcitrância? Não é claro, por razões similares. Por um lado, o que motiva a tese da não-prioridade é exatamente a possibilidade de erro e contestação, como vimos: ela se propõe a mostrar que uma explicação puramente *causal* da relação entre as propriedades responsáveis por uma certa resposta subjetiva e a própria resposta não é satisfatória porque queremos ser capazes de criticar certas respostas como inadequadas. E, de acordo com o defensor da tese da não-prioridade, podemos fazer isto apenas através de uma perspectiva já valorativa – e não de uma perspectiva meramente causal. Esta é a

[51] 2005, p.14. Ver também D'Arms & Jacobson (2006b), onde consideram a possibilidade de o defensor da tese da não-prioridade dizer "que o que torna possível que estejamos discordando um do outro, e não falando sobre coisas distintas, é um sentimento compartilhado de desaprovação moral cuja adequação está em questão. Ele exigiria, assim, uma explicação de desaprovação moral que assegurasse que ambos disputantes tenham um sentimento específico em seu repertório emocional, ou ao menos compreendam-no bastante bem para discutir sobre sua adequação. Porém, a sua explicação de sentimentos parece tornar impossível a satisfação desta exigência. Se desaprovação moral tem que ser individuada por referência ao conceito *errado*, então o sentimento irá herdar a ambiguidade do conceito. Dessa forma, a alegação de que nossos antagonistas estão disputando sobre a adequação do mesmo sentimento não pode ser mais confiável do que a alegação de que estão aplicando o mesmo conceito" (p.207).

razão por que a tese introduz uma caracterização irredutivelmente normativa ou valorativa da resposta subjetiva (como meritória ou adequada).

Porém, pode a tese da não-prioridade realmente explicar erro em tais bases? Aparentemente, ela precisaria dizer que aqueles que cometem um erro de tal natureza não estão respondendo adequadamente ao mundo. Mas, quando tentamos especificar e tornar isto claro, encontramos dificuldade em explicar como uma resposta subjetiva de medo, por exemplo, pode ser tomada como inadequada segundo tal tese. O problema para a tese da não-prioridade é que parece precisar sustentar que aqueles que respondem com medo a uma situação ausente de perigo estão ou julgando mal ou estão simplesmente respondendo causalmente a um estímulo. No último caso, parece óbvio que o defensor da tese da não-prioridade teria que dizer que a resposta em questão não é, de forma alguma, medo. Ela não seria um candidato possível para a aplicação do conceito de medo na medida em que se trata de uma resposta meramente causal. Medo, para a tese em questão, é uma resposta irredutivelmente valorativa ou normativa, de forma que uma explicação causal de tal resposta fracassaria em satisfazer o seu critério para individuação de medo. Mas, talvez, a primeira resposta considerada acima poderia funcionar. Como?

Está longe de ser claro. Se medo é necessariamente uma resposta valorativa ou normativa a uma propriedade, pois ela necessariamente envolve o emprego de um conceito valorativo, então quando alguém responde inadequadamente com medo, parece que precisaria estar cometendo um tipo de erro de *juízo* – por exemplo, julgando algo perigoso quando não é o caso. Mas é inteiramente obscuro como a tese da não-prioridade poderia acomodar evidência equivocada ou algo como a intervenção de um sentimento ou emoção irracional – que são os melhores candidatos para explicar erros de juízo. Uma vez que, para a tese da não-prioridade, evidência e aplicação de conceitos valorativos devem estar inextricavelmente associadas com sentimento, não poderia haver sentimento equivocado diante de evidência correta e aplicação de conceito, nem, conversamente, sentimento correto diante de evidência e aplicação de conceito

equivocados[52]. Assim, parece que, quando quer que alguém sinta medo genuíno, deve necessariamente estar respondendo a periculosidade. Do contrário, não seria medo, mas alguma outra atitude[53]. Parece que seríamos capazes de falar sobre estar em erro em reação a periculosidade apenas se pudéssemos fixar independentemente ou as condições para periculosidade ou para medo. Lamentavelmente, isto é exatamente o que a tese da não-prioridade considera impossível.

É verdade que o defensor da tese da não-prioridade poderia ser capaz de dizer que alguém que não sente medo diante de uma situação perigosa está em erro por sentir, por exemplo, coragem. Neste caso, o agente estaria respondendo de modo totalmente inadequado: isto é, estaria respondendo inadequadamente tanto da perspectiva afetiva quanto da perspectiva cognitiva. Mas, ainda que a tese em questão possa ser capaz de acomodar este tipo de caso, ela não parece ser capaz de ter uma resposta para aquele outro tipo que vimos, em que alguém sente medo inadequadamente mesmo quando (sabe que) a situação não é perigosa. O que quer que alguém *inadequadamente* sinta quando medo é a resposta adequada *não* pode ser medo, dados os compromissos da tese da não-prioridade. Mas isto soa agora bastante insatisfatório. Parece óbvio que queremos garantir a possibilidade de dizer que, em certos casos, indivíduos sentem medo quando não deveriam.

E o que dizer agora da tese judicativa? Sai-se melhor do que a tese da não-prioridade? Ela parece, de fato, ser mais bem sucedida no que diz respeito ao problema da univocidade. Mas parece fracassar de modo similar no problema da recalcitrância. Vejamos por quê.

Uma vez que para a tese judicativa existem critérios objetivos,

---

[52] O mesmo ponto poderia ser colocado da seguinte maneira: dado que a tese da não-prioridade parece compreender sentimentos como atitudes que são ao mesmo tempo afetivas e cognitivas, um erro de juízo pareceria implicar igualmente um erro de afeto.

[53] Isto pode, na verdade, ser um ônus que McDowell esteja disposto a assumir. Afinal, a sua teoria aristotélica da pessoa virtuosa parece endossar o ponto, já que a pessoa virtuosa é necessariamente não-defeituosa (ou equivocada) em suas respostas ao mundo (ver 1998c).

independentes de nossas reações, para se formular juízos valorativos, e dado que a tese caracteriza sentimentos ou emoções como constituídos ou dependentes de tais juízos, parece ser possível explicar discordância entre partes disputantes de uma maneira clara. Quando quer que dois disputantes estejam discordando sobre os critérios para julgar algo perigoso, por exemplo, eles parecem estar discordando sobre o critério de aplicabilidade do conceito em questão. Se eu digo "aquele leopardo é dócil" e você diz "não, é um grande felino e grandes felinos são ferozes", podemos estar em discordância sobre se a situação perante nós é de perigo, assim como sobre a aplicabilidade de 'perigoso'. Da mesma forma que no caso de crenças, estamos apontando para algo no mundo que torna nossas crenças verdadeiras. E, no caso em questão, parecemos estar simplesmente apontando para coisas distintas[54].

É certo que alguns diriam que, em um tal caso, nossa disputa seria melhor explicada não sobre como aplicar 'perigoso' mas sobre o próprio *significado* de 'perigoso'. Se isto fosse o caso, não seria claro que haveria discordância entre as partes no debate. De fato, se o judicativista entendesse 'perigoso' como uma propriedade natural, de modo que a análise do termo seria possível em termos naturais, poderíamos conjecturar que não haveria garantia de que as partes disputantes não estivessem falando sobre coisas distintas, aplicando significados distintos aos mesmos termos em questão. Poderia ser uma mera disputa verbal. Mas, uma vez que o judicativista pode tanto propor uma análise de 'perigoso' que não seja redutiva naquele sentido quanto simplesmente rejeitar a ideia de que 'perigoso' possa ser analisado de forma alguma, concederemos aqui à tese judicativa que ela possui uma resposta para o problema da univocidade. O reducionismo analítico, ou qualquer outra forma de realismo que apele a propriedades objetivas não-analisáveis e independentes de respostas subjetivas, podem ser teses contestáveis por uma série de outras razões, mas talvez não pelo

[54] Uma questão distinta seria perguntar quem está representando corretamente o mundo. Mas que esta é uma outra questão, independente de explicar discordância, fica claro com o fato de que pode haver discordância entre duas partes que julgam, ambas, falsamente.

problema da univocidade[55].

Mas o que dizer da tese judicativa acerca do problema da recalcitrância, isto é, sobre explicar como sentimentos (especialmente recalcitrantes) podem estar em erro diante de juízos contrários? À primeira vista, o judicativista poderia dizer que, uma vez que emoções são estados do tipo de juízos valorativos, pode-se incorrer em tal erro quando quer que se cometa um erro de juízo. Aqui, mais uma vez, os melhores candidatos para explicar este fenômeno parecem ser respostas inadequadas à evidência ou a intervenção de um sentimento irracional. Mas, então, como o defensor da tese judicativa poderia dizer, sobre determinado caso, que sentir medo é algo inadequado porque a situação *não* é perigosa?

Não é claro que o judicativista possa dizer isto. De forma similar ao defensor da tese da não-prioridade, o judicativista não possui um critério para caracterizar sentimentos independentemente de juízos valorativos. Se alguém sente medo com relação a X, deve estar julgando que X é perigoso. Este juízo poderia, claro, estar errado – no sentido de responder mal à evidência disponível. Porém, o problema agora para o judicativista é explicar como alguém abandonaria uma tal resposta quando reconhecesse que o juízo está errado ou equivocado. Parece que o judicativista precisa dizer que nossa racionalidade imediata e naturalmente corrigiria o juízo equivocado. Devemos aqui lembrar da passagem de Nussbaum, citada anteriormente, na qual diz que, ao sentir raiva, "se eu descobrir que a minha crença é falsa – que o aparente mal não ocorreu de fato – irei, abandonando a minha crença falsa, deixar de ficar com raiva"[56]. Isto pode ser verdadeiro até certo ponto. Mas devemos também notar que Nussbaum imediatamente acrescenta àquela passagem que, em um tal caso, após o trabalho intelectual ter sido feito por nossas capacidades cognitivas, "se algum resíduo de

---

[55] Dessa forma, deixamos de lado a objeção inspirada em Hare, segundo a qual, se a tese judicativa apela em última instância para a analiticidade a fim de explicar como empregar conceitos valorativos, isto poderia significar que, afinal, ela não conseguiria explicar discordância (ver HARE, 1952). Uma vez que a tese judicativa não depende do reducionismo analítico, concedemos que a tese pode explicar discordância, pelas razões já fornecidas.

[56] NUSSBAUM, 1990, p.292.

sentimento doloroso permanece, ele não será mais considerado raiva mas, em vez disto, um resíduo de irritação ou excitação irracional"[57]. Mas é exatamente aqui que está o problema. O que dizer sobre ocorrências recalcitrantes de sentimentos ou emoções, apesar de juízos contrários?

A tese judicativa poderia sugerir que, nesses casos, nós simplesmente nos deparamos com óbvia incoerência ou contradição entre crenças: algo como julgar ao mesmo tempo que uma situação é perigosa e não é perigosa. Esta poderia efetivamente ser a descrição correta de alguns casos. Mas não parece ser aplicável a todos. Dizer que estamos respondendo contraditória ou incoerentemente à evidência parece ser uma representação incorreta de vários outros casos. Parece que precisamos também explicar aqueles casos triviais em que temos uma resposta persistente ao mundo (na forma de um sentimento ou emoção) ainda que estejamos convencidos de que é uma resposta inadequada.

Assim, mais precisamente, o problema para o judicativista em tais casos é que teria aparentemente que dizer que julga (ou acredita) que a situação é perigosa (com base em razões e evidência) e também julga que a situação não é perigosa (com base em razões e evidência). O problema aqui não é exatamente sobre justificação, mas sobre *explicação*. Todos estamos dispostos a reconhecer que podemos formular um juízo equivocado e, tão logo tenhamos nos dado conta disto, abandonarmos o juízo. (Assim, um dos dois juízos considerados no caso acima deveria desaparecer.) Porém, ao menos a experiência de sentimentos ou emoções é tipicamente um fenômeno que, às vezes, resiste à justificação racional. Como já mencionamos, um dos melhores candidatos para explicar a irracionalidade ou uma má resposta à evidência é exatamente a intervenção (indesejada) de sentimentos ou emoções. Estes podem, por vezes, nos deixar, por assim dizer, "cegos" com relação à evidência. Mas, lamentavelmente, esta é uma resposta que não parece estar disponível para o defensor da tese judicativa, pois, para ele, a ocorrência de um sentimento ou emoção já deve necessariamente envolver juízo. Este *não* é, portanto, insensível à racionalidade e evidência. Como resultado, o

[57] 1990, p.292.

judicativista precisa explicar como uma má resposta à evidência pode persistir à luz do reconhecimento de seu erro ou equívoco.

Esta é justamente a razão por que alguns defensores (ou simpatizantes) da tese judicativa propuseram um refinamento da tese, de forma a garantir espaço para a possibilidade de sentimentos recalcitrantes apesar de evidência para o contrário. Greenspan, por exemplo, esforça-se para mostrar que a tese judicativa (ou uma versão refinada desta) pode incorporar outros estados mentais (que são do tipo de crenças e juízos, mas não são nem crenças, nem juízos genuínos) na explicação da individuação e caracterização de sentimentos. É apenas ao introduzir esta qualificação que uma tese judicativa poderia explicar a ocorrência de sentimentos ou emoções a despeito de juízos ou crenças contrários. Greenspan diz que:

> […] A tese judicativa [em sua versão mais forte], sustento, não faz justiça à diversidade dos fenômenos mentais. Ela também sugere uma resposta excessivamente simples a questões de justificação, sobre as razões para emoções e seu papel como razões para agir. Tendo crenças como o componente "intencional" da emoção – o componente que é *sobre* algo e, portanto, capaz de representar erroneamente o seu objeto – a "justificação emocional" pareceria ser essencialmente uma justificação feita por e para crenças[58].

Seguindo, então, a sugestão nos termos de Greenspan, o judicativista estaria melhor se sugerisse que sentimentos ou emoções envolvem algo como *pensamento* valorativo em vez de juízo ou crença genuínos – assentimento está presente nos últimos, mas não no primeiro. Dessa forma, o judicativista poderia talvez explicar recalcitrância sugerindo que, embora acreditando ou julgando que X não representa perigo, um indivíduo poderia, apesar disto, *pensar* (valorativamente) que X é perigoso – e, assim, sentir medo com relação a X.

Portanto, como resposta àquele tipo de desafio à tese judicativa, Greenspan propõe introduzir uma forma de tese cognitiva sobre a relação entre sentimento e

[58] 1988, p.3.

valor que compartilha com a tese judicativa seu princípio fundamental – nomeadamente, que sentimentos envolvem conceitos valorativos – enquanto que tenta garantir espaço conceitual para atitudes cognitivas diferentes de juízos e crenças, a fim de explicar a possibilidade de erro de juízo na forma de inadequação sentimental[59]. Chamemos esta atitude diferente de crença e juízo de *pensamento*. Assim, a ideia seria que um *pensamento* emocional em relação a X poderia ocorrer apesar de um juízo ou crença emocional contrário. Ora, se há algo como cognição ao se sustentar um certo pensamento valorativo, seria aparentemente possível para o judicativista propor que um indivíduo pode estar errado em pensar sobre uma dada situação que ela é perigosa (e aplicar conceitos valorativos) ao mesmo tempo que julga que a situação não é perigosa.

Mas esta resposta ao desafio proposto é mesmo satisfatória? D'Arms & Jacobson pensam que não, pelas seguintes razões:

> O quasijudicativista supõe que o que está acontecendo em casos de recalcitrância estável é que um indivíduo está sujeito a pensamentos recorrentes que entram em conflito com os juízos que sinceramente endossa, dos quais estrai inferências e usa em decisões práticas. Isto poderia ser frequentemente verdadeiro, mas é importante perguntar *simplesmente por que a recalcitrância ocorre*[60].

E eles concluem que

> Para o quasijudicativista, o fenômeno da recalcitrância é mais misterioso e recalcitrância estável é especialmente problemática. Por que os supostos pensamentos constitutivos de certas emoções se reafirmam continuamente apesar de conflito com juízos refletidos? Afinal, as pessoas não são geralmente vulneráveis a pensamentos recalcitrantes contrapondo-se a seus juízos bem estabelecidos. Estamos preparados para admitir uma crença recalcitrante não-estável ocasional como possível, assim como

[59] D'Arms & Jacobson (2003) chamam esta versão modificada da tese judicativa de 'quasijudicativismo'.
[60] 2003, p.149, ênfase adicionada.

> percepção recalcitrante estável. [… Mas] tendo em vista que a tradição judicativa está comprometida com definir valorações características da experiência emocional em termos de conceitos disponíveis independentemente, ela é forçada a tratar conflitos entre as emoções de um agente e seus juízos como exercícios conflitantes de pensamento conceitual. Assim, ela também torna indisponível a analogia com a experiência perceptual recalcitrante, tais como ilusões óticas e auditivas. Isto torna a recalcitrância emocional um tipo estranho de fato bruto sobre certos conceitos: que tendemos a ter pensamentos recalcitrantes, influenciados por afetos, que faz uso destes conceitos[61].

D'Arms & Jacobson estão dizendo que, uma vez que a tese judicativa (incluindo a versão refinada de Greenspan) terá que interpretar ocorrências de sentimentos como exercícios de pensamento conceitual – isto é, como o emprego de repertório conceitual valorativo e provavelmente o uso racional de atividades de ordem superior que envolvem capacidades similares a de formular juízos – não é claro como explicar o fenômeno em questão. Em outros termos, já que o pensamento emocional terá que ser interpretado como necessariamente um exercício de um engajamento *racional* (como resultado de um processo de conceitualização), é obscuro como isto poderia envolver qualquer erro, especialmente erro recalcitrante. Afinal, por que continuaríamos aplicando o conceito 'perigoso' a uma dada situação que *não* é perigosa?[62]

Para tornar o ponto mais preciso, será útil trazer um comentário de McDowell sobre a estranheza de se adotar um modelo estritamente perceptual para sentimentos ou emoções. Ele diz que:

> […] a percepção, estritamente falando, não simula o papel da razão no

[61] 2003, p.141-142.

[62] Ao levantar uma objeção similar à tese judicativa, Deigh (1994, p.851-852) diz que "algumas emoções podem nunca se tornar completamente sensíveis à razão, pois as susceptibilidades a elas em certas circunstâncias podem ser tão fixas que elas não se alteram mesmo à medida que se aprende a falar e se desenvolve um sistema de crenças. Estas emoções são, nesse sentido, não-educáveis".

> pensamento valorativo, que parece exigir que consideremos a apreensão de valor como uma questão intelectual em vez de meramente sensorial[63].

Dado que a introdução de uma atitude (emocional) como "pensamento" (similar a crenças ou juízos genuínos) parece ter o objetivo, de certo modo, de simular um modelo perceptual de valoração (ou ao menos tornar possível a analogia de percepção com valoração) – de modo que a tese judicativa pudesse sustentar que uma atitude como "pensamento emocional" poderia operar de maneira similar à percepção, no sentido que provocaria em nós certas respostas mesmo que estas fossem sistematicamente falsas – a tese judicativa parece agora enfrentar o seguinte dilema. Ela precisa interpretar tais exercícios ou como engajamentos intelectuais ou racionais – em cujo caso a analogia perceptual não funcionaria, e não é claro que o problema da recalcitrância seria solucionado satisfatoriamente – ou interpretar tais exercícios de pensamento emocional como uma questão robustamente causal e *quasi*-sensorial – em cujo caso não teríamos uma atitude similar ao juízo ou crença.

Dessa forma, poderíamos dizer sobre a resposta de Greenspan que parece funcionar apenas à medida que supomos que o conflito entre juízo valorativo e pensamento emocional é momentâneo – ou seja, a proposta não parece avançar em relação a uma resposta já disponível à tese judicativa. Se essa tendência se torna estável, então parece ser difícil acomodar o que Greenspan diz em defesa de sua versão refinada da tese judicativa. De fato, Greenspan diz, sobre o caso de Fido (um inofensivo cão) e um agente que tem medo deste, que:

> O medo do agente de Fido, neste caso, parece ser capturado pela ideia de que "sente-se como se" Fido fosse machucá-lo – como evidenciado por sua tendência de entreter aquele pensamento na presença de Fido [apesar de duvidar deste]. Insistir, porém, que ele *acredita* no pensamento quando o aplica a si mesmo, sem apelar às características distintivas de seu próprio caso, parece a mim ser um último recurso do ponto de vista da explicação. Incoerência lógica é possível, obviamente; mas estou pressupondo que o

[63] 1998b, p.132.

> agente está funcionando de modo bastante racional em geral, de modo que a nossa atribuição de crenças a ele deve ser governada pelo princípio da "caridade lógica". Precisamos de um tipo especial de razão [...] para atribuir-lhe um juízo não-reconhecido em conflito com um juízo que ele reconhece. Em vez de supor que suas crenças entram em conflito momentâneo toda vez que Fido está perto, é mais simples e preferível do ponto de vista de uma explicação racional supor que este é um caso em que a emoção se separa do juízo. Ele exibe a tendência das emoções, em contraste com as crenças racionais de um agente, a se prolongar e fixar em objetos que se assemelham aos seus objetos apropriados de maneira acidental[64].

O problema, então, para Greenspan, é que a sua tese parece, por assim dizer, "racionalizar demais". Ao fazer da atitude de pensamento emocional algo que envolve uma atitude similar a de julgar, "como se" o mundo fosse de uma determinada maneira, parece disponibilizar tudo o que precisamos para dizer que, de fato, é inteiramente *racional* sentir-se de tal maneira. Interpretar uma atitude como "como se" rastreasse ou representasse adequadamente o mundo parece ser suficiente para atribuir racionalidade a ela. Mas isto é exatamente o que queríamos evitar: dizer que a atitude em questão é racional em qualquer sentido.

Assim, para explicar a recalcitrância de um sentimento ou emoção, parece que precisamos de uma maneira de caracterizá-la que seja independente de racionalização. Isto é exatamente o que a tese neossentimentalista tenta fazer.

4. Como a tese neossentimentalista pode fazer melhor?

A tese neossentimentalista adota a mesma estratégia geral de resposta aos problemas da univocidade e recalcitrância. Esta estratégia consiste em propor uma caracterização de sentimentos ou emoções independente de conceitos e juízos valorativos, como mencionamos acima. Se esta proposta for bem sucedida, ela significa não apenas um primeiro passo para explicar sentimentos ou emoções

[64] GREENSPAN, 1988, p.18.

como conceitualmente primários em comparação a conceitos e juízos valorativos, mas também como fornecendo os recursos conceituais para responder aos problemas da univocidade e da recalcitrância.

Se há uma forma de identificar e caracterizar sentimento que seja independente de conceitos e juízos valorativos, a tese neossentimentalista será capaz de dizer, sobre o problema da univocidade, que o que torna possível discordarmos sobre uma resposta ser adequada a uma certa circunstância no mundo é exatamente o fato de que podemos compartilhar a resposta. Ou seja, é o fato de que podemos falar sobre uma ocorrência de sentimento de um ponto de vista comum que torna possível apresentarmos razões a favor ou contra o mesmo e, assim, avaliá-lo. O sentimento seria, então, o ponto comum de toque a partir do qual o discurso valorativo poderia ser construído. Uma vez que duas partes disputantes sobre, por exemplo, a aplicabilidade de 'perigoso', poderiam concordar conceitualmente acerca da resposta – caracterizada não-valorativamente – elas teriam a garantia de que estão falando sobre a mesma coisa, ainda que pudessem discordar profundamente sobre qual a resposta adequada ao cenário com o qual se deparam: elas podem, assim, discordar sobre os padrões de avaliação do perigo de algo.

Considere agora a resposta neossentimentalista ao problema da recalcitrância. Se sentimentos ou emoções podem ser caracterizados e individuados (talvez até certo ponto não-conceitualmente) independentemente das operações de nossas atividades mentais de ordem superior, poderíamos dizer que certas situações deflagram a ocorrência de sentimentos, ainda que não exista genuinamente nenhum raciocínio, juízo ou aplicação de conceitos valorativos. Assim, seja qual for a explicação psicológica que tenhamos para isto (nós fornecemos uma cima, baseada em premissas evolucionistas e nos princípios da "revolução afetiva" em psicologia), se formos capazes de identificar o sentimento independentemente de atividades de ordem superior, poderemos facilmente explicar por que o sentimento pode estar em erro ou ser inadequado assim como recalcitrante.

Tomemos um exemplo. Considere que estamos em 2012 e Marcus está em

um Boeing 747. Marcus tem medo de voar de avião. Mas ele também conhece as estatísticas relacionadas a voar de avião em geral (e, em particular, naquele modelo de avião) e confia nelas. Ele sabe que é muito mais seguro do que dirigir um carro, por exemplo. Poderíamos, então, dizer que, embora Marcus tema voar no Boeing 747 em que está agora, ele considera este medo uma resposta inadequada a voar em um Boeing 747.

Considere também a distinção de Gibbard entre dizer que faz sentido (ou é adequado, em nossa terminologia) a um indivíduo *sentir* medo de uma certa maneira e dizer que faz sentido que um indivíduo *sinta* medo de uma certa maneira:

> Se tive um dia ruim e agora enfrento um novo desapontamento, "faz sentido que eu esteja com raiva" – podemos esperar que eu esteja com raiva em tais circunstâncias, por razões que compreendemos – mesmo que não "faça sentido eu estar com raiva" porque o novo desapontamento não é culpa de ninguém. Da mesma forma, faz sentido *que* Cleopatra *estivesse* com raiva do mensageiro [que a trouxe más notícias], mas não faz sentido ela *estar* com raiva dele. Raiva direcionada incorretamente nas circunstâncias em questão era esperada, mas as más notícias não foram culpa do mensageiro[65].

Que interpretação podemos fornecer a essas histórias? Em primeiro lugar, podemos dizer que podem existir excelentes razões evolucionistas para que seres humanos reajam a voar de avião (ou mesmo ao ato de imaginar isto) com medo. Esta pode ser uma boa explicação de por que Marcus tem medo de voar (recalcitrantemente), embora não veja qualquer razão para realmente se sentir desta forma em tal circunstância. Mas, então, se podemos explicar o medo de voar independentemente do emprego de conceitos e da formulação de juízos valorativos sobre voar, poderíamos identificar a ocorrência de um sentimento (e a sua recalcitrância) sem precisar recorrer a algo além de um critério não-valorativo para individuação de sentimentos, ao contrário do que afirmam as teses da não-

[65] GIBBARD, 1990, p.38.

prioridade e judicativa. Como vimos, a teoria evolucionista pode nos ajudar nisto[66].

De modo similar, no caso de Cleopatra, descrito por Gibbard, poderíamos facilmente identificar a raiva de Cleopatra, ainda que ela mesma possa pensar que a raiva não é adequada. Como sugere Gibbard, podemos compreender perfeitamente bem, de um ponto de vista explicativo, por que Cleopatra reagiu daquela maneira, ainda que não fosse a maneira adequada.

Porém, apesar de tudo o que dissemos, nada implica que a tese neossentimentalista negue uma relação estreita entre juízos valorativos e sentimentos ou emoções. O neossentimentalista pretende evitar a "tentação" das teses da não-prioridade e judicativa de assimilar, parcial ou inteiramente, sentimento a juízos. A tese neossentimentalista promete explicar o papel de conceitos e juízos valorativos nos sentimentos como uma tarefa regulativa, sem assimilar os últimos aos primeiros. Como diz Gibbard:

> De onde vem, então, essa tentação de compreender emoções como juízos genuínos, ou como incluindo estes juízos? Quando alguém está com raiva de mim por algo que lhe fiz, sua raiva parece incluir um juízo de que eu deveria ser censurado por isso. [Mas] isto é um equívoco, pois ele pode estar com raiva de mim por algo que fiz e, apesar disso, pensar que eu estava certo. Pode estar com raiva, mas pensar que que não faz sentido estar com raiva. […][67].
>
> A tentação surge em parte do fato de a raiva ter causas e focos típicos. As causas são tipicamente insultos, danos e outros ultrajes morais, e é nestas coisas que a raiva tende a se focar. […] [Podemos explicar por que isto é assim dizendo que] as normas que aceitamos para culpa e raiva se misturam [a outras causas], e elas ordinariamente endossam nossas

[66] Mas, obviamente, não precisa ser a única possibilidade. Como Ekman (1999, p.45-46) sugere, excluindo uma tese construtivista social (que tenderia a rejeitar o apelo a universais para explicar a recorrência de determinado padrão de respostas emocionais), uma tese que seja capaz de explicar constâncias em uma espécie de tal forma que universais possam ser identificados para explicá-las, de modo parcialmente independente da cultura, pode ser bem sucedida na tarefa em questão.

[67] GIBBARD 1990, p.147.

> principais tendências em sentir raiva […]. Por isso, normalmente pensamos que coisas que provocam raiva são condenáveis[68].

Assim, o ponto de Gibbard é simplesmente que juízos valorativos sobre, por exemplo, raiva, tendem a ser incorporados à concepção de raiva porque nós (muito naturalmente) normalmente endossamos manifestações e ocorrências de raiva. Isto é normalmente o caso. Porém, casos de recalcitrância de sentimentos não endossados são exatamente aqueles que não se enquadram em tais condições normais. Podemos ir mais além e dizer que a nossa melhor forma de *articular* (discursivamente) as situações que deflagram o sentimento em questão pode ser feita em termos *valorativos* (isto é, que envolva conceitos e juízos valorativos). Mas, apesar disto, o ponto não implica que conceitos e juízos valorativos devem ser empregados para que sentimentos ou emoções sejam experienciados. Certas experiências de sentimentos ou emoções podem ser simplesmente respostas rápidas, aptas e não-conceitualizadas (no sentido de serem não-racionais e não-intelectuais). Ainda assim, podemos chamá-las de sentimentos ou emoções e sermos capazes de "racionalizá-las" chamando-as *adequadas*.

Esta é a reposta que o neossentimentalista pode oferecer aos problemas que abordamos em nossa discussão anterior. Como deve estar claro agora, a tese neossentimentalista parece ser mais bem-sucedida do que as teses da não-prioridade e judicativa em lidar com eles. Não obstante, a nossa tarefa ainda não está terminada. Alguns adversários da tese neossentimentalista levantam sérias objeções acerca da inteligibilidade e coerência da história que contamos até aqui. Eles podem até conceder que devemos nos manter abertos à possibilidade de que a pesquisa empírica seja capaz de fornecer uma explicação plausível de sentimentos ou emoções em termos não-valorativos. Porém, de acordo com tais adversários da tese neossentimentalista, o problema que esta enfrenta não é empírico. Ao contrário, é um problema *a priori*. Segundo esses críticos, a tese neossentimentalista não se sustenta diante do que devemos esperar de uma

[68] 1990, p.147.

concepção adequada da racionalidade de sentimentos ou emoções. Lidar com esta objeção é agora a tarefa sobre a qual nos debruçamos.

5. Problemas e perspectivas para a tese neossentimentalista

As objeções de McDowell à tese neossentimentalista podem ser colocada de maneira sucinta na seguinte série de passagens. McDowell inicia com uma caracterização precisa da tese neossentimentalista e seus compromissos projetivistas:

> O ponto da imagem da projeção é explicar certas características aparentes da realidade como reflexos de nossas respostas subjetivas a um mundo que, na verdade, não contém tais características. Este rumo extraordinário de conduzir a questão parece exigir uma prioridade correspondente, na ordem do entendimento, entre a resposta projetada e a característica aparente: devemos ser capazes de focar nosso pensamento na resposta sem precisar explorar o conceito da característica aparente que supostamente resulta da projeção da resposta[69].

Como vimos anteriormente, McDowell concorda com o neossentimentalista que "seria obviamente grotesco imaginar que um caso de medo pudesse ser explicado como o resultado de um processo […] mecânico iniciado por um exemplar de 'periculosidade objetiva'"[70]. Porém, McDowell argumenta que a saída neossentimentalista e os seus compromissos projetivistas não funcionarão como uma alternativa viável àquela hipótese "grotesca" da tese judicativa, porque

> […] se aquilo em que estamos engajados é uma "tentativa de compreender a nós mesmos", então explicações meramente *causais* de respostas como medo não serão, de forma alguma, satisfatórias. O que precisamos aqui é de um estilo de explicação que torne inteligível o que é explicado […]. Isto significa que uma técnica para fornecer explicações satisfatórias de casos

[69] McDOWELL 1998a, p.157.

[70] 1998b, p.143-144. Ver também McDowell (1998a, p.156, n. 12), onde ele endossa explicitamente aquele ponto sentimentalista.

> de medo [...] deve acomodar a possibilidade de existir crítica. [...] Assim, explicações de medo que manifestem nossa capacidade de compreensão de nós mesmos neste domínio de nossas vidas simplesmente não serão coerentes com a alegação de que a realidade não contém nada na forma de periculosidade. Qualquer alegação desse tipo eliminaria a inteligibilidade que as explicações conferem a nossas respostas[71].

Dessa forma, McDowell pensa que uma explicação de nossas respostas subjetivas como o apelo do neossentimentalista a mecanismos (por exemplo, evolutivamente adaptativos) como base da manifestação de sentimentos ou emoções não conferiria inteiramente inteligibilidade àquilo que é explicado. Isto, à primeira vista, porque os mecanismos são causais. Devido a isto, segundo McDowell, não podemos nem tornar nossas respostas inteligíveis nem (*a fortiori*) acomodar a possibilidade de crítica às mesmas.

O entendimento deste ponto gera agora dois problemas específicos para o neossentimentalista, que estão expressos nas seguintes passagens:

> Na representação projetivista, ter as respostas éticas ou estéticas [ou valorativas em geral] racionalmente adequadas aos seus objetos é uma questão de ter o mecanismo processador relevante funcionando de modo aceitável. É claro que o projetivista pode acomodar perfeitamente bem a ideia de avaliar o mecanismo processador. Mas ele [isto é, o projetivista] representa o mecanismo como algo que se pode contemplar ele próprio como um objeto. Seria apropriado dizer que é "algo com relação a que se pode dar um passo atrás" não fosse pelo fato de que é preciso usar o próprio mecanismo para avaliá-lo [...] [72].

E todo o sentido de se explicar o discurso valorativo, segundo McDowell, é

> [...] fornecer uma explicação de como vereditos e juízos [valorativos] estão localizados na região apropriada do espaço das razões. Nenhum veredito ou juízo particular é um ponto de partida sacrossanto, supostamente imune

[71] McDOWELL 1998b, p.143-144, ênfase adicionada.
[72] McDOWELL, 1998b, p.148.

> à avaliação crítica, em nossa tentativa de adquirir o direito de dizer que alguns de tais vereditos e juízos têm uma chance de serem verdadeiros. Isto não significa absolutamente dizer que devemos adquirir esse direito a partir de uma posição inicial segundo a qual suspendem-se ao mesmo tempo *todos* os vereditos e juízos, tal como na representação projetivista de uma gama de respostas a um mundo que não contém valores […][73].

Há dois pontos inter-relacionados que merecem ser destacados nessas passagens. O primeiro é uma questão sobre a tese neossentimentalista ser capaz de introduzir um discurso valorativo ou racional (conceitos e juízos) através de uma descrição do funcionamento adequado de um mecanismo caracterizado em termos não-valorativos. Para avaliar o funcionamento de tal mecanismo o neossentimentalista parece precisar apelar a conceitos valorativos. Mas isto, a tese neossentimentalista nos diz, pode ser feito apenas através do exercício do próprio mecanismo que explica, ele próprio, a "invenção" do valor. Como, então, poderíamos aprender algo sobre o discurso valorativo a partir de mecanismos subjacentes às nossas respostas emocionais se parecemos precisar recorrer ao discurso valorativo para avaliar as próprias respostas? Assim, ao que parece, esta é uma questão sobre se o mecanismo que explica o valor como uma projeção pode ser empregado para avaliar um exercício de si mesmo.

O segundo ponto, contido na segunda passagem, questiona se podemos sequer falar de avaliação de uma resposta, se a tese neossentimentalista a explica, em última instância, como uma resposta não-racional – formulada em termos causais. Mais precisamente, a questão aqui é perguntar se estamos legitimamente autorizados a chamar uma resposta de adequada a uma série de características do mundo se podemos fazer isto apenas explicando nossa reação ao mundo. Como o valor surge em tal representação?

Nos termos que adotamos até aqui, podemos reformular as objeções de McDowell da seguinte maneira. A primeira seria: se compreendemos sentimentos como o resultado de um mecanismo causal, amplamente encapsulado e que pode

[73] 1998a, p.163.

operar independentemente de apreciação racional e valorativa, e se é apelando a tal mecanismo que explicamos como podemos dar sentido a valor, qualquer emprego do vocabulário valorativo será explicado como uma operação do próprio mecanismo. Como, então, pode o mecanismo avaliar as suas próprias respostas?

E se interpretamos sentimentos como o resultado de um mecanismo causal, como podemos conferir sentido a dizer que tais sentimentos são adequados ou não – em cujo caso fazemos uso de vocabulário valorativo? Como podemos dizer que são respostas adequadas ao mundo, se, em última instância, são apenas respostas causais deflagradas?

Em resumo, a primeira objeção diz respeito à possibilidade de avaliar uma resposta de um mecanismo recorrendo ao próprio mecanismo. A segunda objeção, por outro lado, diz respeito à possibilidade de avaliar o mecanismo, uma vez que as respostas dadas pelo mecanismo não são racionalmente explicadas, mas causalmente explicadas. Como o neossentimentalista pode lidar com estas objeções?[74]

A mesma estratégia pode ser empregada em ambas as objeções, e consiste em fazer uma distinção entre o que podemos chamar metaforicamente de uma 'reposta bruta' e uma 'resposta rica'. Se fôssemos seres capazes apenas de respostas brutas, nossas perspectivas valorativas seriam bastante limitadas. Elas careceriam de complexidade e sofisticação. Qualquer resposta que déssemos constituiria valor para nós. Do modo como o neossentimentalista explica valor, recorrendo às nossas respostas a um mundo desprovido de valor, se estas respostas carecessem de complexidade, nossas respostas imediatas constituiriam nossos valores. Esta pode ser a representação que estamos dispostos a atribuir a muitos animais[75]. Mas seres humanos não são assim. Nós desenvolvemos uma psicologia altamente sofisticada e complexa, que inclui não apenas a capacidade

[74] Brady (2008) parece levantar pontos muito similares a estes, de McDowell.

[75] Alguns diriam que uma tal caracterização deflacionária de valor implica atribuir valor, ao menos em um sentido mínimo, a animais e crianças e que isto é implausível. Não poderemos abordar esta questão aqui. Mas devemos salientar que não é inteiramente desprovido de sentido falar de valor em animais e crianças em um sentido ao menos rudimentar.

de reflexão e controle sobre nossas próprias atitudes, mas também esquemas conceituais e linguagem. E, apesar de reconhecer isto, não precisamos afirmar que, ao nos movermos para um nível mais elevado em nossa economia mental (que inclui tanto afeto quanto cognição), precisamos ir além de uma representação "mecanística" de nossas reações a fim de conferir sentido a algumas delas como adequadas. Somos capazes não apenas de formular regras gerais, como Adam Smith (1790, parte III) observou, sobre as circunstâncias que aprovamos e desaprovamos, de forma a aplicá-las a casos particulares, mas igualmente de identificar as respostas que são genuinamente representativas de nossas sensibilidades e, ao fazer isto, expressá-las usando vocabulário valorativo de tal forma que funcionem como um guia para nossas condutas e regulem nossas reações futuras[76].

Colocando o ponto em um vocabulário mais próximo do tipo de história evolucionista que contamos anteriormente, podemos seguir Blackburn quando diz que:

> Programas afetivos básicos operam de modo bastante similar a reflexos, produzindo alterações estereotipadas na face e no corpo. Eles explicam emoções imediatas e fenomenologicamente salientes [...]. Mas os programas afetivos básicos não parecem tão centrais a nossas emoções cognitivas de ordem superior, tais como inveja e ciúme, ou a síndromes emocionais de longo prazo, tais como raiva friamente calculada, ou mesmo pesar e amar. Em seu livro recente, Paul Griffths sugere que estas diferem das emoções primárias [ou mais básicas] de quatro modos importantes. Primeiramente, são mais flexíveis do lado do input, não deflagradas por quaisquer tipos específicos de situações funcionais, da mesma forma que medo é deflagrado por medo apreendido, ou raiva por obstáculos e desafios. Em segundo lugar, são persistentes ao longo do tempo, diferentemente de efeitos musculares e fisiológicos dos programas afetivos.

[76] Recorremos aqui amplamente a uma série de pontos que aparecem na discussão de Blackburn sobre as nossas sensibilidades (ver 1984; e 1998, capítulos 5 e 8). Ele diz que "nosso apego aos nossos objetivos de longo termo são eles mesmos estados emocionais" (1998, p.129).

> Em terceiro lugar, não possuem outputs estereotipados, tais como expressões faciais típicas ou estados fisiológicos. E, finalmente, são integradas a outros processos conscientes, e podem conduzir a planos de ação de longo prazo, em vez de simplesmente deflagrar respostas do tipo de reflexos[77].

E isto se dá assim sobretudo porque, no caso de seres complexos como os humanos, estes são "provavelmente geneticamente dotados de disposição *plástica* de segunda ordem para formar diferentes respostas afetivas de vários tipos em vários ambientes sociais"[78].

Ainda assim, embora a plasticidade tenha agora sido adicionada à caracterização do ponto de vista de ordem superior a partir do qual podemos avaliar respostas mais básicas,

> O ponto principal é que não há qualquer dualismo, segundo o qual um [nível de resposta] se move livremente do outro. E […] se olharmos novamente para as quatro razões que separam emoções cognitivas de ordem superior das outras respostas [primárias ou básicas], veremos que […] são bastante compatíveis com a ideia de compreender os estados variáveis culturalmente e de longo prazo essencialmente como disposições cujas manifestações são padrões particulares (e misturas) de ocorrências das emoções primárias[79].

Se Blackburn está na direção certa, não há problema em avaliar as respostas de um mecanismo pelas próprias operações do mesmo mecanismo, dado que estejamos diante de um mecanismo complexo que representa a multiplicidade de respostas possíveis ao mundo, em vários níveis, que podem ser agrupadas, refletidas e controladas por seres como nós.

---

[77] BLACKBURN, 1998, p.128.

[78] 1998, p.147, ênfase adicionada.

[79] BLACKBURN, 1998, p.129. Afinal, Blackburn nos lembra que, para dar sentido a toda esta tese de um ponto de vista naturalista, "devemos supor que nossas manifestações [emocionais] viscerais não são meramente 'percebidas', mas cumprem um papel na determinação daquilo que nos *atrai* e repele ou naquilo que nos tornamos inclinados a fazer" (p.130).

Na verdade, este parece ser um ponto endossado por Wiggins e (ao menos parcialmente) por McDowell. Devemos lembrar de uma passagem de Wiggins citada anteriormente:

> […] frequentemente não haverá como falar das propriedades que esses nomes representam independentemente das reações que provocam. […] Mas igualmente – *ao menos quando o sistema de propriedades e reações se diversifica, se torna complexo e enriquece* – não haverá como falar exatamente sobre *qual* reação uma coisa com a propriedade associada provocará sem direta ou indireta alusão à própria propriedade[80].

E McDowell explicitamente reconhece que o neossentimentalista nos fornece uma representação complexa de várias atitudes agrupadas como respostas a situações no mundo:

> Não há duvida de que uma explicação completa dos fenômenos psicológicos agruparia os itens em termos de suas tendências a produzir tais respostas, mas tais tendências não são propriedades que precisam ser explicadas como projeções das respostas[81].

Assim, McDowell e Wiggins parecem admitir que, na medida em que tornamos nosso sistema mais complexo e rico, passaremos a direcionar nossa atenção a exemplares particulares das respostas ao mundo de uma tal perspectiva complexa e enriquecida que "agrupa os itens em termos de suas tendências a produzir tais respostas". Para Wiggins, uma vez que possuímos tal perspectiva complexa e rica impregnada em nossas práticas, poderíamos muito naturalmente passar a olhar para as instâncias particulares de respostas ao mundo aludindo *primariamente* aos conceitos, juízos e supostas propriedades valorativas relevantes. McDowell, por outro lado, como indica a passagem, não concede isto tão facilmente ao neossentimentalista. Porém, por razões aduzidas em nome do neossentimentalista nos parágrafos anteriores, não é claro por que o

[80] 2002, p.195 (ênfase adicionada).
[81] 1998a, p. 157.

neossentimentalista não possa dizer exatamente aquilo sobre a sua própria explicação da relação entre sentimento e valor.

Talvez o real problema para o neossentimentalista resida na segunda objeção que vimos acima. Talvez a resposta neossentimentalista à primeira objeção não funcione porque a tese neossentimentalista não está legitimada a falar de avaliações de respostas, em primeiro lugar.

Mas o neossentimentalista pode dizer, como resposta à segunda objeção, que ela é simplesmente uma petição de princípio contra a tese neossentimentalista[82]. Esta se propõe a fornecer uma *explicação* de nosso uso da linguagem e do pensamento valorativos. Como tal, a tese simplesmente diz que compreendemos um certo domínio de discurso valorativo compreendendo independentemente respostas (e mecanismos) causais ao mundo. Uma vez que estas respostas são amplamente afetivas e conectadas a ações, a tese neossentimentalista parece ter aqui os elementos básicos para desenvolver uma explicação completa do valor.

Dessa forma, quando valoramos algo, expressamos nossas respostas afetivas (em vários níveis) em relação a um mundo desprovido de valor. Simplesmente não faz sentido, portanto, de acordo com tal modelo, perguntar se uma resposta bruta ao mundo é adequada como tal. Mas, uma vez que não somos seres apenas de respostas brutas, quanto mais acrescentamos complexidade (em vários níveis de atividade mental) ao nosso modelo, mais claro fica como é possível formular juízos sobre respostas particulares e julgá-las adequadas ou não. Assim, é verdade que, sem a complexidade de nossas capacidades conceituais e de atividades mentais de ordem superior relacionadas ao pensamento e discurso, não poderíamos sequer perguntar se uma dada resposta ao mundo é adequada, se incorpora valor. Mas, dado que nossas perspectivas valorativas são complexas e ricamente constituídas, podemos direcionar nossos padrões gerais de respostas às instâncias particulares e julgá-las valorativamente.

Em um certo sentido, o neossentimentalista pode, então, concordar com

[82] Ver Miller (2003, capítulos 4 e 10) para uma discussão similar sobre este ponto.

McDowell quando diz que

> [...] não há nenhum fato encerrado em si mesmo anterior às nossas vidas subjetivas que pudesse fazer parte da explicação projetivista do modo relevante de pensar [valorativamente]; na única resposta relevante, o aparato conceitual que figura no modo relevante de pensar está já presente[83].

Se McDowell está aqui falando de *justificação* de nossas respostas, o neossentimentalista irá concordar inteiramente que estas necessariamente ocorrerão nos limites do escopo do pensamento e discurso valorativos. Porém, se McDowell está falando de *explicação*, o neossentimentalista irá discordar com o que McDowell diz naquela passagem, observando que uma explicação do valor nos termos do tipo de história psicológica que o neossentimentalista conta propõe-se a mostrar como valores e o discurso valorativo podem existir em um mundo natural sem fazer deste mundo o critério de *justificação* de nossas práticas valorativas.

Lamentavelmente, este último ponto é algo que, como deve estar claro que temos razões para suspeitar, não é possível a McDowell conceder ao neossentimentalista. Ele deixa isto claro quando diz que:

> Não há dúvida de que a propensão para rir é, em certo sentido, um fato psicológico prévio e auto-suficiente. Mas diferenciar alguns exercícios daquela propensão não-especificada como casos de *achar graça* é algo que devemos descobrir; e se a sugestão é correta, esta descoberta é indistinguível de descobrirmos algo como cômico[84].

Como deve estar claro, a teoria psicológica de McDowell é totalmente valorativa ou normativa. Isto parece ser um erro que devemos evitar, especialmente devido ao seu alto custo. Problemas como o da univocidade e da recalcitrância mostram que seríamos mais bem-sucedidos se mantivéssemos a

---

[83] 1998a, p.158.
[84] 1998a, p.158.

justificação e a explicação separadas, e não tornássemos a nossa teoria psicológica inteiramente valorativa ou normativa.

Conclusão

Argumentamos que a tese neossentimentalista possui mais recursos do que as teses judicativa e da não-prioridade no que concerne às respostas possíveis aos problemas da univocidade e da recalcitrância.

Nossa resposta em nome do neossentimentalista concentrou-se nas vantagens que uma caracterização independente de sentimentos ou emoções forneceria como resposta aos problemas da univocidade e da recalcitrância. Se pudéssemos identificar sentimentos ou emoções sem recorrer a conceitos e juízos valorativos, poderíamos fornecer uma resposta clara a tais problemas.

Como propusemos, para que essa estratégia neossentimentalista funcione, ela parece precisar do suporte de uma teoria empírica que confirme que é possível caracterizar sentimentos ou emoções independentemente do pensamento e discurso valorativos. Indicamos uma maneira de fazer isto trazendo à tona a "revolução afetiva" em psicologia e o tipo de história evolucionista que normalmente a acompanha. Por outro lado, não fornecemos qualquer defesa desta hipótese na literatura psicológica. Nesse sentido, nossos resultados aqui são condicionais à confirmação da verdade daquela hipótese.

Por último, formulamos uma resposta neossentimentalista a algumas objeções sérias levantadas por McDowell (et al) aos compromissos gerais da tese neossentimentalista. Basicamente, como vimos, o problema do neossentimentalista é explicar como tornar possível falar de valor a partir de uma representação inteiramente naturalista, representada por mecanismos causais subjacentes às nossas respostas ao mundo e que se tornam constitutivas de valor. Sugerimos, assim, que questões de valor e adequação fazem sentido apenas com base em uma perspectiva complexa e enriquecida (com níveis distintos de atitudes) de sentimentos ou emoções, a partir da qual manifestações particulares de respostas ao mundo são julgadas adequadas ou não na medida em que são coerentes com os padrões gerais das perspectivas enriquecidas de ordem

superior. Se isto funcionar, o neossentimentalista será capaz de fornecer uma *explicação* da origem do pensamento e discurso valorativos em termos causais e, ao mesmo tempo, sustentar que, uma vez que tornamos nosso pensamento e discurso valorativos enriquecidos e complexos o bastante, de modo que se tornem partes essenciais de nossas interações e interpretações do mundo, nossas práticas tornar-se-ão ineliminavelmente valorativas ou normativas[85].

Referências

BLACKBURN, S. *Ruling passions: A theory of practical reasoning*. Oxford: Clarendon Press, 1998.

_____. *Spreading the word: Groundings in the philosophy of language*. Oxford: Oxford University Press, 1984.

BRADY, M. S. "Value and fitting emotions". In: *The Journal of Value Inquiry* 42, p.465-475, 2008.

BROAD, C. D. "Emotion and sentiment". In: *Journal of Aesthetics and Art Criticism* 13, p.203-214, 1954.

D'ARMS, J.; JACOBSON, D. "Expressivism, morality, and the emotions". In: *Ethics* 104.4, p.739-763, 1994.

_____. "Sentiment and value". In: *Ethics* 110.4, p.722-748, 2000a.

_____. "The moralistic fallacy: On the 'appropriateness' of emotions". In: *Philosophy and Phenomenological Research* 61.1, p.65-90, 2000b.

_____. "The significance of recalcitrant emotion (or anti-quasijudgmentalism)". In: A. Hatzimoysis (ed.), *Philosophy and the emotions*. Cambridge: Cambridge University Press, 2003.

[85] Agradeço a Cinara Nahra por comentários em uma versão anterior deste artigo.

_____. “Anthropocentric constraints on human value”. In: R. Shafer-Landau (ed.), *Oxford studies in metaethics*, 1. Oxford: Oxford University Press, 2006a.

_____. “Sensibility theory and projectivism”. In: D. Copp (ed.), *The Oxford handbook of ethical theory*. Oxford: Oxford University Press, 2006b.

D’ARMS, J. “Two arguments for sentimentalism”. In: *Philosophical Issues* 15, p.1-21, 2005.

DEIGH, J. “Cognitivism in the theory of emotions”. In: *Ethics* 104.4, p.824-854, 1994.

EKMAN, P. ; FRIESEN, W. V. “Constants across cultures in the face and emotion”. In: *Journal of Personality and Social Psychology* 17.2, p.124-129, 1971.

_____. “All emotions are basic”. In: P. Ekman & R. J. Davidson, *The nature of emotion: Fundamental questions*. Oxford: Oxford University Press, 1994.

_____. “Basic emotions”. In: Tim Dalgleish & Mick J. Power (eds.), *Handbook of cognition and emotion*. London: John Wiley & Sons, Ltd, 1999.

FODOR, J. A. *The modularity of mind*. Massachusetts: MIT Press, 1983.

FOOT, P. *Virtues and vices and other essays in moral philosophy*. Oxford: Blackwell, 1978.

GIBBARD, A. *Wise choices, apt feelings: A theory of normative judgement*. Cambridge, Mass.: Harvard University Press, 1990.

GREENSPAN, P. *Emotions and reasons: An inquiry into emotional justification*. London: Routledge, 1988.

GRIFFITHS, P. E. *What emotions really are: The problem of psychological categories*. Chicago and London: The University of Chicago Press, 1997.

HARE, R. *The language of morals*. Oxford: Clarendon Press, 1952.

HELM, B. W. *Emotional reason: Deliberation, motivation, and the nature of value.* Cambridge: Cambridge University Press, 2001.

HUME, D. *An enquiry concerning the principles of morals*, L. A. Selby-Bigge (ed.), 3rd edition revised by P. H. Nidditch. Oxford: Clarendon Press, 1751/1975.

KUNST-WILSON. W. R.; ZAJONC, R. B. "Affective discrimination of stimuli that cannot be recognized". In: *Science* 207.1, p.557-558, 1980.

LAZARUS, R. S. "Thoughts on the relations between emotion and cognition". In: *American Psychologist* 37.9, p.1019-1024, 1982.

_____; LAZARUS, B. N. *Passion and reason: Making sense of our emotions*. Oxford: Oxford University Press, 1994.

LOGUE, A. W.; Ophir, I.; Strauss, K. E. (1981). "The acquisition of taste aversion in humans". In: *Behav. Res. & Therapy* 19, p.319-333.

McDOWELL, J. "Projection and truth in ethics". In: *Mind, value, and reality*. Cambridge, Mass.: Harvard University Press, 1998a.

_____. "Values and secondary qualities". In: *Mind, value, and reality*. Cambridge, Mass.: Harvard University Press, 1998b.

_____."Virtue and reason". In: *Mind, value, and reality*. Cambridge, Mass.: Harvard University Press, 1998c.

MILLER, A. *An introduction to contemporary metaethics*. Cambridge: Polity Press, 2003.

NESSE, R. M. "Evolutionary explanations of emotions". In: *Human Nature* 1.3, p.261-289, 1990.

NUSSBAUM, M. C. *Love's knowledge: Essays on philosophy and literature*. Oxford: Oxford University Press, 1990.

_____. *The therapy of desire: Theory and practice in hellenistic ethics*. Princeton: Princeton University Press, 1994.

RABINOWICZ, W.; RONNOW-RASMUSSEN, T. “The strike of the demon: On fitting pro-attitudes and value”. In: *Ethics* 114.3, p.391-423, 2004.

ROBERTS, R. C. *Emotions: An essay in aid of moral psychology*. Cambridge: Cambridge University Press, 2003.

SAUER, H. “The appropriateness of emotions: Moral judgement, moral emotions, and the conflation problem”. In: *Ethical Perspectives* 18.1, p.107-140, 2011.

SOLOMON, R. *Not passion’s slave: Emotions and choice*. Oxford: Oxford University Press, 2007.

SHAFFER, J. “An assessment of emotion”. In: *American Philosophical Quarterly* 20.2, p.161-173, 1983.

SMITH, Adam. *The theory of moral sentiments*. New York: Dover, 1790/2006.

TOOBY, J.; COSMIDES, L. “The past explains the present: Emotional adaptations and the structure of ancestral environments”. In: *Ethology and Sociobiology* 11, p.375-424, 1990.

WIGGINS, D. “A sensible subjectivism?”. In: *Needs, values, truth: Essays in the philosophy of value.* Third edition, amended. Oxford: Oxford University Press, 2002.

WILSON, W. R. “Feeling more than we can know: Exposure effects without learning”. In: *Journal of Personality and Social Psychology* 37.6, p.811-821, 1979.

ZAJONC, R. B. “Feeling and thinking: Preferences need no inferences”. In: *American Psychologist* 35.2: p.151-175, 1980.

10

SENTIMENTOS MORAIS, INTERNALISMO MOTIVACIONAL, SENTIMENTALISMO

Wilson Mendonça

1. Três teses sobre os sentimentos morais

Filósofos se referem frequentemente a sentimentos morais em um sentido estreito. Exemplos paradigmáticos de sentimentos morais nesse sentido são os estados emocionais de culpa e indignação imparcial, simpatia estendida, sentimentos de aprovação e desaprovação. Algumas vezes, porém, a expressão "sentimentos morais" é usada em um sentido mais amplo para denotar, por exemplo, atitudes, pró-atitudes, intenções, desejos de primeira ordem, desejos de segunda ordem, endosso de normas, planos contingentes. Algumas teorias morais consideram os sentimentos morais como estados afetivos e conteudísticos (isto é, sentimentos em um sentido mais amplo) que têm emoções (isto é, sentimentos em um sentido estreito) como seus conteúdos. Outras teorias morais misturam os sentidos estreito e amplo dos sentimentos de outros modos. Assim, o expressivismo de Alan Gibbard afirma que o juízo moral expressa o endosso de uma norma que justifica racionalmente os sentimentos de culpa e ressentimento em certas circunstâncias, ao passo que o "emocionismo" de Jesse Prinz identifica os sentimentos com "disposições cujas ocorrências manifestas […] são emoções"[1].

Sentimentos morais, amplos ou estreitos, desempenham comumente um papel importante em teorias morais—mesmo em teorias fundamentalmente antissentimentalistas. O racionalismo moral de Kant é um exemplo significativo. Após afirmar que "dever é a necessidade de uma ação por respeito à lei"[2], Kant

---

[1] PRINZ 2007, p.84.
[2] KANT 1785/2005, p.31.

não deixa dúvidas que "o respeito é um sentimento." Ele pode ser diferente de todos os outros sentimentos "que se podem reportar à inclinação ou ao medo." Apesar disso, como deve ser visto como "o *efeito* da lei sobre o sujeito e não a sua *causa*"[3], o respeito kantiano é uma resposta subjetiva, afetiva à realidade moral, isto é, um sentimento em um ou outro dos sentidos distinguidos acima. Na medida em que explica como o sujeito pode estar disposto a agir de acordo com os fatos morais apreendidos pela razão, o sentimento do respeito desempenha na teoria de Kant um papel *motivacional*. Visto de perspectiva mais geral, esse é somente um dos papéis que os sentimentos podem desempenhar em explicações metaéticas. De fato, é possível distinguir pelo menos três teses sobre sentimentos, que correspondem a três diferentes papéis que as teorias filosóficas estão dispostas a atribuir a sentimentos.

1. *A Tese Motivacional:* É bastante natural conceber os sentimentos morais como *links* entre nossas convicções morais, por um lado, e as ações correspondentes, por outro: eles têm a direção de ajuste "certa," isto é, a direção de ajuste oposta à das meras crenças. A natureza da conexão entre os juízos morais e os sentimentos motivadores é objeto de disputa entre os defensores da tese motivacional. De acordo com uma versão forte desta tese, a conexão é *interna* ou necessária, no sentido da necessidade conceitual ou *a priori*. Uma versão mais fraca reconhece somente um *link externo* (contingente) dos juízos morais aos sentimentos.

2. *A Tese Epistemológica:* Sentimentos morais são indicadores mais ou menos confiáveis da presença de valor moral. Alternativamente, especialmente os sentimentos negativos (culpa e indignação) são tratados por psicólogos morais como sinalizando o fato de que uma norma moral foi violada. O valor moral *per se* ou o conteúdo da norma moral pode não ter coisa alguma a ver com os sentimentos. Nesse caso, os sentimentos morais que desempenham um *papel epistemológico* não indicariam nem sentimentos, nem entidades constituídas por

[3] KANT 1785/2005, p.32.

sentimentos. Isto quer dizer que a tese epistemológica é independente da terceira tese.

3. *A Tese Constitutivista:* Esta é o sentimentalismo propriamente dito. O que é afirmado aqui é que os sentimentos entram na constituição dos valores morais (sentimentalismo metafísico) ou que os juízos morais são (construídos com) sentimentos (sentimentalismo de juízo). O emotivismo de Ayer é um exemplo "tradicional" da tese constitutivista. Os sentimentalistas contemporâneos atribuem um *papel constitutivo* aos sentimentos na medida em que adotam um ponto de vista segundo o qual o conceito de valor moral (ou de juízo moral) é um conceito dependente das nossas respostas (*a response-dependent concept*): o que significa que alguma coisa é um valor moral (ou um juízo moral) só poderá ser especificado se nós recorrermos às respostas afetivas em nós às características moralmente relevantes do mundo—se nós recorrermos aos sentimentos subjetivos causados pelos eventos mundanos objetivos.

O alvo principal do presente trabalho é uma certa visão das relações de dependência entre a tese motivacional e a tese constitutiva. As próximas seções (2 e 3) discutem a questão do internalismo *versus* externalismo na teoria da motivação. Na seção final (4), os resultados desta discussão são aplicados ao tema principal do trabalho.

2. Internalismo e motivação moral

Para muitos filósofos existe uma conexão necessária (interna) entre o juízo moral (a aceitação de uma sentença moral) e a motivação para a ação. De acordo com uma versão desse princípio, quem pensa que φ é certo está necessariamente motivado a realizar φ. A motivação é uma disposição para agir que pode não se manifestar em uma circunstância ou outra, mas ainda assim ela está presente se o agente mantém o juízo moral correspondente. A motivação pode ser superada por uma motivação contrária. Uma versão mais fraca deste princípio condiciona a conexão necessária à racionalidade do agente: o agente que endossa a sentença "φ é certo" está necessariamente disposto a realizar φ, sob pena de irracionalidade prática. Essa versão abre espaço para a ação acrática, que é normalmente

considerada como manifestação da irracionalidade. A tese da conexão necessária (condicionada ou não à racionalidade do agente) entre o juízo moral e a motivação adequada é normalmente conhecida como "internalismo de juízo," de acordo com uma sugestão terminológica de Darwall (1983). Dela se distingue a afirmação de que há uma conexão interna ou necessária entre as *razões normativas* de um agente e os elementos do seu "conjunto motivacional" atual (ou simplesmente dos seus "desejos" atuais), onde razões normativas são considerações que, independentemente da percepção por parte do agente, isto é, independentemente das *razões motivacionais* subjetivas do agente, falam objetivamente a favor da ação realizadora de φ. Este é o "internalismo de existência" na terminologia de Darwall. Aqui serão utilizadas as expressões "internalismo motivacional" e "internalismo de razões." As implicações metaéticas do primeiro tipo de internalismo serão tratadas na presente seção.

Um argumento muito comum a favor do não-cognitivismo mobiliza o internalismo motivacional como sua premissa principal.

O Argumento Internalista a Favor do Não-cognitivismo:

1. Juízos morais motivam necessariamente. [internalismo motivacional]
2. Estados mentais têm direção de ajuste mente-mundo *ou* mundo-mente.
3. Somente estados mentais com direção de ajuste mundo-mente são essencialmente motivadores. [teoria humiana da motivação]

Logo,

4. Juízos morais têm direção de ajuste mundo-mente. [não-cognitivismo moral]

A segunda premissa desse argumento não é muito controversa. Ela foi negada pelos defensores de estados mentais "híbridos," isto é, estados com dupla direção de ajuste. As teorias híbridas não serão consideradas aqui. Portanto, a segunda premissa não será questionada. A defesa mais plausível da terceira premissa foi formulada por Smith, que (i) situa a motivação para agir em uma

estrutura teleológica—"estar motivado é basicamente ter uma meta"—e (ii) identifica o estabelecimento de uma meta com um estado de direção de ajuste mundo-mente—"ter uma meta é estar em um estado ao qual o mundo deve ajustar-se"[4]. Embora tenha sido alvo de considerações críticas por alguns filósofos, a terceira premissa será considerada aqui como estabelecida. Isto permite isolar a primeira premissa como o ponto crucial do argumento reproduzido acima. Quão forte é o argumento que leva do internalismo motivacional ao não-cognitivismo?

Antes de abordar esta questão, é conveniente considerar outro argumento que combina o internalismo motivacional com o cognitivismo moral para concluir com a afirmação de que os fatos morais são inexistentes ou pelo menos "bizarros" (*queer*).

O Argumento Internalista a Favor do Irrealismo Moral:

1. Juízos morais motivam necessariamente. [internalismo motivacional]
2. Juízos morais são estados mentais de direção de ajuste mente-mundo, isto é, crenças. [cognitivismo moral]

Logo,

3. Os fatos morais (os objetos de nossas crenças morais) são intrinsecamente motivacionais, no sentido de que as representações de tais fatos (as crenças correspondentes) são necessariamente motivadoras.
4. Os fatos não-morais com que estamos familiarizados não são intrinsecamente motivacionais.

Logo,

5. Fatos morais são bizarros, muito provavelmente inexistentes.

A quarta premissa desse argumento é uma extensão da teoria humiana da motivação. Uma crença comum (não-moral)—a representação subjetiva dos fatos mundanos—faz parte dos antecedentes causais motivadores das nossas ações.

[4] SMITH 1994, p. 116.

Intrinsecamente, porém, a crença comum é inerte: somente em combinação com um estado ativo independente e de direção de ajuste mente-mundo (desejo) ela contribui para a motivação. Como na exposição do primeiro argumento esquematizado no início desta seção, a quarta premissa pode ser considerada aqui como estabelecida. A segunda premissa (o cognitivismo moral), por outro lado, pode ser vista como uma assunção. Nesse caso, o argumento leva do internalismo motivacional ao irrealismo moral. Na opinião de muitos filósofos, essa é a estrutura do argumento da bizarrice (*argument from queerness*), no qual Mackie baseia a teoria do erro[5]. De qualquer forma, a primeira parte do argumento pode ser vista como a base de uma *reductio*: se, depois da conclusão da primeira parte, acrescentarmos a premissa plausível de que não existem crenças essencialmente motivadoras, a falsidade do cognitivismo moral poderá ser inferida com alguma segurança.

O que foi dito até aqui justifica a relevância atribuída ao internalismo motivacional no debate metaético. Os não-cognitivistas veem comumente no internalismo motivacional um aliado poderoso, ao passo que os cognitivistas o consideram um adversário. Não por acaso, os cognitivistas são, via de regra, também externalistas motivacionais que (i) não negam a existência de uma conexão regular e relativamente estreita entre o juízo moral e a motivação correspondente, mas (ii) criticam os internalistas por terem aparentemente confundido uma contingência recorrente com uma relação necessária. Para o cognitivista externalista típico é perfeitamente concebível e coerente a descrição de um "amoralista," isto é, de um agente racional no qual a perda da motivação correspondente deixa ainda intacta sua crença moral. Todavia, alguns cognitivistas não gostariam de adotar o externalismo motivacional. Estes veem na compatibilização do internalismo motivacional com o cognitivismo—e a consequente refutação da hipótese amoralista—a solução do problema mais central da metaética, o que Smith chama de "problema moral"[6].

---

[5] 1977, p.38-42.

[6] 1994, p.4-13. Cf. também McNAUGHTON, 1988, p.23.

3. Variantes do internalismo motivacional

Jon Tresan observou em alguns ensaios recentes (Tresan 2006, 2009a e 2009b) que o debate sobre o internalismo motivacional repousa em uma ambiguidade. A ideia de uma conexão necessária entre juízos morais e motivações pode ser interpretada em dois sentidos excludentes. De acordo com a interpretação *de dicto*, a necessidade é atribuída a uma proposição: necessariamente, juízos morais são motivadores—ou talvez: necessariamente, juízos morais são acompanhados de motivações. (Em termos mais formais, o internalismo motivacional *de dicto* pode ser simbolizado como um condicional material universalmente quantificado e precedido do operador proposicional de necessidade.) De acordo com a interpretação *de re*, a necessidade se aplica aos juízos morais e não a proposições sobre os juízos morais: juízos morais são necessariamente motivadores—ou talvez: juízos morais são necessariamente acompanhados de motivações. (Em termos mais formais, o internalismo motivacional *de re* pode ser simbolizado como um condicional material universalmente quantificado onde o consequente do condicional é precedido do operador de necessidade.) A distinção pode ser esclarecida com base em um exemplo trivial. *De dicto*, é necessariamente verdadeiro que pais têm filhos. Mas não é verdadeiro que pais, *de re*, necessariamente têm filhos. Os pais no mundo atual podem existir em outros mundos possíveis sem filhos. Eles não seriam pais nesses mundos, mas ainda assim seriam os mesmos seres que no mundo atual são pais e existem, em alguns mundos possíveis, sem filhos. Isso é o mesmo que dizer que os pais atuais não são essencialmente pais. Em todos os mundos possíveis, porém, os pais nesses mundos têm filhos. No que diz respeito à relação entre pais e filhos, nós devemos ser internalistas *de dicto*, mas externalistas *de re*. Qual deve ser a nossa posição quando o interesse é dirigido não à relação pais e filhos, mas à relação entre juízos morais e as motivações correspondentes?

Aqui é importante observar que a validade dos argumentos esquematizados acima depende da interpretação *de re* da conexão necessária entre juízo moral e motivação. As premissas relevantes afirmam que a crença moral (a representação

dos fatos morais) não pode deixar de ser motivadora. Contudo, se a necessidade em questão fosse *de dicto*, a crença moral *poderia* não motivar, como qualquer pai atual poderia não ter filhos. Em um mundo possível onde um pai atual não tem filhos, ele não seria pai, mas ainda a mesma pessoa com todas as suas características essenciais. Da mesma forma a representação dos fatos morais sem a propriedade de motivar ou desacompanhada de motivação não seria uma crença *moral*, mas ainda assim a mesma crença com todas as suas propriedades essenciais. O que subjaz a este raciocínio é a ideia independentemente plausível de que a propriedade de motivar para a ação é uma propriedade não-essencial das crenças morais, assim como ser pai não é uma propriedade essencial dos que são pais. O resultado é que somente a interpretação *de re* da conexão necessária entre juízos morais, por um lado, e estados de direção de ajuste mundo-mente, por outro, tornaria válidos os argumentos acima. Mas talvez devamos ser internalistas motivacionais *de re*.

A ambiguidade notada por Tresan se manifesta na discussão sobre o internalismo motivacional—não na formulação das consequências do internalismo motivacional. Nesse contexto a interpretação relevante é a *de dicto*. É a verdade do internalismo motivacional *de dicto* que está em jogo na discussão metaética. Argumentos pró e contra o internalismo motivacional giram em torno da figura do amoralista e apelam às nossas intuições. Defensores do externalismo afirmam ser natural e intuitiva a interpretação dos proferimentos do amoralista como genuinamente morais, apesar da ausência de toda e qualquer motivação. Os internalistas, ao contrário, consideram que, na pior das hipóteses, os proferimentos do amoralista não são atos de fala significativos, não expressam mais juízos—o amoralista deixou de *falar*. Na melhor das hipóteses, os internalistas consideram que os proferimentos do amoralista devem ser interpretados como expressando juízos não-morais, por exemplo, sobre o que seus ouvintes (mas não ele próprio) julgam ser certo ou errado. O sentido do proferimento do amoralista "φ é errado, mas eu não estou motivado a fazer φ" é o da sentença "φ é o que vocês chamam de errado, mas eu não estou motivado a fazer φ." De acordo com esta interpretação do proferimento do amoralista, este não é verdadeiramente um

*a*moralista, mas um não-conformista moral cuja motivação está perfeitamente alinhada com o juízo moral negativo sobre as convenções locais. Evidentemente, os internalistas motivacionais pensam que essa é a interpretação que está mais de acordo com nossas intuições. Qualquer que seja a decisão sobre esse conflito de interpretações, o que está sendo discutido aqui são as condições de aplicabilidade do conceito de juízo moral. O internalista motivacional e o externalista não estão discutindo, com referência ao amoralista, se um juízo moral no mundo atual é acompanhado de motivação em qualquer outro mundo possível. O que o internalista motivacional e o externalista discutem é se o juízo proferido pelo amoralista, *no mundo possível do amoralista*, é ou não um juízo *moral*. O internalista insiste — e o externalista nega — que a presença da motivação adequada seja uma condição para a aplicabilidade do conceito de juízo moral ao proferimento do amoralista. Nesse sentido, a tese internalista estabelece uma conexão necessária e *a priori* entre o conceito de juízo moral, por um lado, e o de motivação, por outro, não uma relação *de re* entre juízos morais, *independentemente da descrição como morais*, e motivações. Como Tresan define o internalismo motivacional *de dicto*: "moral beliefs are accompanied by conations, and *are conceived by us* precisely *as* so accompanied (perhaps inter alia)"[7].

4. Sentimentalismo sobre juízos morais e a tese motivacional

Consideremos mais uma vez o argumento internalista a favor do não-cognitivismo que foi esboçado acima (seção 2). O argumento não é válido independentemente de como deve ser concebida a suposta conexão necessária entre juízo moral e motivação. Somente sob a suposição de que esse é um caso de necessidade *de re*, a direção de ajuste mente-mundo constitutiva de um dos *relata* (a motivação) entra também na constituição do outro *relatum* (o juízo). Mesmo que sejam necessariamente conectados, a motivação e o juízo podem ter direções de ajuste opostas, desde que a necessidade em questão seja do tipo *de dicto*. Isto quer dizer: se a premissa internalista expressar uma relação de *dicto*, o

[7] TRESAN, 2006, p.147, grifo nosso.

argumento é simplesmente inválido. Isto nos obriga a reconsiderar a justificação da premissa internalista. Como foi argumentado acima (seção 3), se alguma forma de internalismo motivacional puder ser justificada no debate entre internalistas e externalistas, essa forma será a do internalismo *de dicto*. A verdade do internalismo motivacional *de re* não é o objeto da discussão centrada na possibilidade (ou impossibilidade) do amoralista. De fato, o internalismo motivacional *de re* parece ser uma posição implausível.

Sob a interpretação *de dicto*, existem mundos possíveis nos quais o mesmo juízo, que vem acompanhado, no nosso mundo, da motivação "certa," existe sem motivação—esses são os mundos do amoralista—ou com a motivação "errada"—esses são os mundos do agente "satânico," que está deliberadamente disposto a agir contrariamente ao que ele julga ser correto[8]. Isso requer que nós formulemos claramente as condições de identidade do juízo através dos mundos possíveis: ele deve ser o *mesmo* juízo em todos os mundos possíveis onde a motivação varia. E isso, por sua vez, requer que façamos a distinção entre propriedades essenciais e propriedades acidentais dos juízos. A tarefa é difícil. Mas o que é factível como forma de abordagem do problema é avançar a tese segundo a qual o termo "moral" na expressão composta "juízo moral" não designa uma propriedade essencial dos juízos. Isso nos permite dizer que um juízo é moral quando ele é, entre outras coisas, acompanhado, no mundo real, pela motivação "certa." Ele deixa de ser moral, mas ainda assim sem perder sua identidade essencial, quando, em uma situação possível, existir desalinhado da motivação "certa." É desse modo que Jackson afirma a compatibilidade do internalismo motivacional e o cognitivismo moral: "a moral judgment is accompanied by the relevant pro-attitude … And we can accommodate this view by refusing to *call* something a *moral* belief unless it is accompanied by the relevant pro-attitude"[9].

É um erro, portanto, afirmar que uma concepção "robusta," isto é,

---

[8] O Satã de Milton diz: "So farewel Hope, and with Hope farewel Fear,//Farewel Remorse: all Good to me is lost;//Evil be thou my Good [...]" (MILTON 1667, (IV, 109-111).
[9] JACKSON 1998, p.161.

antissentimentalista, dos fatos morais *qua* fatos independentes de respostas subjetivas (*response-independent facts*) não pode coexistir com uma psicologia moral internalista, de tal forma que um realista moral estaria racionalmente obrigado a adotar uma visão fraca ou externalista da conexão entre juízos morais e motivações. David Brink parece cometer esse erro quando escreve:

> It is sometimes thought difficult for a realist to explain the practical character of morality. How can moral considerations influence conduct in the appropriate way if, as the realist seems to claim, they merely state facts?
> This might be a problem for the realist if one had to represent the connections between morality and motivation […] as internal, conceptual connections, since, it seems, purely cognitive states can be motivationally inert. […] An externalist moral psychology […] allows the realist to avoid this objection […][10].

Mais relevante para os propósitos do presente trabalho é a tese segundo a qual os sentimentalistas cometeriam estruturalmente o mesmo erro se argumentassem a favor da tese constitutiva a partir da tese motivacional. Eles estariam seguindo a sugestão que, como os sentimentos desempenham um papel motivacional e como esse papel supostamente reflete uma conexão interna entre estados motivacionais e pensamentos morais, esses pensamentos devem ser constituídos por sentimentos, no sentido em que os conceitos morais (os componentes dos pensamentos morais) ou as propriedades morais (os conteúdos dos pensamentos morais) são essencialmente dependentes de respostas (*response-dependent*), onde as respostas relevantes são evidentemente sentimentos. Como as considerações acima devem ter mostrado, essa sugestão é enganadora. Ademais, há algumas evidências que os sentimentalistas sobre juízos morais foram levados ao erro por essa sugestão. Antii Kauppinen (2012) afirma que os proponentes da versão não-cognitivista do sentimentalismo de juízo—a visão que os juízos morais são *constituídos* por emoções—apelam frequentemente a considerações relacionadas ao "Argumento a partir do internalismo de juízo," que

[10] BRINK 1989, p.8.

é essencialmente o mesmo argumento internalista a favor do não-cognitivismo descrito acima. Todavia, o argumento a favor da versão cognitiva do sentimentalismo de juízo — a visão que os juízos morais são constituídos por crenças sobre emoções — também depende parcialmente, de acordo com Kauppinen, da tese segundo a qual "it is not entirely implausible that beliefs with such content [beliefs about what merits moral sentiments] have an intimate link to motivation, either by themselves or by way of a necessary connection to desires"[11]. Ao tentar aparentemente inferir a tese constitutiva da tese motivacional, os sentimentalistas podem estar cometendo o erro complementar ao dos realistas *à la* Brink, que pensam que o comprometimento com conceitos ou propriedades morais independentes de resposta (*response-independent moral concepts or properties*) implica um comprometimento com uma conexão fraca ou externalista entre motivações e juízos morais.

É claro que a tese constituinte e a versão forte ou internalista da tese motivacional podem ser *ambas* verdadeiras. É um fato contingente que os sentimentalistas sobre o juízo moral também endossam a tese motivacional. E eles podem estar certos nos dois lados. Mas eles tornam sua posição objetável se não reconhecerem a necessidade de argumentos *independentes* para a tese constitutiva *e* a tese motivacional.

Referências

BRINK, David. *Moral realism and the foundations of ethics*. Cambridge, Mass.: Cambridge University Press, 1989.

DARWALL, Stephen. *Impartial reason*. Ithaca: Cornell University Press, 1983.

JACKSON, Frank. *From metaphysics to ethics. A defense of Conceptual Analysis*. Oxford: Oxford University Press, 1998.

KANT, Immanuel. *Fundamentação da Metafísica dos Costumes*. (Traduzido por Paulo Quintela). Lisboa: Edições 70, 1785/2005.

[11] KAUPPINEN 2012, p.12.

KAUPPINEN, Antti. “Sentimentalism”. In: *The International Encyclopedia of Ethics,* 2012. Available at: <http://tcd.academia.edu/AnttiKauppinen/Papers/1309568/Sentimentalism_International_Encyclopedia_of_Ethics_>.

MACKIE, J. L. *Ethics: Inventing right and wrong*. London: Penguin, 1977.

McNAUGHTON, David. *Moral vision: An introduction to ethics*. Oxford: Blackwell, 1988.

MILTON, J.*Paradise Lost*. London: S. Simmons, 1667.

PRINZ, Jesse. *The emotional construction of morals*. Oxford: Oxford University Press, 2007.

SMITH, Michael. *The moral problem*. Oxford: Blackwell, 1994.

TRESAN, Jon. “De dicto internalist cognitivism”. In: *Noûs* 40, p.143-165, 2006.

_____. “Metaethical internalism: Another neglected distinction” In: *The Journal of Ethics* 13, p.51-72, 2009a.

_____. “The challenge of communal internalism”. In: *The Journal of Value Inquiry* 43, p.179-199, 2009b.

11

ELEMENTOS DA METAÉTICA: COGNITIVISMO INTERNALISTA E A TEORIA VOLITIVA

Idia Laura Ferreira

Introdução

Ao fazer o retrospecto da discussão sobre motivação moral a partir dos problemas apontados pelas abordagens dominantes, é possível determinar por que teorias volitivas têm sido trazidas novamente ao debate e, a partir de novos conceitos sobre intenção e vontade, buscar novos rumos para uma teoria unificada que acomode de maneira plausível toda a complexidade envolvida na agência moral humana. Em consequência da centralização do debate nas teorias cognitivistas, de um lado, e não-cognitivistas, de outro, por razões que podem ser rastreadas nas discussões das últimas décadas, foi paulatinamente deixado de lado o enfoque na colaboração das noções de *intenção* e *vontade*, em detrimento das noções básicas de crenças e desejos. Quando reabriu o debate das teorias internalistas da agência, Gary Watson argumentou (i) que a reformulação da decisão (*making up one's mind*) é o *locus* primário da agência humana e (ii) que a vontade é uma instância genuína dos juízos práticos. Watson considera se deveríamos dizer que a vontade, e, portanto, a agência, tem uma existência além do domínio prático, uma existência também na esfera cognitiva. Afinal, argumenta que, juntamente ao fenômeno de "decidir-se a", o fenômeno "decidir que" também deveria ser classificado como um fenômeno ativo e, daí, como um modo de agência. Tendo como ponto de partida a articulação da distinção entre cognitivismo e não-cognitivismo moral e a questão psicológica sobre a natureza dos estados mentais expressos por enunciados morais, o trabalho busca, em primeiro lugar,

traçar o curso da discussão sobre o internalismo motivacional e, em segundo lugar, discutir uma reconceituação da noção de intenção que fuja à abordagem tradicional, para servir à nova noção de vontade como instância decisional ativa. Para isso, o trabalho mostra que há um hiato entre juízos morais e a ação moralmente motivada. O argumento central defende duas teses: (i) o internalismo motivacional se apoia em noções *ad hoc* sobre juízos morais, e (ii) a deliberação moral envolve instâncias que fogem à descrição comumente aceita pela teoria metaética apartada da teoria volitiva da ação, isto é, a descrição que reduz os estados mentais envolvidos na ação a crenças e desejos.

Grande parte da discussão metaética mais atual se situa no campo da psicologia moral. A discussão tem sido dominada por dois pontos de vista aparentemente excludentes: o cognitivismo e o não-cognitivismo. Cognitivistas e não-cognitivistas propõem explicações divergentes acerca do *status* dos juízos morais. A questão decisiva diz respeito ao tipo de estado mental expresso pelo proferimento moral. Para o cognitivista, o estado metal em questão é uma crença genuína, isto é, um estado mental de direção de ajuste mente/mundo que pode ser avaliado como verdadeiro ou falso. Crenças são estados intencionais de conteúdo proposicional: crenças representam o mundo ou almejam representar o mundo como ele é. Se a percepção do mundo revelar que a representação está errada, a crença deve ser abandonada e substituída por uma crença adequada ao mundo. A crença se ajusta ao mundo, pelo menos nos casos em que o sujeito da crença é racional. Outros estados mentais intencionais não se ajustam ao mundo. Eles têm uma direção de ajuste mundo/mente. O caso paradigmático aqui é o dos desejos. Não é irracional não abandonar ou substituir um desejo simplesmente por que seu conteúdo não reflete a realidade do mundo. O não-cognitivista afirma que o que é expresso por um juízo moral é um estado mental desse tipo.

Visões de internalismo moral podem ser de dois tipos: (i) internalismo quanto a juízos morais, que afirma a existência de conexões necessárias entre julgamentos morais, por um lado, e motivos, por outro; e (ii) internalismo quanto à motivação dada por obrigações morais; estas nos dariam necessariamente boas razões para a ação, isto é, razões objetivas que não podemos ignorar, sob pena de

irracionalidade. Essa distinção coloca, de um lado, o *internalismo motivacional de juízo* (ou simplesmente *internalismo motivacional*) e, de outro, o *internalismo de razões*. Dependendo do tipo de internalismo adotado, o cognitivista formula diferentes respostas ao desafio representado pelo argumento da psicologia moral.

O argumento da psicologia moral apresenta-se da seguinte maneira: Suponhamos que juízos morais expressem crenças (tese cognitivista). Estar motivado a fazer algo ou perseguir uma linha de ação é sempre uma questão de ter uma crença e um desejo, ou seja, eu estou motivada a fazer algo porque acredito que praticando essa ação alcançarei algo que desejo (teoria humiana da ação). Mas é um fato interno e necessário sobre qualquer agente que, se ele julga sinceramente que X é bom, no sentido moral, está motivado a perseguir o curso de ação para X (internalismo motivacional). Assim, se o juízo moral for uma crença, teria de ser uma crença que mantém uma conexão interna e necessária a um desejo: teria de ser uma verdade necessária que um agente que possui essa crença possuiria também *inter alia* um desejo (anexado). Segundo a teoria humiana da motivação, contudo, nenhuma crença é necessariamente conectada a um desejo. Como Hume afirmou, "crenças e desejos são existências distintas" e é impossível haver uma conexão necessária entre existências distintas. Assim, continua o argumento, não pode ser o caso que juízos morais sejam crenças. Juízos morais não são passíveis de avaliação como verdadeiros ou falsos (*truth-apt*), mas estados conativos de valor prático.

O internalismo motivacional desempenha um papel relevante nos argumentos a favor do não-cognitivismo e contra o cognitivismo. O internalismo é um aliado natural do não-cognitivismo, em geral, e do expressivismo, em particular, e, em princípio, pensam estes, um problema para o cognitivismo. O internalismo é considerado a premissa forte do argumento da psicologia moral (não-cognitivista) porque, antes de tudo, é adotado como um princípio que, juntamente com o princípio do não-factualismo, reduz a moralidade ao seu aspecto prático. Nesta teoria, o estado mental envolvido na moralidade ou está *necessariamente* conectado à motivação ou é identificado com ela: expressar um juízo moral *é* estar motivado a seguir o curso de ação apontado por ele. O internalismo motivacional é

sempre apresentado como a representação da natureza ou conteúdo dos estados mentais morais, devido aos quais estes são necessariamente conectados a motivações.

Há, porém, como aponta Jon Tresan (2006), duas leituras possíveis da necessidade envolvida no internalismo motivacional: a leitura *de re* do internalismo é a afirmação de que, se algo é uma crença moral, então essa crença não pode ser desacompanhada de motivação. O internalismo *de dicto*, por outro lado, é a afirmação de que, se algo é uma crença moral, então essa crença efetivamente está acompanhada de motivação. Esta afirmação permite a possibilidade de crenças que são efetivamente crenças morais virem a ser desacompanhadas de motivação e, neste caso, não serem crenças morais.

Tresan nos lembra que é o internalismo *de re* que apresenta um desafio para o cognitivismo. De acordo com o cognitivismo, crenças morais são crenças normais, e é, com certeza, possível qualquer crença normal estar acompanhada de motivação. Mas o problema é que, uma vez que crenças não são estados mentais motivacionais, toda e qualquer crença pode vir a ser desacompanhada de motivação, inclusive as morais. O internalismo *de dicto* não diz que a razão da conexão necessária entre crenças morais e motivação tem de dar-se porque crenças morais têm certo conteúdo ou são de uma natureza especial; a razão pode ser que nosso *conceito* de *crença* moral exige que qualquer crença moral seja *efetivamente* acompanhada de motivação.

Uma crença pode normalmente não estar acompanhada de motivação, mas uma crença do tipo acima, que é caracterizada como uma crença moral ou 'crença acompanhada de motivação para x' não pode. Esse tipo de internalismo pode ser descrito como internalismo condicional, porque existe um elemento condicional que faz parte do conceito que temos de juízo moral. Uma forma de explicar esse elemento condicional que faz parte do juízo baseado em, ou sendo, uma crença moral, é a tendência de usarmos os conceitos com extensões que sejam importantes ou nos interessem de alguma maneira e que, enfim, possam ser referidas diretamente. No caso de crenças morais, poderia ser uma explicação para tal elemento condicional a maneira de identificar, dentre outras, aquelas

crenças que são morais.

Em todas as variantes do argumento não-cognitivista, a premissa, ou premissas (no caso da versão apresentada neste texto), que contêm a afirmação do internalismo, como demonstrado acima, a necessidade motivacional interna é lida *de dicto.* A questão formal que surge é que no debate sobre o externalismo, por exemplo, a interpretação utilizada pelo não-cognitivista é a *de dicto*, quando somente a interpretação *de re* tornaria válido o argumento contra o externalismo (conexão contingente entre juízos e motivação), o colocando em xeque. A premissa relevante afirma que a crença moral não poderia deixar de ser motivadora; contudo, se a necessidade em questão é *de dicto*, a crença moral poderia não motivar e, no caso do argumento apresentado, motivações teriam de ser propriedades essenciais das crenças morais.

A forma comumente utilizada pelo externalista para responder ao internalismo é mostrar que a figura de um amoralista é concebível. Para explicitar o erro, ou a dificuldade contida na necessidade do internalismo quando lida *de re*, essa figura já clássica é o melhor exemplo. Quando o externalista concebe o amoralista, quer mostrar que não é impossível, como afirmam os internalistas motivacionais puros, que um agente possa, hipoteticamente, fazer juízos morais e não estar motivado a agir de acordo. A simples possibilidade de conceber um tal agente já é prova contra a leitura *de re* da necessidade internalista. E não é preciso que este agente seja real, basta que seja concebível. De fato, ao desafio representado pelo amoralista, o internalista responde que os juízos que ele tem não são juízos morais verdadeiros; mas essa é a resposta da leitura *de dicto* da necessidade internalista – para a resposta *de re* o amoralista teria que ser um sujeito inconcebível. Deste modo, a leitura *de re* da necessidade internalista está fora do debate, haja vista que não só concebemos a figura do amoralista, mas também nos perguntamos se não poderia ser, de fato, real.

O internalismo motivacional está assentado numa intuição amplamente aceita de que 'obrigações morais necessariamente motivam'; esta intuição traduz uma pressuposta ideia de prescritividade normativa característica da moralidade. O externalismo apenas nega a ligação juízo/motivação como sendo conceitual, para

ele a ligação é contingente, a distinção é conceitual. Mas defender a noção do amoralista não precisa necessariamente ser uma defesa do externalismo. Mais do que uma negação do internalismo, ou apenas isso, é mostrar que a questão não é conceitual, mas pode estar aberta à observação empírica. O internalismo simplesmente não pode ser visto como uma generalização universal para agentes racionais, como seus defensores afirmam. Brook Sadler, quando defende a noção do amoralista, aponta duas considerações que sugerem que não há uma verdade generalizável como a de que juízos morais necessitam certo tipo de motivação: (i) não existe essa verdade generalizável sobre a relação entre juízos morais e motivação porque, para diferentes agentes e diferentes momentos de um mesmo agente, pode haver variação para o caso que se apresenta; as motivações para um mesmo agente podem sofrer variações: um agente pode ser mais facilmente motivado por demandas morais pertinentes a certos aspectos de sua vida e menos motivado por outras demandas morais, não pode haver uma verdade generalizável sobre essa relação; (ii) é possível detectar um problema epistêmico na determinação de escolhas polarizadas em dois campos como o internalista e o externalista: se for perguntado a um agente por que agiu de certa maneira possivelmente não haverá uma resposta precisa ou definitiva, o agente pode ter múltiplas razões ou motivos para a sua ação e nenhuma noção clara de como elencá-las ou hierarquizá-las, e pode inclusive compreender sua ação como o resultado de um aglomerado de razões, algumas delas não morais.

Sadler aponta que há três tipos de questões que podemos imaginar que agentes podem vir a enfrentar quando deparados com uma obrigação moral (é moralmente certo ϕ em C): (1) Eu (tenho razões) *prefiro* fazer ϕ? (2) Eu tenho razões para *escolher* ϕ? (3) ϕ está *justificado*? Estas questões identificam ou enfatizam um aspecto diferencial de como o agente pode entender uma obrigação moral no contexto de outras considerações. Avaliar essas possibilidades significa analisar se na experiência moral há ou não espaço para agentes racionais cujas razões morais não motivam quando estão em conflito ou competindo com outras razões não morais e outras motivações. O que significa que, se há espaço para o amoralista em termos de deliberação como um processo complexo de razões e

motivações, então o internalismo não é, claramente, uma verdade universal para agentes racionais.

E por que, afinal de contas, o estado psicológico da motivação é utilizado pelos internalistas como critério de proficiência de termos morais? Segundo estes, os termos morais proferidos pelo amoralista devem ser entendidos entre aspas, ou seja, não são verdadeiros termos morais, uma vez que não estão acompanhados da devida motivação para agir de acordo; a menos que haja um argumento independente para a relação entre juízo moral e motivação, definir um pelo outro é apenas petição de princípio. A tentativa de fornecer esse argumento novamente apela a uma aparente fenomenologia da prática moral – o fato corriqueiramente observado de que um agente sofre uma mudança na motivação assim que ocorre uma mudança em seu juízo moral, ou seja, a mudança no juízo moral é seguida por uma mudança na motivação. E o desafio é que, a menos que o externalista apele para um desejo moral *de dicto* (uma disposição moral mais básica de fazer a coisa certa), não poderia explicar essa mudança na disposição moral (Sadler, 2003).

Mais do que um desafio ao internalismo, portanto, a figura do amoralista mostra que está aberta a questão sobre as noções que sustentam as posições padrão que seguem a teoria da ação baseada em crenças e desejos. Justamente porque tais noções estão calcadas em intuições, e nada mais.

Teoria volitiva da ação

A filosofia e a psicologia morais têm interesse em teorias da ação porque ações definem aqueles comportamentos que estão sob nosso controle, em oposição a comportamentos que, em algum sentido, simplesmente acontecem. Neste contexto amplo, a teoria da ação tem implicações na teoria da moralidade, na medida em que podemos buscar na primeira explicações que iluminem ou corroborem com a compreensão e as explicações da segunda. Historicamente, a teoria volitiva foi uma das teorias da ação que liderava o debate. Volição, no sentido mais simples, refere-se a um ato da vontade. Paulatinamente, esta teoria foi substituída pela teoria da ação chamada de teoria crença/desejo, e é assentada

nesta teoria que a metaética, e a psicologia moral, especificamente, vem se desenvolvendo nas últimas décadas.

Em consequência da centralização do debate nas teorias cognitivistas, de um lado, e não-cognitivistas, de outro, por razões que podem ser rastreadas nas discussões das últimas décadas, foi aos poucos deixado de lado o enfoque na colaboração das noções de *intenção* e *vontade*, ao menos como noções importantes *sui generis*. O debate cognitivismo/não-cognitivismo, nos moldes polarizados que conhecemos, parece ter chegado ao ponto de petições de princípio para ambos os lados. Essa impressão fica evidenciada justamente no debate sobre a motivação, como já foi demonstrado, quando depende da noção adotada como princípio, baseada em uma intuição, sobre o conceito de juízo moral, por um lado, e motivação, por outro, em que está assentada a recente discussão.

Embora alguns filósofos prefiram simplesmente negar sua existência, ao menos enquanto questão pertinente, a noção (e devida explicação) da chamada 'fraqueza da vontade' não pode ser descartada. Ao contrário, parece não só relevante como pertinente que essa falha na motivação seja compreendida à luz da metaética, sobretudo pelas implicações trazidas pela recente discussão metaética sobre a ação, e consequente agência moral, e sua explicação. Quando comparada à discussão clássica sobre *akrasia*, alguns filósofos defendem que a falha motivacional chamada fraqueza da vontade, e que por muito tempo foi identificada àquela, é potencialmente diferente, e sendo diversa pode apontar para outros desdobramentos. No primeiro caso, o conceito básico é que se trata do fenômeno de um agente agir *contra* seu melhor juízo, e, no segundo, no fenômeno de um agente falhar em *seguir* seu melhor juízo. Ambas as noções podem ser compreendidas complementarmente como pertencentes ao campo da discussão sobre a motivação, mas, assim como não são idênticas, não podem ser reduzidas ao âmbito restrito da questão motivacional.

Sobretudo para a metaética, a explicação dos casos de falha na motivação e ação morais serve para esclarecer também os casos de sucesso. Essa estratégia tem sido seguida por vários teóricos, e alguns desdobramentos apontam

para noções mais complexas, como intenção e volição. Os agentes racionais devem possuir um intelecto prático ou uma *capacidade deliberativa*, ou seja, formar juízos práticos sobre ações que sejam racionalmente justificados e, além disso, possuir uma *capacidade executiva* para aplicar esses julgamentos em suas ações (o que não é contraditório com o agir irracionalmente, posto que esta pode simplesmente ser vista como uma falha puramente executiva e não uma alternativa conceitual à agência). Mas, além dessas capacidades, o agente racional deve ser tido como possuidor da capacidade de estar motivado a agir através de decisões racionais, formando desse modo intenções para agir. As decisões e intenções terão de estar baseadas, a seu turno, em deliberações sobre a justificativa dessas ações. É neste ponto que se encontra a discussão apontada: é com esta capacidade de tomar decisões que a discussão sobre a faculdade da vontade sempre esteve relacionada.

Os teóricos que descrevem o aspecto prático da moralidade normalmente explicam esse aspecto pelo sucesso dos juízos morais para levar à ação correspondente. Visto por este ângulo, juízos morais são tomados como essencialmente motivadores. Essa perspectiva levou as conclusões sobre o entendimento da dinâmica da psicologia moral à noção de que o estado mental expresso nestes juízos é um estado caracteristicamente da classe dos desejos, como comentado anteriormente. Se a explicação da psicologia moral dependesse apenas dos casos de sucesso, as teorias não-cognitivistas teriam uma explicação direta. É por essa razão que casos de falha são importantes, porque neles se encontra a resposta sobre qual, de fato, deve ser o tipo do estado mental prioritariamente envolvido nos juízos que fazemos sobre questões morais e, assim, o tipo de papel que a motivação realmente exerce. Não se trata de investigar empiricamente essas falhas; para a metaética, falhas conceitualmente concebíveis cumprem o propósito da investigação, embora todos nós sejamos aptos a pensar em falhas que cometemos todos os dias sobre nossos juízos morais.

Vimos anteriormente as duas principais estratégias da teoria metaética para explicar os casos de sucesso motivacional, o internalismo e o externalismo, e suas versões para explicar casos de falha. Será introduzida agora a interpretação para a

fraqueza da vontade como falha motivacional radical contida na explicação de 'incapacidade volitiva' de Gary Watson (2004). Esta perspectiva expõe uma interpretação que parece desnudar aspectos que não são contemplados por teorias internalistas que identificam juízos morais a estados conativos, nem por formas mais sofisticadas do expressivismo, como o de Allan Gibbard (2003).

A incapacidade volitiva descrita por Watson não é nem uma falha de julgamento nem de autocontrole. É um tipo de falha que nos força a distinguir entre juízo e vontade, pois alguém é incapacitado volitivamente quando seu juízo moral autodirecionado não resulta em uma intenção correspondente. Diferentemente do fraco da vontade que não resiste a uma tentação, por exemplo, indo de encontro a seu melhor juízo (poderíamos pensar em um fumante que precisa parar de fumar ou um obeso mórbido que precisa abrir mão de doces, que adota/subscreve essa crença mas falha em segui-la), a falha descrita por Watson envolve a incapacidade de formar o próprio juízo (o fumante que, contra seu melhor interesse e crença, nem planeja parar de fumar). O conflito interno apontado por esses dois tipos de falhas representa dois tipos distintos de falha motivacional: i) o hiato entre vontade e ação; e ii) o hiato entre juízo e vontade (Kauppinen, 2006.)

Casos de falha motivacional podem ser de vários tipos, e diferenciar alguns deles pode ser ilustrativo para mostrar a lacuna entre vontade e ação. Primeiro, podemos apontar para aquela falha que surge quando alguém abandona uma primeira intenção original e forma outra, como ao deixar de lado um plano e adotar outro, sem uma mudança em sua mente sobre razões para agir. Ou, descrito em termos causais, representada pelo agente que se propõe a fazer algo mas um desejo de fazer outra coisa é causalmente mais forte, levando-o a formar uma intenção proximal[1] (estado psicológico que causa e mantém os movimentos corporais constitutivos das ações intencionais) diferente e, assim, a fazer essa outra coisa intencionalmente, a menos que algum obstáculo externo o impeça. Ambas as descrições encaixam-se na noção de fraqueza da vontade, a rigor, na noção que envolve uma vontade que pode ser forte ou fraca, e não apenas na

[1] Ver a noção de intenção proximal de Alfred Mele (1992).

descrição frequentemente adotada de agir contra o juízo sobre o que é melhor. É importante reconhecer os dois fatores presentes nesta descrição, porque, ao lado da noção comum de que é regra seguirmos o que julgamos ser o melhor, também há a vontade de fazê-lo envolvida. O agente fraco da vontade não apenas age contra seu melhor juízo, mas também contra sua vontade, e essa afirmação pareceria uma tautologia, não fosse a distinção cuidadosa entre juízo e vontade. Por essa razão, a descrição de fraqueza da vontade é mais bem empregada apenas nos casos em que o agente já formou uma vontade, ou seja, já se propôs a determinada ação.

Há o tipo de falha da vontade muito mencionada na literatura filosófica, que é caracterizada por torpor ou indiferença, onde o agente propõe-se a fazer algo mas é incapaz de implementar qualquer esforço para alcançar esse objetivo. Esta condição está mais ligada às noções de patologias como depressão do que propriamente de fraqueza da vontade, ao menos no sentido conceitual que é adotado na metaética. E há o caso do agente que faz o que se propôs a fazer, mas a ação é de fato causada por outra fonte de motivação do que sua aparente proposta voluntária; esse tipo de "causação desviante" que é uma falha de autocontrole é relevante numa avaliação do agente, quando surge a questão de o que ele faria em uma situação contrafactual onde o desejo causalmente eficaz estivesse ausente. Todas essas situações são exemplos em que, em um curso de ação, o agente formou uma *intenção* de agir de um certo modo, um compromisso que implica a formação de uma certa 'intenção proximal' causalmente efetiva no momento da ação, e esse compromisso tem uma 'direção de ajuste' mundo/mente (como um estado cognitivo tem) e um papel causal funcional característico na condução da intenção proximal (Mele, 1992.)

Essa descrição traduz ideias defendidas por Richard Holton, sobre a *fraqueza da vontade* ser mais uma questão que se refere à desistência das próprias intenções do que a agir contra o melhor juízo, e por Alfred Mele, sobre "ter uma atitude executiva a respeito de um plano" (Mele 2003) como ideia de *intenção*. Se, por um lado, a fraqueza da vontade nos mostra que valorar e desejar são separáveis, ao menos no modo, a incapacidade volitiva mostra, por outro lado, que

valorar e estados conativos, tais como planejar, são modalmente separados e não podem ser identificados, como os expressivistas afirmam. Desta forma, este é mais um indício de que a conexão interna entre juízo moral e motivação deve ser explicada em termos de crenças firmadas racionalmente, como pregam os internalistas cognitivistas.

Referências

BRINK, D. *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press, 1989.

GIBBARD, A. *Thinking how to live*. Cambridge, Mass.: Harvard University Press, 2003.

HOLTON, R. *Willing, wanting, waiting*. Oxford: Clarendon Press, 2009.

_____. "Intention and weakness of will". In: *Journal of Philosophy* 96, p.241-262, 1999.

KAUPPINEN, A. "Moral Judgment and Volitional Incapacity" In: O'Rourke, M. (org.) (2010). *Topics in Contemporary Philosophy* vol.7. Massachusetts: MIT Press, 2006.

MELE, A. *Springs of action*. New York: Oxford University Press, 1992.

_____. *Motivation and agency*. New York: Oxford University Press, 2003.

SADLER, B. J. "The possibility of amoralism: A defense against internalism". In: *Philosophy* 78, p.63-78, 2003.

SMITH, M. *The moral problem*. Oxford: Blackwell, 1994.

_____.(Org.). *Meta-ethics*. Aldershot: Dartmouth Publishing Group, 1995.

_____."Moral realism". In: H. LaFollette (Org.), *The Blackwell guide to ethical theory*. Oxford: Blackwell, 2000, p.15-37.

TENENBAUM, S. “Direction of fit and motivational cognitivism”. In: Shafer-Landau, R. (org.) *Oxford Studies in Metaethics,* 1. New York: Oxford University Press, 2006, p.235-264**.**

TRESAN, J. “De dicto internalist cognitivism”. In: *Noûs* 40, p.143-165, 2006.

WALLACE, R. J. *Normativity & the will: Selected essays on moral psychology and practical reason*. Oxford: Clarendon Press, 2006.

WATSON, G. *Agency and answerability: Selected essays*. Oxford: Clarendon Press, 2004.

WILLIAMS, B. “Internal and external reasons”. In: Bernard Williams, *Moral luck*. Cambridge: Cambridge University Press, 1981, p.101-113.

12

MEMÓRIA, AUTOCOMPREENSÃO E AGÊNCIA*

Marina Oshana

1. Do que (e do que não) trata este capítulo

A tarefa deste capítulo é determinar quais aspectos da identidade ou individualidade de uma pessoa tem de lhe estar disponíveis, e de que maneira precisam lhe estar disponíveis, a fim de que possa operar como um agente autogovernado. Para fins desta discussão, entenderei a agência como um estado razoavelmente robusto. Assim entendida, um agente é um ser que delibera, reflete, decide, planeja e produz estados de coisas. Idealmente, um agente estabelece para si padrões de comportamento e está no controle de si, diferentemente de um paciente, ou de um ser que é influenciado, manipulado ou levado a adotar vários estados. Um paciente pode bem ter desejos e objetivos, e uma concepção própria do bem que deseja realizar. Ele pode ter a boa sorte de vê-los realizados, talvez em seu nome, por outros. Mas somente *enquanto* agente uma pessoa está em condições de realizar esses objetivos por seu próprio esforço. Ser um agente *autogovernado* exige uma característica adicional. Essa característica é a capacidade de representar a si mesmo no mundo por seus próprios meios – e fazê-lo por decreto próprio. Ser um agente autogovernado é antecipar as próprias intenções como condutoras à ação mediante um comportamento de automonitoramento. No que segue, sugerirei que essa forma agência robusta está ausente das vidas de pessoas que sofrem de um espectro de distúrbios da memória e de senilidade.

Como advertência, devo salientar que a posição que estou esboçando não

---

*Traduzido por Alessandra Lessa dos Santos e Flavio Williges.

pretende depender de comprometimentos metafísicos fortes acerca do *eu*. Quero manter certa distância do debate entre os defensores da tese de que o eu é uma das categorias ontológicas fundamentais, o sujeito de nossas vidas mentais interiores e a fortaleza interior do livre agir, por um lado, e os críticos dessas concepções, por outro, como Hume e aqueles cujas posições sobre o assunto refletem sua influência[1]. Ao mesmo tempo, considero seriamente que há algo a que nos referimos sensatamente como "o eu (self) que alguém é" e considero que isso não pode ser equiparado conceitualmente ao ser humano que alguém é ou à vontade que alguém manifesta em escolhas e ações. "O ego que alguém é" é qual(is)quer parte(s) de uma economia psíquica, corporificada e social que cumpre(m) o papel de sujeito prático do autogoverno.

Por razões similares, é importante que eu afaste a discussão empreendida neste capítulo de alegações sobre a metafísica da identidade e também de alegações sobre o interesse especial que as pessoas têm pelas frações de tempo de suas vidas das quais estão conscientes[2]. A razão é a seguinte: considere a questão da metafísica da identidade. Suponhamos que a totalidade (ou uma quantidade substancial) de informações no banco de memórias de uma pessoa A em t1 seja apagada. (Estipularei que os meios pelos quais isso ocorre são irrelevantes.) Sem que A saiba, informações novas são substituídas em t2. O ser em t2 – vamos chamá-lo de B – não experiencia a perda de memória. Mais precisamente, suas memórias são "como-se" um certo conjunto de experiências fossem dele próprio, quando de fato, não são – elas foram inteiramente fabricadas. Agora, parece sensato supor que a veracidade da memória é certamente importante para a preservação da identidade. Na medida em que a memória é *tudo* o que importa à identidade ou que a memória, nas suas várias encarnações, é essencial para identidade[3], parece correto dizer que, nesse caso, a pessoa A deixou abruptamente de existir em t1 e uma nova pessoa, B, assumiu

[1] Agradeço a Richard Kraut por instar-me a tornar esse ponto explícito.

[2] Henry Richardson e Gideon Rosen insistiram nesses pontos; Michael Bratman notou que essas questões são tangenciais em relação àquelas que são centrais a esse artigo.

[3] Ver a discussão dos tipos de memória abaixo.

ininterruptamente seu lugar em t2. Essa substituição não provoca qualquer rebaixamento nas qualificações de B como um agente, é claro, e as ações de B poderiam ainda parecerem congruentes com as de A. Mas elas não fariam sentido como uma continuidade da vida do mesmo indivíduo, nem poderiam contar como expressões de atividades que estão sob o controle de um agente temporalmente contínuo. Isso não precisa ser mencionado como uma deficiência ou perda na vida de B, na medida em que não tenha nenhum efeito mensurável sobre a atividade prática de B. Mas parece correto dizer que o conteúdo mental necessário para a agência autogovernada tal como a descrevi está, nesse caso, comprometida, porque, o que quer que B faça, suas escolhas e ações não são tais que representam a continuidade de uma concepção de si, e os planos e projetos que essa concepção de si envolve. Como agentes autogovernados, operamos na expectativa de que o nosso comportamento carregue a marca de um ser temporalmente estendido, capaz de dar sentido à vida presente em termos da trajetória que o levou a esse ponto. Operamos na expectativa de que seremos capazes de justificar comportamentos subsequentes como contínuos com uma concepção verídica de nós mesmos.

Nada disso deve sugerir que questões de real importância não dependem da conservação da nossa identidade como agentes temporalmente estendidos. Entre essas questões, existe uma preocupação prática com a qualidade de nossas vidas futuras, e como planejamos nossas vidas é uma função daquilo com que nos importamos e do que prevemos para nós mesmos. Se me dizem, em t1, que serei torturada em algum momento futuro, mas que qualquer recordação dessa experiência será erradicada de meu banco de memórias, de modo que a continuidade psicológica não seja preservada, tenho razões para temer essa experiência futura? Parece evidente que tenho algumas boas razões para temer o que irá acontecer. Talvez essa preocupação seja uma perda de tempo e fosse mais apropriado se eu fosse configurada de um modo tal que minhas preocupações fossem redirecionadas para um futuro psicologicamente contínuo com o presente. Mas, mesmo se essa preocupação for justificada, ela só levanta questões pertinentes àquelas que são o foco deste capítulo somente na medida em que gera

(ou sinaliza) problemas para a agência prática e para a atividade de planejamento. Sendo mais precisa, uma expressão natural do meu autogoverno será o fato de que planejo as experiências às quais o meu corpo pode vir a se submeter no futuro. Mas o fato de que o faço – o fato de que certos estados são objeto de minha preocupação autointeressada – não estabelece, por si mesmo, que eu esteja planejando a vida da mesma pessoa que sou hoje e não estabelece, certamente, que a mesma pessoa que sou agora irá, de fato, continuar. Disso não se segue nada sobre a situação metafísica de tais estados.

2. Arquivos de identidade

Comecemos supondo que existe algo chamado "o eu". O eu experimenta o mundo através das estruturas sensoriais de uma pessoa e tem um efeito sobre o mundo por meio dos seus sistemas motores. Mas certamente o eu que se identifica com um ser particular e que opera como o centro conceitual da agência envolve mais do que isso. Para motivar esse pensamento, farei uso da ideia de John Perry de um arquivo da própria identidade. Perry oferece uma crítica a Thomas Nagel, que afirma que a moralidade exige que pensemos sobre nós mesmos de forma impessoal, como apenas uma pessoa entre muitas[4]. O argumento de Perry é que, embora esse ponto de vista impessoal possa ter importância moral, o pensamento indexical é um componente essencial e ineliminável de nossa caixa de ferramentas teóricas, sem o qual estaríamos mal equipados para nos envolvermos com o mundo real e nos engajarmos no raciocínio prático.

Seguindo Perry, suponhamos que o termo "Marina Oshana" refere-se a um ocupante do mundo e que um eu particular adota a sua perspectiva. O eu pode ter uma visão de si mesmo como um entre vários outros, certamente. Ou seja, Marina Oshana está entre as coisas que aparecem no mundo para Marina Oshana:

> Na minha concepção impessoal do mundo, há uma representação de [Marina Oshana], uma representação inteiramente completa e robusta, uma

[4] Thomas NAGEL, *The view from nowhere*. Oxford University Press, 1986.

> vez que sei muito sobre ela […]. Mas sua situação, no âmbito da representação impessoal, está no mesmo nível que minhas representações de quem quer que seja. Normalmente tenho outra representação bastante intrusiva de mim mesma, uma representação que eu poderia manter mesmo se esquecesse quem fui, que se baseia em informações atuais sobre mim mesma que posso adquirir através de sentimentos e percepções, e que está ligada à palavra "eu" . Mas nesse ânimo filosófico [de Nagel], essa representação fica atenuada; eu isolo a maior parte dessas informações e foco na concepção impessoal. Mas não posso romper totalmente a conexão, pois enquanto o que se concebe pode ser objetivo e impessoal, a concepção mental propriamente dita pertence a apenas uma das pessoas nela representadas. Essa pessoa tem um modo especial de lidar com ela, o qual lhe permite pensar nela como *isto*[5].

Perry oferece uma história bem complicada sobre como ocorre o autorreconhecimento quando "informações representadas objetivamente sobre pessoas em várias posições em relação a mim" são anexadas aos "papéis relativos ao agente que ocupam"[6]. Grosso modo, a concepção de Perry é que, quando as informações são reunidas no "arquivo temporário [*buffer*] da concepção de si" do indivíduo, e postas em uso de um modo particular, as informações passam a ser reconhecidas pelo indivíduo como autorrepresentativas. O conteúdo do arquivo passa a expressar não apenas uma relação de identidade, mas uma concepção de si próprio. A abordagem de Perry da gênese da autocompreensão fornece um ponto de partida útil para uma apreciação da agência. Mas ela não será suficiente para este projeto. Para começar, muita coisa está incluída num arquivo de autoconhecimento que é de pouca importância para a agência. Minha concepção de mim mesma inclui os fatos de que nasci em um determinado ano, morei em certa residência durante a minha infância, e de que não gosto mais de *sundaes* de caramelo, nenhum dos quais me parece ser um fato do qual a agência

---

[5] John PERRY, "The Sense of Identity". In: Perry, *Identity, personal identity, and the self*. Indianapolis: Hackett Publishing, 2002, p.217-218.
[6] 2002, p.228.

e o autogoverno dependem. Não estou defendendo ou comprando a história de Perry em sua totalidade. Estou apenas interessada na ideia de que uma parte das informações que geram asserções sobre a própria identidade têm de ser adquiridas e postas em uso de maneiras particulares, se for para essas informações servirem para o autogoverno agencial.

Os aspectos dos arquivos do estilo de Perry nos quais estou interessada são os que embasam a agência. Dado que nem todos os fatos sobre a identidade de uma pessoa situados no arquivo da concepção de si são cruciais para sua agência, precisamos começar perguntando quais informações em primeira pessoa, entre a massa de fatos incluídos em um arquivo da concepção de si, são vitais. Obviamente, esses fatos incluem informações em primeira pessoa sobre as crenças, valores, disposições, relações, história de uma pessoa, e assim por diante. Mas de algum modo essa informação tem de permitir que se trace uma distinção entre os fatos objetivos sobre o mundo que fornecem as condições de verdade para o enunciado "A mulher no espelho é Marina Oshana e ela acredita que X e valoriza Y" e aquelas condições que tornam as palavras "Eu sou Marina Oshana e eu acredito que X e valorizo Y" verdadeiras quando as enuncio. O autorreconhecimento é um fenômeno essencialmente indexical[7]. É reconhecer-me como eu mesma; não é reconhecer Marina Oshana. Além de qualidades práticas bastante genéricas que de fato têm de estar presentes para que uma pessoa seja um agente autogovernado, tal como a qualidade de ser um deliberador minimamente racional, habilidades e uma compreensão autorreflexivas de uma variedade mais indexical têm de estar ao alcance do suposto agente. Uma habilidade, capacidade ou proficiência é autorreflexiva quando é direcionada ao eu pelo eu. A compreensão é indexical quando indica algum estado; nesse caso, algum estado do agente.

Quero explorar a ideia de que os elementos relevantes da individualidade, quando focamos em questões sobre a agência, são aqueles que tornam o autorreconhecimento possível. Mais precisamente, os tipos relevantes de

[7] John PERRY, "The problem of the essential indexical". *Noûs* 13.1, p.3-21, 1979.

habilidades e informações de primeira pessoa são aqueles que permitem a alguém *pensar em si mesmo como um agente e conduzir a si mesmo como um agente que possui uma identidade própria*. Para que isso aconteça, a informação tem de estar conectada à pessoa de um modo particular. Comecemos notando que certos aspectos do arquivo de identidade de uma pessoa estão, de fato, sujeitos à direção reflexiva e podem ser cruciais para a agência, tal como a defini. Por "sujeitos à direção reflexiva", quero dizer que nós decidimos e explicamos nossa atividade com base em fatores como esses. Certos aspectos do arquivo de identidade fornecem à pessoa razões para o que ela faz, e são razões que ela considera conhecidas. Estas operam como razões por meio das quais atribuímos o nosso comportamento a nós mesmos, tanto no presente quanto em tempos passados, devido a conexões causais com nossa própria atividade. Tipicamente, os elementos do eu sujeitos à direção reflexiva são fenômenos que uma pessoa pode facilmente acessar e reter na memória, e são de um tipo que influencia os compromissos públicos da pessoa, afetando os motivos pessoais para a ação e o comportamento.

3. Memória

A concepção ortodoxa tem sido que memórias autobiográficas, consistindo em recordações explícitas de episódios passados da própria vida, são centrais para suprir os elementos constitutivos do eu. Eu endosso essa concepção como uma concepção que enuncia condições necessárias, porém insuficientes, para a agência. A razão pela qual a memória explícita episódica é insuficiente para a agência é suficientemente manifesta. A memória episódica corresponde a eventos específicos. É geralmente declarativa e proposicional, envolvendo crenças sobre algum fenômeno experienciado conscientemente pela pessoa. Mas a memória implícita, ou a memória sem recordação explícita, também é um fator na autobiografia. Uma variedade de memória implícita é de natureza semântica ou conceitual. A memória semântica serve, com frequência, para representar atributos gerais de pessoas e experiências. É uma memória que deriva do que afetou uma pessoa em sua experiência passada. A memória semântica de uma emoção

associada a uma experiência – de frustração ou dor, por exemplo, ou a memória semântica de uma qualidade particular associada a uma pessoa ou a um evento – bom humor, ou prazer gustativo, ou dos maneirismos e convenções distintas de certas funções, tais como a de uma garçonete do Hooters, por exemplo, é muitas vezes implícita. A memória implícita também inclui a memória procedimental de tipo "como-fazer", o tipo que retemos em nosso saber como andar de bicicleta ou fritar um ovo, por exemplo. Tudo isso pode figurar com vários graus de saliência na memória autobiográfica e, assim, todas podem embasar nossa aptidão como agentes dotados de identidades próprias.

As informações proposicionais sobre o ambiente que uma pessoa acumulou na memória juntamente com as habilidades aprendidas lhe permitem exercer controle sobre suas circunstâncias tal como a agência prática exige. A questão sobre que tipo de memória ou que combinação desses tipos de memória é crucial para a agência é mais difícil de resolver. E também fica em aberto a questão sobre que tipo de coisas sobre si mesmo alguém tem de lembrar – qual tem de ser o conteúdo dos estados de memória – para que se sinta no controle de sua vida. Pode-se progredir em ambas as questões examinando-se a situação de indivíduos acometidos por várias manifestações no espectro de perda de memória. O que faremos com a individualidade agencial de pessoas para as quais elementos de sua história se perderam pela falta de informações autobiográficos, proposicionais ou implícitas previamente presentes?

Considere o caso de Henry Gustav Molaison – o famoso "H.M.", cujo desenvolvimento de amnésia anterógrada grave (uma incapacidade de formar memórias de longo prazo), após uma cirurgia cerebral experimental para tratar uma epilepsia grave, tornou-o objeto de estudo médico por cinquenta e cinco anos, até a sua morte em dezembro de 2008, aos oitenta e dois anos[8]. Argumentei, em

[8] O estudo seminal de H.M. foi conduzido por Brenda Milner da McGill University. O estudo de Milner envolvia uma tarefa de coordenação motora em que H.M. aprendia a traçar uma linha entre dois contornos de uma estrela de cinco pontas refletida num espelho. Em cada ocasião dos três dias do estudo, a tarefa impressionou H.M. como uma experiência inteiramente nova. E ainda assim ele se tornou habilidoso com a prática. "As descobertas levaram Milner a especular que certas habilidades

outro lugar, que a amnésia de Molaison deixou-o com um eu tão efêmero a ponto de ser inadequado para agência robusta[9]. Ainda tenho essa concepção, mas creio que ela precisa de algum esclarecimento. Deixe-me esmiuçar a natureza da amnésia de Molaison. Para começar, Molaison não se esqueceu de si mesmo, no sentido rudimentar de erros no reconhecimento físico, pois era capaz de reconhecer que a imagem da pessoa refletida no espelho era dele mesmo. Ele era capaz de diferenciar entre si mesmo e os outros, e compreendia conceitos reflexivos e de primeira pessoa. Apesar de sua incapacidade de reter informações novas adquiridas, Molaison preservou algumas memórias autobiográficas episódicas e semânticas prévias, bem como um corpo de valores, preferências e idiossincrasias persistentes.

Algumas dessas memórias exibiam um alto grau de saliência em suas relações interpessoais. Um observador poderia citá-las como indício de que Molaison continuava a ter propósitos gerais e mesmo planos que estruturassem suas ações[10]. Por exemplo, Molaison permaneceu genial e articulado, e interessado na vida daqueles que conheceu. Era dedicado às palavras cruzadas e foi um leitor ávido de jornais. Era capaz de orientar seus dias em torno de rotinas, como preparar almoço, cortar a grama e assistir televisão. E, estranhamente, era

---

motoras podem se desenvolver independentemente do sistema no lobo temporal médio. A grande descoberta provou que o cérebro não é simplesmente governado por um sistema solitário de memória, um conceito revolucionário nos anos 1950". O trabalho de Milner sobre a memoria rendeu-lhe o apelido de "fundadora da neurociência cognitiva" (Um petisco: Milner, britânica de nascimento, ganhou uma bolsa em Cambridge, onde pretendia estudar matemática. Percebendo que "não queria ser uma matemática medíocre que apenas arranha", ela cogitou mudar para a filosofia. "Mas as pessoas em Cambridge disseram 'Brenda, você não precisa ganhar a vida, afinal? Bem, ninguém ganha a vida com filosofia'"). *Profile: Brenda Milner*, McGill Reporter, volume 40: 2007-2008, December 6, 2007. Disponível em: <http://www.mcgill.ca/reporter/40/08/milner/>.

[9] OSHANA, M, *The importance of how we see ourselves: Self-identity and responsible agency*. Lexington Books, Rowman and Littlefield: Lanham, MD, 2010.

[10] Como faz William HIRST. Ver "The remembered self in amnesies" (sic). In: *The remembering self: Construction and accuracy in the self-narrative*. Ed. Ulric Neisser e Robyn Fivush, Cambridge, U.K: Cambridge University Press, 1994, p.252-277.

capaz de formar novas memórias explícitas quando o conteúdo possuía alguma ressonância emocional especial. Por exemplo, ele era capaz de dar detalhes únicos, detalhes que permitiam a identificação de um grupo seleto de pessoas famosas. Ele descrevia John Glenn como "o primeiro operador de foguetes", e Lee Harvey Oswald como "o homem que assassinou o presidente". Ele disse, de Liza Minnelli, "Ela é cantora e também dançarina" e era tão aficcionado pela comédia de televisão *All in the family* que sabia que Archie Bunker se referia ao seu genro como 'Cabeça oca'. Conklin descreve isso como "surpreendente" e conjectura que se poderia traçar "um componente emocional" do conhecimento de Molaison sobre essas pessoas "porque eram pessoas de quem gostava ou que estavam associadas a um evento violento, como o assassinato de Kennedy. Acredito que esse processamento extra do componente emocional fez com que se fixasse melhor em sua memória"[11]. Talvez mais incisivamente, Molaison tinha uma sensação clara e persistente de si mesmo como alguém incapaz de guardar as experiências na mente por qualquer período de tempo, bem como do que o futuro lhe reservaria em consequência disso[12]. Ele exibia o tipo de interesse prático em primeira pessoa no bem-estar de seu eu futuro que é característico de pessoas normais.

---

[11] Corkin relata que "O que [Molaison] não podia fazer era dizer o que ocorreu em um tempo e lugar particular. Não conseguia dizer: 'Lembro que no meu aniversário de 10 anos derramei chocolate em minhas calças brancas e minha mãe ficou furiosa comigo'. Nós tentamos, tentamos e tentamos obter dele essas memórias episódicas específicas e detalhadas – algo que ocorreu nas férias, aniversário ou o onde quer que fosse. Ele não podia oferecer uma única memória episódica, com uma exceção: em um de seus aniversários, [ele lembrava] ter embarcado em um pequeno avião e voado ao redor de Hartford. Obviamente isso tinha um enorme impacto emocional sobre ele". CORKIN, "*The man who couldn't remember*". Disponível em: <http://www.pbs.org/wgbh/nova/body/corkin-hm-memory.html>.

[12] De fato, Molaison era impaciente com que seu cérebro fosse preservado depois de morte para estudos científicos. Ver a citação de Milner em <http://www.guardian.co.uk/science/2009/feb/05/obitury-henry-molaison>. Acesso em 9 de setembro de 2009. Ver também Dr. Suzanne Corkin em "The man who couldn't remember", entrevistada em fevereiro de 2009 por Sarah Holt, produtora de "How memory works", NOVA Online. Disponível em: <http://www.pbs.org/wgbh/nova/body/corkin-hm-memory.html>.

Ainda assim, não estou convencida de que essas regularidades sirvam como indícios de atividade de planejamento, nem de que isso baste como prova do tipo de personalidade sobre a qual repousa a agência. O caso de Molaison sugere que concepções como a de Perry que tratam um estoque de conhecimento em primeira pessoa como a base da individualidade são inadequadas para a agência. O uso gramatical correto de "eu" é insuficiente para ter um sentido do eu ou para uma capacidade de agência. Muitos dos elementos essencialmente indexicais do arquivo da concepção de si de uma pessoa estão disponíveis mesmo para pessoas com danos agudos na memória. Henry Molaison possuía um estoque razoavelmente extenso de conhecimento em primeira pessoa, ao mesmo tempo em que aspectos-chave de sua individualidade tinham se perdido. Penso que a lição é que as qualidades da individualidade que estão disponíveis para as pessoas e que consolidam suas atividades de agentes são aquelas que pertencem ao indivíduo de um modo especial, sendo esse modo especial marcado pela forma de acesso e pelo modo de apresentação. Notavelmente, os danos de Molaison roubaram-lhe uma variedade importante de autorreconhecimento transtemporal, a saber, o autorreconhecimento transtemporal de um tipo que exige que coloquemos a nós mesmos no passado e que esse pensamento de autorreconhecimento persista para nós como somos agora, com planos que pretendemos vir a desfrutar no futuro[13]. Quero sugerir que a agência necessária para o autogoverno e para a responsabilidade exige essa variedade rica de autorreconhecimento transtemporal e, portanto, uma habilidade de reconhecer a si próprio de maneiras menos restritivas e mundanas do que as que estavam ao alcance de Molaison.

A ficção também fornece retratos vívidos de uma memória corrompida. Uma das minhas favoritas é a de Leonard Shelby, o protagonista do filme *Memento*, de Christopher Nolan. Após o assassinato de sua esposa, Shelby desenvolveu amnésia anterógrada. Assim como ocorreu com Molaison, Shelby não perdeu

[13] Ver Christine KORSGAARD, "Personal identity and the unity of agency: A Kantian response to Parfit". *Philosophy and Public Affairs* 18.2 (1989). Reimpresso em D. Kolak e R. Martin (eds.), *Self & identity*, New York: Macmillan, 1991, p.326.

certas memórias de identidade salientes. Ele sabe seu nome, tem familiaridade com os nomes e usos da maioria dos objetos, e tem impressões intensas, embora limitadas, de sua esposa falecida e de capítulos de sua vida. Ele não perdeu de vista o fim primordial que deseja realizar (a saber, descobrir quem assassinou sua esposa). Além disso, a memória procedimental de Shelby não foi danificada. Ele ainda pode dirigir um carro, por exemplo, e adquiriu a habilidade de se tatuar. Ele parece estar consciente do eu que era antes do início de sua amnésia e é capaz de representar a si mesmo em primeira pessoa, embora a lucidez de sua concepção de si e a exatidão dos pensamentos que pode ter sobre si mesmo sejam duvidosas, assim como eram para Molaison.

No entanto, a memória semântica de Leonard Shelby foi prejudicada, com o resultado de que não é capaz de detectar e, assim, de relatar de modo confiável alterações em sua personalidade. Também não é capaz de decifrar as motivações dos outros para períodos extensos de tempo, embora, estranhamente, possa recordar, em certa medida, a identidade das outras pessoas. "Teddy" e "Natalie" lhe são familiares, mas nunca está completamente certo do que estão fazendo em sua vida.

4. Agência

O que danos à memória e, mais amplamente, à cognição, tais como a demência, interrompem? O que Molaison e Shelby (bem como minha sogra senil) não têm e que é necessário para a agência plena? As pessoas com amnésia anterógrada (e outras formas de danos cognitivos) são menos capazes de adquirir conhecimento de suas próprias ações? O que é preciso experienciar e de que é preciso estar consciente para se ter a experiência da própria agência? Que forma essa experiência precisa ter?

Num nível bem básico, Molaison e Shelby são agentes. Eles têm estados qualitativos do tipo que é "parte de uma unidade coerente que é constitutiva de mim mesmo e experienciado como eu mesmo"[14]. Nenhum deles (incluindo minha

[14] John SEARLE, "Review of Self comes to mind: Constructing the conscious brain by Antonio

sogra, pelo menos na maior parte do tempo) sofre do transtorno dissociativo de identidade. A agência básica exige capacidades ativas, tais como decidir-se, e isso é algo que Molaison e Shelby podem fazer, mesmo que não possam lembrar o que resolveram fazer. Uma expectativa mais exigente é que os agentes exerçam controle sobre suas ações; o controle exige posse, e posse exige consciência de que o ato está ocorrendo e é um ato de si próprio. Com efeito, boa parte do comportamento de Molaison e Shelby conforma-se à concepção de senso comum da ação intencional. A ação intencional exige o conhecimento da própria agência. Conhecer a própria agência significa estar ciente do que se está fazendo ao agir, saber se o que se está fazendo está em conformidade com o que se pretende fazer, e saber o que se deveria estar fazendo, dadas essas intenções[15]. Supomos que a ação de uma pessoa é menos sua quando não está consciente ou não é capaz de tornar-se consciente do que pretende fazer. Como Anne Newstead escreve,

> [a] ideia é que, enquanto alguém está agindo intencionalmente, terá (potencialmente) acesso à própria intenção e interpretará a ação à luz dessa intenção. […] Os movimentos que alguém empreende mas não pode interpretar como realizando alguma intenção própria não são ações intencionais suas, mas meros *acontecimentos*[16].

Mas se pode ter esse tipo de consciência, argumenta-se, simplesmente quando as ações são produto do próprio raciocínio prático. Ademais, em casos normais, muitas características da [ação] intencional operam de modo não-consciente e resultam do processamento implícito de informação de um tipo que os anterógrados são capazes de realizar. Novamente, quando a ação ocorrente está

Damasio", NYRB, 58.10 (2011), p.50.

[15] Ver Naomi EILAN e Johannes ROESSLER, "Mechanisms and epistemology". In: *Agency and self-awareness: Issues in philosophy and psychology*, Oxford: Clarendon Press, 2003, p.15.

[16] Anne NEWSTEAD, "Knowledge by intention? On the possibility of agent's knowledge". Capítulo 12 de S. Hetherington (ed.), *Aspects of knowing: Epistemological essays, perspectives in cognitive science*, Amsterdan: Elsevier Press, 2006, p.2.

em questão, Molaison e Shelby satisfazem essas condições do comportamento intencional. Os danos à memória não precisam comprometer a familiaridade não-consciente com as próprias ações[17].

A questão é se isso basta para a ação intencional de um tipo que sustenta a agência autogovernada ou uma agência unificada plena. Acredito que as vidas de Henry Molaison e Leonard Shelby indicam que existem certas coisas sobre nós mesmos que precisamos lembrar se for para operarmos como agentes unificados plenos. Adicionalmente, precisamos ser capazes de lembrar de um modo particular, de modo a operarmos como agentes. Em especial, a agência exige mais do que memória implícita aparente na preservação de preferências, idiossincrasias e habilidades práticas. Ela pede mais do que respostas adquiridas a estímulos do ambiente e sinalizações sociais. Demanda ter mais do que a habilidade de continuar uma narrativa autobiográfica, mesmo quando nossa motivação nesse empreendimento seja um desejo de dar sentido ao que reconhecemos (ainda que penas visceralmente) como o empobrecimento de nossas circunstâncias, relações de dependência e autogovernança comprometida. Pede mais do que a habilidade de empregar apropriadamente conceitos indexicais. Finalmente, mesmo que amnésicos anterógrados possam ter conhecimento não-observacional sobre suas ações atuais sem indícios adicionais, e mesmo que saibam o que estão fazendo simplesmente por agir e confiar que o que fazem se ajusta, sob alguma descrição, às suas intenções, sugerirei que esse nível de conhecimento não é suficiente para a agência autogovernada.

Lembre que agentes autogovernados precisam ser capazes de antecipar suas intenções como condutoras à ação por meio do comportamento de automonitoramento. Ser um agente autogovernado é operar de uma maneira especial. O eu do agente autogovernado pode assumir as atividades práticas que esperamos de deliberadores autônomos. E, se estou certa, pode-se oferecer uma

[17] Eilan e Roessler ressaltam a tese de Marc Jeannerod de que "mesmo sob circunstâncias perfeitamente normais, muitos aspectos das ações intencionais são controlados com base em informações inconscientes, 'processadas implicitamente'". Eilan and Roessler, p.24.

defesa apropriada de que pessoas que apresentam danos cognitivos do tipo que a amnésia anterógrada produz estarão mal equipadas para assumir tal função. Essas pessoas farão tremendo esforço para responder transtemporalmente a questões sobre seus planos futuros, apelando ao que haviam se determinado a fazer como uma predição. Por exemplo, embora Molaison estivesse ciente de si mesmo como tendo tido um passado e se confrontando com um futuro, não havia um ponto de vista futuro a partir do qual poderia testemunhar que "as intenções que [formei] hoje virão a ser executadas"[18]. Se a ação intencional do tipo do qual a autogovernança depende é mais do que simplesmente um modo de antecipação em que uma pessoa projeta-se a si mesma numa perspectiva futura, de modo a tornar suas ações atuais inteligíveis para si mesma – se exige predizer o comportamento futuro com base na decisão do que fazer – amnésicos anterógrados são desafiados nesse quesito. Eles não terão uma perspectiva transtemporal de um tipo que sustente a continuidade do comportamento e permita a persistência de projetos transtemporais. A memória faz precisamente isso. Ela organiza os perfis autobiográficos ao apresentar o presente em termos que lhe dão sentido como a extensão de uma vida. Mas a memória não assiste aos amnésicos anterógrados (e pessoas como a minha pobre sogra) a esse respeito. Suas capacidades de deliberação prática também estão enfraquecidas.

Amnésicos anterógrados ficam aquém de controlar suas ações por meio do comportamento de automonitoramento em outros aspectos óbvios. Nossa agência depende da contribuição que damos para levar nossas vidas. Por carecerem de um mecanismo interno para manterem a si mesmos – seus caráteres, compromissos, planos e valores – presentes aos olhos da mente, anterógrados como Leonard Shelby não podem dar testemunho de si mesmos confiavelmente[19]. A determinação de Shelby de manter um diário com dados importantes tatuando a

---

[18] J. David VELLEMAN, "Self to self". In: *Self to self. Selected essays*, New York: Cambridge University Press, 2006, p.172.

[19] Não tenho a menor ideia do que dizer sobre a compreensão interna de si mesmo como um ser temporalmente estendido que poderia estar disponível para alguém com amnésia. Pergunto-me como é sentido "de dentro" ser o mesmo ego ao longo do tempo.

si mesmo não é confiável, pois ele precisa lembrar por que uma tatuagem particular é importante. O fato de que precisa readquirir, às vezes repetidamente, informações significativas sobre sua vida dificulta sua capacidade de ação, assim como o fato de que é dependente em muitas dessas informações da confiabilidade dos outros. A capacidade de Shelby de conquistar o objetivo que busca *por sua própria vontade, através do comportamento de automonitoramento* está operacionalmente danificada. E embora não tenha perdido a ação intencional inteiramente, dado que consegue tomar algumas decisões sobre o que fazer e, para algumas ações, consegue manter a decisão em mente por tempo suficiente para finalizar sua ação – ele consegue, por exemplo, vasculhar um carro em busca de um vestígio – não é capaz de planejar ações futuras transtemporalmente estendidas. Sua aflição rompe com sua proficiência como um agente – como uma entidade capaz de decidir-se por um curso de ação, inventar um estratagema que faça sentido, dados esses objetivos, e capaz de oferecer razões para seu comportamento à luz desses mesmos objetivos.

Afirmei acima que a tarefa da memória é organizar autobiografias de modo a dar sentido ao presente como uma continuação do passado. Alguma quantia de continuidade autobiográfica pode ser preservada por meio de lembranças compartilhadas, com as quais memórias de eventos importantes na vida de alguém são oferecidas por terceiros e por documentação material. Que Molaison e Shelby não possam ressuscitar memórias prévias por si mesmos não significa que o acesso à informação que a memória proposicional proporciona não possa ser fornecido por outras fontes. Os outros podem lhes construir uma narrativa verídica de memórias episódicas e, talvez, até de memórias semânticas. Eu questiono se testemunhar a si mesmo dessa maneira indireta, em terceira pessoa, permitirá a uma pessoa reviver uma concepção de si em primeira pessoa. Mas mesmo se o fizer – mesmo se, digamos, Shelby acreditar na narrativa que construímos e vier a se reconhecer entre as informações, possuir essa concepção de si mesmo em primeira pessoa é insuficiente para a agência.

Uma vez que compreendamos o que Molaison e Shelby perderam, estaremos em condições de identificar uma característica fundamental da agência.

No fim das contas, realmente não importa quanto material em primeira pessoa está no arquivo de si mesmo. Tanto Molaison, quanto Shelby mantiveram um corpo razoavelmente extenso de conhecimento em primeira pessoa. O que importa é se o material que está no arquivo de si permite a alguém tornar as suas ações inteligíveis para si mesmo. É sob a aparência do eu enquanto agente que adquirimos um interesse imediato em componentes particulares do arquivo da concepção de si. O que é crucial para a agência é que se retenha a capacidade de reconhecer a si mesmo em primeira pessoa – não meramente como "Marina Oshana", mas como "eu" – e que se retenha a memória em primeira pessoa de coisas que fornecem a si um sentido de direção na vida e um sentido de identidade transtemporal. Essas duas coisas são necessárias para possibilitar a alguém dar sentido a suas ações como próprias (suas) e como constitutivas da própria vida. As experiências de Henry Molaison e Leonard Shelby expõem a importância do autorreconhecimento e a importância de reconhecer a si mesmo nas próprias ações.

O fato de que elementos seletos do arquivo identidade ancoram o autorreconhecimento e são vitais para a agência não-corrompida é evidente em vários aspectos. Primeiro, quando eles se tornam acessíveis para a pessoa através da reflexão ou através da atenção lançada a eles por perspectivas externas, uma pessoa registrará essas características do arquivo de identidade como salientes para sua concepção de si. Por exemplo, algumas memórias são mais centrais para o arquivo de identidade do que outras, à medida que formam um componente fundacional da concepção de si da pessoa. Frequentemente se encontra uma indicação do grau de saliência [das memórias] no fato de que a pessoa estará disposta a responder pela saliência que elas têm para ela[20].

[20] Minhas observações aqui foram inspiradas na abordagem da autonomia como "responsabilidade de si" de Andrea Westlund. Westlund argumenta que "a autogovernança da escolha e da ação exige a autogovernança do raciocínio prático que resulta na escolha e na ação, e que a autogovernança do raciocínio prático exige uma disposição para imputabilidade dialógica". Penso que Westlund está oferecendo uma abordagem mais da agência responsável do que da agência autônoma. Não obstante, acredito que a agência "exige uma forma de reflexividade que é, quanto à forma,

Poderíamos dizer que uma pessoa está disposta a "apropriar-se" dessas características identitárias, preservando-as tanto quanto as circunstâncias permitam. Em segundo lugar, elementos seletos do arquivo de identidade dão forma à interpretação do mundo da pessoa. Eles oferecem um quadro de referência para a escolha e a ação, por assim dizer. Estando eles ausentes, seria bastante difícil para uma pessoa viver de uma maneira que dê expressão a seus valores e preocupações, se não por alguma outra razão, pelo menos porque teria uma compreensão tênue de quais são os seus valores e preocupações. Consequentemente, uma pessoa seria menos capaz de permanecer como um parceiro confiável e genuíno nas trocas sociais com os outros.

Não estou qualificada para enunciar com qualquer grau de exatidão quais componentes da psicologia motivacional de uma pessoa, especificamente, são aqueles de que está disposta a se apropriar, nem quais elementos permitem a uma pessoa permanecer um parceiro ao lado de outros. Apesar disso, estou preparada para enunciar que aqueles que o fazem, adquirem proeminência na medida em que guiam o pensamento prático dando forma às razões de uma pessoa para agir e ao seu estilo de atuação. Os elementos do arquivo de identidade que estão no coração do autorreconhecimento e que são cruciais para a agência não danificada adquirem proeminência em circunstâncias em que as habilidades de lidar com situações difíceis (*coping*) são testadas. Eles tendem a ser características das quais outras facetas de si mesmo dependem, como preferências por estados mentais e cursos de ação. Os aspectos do eu que prevalecem na psicologia motivacional de uma pessoa tendem a ser aqueles que a permitem "ser ela mesma", e que permitem àqueles com quem interage reconhecê-la "do jeito que ela é". Agora, ser o que somos ou ser do modo como normalmente somos pressupõe certa constância de identidade, bem como uma constância da concepção de si. Pressupõe um fundamento em que a unidade transtemporal do eu possa se assentar. Com base nisso, as pessoas são capazes de se engajar

---

irredutivelmente dialógica". Ver WESTLUND. "Rethinking relational autonomy". In: *Hypatia* 24.4, p.26-49, 2009.

num comportamento intencional que é inteligível para si mesmas e estão melhor equipadas para relações interpessoais. Nos termos de Jan Bransen, o que está em questão nessas ocasiões é "como determinar a melhor alternativa de si mesmo"[21]. Uma alternativa de si mesmo é uma "continuação de si mesmo", determinada pela escolha entre diferentes modos de ser, cada um dos quais está de acordo com o próprio perfil avaliativo e motivacional, geralmente "naqueles momentos em que viver essa vida exige uma interpretação da direção que a vida deveria tomar"[22].

As pessoas que estão em circunstâncias parecidas com as de Molaison e Shelby podem escolher entre continuações de si mesmas nesse sentido? Que tipo de memória ou modos de conhecimento transmitem a informação em primeira pessoa de que uma pessoa é ela mesma ou que ela é como deveria ser? Richard Kraut chamou minha atenção para uma passagem de Anthony Ashley Cooper, o terceiro Conde de Shaftsbury (1671-1713), no seu "Philosophical Regimen"[23]. O que Shaftesbury escreve desafia a posição que vim esboçando. Eis a citação:

> Os metafísicos e raciocinadores notáveis, acerca da bela questão da identidade, afirmam que, se a memória fosse suprimida, o eu desapareceria. [Mas] o que importa a memória? O que tenho a ver com essa parte? Se, *enquanto sou*, sou como deveria ser, o que mais me importa? E assim me deixem perder o *eu* a cada hora e ser vinte eus sucessivos, ou novos eus; é tudo um só para mim: então não perco minha opinião. Se o carrego comigo, isso sou eu; e está tudo bem… – o *agora*; o *agora*. Lembre disto: nisso está tudo.

Kraut cita essa passagem do livro recente, *Selves*, de Galen Strawson, que usa a passagem para sustentar sua própria tese[24]. A tese de Strawson é a que "a

---

[21] Jan BRANSEN, "Alternatives to oneself: Recasting some of our practical problems", *Philosophy and Phenomenological Research* 60.2, p.381-400, 2002.

[22] Jan BRANSEN. "Personal identity management". In: *Practical identity and narrative agency,* New York: Routledge, 2007, p.109.

[23] *The life, unpublished letters, and philosophical regimen of Anthony, Earl of Shaftesbury*, Ed. Benjamin Rand, London, Swan Sonneschein: Macmillan, 1900.

[24] Galen STRAWSON. *Selves: An essay in revisionary metaphysics.* Oxford: Clarendon Press, 2009,

perspectiva geral, o caráter essencial e a identidade moral"[25] que garantem nossas vidas práticas não dependem da preservação do autorreconhecimento temporalmente estendido. Ele argumenta que a maioria das pessoas de fato não experienciam a si mesmas diacronicamente, como seres cuja pessoalidade estende uma ponte para o passado e uma projeção para o futuro. Embora a autoconsciência humana seja consciência de si mesmo "especificamente como uma presença mental interna ou eu", essa consciência não é marcada pela experiência da vida mental como sendo de eventos que aconteceram para um *eu*. Strawson sustenta que muitas pessoas, incluindo ele próprio, interpretam suas identidades episodicamente, com "pouco ou nenhum entendimento de que o eu que somos estava ali no passado remoto e estará ali no futuro, embora estejamos perfeitamente conscientes de que temos continuidade de longo prazo quando considerados como um ser humano completo"[26]. Em suma, não há necessidade de representar a própria vida como uma narrativa temporalmente estendida para "ter uma apreensão perfeitamente boa de si mesmo como tendo certa personalidade"[27]. Uma ciência momentânea de si será suficiente para a autoconsciência contanto que se mantenha constante a concepção valorativa. Num espírito de pura confissão, permitam-me dizer que considero a concepção do eu como "pérolas da consciência", de Strawson, inadequada para a tarefa em mãos, a saber, oferecer condições de verdade para afirmações sobre a agência pratica e o autogoverno[28]. Mesmo se estivermos dispostos a reconhecer que

---

p. 202.

[25] Strawson diz que isso é o que Shaftesbury quer dizer por "opinião".

[26] STRAWSON. "Against narrativity", *Ratio* 17.4, p.428-452, 2004. Essa citação aparece na p.430. Strawson também rejeita a alegação normativa de que "uma perspectiva abundantemente Narrativa é essencial para uma vida bem vivida, para uma personalidade verdadeira ou plena".

[27] STRAWSON, "Against narrativity", p.438.

[28] "Muitos egos mentais existem, um por vez, e um depois do outro, como pérolas em uma corrente, no caso de algo como um ser humano. De acordo com a concepção de pérolas, cada um tem uma existência distinta, uma coisa física ou objeto individual, embora possam existir por extensões de tempo consideravelmente diferentes". STRAWSON. "The self and the SESMET", *Journal of Consciousness Studies* 6.4, p.99–135, 1999.

Shelby e Molaison são capazes de uma variedade mínima de autorreconhecimento transtemporal, nenhum deles é capaz reter uma imagem rica de si mesmo diante dos olhos de sua mente por qualquer tempo maior do que uns poucos minutos e, mais tipicamente, por não mais do que 30 segundos. (Também não estou convencida de que a constância da perspectiva valorativa seja suficiente para agência robusta).

Mas voltemos a Shaftesbury. Poderíamos perguntar – como Kraut me perguntou – se uma vida vivida no presente eterno seria, necessariamente, tão ruim: "Se a experiência presente de alguém for suficientemente rica, por que teria de ser rebaixada simplesmente por não ser experienciada como algo conectado a um passado ou futuro?"[29] Bem, não estou certa de que seria. Como Kraut ressalta, isso seria parecido com a vida exaltada pelo budismo. E suponho que se poderia obter satisfação e um sentimento de gratificação de uma vida vivida de momento a momento. Henry Molaison, como vimos, viveu muito feliz. Mas duas observações precisam ser feitas. Primeiro, não tenho certeza sobre como as experiências de alguém que viveu de tal modo carregariam um sentido de como essa pessoa deveria viver, como cobra Shaftesbury, nem sobre como poderia manter consigo sua perspectiva geral, seu caráter ou sua identidade. Fazê-lo parece exigir reconhecimento e identificação de si mesmo como um agente transtemporal. Segundo, estou razoavelmente convencida de que mesmo se, digamos, a perspectiva geral e o caráter essencial de Molaison pudessem ser preservados, como Shaftesbury e Strawson alegam, sua agência poderia não ser. A agência exige uma perspectiva geral ou caráter essencial que proporcione uma conexão psíquica e prática com o próprio passado e futuro.

Enfatizei o papel que a memória de uma pessoa de um subconjunto de crenças, valores, preferências e similares desempenha para oferecer uma bússula para a vida. A memória autobiográfica apoia a identidade prática ao equipar a pessoa com alguns dos recursos necessários para construir uma concepção de si

[29] Richard KRAUT, "Comments on Marina Oshana, 'Memory, self-understandig, and agency'", *Riquewihr*, 30 de junho a 2 de julho de 2011.

como um ser temporalmente estendido, alguém em condições de dar atualmente sentido à sua vida em termos da trajetória que a conduziu ao seu estado atual, e capaz de justificar comportamentos subsequentes como continuações de sua concepção de si. A memória nos diz o que lembramos de nós mesmos e a antecipação revela o que esperamos de nós mesmos. Mais precisamente, a memória integra o processo de subscrever nossas capacidades normativas como agentes intencionais. Sem recordar quem fomos e o que fizemos, somos incapazes de saber o que devemos fazer – qual plano de ação deveríamos desenvolver – em congruência com as necessidades e interesses que levamos em consideração.

Ter acesso a essas memórias permite às pessoas possuir o tipo de identidade – identidade agencial – que é geralmente estimado. Certamente nem todas as coisas que acabam por nos dar uma sensação de direção na vida e uma sensação de identidade transtemporal precisam ser recordadas de uma perspectiva de primeira pessoa, ou recordadas pura e simplesmente, para que uma pessoa tenha uma sensação de si como um agente. Tenho memórias vívidas do senhor Baines e da senhorita Flax, meus professores favoritos no primário. Se essas memórias antigas se perdessem para mim, a perda seria perigosa – ambos os professores cultivaram meu amor pelo aprendizado. Mas a perda não forçaria uma reconstituição de meu arquivo da minha própria identidade. Eu poderia continuar tendo um sentido de mim mesma como uma educadora com um estilo pedagógico particular. Mas a perda de outras memórias teria um efeito mais profundo. Perdendo-as, chegaria um ponto em que eu poderia viver apenas de um modo atemporal, ou no momento presente metaforicamente congelado[30].

A perda de memória confronta a todos nós conforme envenhecemos. (Eu já

[30] Pode ser que as memórias com que mais me importo sejam desse tipo necessário para que a minha agência prospere. Eu estimo memórias "essenciais" – das expressões faciais do meu pai, do brilho do seu trompete, do cheiro do calor do verão em nosso minúsculo apartamento, dos ricos aromas inebriantes da paleta de tintas da minha mãe. Uma incapacidade de lembrar dessas e outras memórias semânticas similares poderia ser tão devastadora para o sentido de "ser eu mesma" quanto a demência.

disse isso antes?) Comumente, essa perda é de um tipo suave e benigna. Certamente a maioria de nós nunca foi forçada a lidar com a confusão psíquica profunda causada pela amnésia anterógrada. Ainda, tememos sucumbir a uma senilidade de tipo mais comum. David Copp me lembrou que parte do que tememos sobre o prospecto de tornarmo-nos vítimas da demência é essa espécie de existência atemporal. A demência exemplifica uma perda do sentido do eu que nós todos valorizamos – o reconhecimento de nós mesmos em nossas ações combinado com a garantia de que somos reconhecidos como nós mesmos pelos outros. No coração dessa perda de identidade e da identidade do eu está a perda do governo de si. A demência rouba das pessoas a capacidade de antecipar seu comportamento e sujeitar seu comportamento a um direcionamento reflexivo. Ela diminuiu a capacidade de ter as atitudes parecidas com planos ou intenções que, como Michael Bratman argumentou, funcionam como politicas oficiais *vis a vis* com a perspectiva motivacional do agente sobre seus próprios desejos[31]. Nós nos importamos com os traços de caráter e as memórias que se perdem com a demência devido ao papel central que desempenham na constituição e suporte de nossa permanência como sujeitos unificados, ativos e independentes, que possuem um arquivo da própria identidade rico, em evolução, geralmente verídico e acessível em primeira pessoa.

Conclusão

A agência autogovernada exige que as informações da concepção de si às quais uma pessoa tem acesso sejam aquelas que tornam possível o reconhecimento de si como um ser temporalmente estendido. O reconhecimento desse tipo tem de ser essencialmente indexical e de um material informativo que forneça razões para o que a pessoa faz, um material que a pessoa possa utilizar

[31] Como nota Bratman, "Esse papel característico de tais políticas lhes dá uma solicitação para falar pelo agente, para ajudar a estabelecer onde o agente está com respeito a uma forma particular de motivação". Ver o seu "Hierarchy, circularity and double reduction". In: *Contours of agency*: *Essays on themes from Harry Frankfurt,* ed. Sarah Buss e Lee Overton, Cambridge, Mass: MIT Press, 2002, p. 76.

para dar sentido ao que faz como seus próprios feitos. Para que o reconhecimento de si nesse nível aconteça, as manobras normais de monitoramento e representação de si devem estar operativas. Essas manobras envolvem fortemente a memória.

Sugeri que a ideia de um registro informacional da concepção de si serve como um ponto de partida útil – mesmo que meramente heurístico – para explorar questões acerca da natureza da agência prática. Três questões, em particular, motivam essa exploração. Primeiramente, que elementos do registro da concepção de si tornam possível o autorreconhecimento? Em segundo lugar, que elementos do registro nos são mais importantes e por quê? E, em terceiro lugar, que elementos são cruciais para a agência e para quais tipos de agência? Minha proposta é que nos três casos olhamos para um material do registro que a pessoa pode acessar, pode sujeitar à orientação reflexiva e pelo qual está disposta a se responsabilizar. O material é indexical e consiste num conjunto de crenças, valores, idiossincrasias, memórias e habilidades que embasam a psicologia motivacional da pessoa. O material nos importa porque, em circunstâncias normais, fornece à pessoa uma conexão psíquica com sua atividade passada, permitindo à pessoa pensar de si mesma, tratar-se e ser tratada como um agente cuja vida se estende para o futuro, mesmo que seja para um futuro a quinze minutos dali.

13

DELICADEZA E TERNURA: HUME E O FEMININO

Lívia Guimarães

Introdução

Neste capítulo, apresento a maneira como Hume estabelece diferenças de gênero, como ele as transgride e, por fim, como as abandona. Em minha visão, por pensar tanto dentro, quanto fora dos limites de gênero, e por exaltar o feminino, Hume pode-se dizer autor não apenas de uma filosofia, mas também de uma utopia feminista. 'Ternura' (*tenderness*) é um foco principal de sua filosofia. Esta é uma excelência feminina. Quando supera restrições de gênero, a "ternura ampliada" é benéfica para toda vida pessoal, social e política. E, se prevalecesse, sem oposição, na natureza humana, a ternura sedimentaria o mundo utópico imaginário de Hume.

Após examinar Hume, o homem, passo ao pensador, abordando quatro tópicos, que considero fundamentais em sua filosofia e aos quais teorias feministas são sensíveis. São eles: circunstância, corporeidade, atenção e inatenção ao conceito de gênero. Minhas anotações inspiram-se nas intuições originais de Annette Baier e se propõem como um ponto de partida para futuros diálogos.

Consciente das várias importantes contribuições teóricas do pensamento feminista moderno, parece-me que o texto de Hume propõe um desafio peculiar. Ao contemplarmos a tradição filosófica, não raro, inclinamo-nos à crítica, até mesmo à rejeição de autores cujos pensamentos, de algum modo, possuem preconceitos de gênero, seja nas pressuposições, seja nos conteúdos de suas teorias; se nos dispomos ao diálogo com autores do passado, este vem, muitas vezes, tingido de prudente cautela.

Hume parece ser uma rara exceção – ele jamais considera a razão, senso,

virtude e acesso à esfera pública privilégios ou direitos exclusivos dos homens. Seus heróis e heroínas igualam-se em excelência, pois excelência não é qualidade específica de gênero. Seu fascínio pela psicologia das paixões, e ocasional aparte humorístico diante do absurdo a que chegam as circunstâncias da existência humana, colocam-no em uma posição onde sexismo não cabe mais. Sensibilidade e sensitividade, muito mais do que gênero, comandam o melhor de sua atenção. Assim, com certa audácia, eu diria que Hume cria seu particular refúgio feminista, no século XVIII escocês[1].

1. Annette Baier

Possivelmente, o que é distintivo na peculiar relação que parece haver entre as teorias de Hume e a teorização feminista está no fato de que, naquelas, já se encontra muito daquilo que esta, no século XX, com imenso esforço, chegou a estabelecer. Com a ajuda de Annette Baier em "Hume, the Women's Moral Theorist?", "Hume, the Reflective Women's Epistemologist?", *A Progress of Sentiments*, e seu extenso trabalho acerca da 'confiança' (*trust*), aumenta a suspeita[2]. Baier nos mostra que Hume foi ou, ao menos, poderia ter sido, feminista. Quando se pensa sobre Hume e o feminismo, é preciso, portanto, começar com ela.

Tomemos, por exemplo, seu trabalho sobre o conceito de 'confiança'. Baier observa que vários filósofos, embora reconheçam que confiança mútua é necessária para a manutenção de uma sociedade, não percebem que confiança começa na relação entre a mãe (ou aquela) que cuida) e a criança, que depende

---

[1] A bibliografia inclui: David Hume, *Tratado da natureza humana*, tradução de Déborah Danowski (São Paulo: Unesp, 2001), em referências futuras, THN. *Investigações sobre o entendimento humano e sobre os princípios da moral*, tradução de José Oscar de Almeida Marques (São Paulo: Unesp, 2003), em referências futuras, EPM. *Essays, Moral, Political, and Literary*, ed. Eugene F. Miller (Indianapolis: Liberty Fund, 1987), em referências futuras, E. *The Letters of David Hume*, 2 Vols., ed. J. Y. T. Greig (Oxford: Oxford University Press, 1932), em referências futuras, LDH. *The History of England*, 6 Vols. (Indianapolis: Liberty Classics, 1983), em referências futuras, HofE.

[2] Cf. Annette BAIER. *Moral Prejudices*. Cambridge: Harvard University Press, 1994; e *A Progress of Sentiments*. Cambridge: Harvard University Press, 1991.

dela para sobreviver e bem viver. Assim, um poderoso suporte da sociedade política repousa na mais comum e primitiva experiência humana – base logo esquecida nas alturas da especulação política, científica e filosófica. Neste sentido inesperado, inovador, Baier contribui para que, de fato, estabeleçamos uma relacão entre o pessoal e o político. Partindo de bases similares, ela deriva também outros elementos centrais da filosofia de Hume: a associação de ideias por causalidade, por exemplo, relaciona-se aos laços de família. Ademais, ela interpreta a epistemologia de Hume como social e intersubjetiva – como análoga ao trabalho feminista contemporâneo em prol de uma epistemologia sensível a situação, contexto, ponto de vista. Prosseguindo nesta perspectiva, Baier defende corporeidade, contra abstração rarefeita. Em um de seus mais notáveis ditos, ela afirma que, para Hume, somos mamíferos, trazendo conosco todas as possibilidades e limites constitutivos desta condição.

O que se segue abaixo, segue intuições de Annette Baier. Eu gostaria, por julgá-lo apropriado, de adotar um percurso que progride da vida aos escritos de Hume.

2. Vida

Algumas das mais significativas personagens na vida (e, como não dizê-lo?) aventuras de Hume, são mulheres. Hume foi educado por sua mãe, em perfeita harmonia dividiu a casa com a irmã e encontrou, em Margaret Irvine, uma ajudante capaz de cozinhar, para sua grande satisfação, algumas de suas iguarias francesas favoritas. Entre seus encontros casuais, os biógrafos narram o que se deu com uma mulher desconhecida, certa vez, ao escorregar no pântano, a caminho da Cidade Nova em Edimburgo. Quando Hume lhe pediu socorro, ela, tendo-o reconhecido, respondeu-lhe que aquiesceria, mas somente depois de ele recitar o Pai Nosso do começo ao fim – condição à qual, obviamente, Hume humildemente se submeteu. Com bom humor, e não sem alguma razão, ele a declarou "o melhor teólogo" que havia conhecido – ali, Hume se curva à sagacidade feminina. Um outro episódio conta da esposa de um fabricante de velas, que o visitou, em seu leito de morte, com a firme intenção de convertê-lo, de

salvar o infiel da danação eterna. A entrevista, diz-se, teve resolução bastante satisfatória: ela se foi, levando consigo a encomenda de um grande número de velas moldadas por seu marido. Nestas histórias, há um sentimento de fácil convívio entre Hume e mulheres comuns.

Com respeito às mulheres da nobreza, encontra-se um registro no testamento – Hume deixou para sua irmã um legado de dinheiro, mas também cem livros a sua escolha – livros, seria supérfluo dizer, contariam, para Hume, entre as suas mais preciosas posses. E, para Hume, como se vê, uma mulher tem pleno direito a eles. Na França e em Edimburgo, Hume cultivou amizades sinceras e correspondência regular com mulheres, nas quais confiou e com as quais se aconselhou. Ao se decidir sobre como agir na crise de Rousseau, tivesse ele seguido o conselho das amigas, o resultado teria sido incomparavelmente melhor. Com seu derradeiro amor, a jovem Nancy Orde, sua relação continua, como sempre, livre e travessa. Mas é com a Condessa de Boufflers, objeto de seu amor mais profundo, que descobrimos incidentes particularmente reveladores. Uma passagem de suas cartas, eloquente e tocante, diz:

> Among other obligations, which I owe you, without number, you have saved me from a total indifference towards everything in human life. I was falling very fast into that state of mind, and it is perhaps worse than even the inquietudes of the most unfortunate passion: how much, then, is it inferior to the sweetness of your commerce and friendship! (LDH 1. 451).

Idolatrado na França, feliz em suas amizades e em suas atividades intelectuais, Hume, nesta passagem, faz a mais apaixonada declaração de amor: “antes de te encontrar”, ele diz, “minha vida não tinha sentido”.

Hume é o amante ardente e impetuoso:

> You may cut me to pieces, limb by limb; but like those pertinacious animals in my country, I shall expire still attached to you, and you will in vain attempt to get free (LDH 1.457).

É o amante submisso:

> Good god, how much am I fallen from the airs which I at first gave myself! […] now, I throw myself at your feet, and give you nothing, but marks of patience and long suffering and submission (LDH 1.459).
> Have you ever had any experience of the situation of your mind, when we are very angry with the person whom we passionately love? You have, surely; can anything be more tormenting and more absurd […] but I then reflected, is this the person for whose welfare I would sacrifice my existence, and can I now think of taking pleasure in her pain and uneasiness? […] the very sight of your handwriting, I own began the cure (LDH 1.462).
> Softness, I beseech you, dear madam, continue to like me a little, for otherwise I shall not be able in a little time to endure myself (LDH 1.463).

É o amante pronto a se entregar, inteiro, à amada:

> I shall never, I hope, be obliged to leave the place where you dwell […] This long absence convinces me more fully than ever before, that no society can make me compensation for the loss of yours, and that my attachment to you is not of a little or common nature (LDH 1.475).

E é o amante sonhador, que fantasia a amada, ausente, a distrair-se, assim como ele, em diversões pastorais:

> If you have been so happy, as to execute your purpose, you are almost in the same state as myself, and are at present wandering along the banks of the same beautiful river, perhaps with the same books in your hand, a Racine, I suppose, or a Virgil, and despise all other pleasure and amusement. Alas. Why am I not so near you, that I could see you for half an hour a day, and confer with you on such subjects? (LDH 1.449)

Por fim, sua descrição da condessa dota-a de excelências que incluem, como era de se esperar, graças e encantos, mas também caráter e entendimento – seu perfeito ideal:

> Should I meet with one in the future, in any time future, for, to be sure, I know of none such at present, who was endowed with graces and charms beyond all expression, whose character and understanding were equally an

> object of esteem, as her person was of tenderness; I ought to fly all company, to avoid all connexion with her, even such as might bear the name of friendship; and to endeavour to forget her as soon as possible… I know not if it would be prudent even to bid her adieu; surely, it would be highly imprudent to receive from her any testimonies of friendship and regard, but who, in that situation, could have resolution to reject them? Who would not drink up the poison with joy and satisfaction? (LDH 1.451)

Os fragmentos acima retratam sentimentos sinceros, em uma disposição romântica, arrebatada pela paixão, temerosa da perda. Depois das mudanças na vida da condessa (o falecimento do esposo, a esperança de matrimônio com o Príncipe de Conti), Hume, generoso, tornou-se seu conselheiro e leal amigo. Em 20 de agosto de 1776, ele escreveu-lhe uma carta de despedida. Foi uma de suas últimas cartas, e a única delas que não tratou de negócios.

Por sua vez, a condessa de Boufflers provou ser uma amiga leal, reflexiva e cuidadosa. Acredita-se ter sido ela quem procurou e, inclusive, obteve a nomeação de Hume para o posto na embaixada Britânica na França. Depois do retorno à Grã-Bretanha, ela quis trazê-lo de volta e fez planos para lhe proporcionar um ambiente ao seu gosto, agradável à sua disposição estudiosa. No conflito com Rousseau, opôs-se à publicação do relato de Hume em defesa própria e, a Rousseau, escreveu uma carta severa, condenando a conduta ingrata e insana. Ela não foi, e não foi vista por Hume, como uma mulher estereotipada. Findo o amor romântico (se é que chegou a findar), permaneceram a ternura, admiração e respeito.

Após este vislumbre de Hume, o homem, como prosseguir até Hume, o pensador? Proponho, endereçando os tópicos antes mencionados: circunstância, corporeidade, gênero e não-gênero.

3. Circunstância: método e resultados – corporeidade e situação

Depois de havermos testemunhado alguns encontros reais de Hume com mulheres, cabe agora ensaiarmos uma aproximação inicial a seu pensamento, partindo da perspectiva de como este aborda situação, contexto e condições sociais, em suma, do que, em seu vocabulário, ele denomina “circunstâncias”. Por

meio destas, introduz-se, na análise de Hume, a noção de que nossos juízos, crenças, paixões, agrados e desagrados dependem, em grande medida, das condições particulares e contingentes da vida humana.

Como sabemos, na introdução ao *Tratado*, Hume assume a tarefa de fundar a 'ciência do homem', ela própria, fundamento de todas as demais ciências. Ele pretende descobrir a essência de mente – seus poderes e qualidades – pelo método da observação e experiência. O modelo é newtoniano. Em sua variação positivista, que aspira a verdades objetivas, necessárias, neutras e universais e que, desse modo, oculta o agente do conhecimento em sua situação particular e parcial, este é um método fortemente criticado por várias pensadoras feministas. Creio, contudo, que a descrição positivista não se aplica nem à ciência, nem ao método de Hume.

Sua ciência é falibilista e probabilística, não ambiciona mais que princípios gerais, matizados de um ceticismo moderado. Quanto ao sujeito cognoscente, na própria "Introdução", ele aponta uma diferença importante entre sua ciência e a de Newton. Na filosofia moral, contrariamente à filosofia natural, não se podem realizar experiências com premeditação e exato controle das variáveis. E, dado que o realizador do experimento é simultaneamente seu objeto, a mera consciência desta condição altera o comportamento observado.

Ademais, o método consiste em um conjunto diverso de estratégias investigativas, que muito ultrapassa a breve descrição inicial. Experimentar, para Hume, consiste em uma prática rica e variada. Em alguns casos, as experiências efetivamente ocorrem, enquanto, em outros, são apenas imaginadas. Algumas são irrealizáveis, ou mesmo inconcebíveis – o que, em si, já é instrutivo. A investigação moral de Hume vai às mais diversas fontes: às tradições poética e literária – a poesia de Ovídio, Horácio e Homero, a tragédia de Eurípides e mesmo fábulas não contribuem menos para a sua ciência que as teorias dos filósofos e, entre estes últimos, os antigos não contribuem menos que os modernos. Outros exemplos são a vida familiar, fatos históricos, análises conceituais, analogias, experiências de pensamento, contrafatuais e memórias pessoais. A ciência experimental do homem não parece apresentar um método unificado em definitivo. Pode-se quase dizer

que a escolha de seus componentes obedece, ao menos em parte, ao gosto e inclinações do praticante.

O que é verdadeiro do método, aplica-se também aos resultados. Hume não possui inabalável certeza de que a política, por exemplo, admite verdades gerais e pode ser reduzida a uma ciência. Mas algumas coisas ele sabe, com firme certeza: a longa e indefesa infância dos seres humanos exige a combinação dos pais para a subsistência da prole; se as condições da vida humana fossem diferentes, a prática da justiça poderia ser desnecessária; ao princípio geral que afirma que a natureza humana não pode subsistir sem a associação dos indivíduos e que, a fim de coordenar suas ações em proveito do bem comum, é necessária a existência do governo, Hume aduz a observação que a humanidade, todavia, constitui as mais diversas sociedades, com os mais diversos fins (EPM 4). Os exemplos proliferam: repúblicas favorecem o progresso das ciências e monarquias, o das artes (EPM 6, e "Da origem e progresso das artes e ciências"); a Inglaterra, onde a vida doméstica prevalece, valoriza antes o útil, mas a França, onde a vida social tem predominância, dá maior valor ao agradável (EPM 8).

Do ponto de vista de Hume, raramente julgamos as coisas por seu valor e mérito intrínsecos. Isto aplica-se aos objetos materiais (mais valiosos, se raros), à dor e à doença (que, por serem comuns a todos, não são causa de orgulho e humildade nas pessoas jovens, enquanto o são nas idosas e nas sofredoras de doenças contagiosas ou geneticamente transmissíveis), aos objetos de estima (menos apreciados, se se apresentam em grau da perfeição menor que ao que se está habituado), à estimativa que se faz da própria felicidade ou miséria (susceptível à comparação com os outros), à força e vividez de uma ideia (tendo vantagem sobre o que aparece obscurecido), à influência superior da contiguidade, e assim por diante.

A influência de circunstâncias particulares não é sempre obscurecimento do valor intrínseco. A virtude da coragem, Hume diz, é muito mais admirada em sociedades incultas e, nas cultivadas, a beneficência, justiça, e as virtudes sociais colocam-se acima dela (EPM 7). Acrescento, é assim mesmo que deve ser. Na segunda *Investigação*, Hume pergunta, o que é mais valioso?

> se uma compreensão rápida ou outra mais demorada; se alguém que pode avançar muito em um assunto já à primeira vista, mas nada consegue realizar por meio do estudo, ou uma personalidade oposta, que progride em tudo apenas à custa de muito esforço e aplicação; se um cérebro claro ou uma invenção copiosa; se um gênio profundo ou um julgamento infalível; quando se pergunta, em suma, qual caráter ou traço peculiar do entendimento é superior a outro,

Sua resposta:

> é evidente que não podemos responder a nenhuma dessas questões sem considerar qual dessas qualidades capacita melhor uma pessoa para viver no mundo e a leva mais longe em qualquer empreendimento (EPM 6.17).

Os costumes, as situações e os acidentes alteram *verdadeiramente* a utilidade e o mérito das qualidades. Nas palavras de Hume:

> Aquele que possui os talentos e as habilidades que convêm a seu *status* e profissão será sempre digno de maior estima do que outro a quem a fortuna prejudicou ao atribuir-lhe seu quinhão. As virtudes privadas ou voltadas para o interesse próprio são, nesse aspecto, mais arbitrárias do que as de natureza mais pública. Em outros aspectos, elas são, talvez, menos expostas à dúvida ou controvérsia. (EPM 6.20)

Ainda mais radicalmente, defeitos, quando associados a qualidades, podem deixar de ser defeitos. Inversamente, genuínas qualidades, em circunstâncias erradas, no mínimo, perdem um pouco de seu valor. Dois exemplos da *História de Inglaterra*, onde tais valorações e desvalorações complexas ocorrem com frequência: a constância, às vezes, confere um certo peso a medidas erradas (HofE 2.29); a avareza é sinal de uma mente não estreita, quando suas aquisições são pretendidas por serem instrumentais "para se alcançar maior grandeza e poder futuros" (HofE 2.54). Por fim, em um grande número de questões, é impossível aspirar-se a um padrão universal: o gosto sempre diferirá, devido aos humores diferentes de pessoas particulares e das diferentes maneiras particulares de seu lugar e tempo (E "Do padrão do gosto").

Há, contudo, exceções. Segundo Hume, a "*discrição, cautela, iniciativa, diligência, assiduidade, frugalidade, economia, bom senso, prudência, discernimento*" não se pode jamais negar "louvor e aprovação". Do mesmo modo, a "*temperança, sobriedade, paciência, constância, perseverança, providência, cortesia, reserva, método, persuasão, decoro, presença de espírito, rapidez de compreensão, facilidade de expressão*" não se negará o título de "excelências e perfeições" (EPM 6.21). Para Hume, certamente, muito acaba se provando relativo – mas não tudo.

Ao considerar circunstâncias, o pensamento de Hume passa, necessariamente, a se mover em meio a homens, mulheres, governantes, governados, crianças, jovens, anciãos, citadinos, camponeses, senhores, servos, escravos, nacionais, estrangeiros, parentes, pais, amigos, amantes etc. Isso significa, mais uma vez, como já vividamente apontado por Annette Baier, que seu pensamento retorna à corporeidade.

As mulheres estão entre os interesses e atores principais no mundo que Hume esboça e que inclui seus papéis, seus cuidados e suas disposições. O amor romântico é a última paixão a que se dedica o livro 2, parte 2 do *Tratado*. O amor sexual e o amor da prole são inatos – nas palavras de Hume, a natureza infundiu em todos os animais "o apetite geral entre os sexos" e "um similar preconceito a favor de sua prole" (E "O cético" 162). Ainda mais interessante, Hume não teme a sexualidade das mulheres. Ao contrário, deleita-se nela, e chega a pausar, no texto, para refletir sobre a infelicidade da impotência masculina:

> Quanto escárnio e desdém, por parte de ambos os sexos, acompanham a *impotência*! O infeliz indivíduo é visto como privado de um prazer essencial na vida e, ao mesmo tempo, incapaz de proporcioná-lo a outros. (EPM 6.27)

Em seu progresso, Hume desenha círculos dentro de círculos do pertencimento de mulheres corpóreas e situadas: sexo, amor, família, nação, lugar, tempo, humanidade, senciência.

4. Gênero: traços femininos

Se tomamos corporeidade e diferença sexual de um lado, e condições sociais ou situação de outro, sempre encontraremos 'gênero' no texto de Hume – um texto livre da dicotomia entre natureza e cultura.

A ênfase, em várias passagens, está na diferença. Os principais traços das mulheres são suavidade (*softness*) e ternura (*tenderness*). Falta-lhes força "para os extenuantes esforços da filosofia mais abstrusa" (note-se que, em Hume, a expressão "filosofia mais abstrusa" nunca escapa de certa ambivalência, quando não de franca hostilidade), elas têm corações ternos, ao ponto de serem mais propensas do que os homens à simpatia (mesmo por ladrões, se belos); elas amam intriga e romance e são, apenas fisicamente, inferiores aos homens. Na maior parte da vezes, aparecem, no texto, sob uma luz favorável e sob apreciação positiva. São elas as estudantes eleitas da história (que é, afinal, o grande laboratório da ciência do homem), as senhoras do gosto, as soberanas do mundo da conversação.

Diferença, quando existe, não impede igualdade. Em uma passagem eloquente, Hume diz:

> Em muitas nações, os membros do sexo feminino estão reduzidos a uma condição próxima da escravidão e não podem ter nenhuma propriedade, ao contrário de seus senhores. Mas, embora os indivíduos do sexo masculino, quando aliados, tenham em todos os países força corporal suficiente para manter essa severa tirania, são tais as insinuações, langores e encantos de suas belas companheiras que as mulheres são geralmente capazes de romper essa aliança e compartilhar com o outro sexo de todos os direitos e privilégios da sociedade (EPM 3.19).

Como podemos ver, a especificidade do gênero feminino "nas insinuações, langores e encantos" (*insinuation, address, and charms*) não é aqui percebida como condenável; na verdade, estes traços servem bem como meios aos fins desejados. Eu me arriscaria mesmo a dizer que são preferíveis a outros meios (os fisicamente violentos), pois repousam em qualidades admiradas por Hume, as

qualidades que devem ser parte da sociedade polida e iluminada de seus sonhos.

5. Conceitos

Hume, provavelmente muito mais do que a maioria dos filósofos, recorre a metáforas femininas para designar conceitos centrais de seu pensamento. A virtude, para ele, é uma mulher – delicada, caridosa, afável e, mesmo, alegre:

> Mas quais verdades filosóficas poderiam ser mais vantajosas à sociedade do que as que aqui apresentamos, que representam a virtude com todos os seus mais genuínos e atraentes encantos e fazem-nos aproximar dela com desembaraço, familiaridade e afeto? Caem por terra as lúgubres roupagens com as quais muitos teólogos e alguns filósofos a cobriam, e o que surge à vista é apenas gentileza, humanidade, bondade, e até mesmo, a intervalos apropriados, divertimento, júbilo e alegria. Ela não fala de inúteis rigores e austeridades, sofrimentos e abnegações. Ela declara que seu único propósito é fazer que seus adeptos e toda a humanidade se tornem alegres e felizes em todos os momentos de sua existência; e não descarta voluntariamente nenhum prazer a não ser com a perspectiva de uma ampla compensação em algum outro período de sua vida. O único esforço que ela demanda é o de um cálculo correto e uma firme preferência por um máximo de felicidade. E se dela se aproximam austeros pretendentes, inimigos da alegria e do prazer, ela ou os rejeita como hipócritas e impostores ou, se chega a admiti-los em seu séquito, atribui-lhes um lugar entre os menos favorecidos de seus devotos (EPM 9.15).

Para além da metáfora, os conceitos mais fundamentais da filosofia de Hume assimilam as características relativas ao gênero feminino, fundindo-as na, ou as elevando a princípios gerais da natureza humana. Razão, simpatia e sentimento são exemplares. Conceitos acessórios também se espelham, com frequência, em qualidades femininas. Como exemplos, temos a modéstia, complacência, delicadeza, e ternura. Similarmente à 'confiança' como compreendida por Baier, todos compartilham desta origem terrena e mundana – a excelência da mulher, ou feminina.

Assim, no que concerne ao conteúdo conceitual, uma *primeira lição* da

abordagem de Hume é a transgressão da divisão binária de traços de gênero, apagando a linha divisória tradicional que recita masculino e feminino como mesmo, outra, par, ímpar, luminoso, obscura, ativo, passiva, mente, corpo, cultura, natureza, razão, paixão.

Para Hume, na crença causal sobre questões de fato, a razão inclui o sentimento de determinação e necessidade na mente, que vem da formação de costume ou hábito na experiência de conjunção constante, e que é projetado no mundo na ideia de conexão necessária. Concomitantemente, os juízos de gosto, morais e estéticos, são expressões de sentimentos, mas sentimentos que agregam habilidades e operações racionais, reflexivas, cognitivas.

No *Tratado*, Hume observa que a "natureza humana se compõe de duas partes principais, requeridas para todas as sua ações, ou seja, os afetos e o entendimento" e que "podemos considerar separadamente os efeitos resultantes das operações de cada uma dessas duas partes que compõem a mente". Logo, que

> [p]ode-se conceder aos filósofos morais a mesma liberdade concedida aos filósofos naturais", a saber, a liberdade de considerarem "como composto e consistindo em duas partes separadas", embora reconhecendo que "em si mesmo, esse movimento é simples e indivisível (THN, 3.2.2.14).

Se a observação de Hume é verdadeira, a distinção entre o entendimento e as paixões limita-se a um dispositivo explanatório no estudo da filosofia moral. Por mais que uma descrição possa empregar a distinção instrumental entre a razão e as paixões, uma visão mais acurada reconhece sua artificialidade. E, das duas categorias, a segunda, isto é, as paixões, parece ser mais fundamental, pois elas são originais, enquanto as ideias, de que se ocupa a razão, são suas cópias.

Além de fusões transgressoras, tais como a que acabo de citar, entre paixão e razão, da qual resulta 'sentimento', encontramos no pensamento de Hume, como já notei, a expansão de conceitos, antes restritos ao gênero feminino, a esferas mais amplas de domínio e significado. Hume diz da delicadeza do gosto: "A própria sensibilidade a estas belezas, ou um *refinamento* (*delicacy*) do gosto, já

constitui por si só um belo traço em qualquer caráter, ao propiciar o mais puro, o mais durável, e o mais inocente de todos os prazeres" (EPM 7.28). A modéstia, por seu turno, começa no comportamento casto feminino, mas se torna a virtude social que mitiga os excessos do orgulho (EPM 8, THN 3.3.2) e a virtude filosófica que mitiga o dogmatismo.

Assim, a *segunda lição* de conteúdo, em uma perspectiva marcada por gênero, é que, se as qualidades de um gênero hão de prevalecer, estão serão as femininas. Não há barreira impedindo a livre circulação das mulheres entre as esferas doméstica e pública. E, como sugeri, as virtudes femininas são especialmente apropriadas à civilização. Hume coloca as virtudes sociais acima das virtudes viris de um guerreiro. Na política, Hume recomenda brandura e moderação – assim incorporando boas maneiras à política. Na sociedade, Hume argumenta que a presença de mulheres e o livre intercurso entre os sexos permite relações mais vívidas, polidas e refinadas. Finalmente, no nível individual, o alvo a se buscar, de acordo com Hume, são temperamentos suaves e mentes cultivadas.

Em um paráfrase ligeiramente modificada, podemos dizer que, de Hume, é a época do sentimento e, sua, é a nação sentimental – do sentimento reflexivo, elaborado, sofisticado, ao mesmo tempo em que simples e natural; e do sentimento presidindo sobre a moral, a estética e o conhecimento. Assim, quando há diferença de gênero operando em sua filosofia, ela opera, no mais das vezes, a favor das mulheres, concedendo-lhes eminência. Não, observem, de modo casual, mas em uma maneira sustentada por todo o rigor da investigação.

6. Não-gênero

Em uma outra perspectiva, e esta seria a *terceira lição* de Hume, a categoria de gênero acaba por se mostrar desnecessária. Com isso, em um sentido, quero dizer que, em diversas passagens, Hume interessa-se mais por semelhanças do que por dessemelhanças. Nas famílias que ele concebe, não há distinção de propriedade. Suas mulheres, assim como seus homens, ambicionam o poder e, se frustradas por esposos estritos e dominadores, transformam-se em tiranas domésticas. No amor, devem ser companheiras, em uma relação de

igualdade, onde a paixão, a amizade, a bondade e a ternura sexuais são cimentadas por interesses comuns e pelo interesse no bem estar da prole. Hume diz:

> O amor entre os sexos gera um contentamento e afeto muito distintos da satisfação de um apetite. A ternura pela sua prole, em todos os seres sensíveis, é usualmente capaz por si só de contrabalançar as mais fortes motivações do amor de si mesmo, e em nada depende dessa afecção. Que interesse pode ter em vista uma mãe amorosa que põe sua saúde a perder pelos cuidados infatigáveis com seu filho doente, e em seguida definha e morre de tristeza quando libertada, pela morte da criança, da escravidão imposta por esses cuidados? (EPM Appx. 2.9)

As mulheres são livres para governar no domínio público. Negar-lhes este direito é uma clara evidência de barbarismo:

> But though positive law seems wanting among the French for the exclusion of females, the practice had taken place; and the rule was established beyond controversy on some ancient as well as some modern precedents. During the first race of the monarchy, the Franks were so rude and barbarous a people, that they were incapable of submitting to a female reign; and in that period of their history there were frequent instances of kings advanced to royalty in prejudice of females, who were related to the crown by nearer degrees of consanguinity. These precedents, joined to like causes, had also established the male succession in the second race; and though the instances were neither so frequent nor so certain during that period, the principle of excluding the female line seems still to have prevailed, and to have directed the conduct of the nation (HofE 2.197).

Na mesma linha, quando Hume se acerca do sublime moral, que consiste na grandeza de espírito e dignidade de caráter, ele põe lado a lado Alexandre Magno, Ajax, Vitélio, e uma mulher, Medeia – cuja magnanimidade consiste no sentimento de autonomia e segurança em si mesma (EPM 7).

Esta espécie de 'não-gênero', que desconstrói, ignora dicotomias, é especialmente visível na *História da Inglaterra*. As mulheres são, tanto quanto os

homens, protagonistas da história, as qualidades e caráter de personagens, geralmente, não se apresentam com etiquetas de gênero. Tantos reis perdem-se por suas afeições e temperamentos demasiado brandos, quantos por temperamentos impulsivos e violentos; tantos por seu amor imoderado da glória, quantos por sua fraqueza e indecisão; tantos por espíritos delicados, sensíveis e clementes, quantos por leviandade e capricho. A valentia das mulheres pode se mostrar tanto no campo de batalha, quanto no parlamento, quanto no submisso ato de se ajoelhar e implorar por uma boa causa. Na *História*, pode-se ao mesmo tempo possuir coragem e benevolência, vigilância e afabilidade. É assim que Hume apaga a presumida especificidade de gênero nas paixões e disposições humanas.

Em um outro sentido, 'não-gênero' alude ao fato de que a ciência de Hume aspira ao conhecimento da humanidade em geral. Acima e além das diferenças, aspira a determinar uma comunalidade inclusiva de todos os seres humanos, extensiva a todos os seres sensíveis. Ainda em um outro sentido, que pretendo sugerir adiante, na conclusão deste ensaio, a exaltação do feminino, de certo modo, marca a utópica abolição do gênero.

Para Hume, os seres humanos concordam em certas preferências e distinções. Sem isso, a linguagem não seria compreensível, e a tradução seria impossível. Hume admite que participamos mais prontamente nos sentimentos que se assemelham ao nossos próprios. Mas este não é, para ele, um limite intransponível. Em suas palavras: "nenhuma paixão, quando bem representada, pode ser-nos inteiramente indiferente, porque não há nenhuma da qual cada pessoa já não tenha dentro de si pelo menos as sementes e os primeiros princípios" (EPM 5.30).

Todos temos uma propensão à simpatia, mesmo que em grau ínfimo. Pela simpatia, somos suscetíveis às emoções de outras pessoas; a simpatia estende nosso interesse além de nosso círculo imediato de relações; ela causa a benevolência, a compaixão e outras virtudes sociais; por conseguinte, é uma causa de nossa existência mesma em sociedade, uma vez que a fonte das relações sociais são paixões compartilhadas. A simpatia cria a sociedade

estabelecendo laços afetivos, padrões de comportamento, experiências e conceitos compartilhados por seus membros. Uma operação natural da mente humana, ela é igualmente uma condição necessária da moral. Os sentimentos morais da aprovação dependem da simpatia isenta das flutuações devidas à proximidade e à distância. Isso nós alcançamos ao assumirmos pontos de vista gerais, em que assumimos posições imaginativas próximas de alguém e das pessoas que a cercam, para sentirmos, pela simpatia, os efeitos de suas qualidades e ações. O prazer que ela causa, induz nossa aprovação, e a dor induz a desaprovação. Em nossas distinções morais, alcançamos um padrão comum através de trocas sentimentais – estas, em princípio, independem de gênero. Em resumo, a moral vem "de um sentido interno que a natureza fez universal na espécie inteira".

7. Utopias

Este ensaio começou em uma rememoração das contribuições do pensamento feminista. Um efeito curioso desse movimento foi a produção de romances utópicos e distópicos. Interessantemente, o gênero literário data pelo menos do início da modernidade. Naquele período, protofeministas contribuiriam à literatura ora com tratados sobre a instrução das mulheres, ora com romances sobre sua condição. Dois exemplos destes últimos estão nas obras de Margaret Cavendish e Sarah Fielding. De Sarah Fielding, temos *The Governess; or, Little Female Academy*, uma pastoral onde as jovens alunas, em suas conversas, contam histórias imaginadas ou suas próprias histórias e, juntas, avançam no caminho da virtude, ensinam e aprendem a bondade e o reto governo das paixões. Elas estão no processo de se cultivarem como exemplos para todas as mulheres. Em *The Adventures of David Simple* e sua continuação, Fielding contempla um plano mais ambicioso: uma sociedade utópica, na qual laços de amor e amizade reúnem dois casais (David e Camilla, Cynthia e Valentine), um de seus pais, e seus filhos. Confiança mútua, ternura, altruísmo governam esta sociedade feliz. O pequeno círculo é incorruptível. Seus princípios sobrevivem a várias provações (pobreza, doença, separação), antes de sua derrocada final. O círculo se desfaz

sob a forte pressão externa, vinda de um mundo de inveja, perfídia, indiferença. O custo é altíssimo: a vida de quase todos. Fielding cria um nicho utópico, ameaçado pelas forças destrutivas da sociedade mais ampla, que é distópica. Há sofrimento e real tragédia em *David Simple*. Mas na narrativa moral de Fielding pode-se antever uma sociedade, no porvir, onde as virtudes ternas e femininas governam homens e mulheres, indistintamente[3].

Quando penso em Margaret Cavendish, recordo-me de que, não satisfeita com um aposento só seu, ela cria para si *The Blazing World*[4]. Neste conto fantástico, uma mulher levada, por acidente, do nosso a um outro mundo, é conduzida ao imperador, que a desposa e lhe confere "poder absoluto para reger e governar esse mundo como lhe aprouver" (BW 132). Ela se revela uma "Blazing" Imperatriz. Dotada de espírito generoso e sagaz, reina com sabedoria, comanda investigações científicas em todas as matérias, e é a juíza suprema de seus resultados, que parecem às vezes inconsequentes, ou contraditórios e, sempre, ligeiramente estranhos. Elimina um campo de pesquisa inteiro na matemática, salva os oradores dos discursos grandiloquentes, mas vazios, e confina os lógicos a limites bastante estreitos. Debate fluentemente a teologia com seres espirituais e, quando seu Velho Mundo está sob ataque, alça velas e derrota seus inimigos em combate naval. Na conclusão do livro, tendo se aproximado de Cavendish, diz-lhe que ela, também, pode reinar absoluta sobre um mundo de sua criação. Esta espirituosa fantasia subverte todas as convenções de gênero, provocando um estado de perplexidade, confusão e, claro, grande divertimento na leitora.

Não possuo evidência de que Hume tenha lido as obras destas romancistas. Provavelmente, não as leu. Mas, se consideramos seu gosto literário, é certo que teria rido um tanto e pensado uma ou duas ideias arteiras, diante do conto fantástico de Cavendish, e que se emocionaria com a ternura de Fielding. A

---

[3] Cf. Sarah FIELDING. *The Governess; or, Little Female Academy*. U.S.A.: Hard Press; *The Adventures of David Simple and Volume the Last*, ed. Peter Sabor.Lexington: The University Press of Kentucky, 1998.

[4] Cf. Margaret CAVENDISH, *The Blazing World & Other Writings*, ed. Kate Lilley (London: Penguin Books, 1994), em referências futuras, BW.

primeira iria lembrá-lo das mulheres não tão modestas, mas intrigantes, que conhecia (Alison Cockburn!). O mundo imaginário da última é similar ao seu próprio mundo real, como deve lhe ter aparecido, ao oscilar entre disposições otimistas e céticas.

Hume não fala do homem como o ser humano universal. Não escolhe a faculdade da razão como a qualidade humana essencial e específica. Não atribui um papel subordinado às mulheres. Põe-se acima dos humilhantes estereótipos que tanto indignaram as feministas do século XX. Abraça a diversidade e rejeita o dualismo. Quando se mantém dentro dos confins da diferença de gênero, favorece o gênero feminino, por suas qualidades mais 'humanas' (*humane*). Quando transgride estes limites, considerando ser uma impressão da reflexão, ou sentimento, a causa determinante do conhecimento, moral e estética e, portanto, tornando a razão, por assim dizer, "escrava das paixões", exalta, mais uma vez, o que se costumava predicar do gênero feminino. E quando elimina gênero por completo, falando da e para toda a humanidade, descobrindo traços comuns a todos, podemos dizer que põe o acento nas afeições femininas, mais suaves e mais ternas. A aprovação moral é uma terna simpatia com os outros, um generoso cuidado com os de nosso tipo e espécie. É uma espécie de amor e de estima. A primeira das virtudes sociais, a benevolência, Hume descreve como terna e humana. As virtudes sociais, na *Investigação*, trazem ordem à sociedade, felicidade à família e à humanidade, apoio mútuo entre amigos, e um delicado domínio sobre os corações dos homens.

Por tudo isso, eu gostaria, neste parágrafo de conclusão, de nomear um outro membro à lista de utopistas feministas – o de Hume. Sua utopia, como já insinuada: ternura; sua fonte principal? Passagens sobre a justiça, onde contempla cenários que tornariam a prática desnecessária. Se a natureza fosse abundante e providente, não haveria nenhuma necessidade de se distinguir entre "o meu e o seu". Ou, então, se a natureza humana acabasse em ternura e cuidado pelos outros, tampouco haveria necessidade da distinção. Fora da utopia, como mostra o exemplo da benevolência, a ternura encontra-se em tudo que importa para Hume. Encontra-se implícita até mesmo nos princípios fundadores de sua filosofia, pois

bem podemos perguntar: como pode alguém ser receptivo a impressões e ideias, e responsivo a elas, se não for terno ou tenro?

14

KANT SOBRE O SENTIMENTO MORAL

Carlos Adriano Ferraz

O problema da justificação é essencial à filosofia moral em geral e à filosofia kantiana em particular[1]. Afinal, como justificarmos, por exemplo, que mentir é moralmente errado? Como diz a sabedoria vulgar, uma coisa é *saber* o que é certo. Outra, *fazer* o que é certo. Colocado em outros termos, ainda que reconheçamos o valor objetivo da lei, isso não significa que seremos, por essa razão, impulsionados a segui-la.

Assim, é parte do projeto kantiano "provar" que os dilemas morais são passíveis de resolução. Ou, ainda, que há uma resposta precisa e válida necessariamente[2] para os conflitos morais. Entretanto, para desconforto de alguns, tal questão não é exatamente clara em Kant, seja na *Fundamentação da metafísica dos costumes* (*Grundlegung zur Metaphysik der Sitten*, 1785)[3], seja em sua *Crítica da razão prática* (*Kritik der praktischen Vernunft*, 1788, doravante 'KpV'). Aliás, quanto a este problema, estas obras mesmas não se mostram de acordo. Isso porque, na GMS, notadamente em sua terceira seção, Kant parece tentar, desafortunadamente, a justificação mediante um argumento teórico para a liberdade, um argumento a partir do qual poderíamos, então, deduzir a lei moral. Na KpV, por seu turno, Kant recorre à inicialmente obscura ideia de um *factum der Vernunft*, um "fato da razão" do qual poderíamos inferir a liberdade e, dessa forma,

---

[1] Tal "problema" diz respeito à questão da dedução, a qual se manifesta também em outros âmbitos da filosofia kantiana: em sua epistemologia e em sua estética, por exemplo. Nesse sentido, tal ponto pode ser plenamente compreendido a partir dos valiosos estudos editados por Eckart Förster (1989). Tais estudos abordam o problema da dedução nas três *Críticas* e no *Opus Potumum*.

[2] Apodítica, portanto.

[3] In: Kant (1998). Doravante 'GMS'.

justificá-la (juntamente com a lei moral).

Com efeito, o projeto kantiano de uma "fundamentação" de sua "metafísica dos costumes" parece não lograr sucesso naquilo que lhe é elementar: apresentar precisamente a justificação de ações realizadas por dever (*aus Pflicht*). Nesse sentido, um dos principais méritos da GMS não seria o de fundamentar uma metafísica dos costumes, mas, sim, o de apresentar aqueles elementos basilares para uma fundamentação, a qual ocorreria, ainda que obscuramente, na KpV. Em suma, somente nesta última teríamos uma resposta àquela questão inicial: por que mentir é moralmente errado?

Sucintamente, poderíamos dividir os temas das três seções da GMS da seguinte forma: na primeira seção, na qual Kant se vale do método analítico[4], parte-se do que ele denomina "senso comum", isto é, de um entendimento pré-filosófico acerca da moralidade[5], demonstrando que a boa vontade (*das gute Wille)* é a única coisa boa irrestritamente, ligando a ela o conceito de dever, uma vez que o conceito de boa vontade conduz ao conceito de dever em virtude deste lhe ser inerente. Na segunda seção, ainda se valendo do método analítico, Kant apresenta diversos elementos fundamentais de sua filosofia moral, especialmente ao partir da ideia de um agir racional para chegar à fórmula do imperativo categórico. Ou, ainda, ele demonstra que para seres capazes de racionalidade o imperativo categórico é o primeiro princípio da ação. Por fim, na terceira e mais problemática seção, Kant intenta demonstrar que o homem é capaz de agir moralmente. Aqui, utilizando-se do método sintético[6], ele pretende justificar a ideia exposta nas

---

[4] Tal método parte de verdades incontestáveis para, então, encontrar o seu fundamento. Ele é utilizado, sobretudo, na 'descoberta' da verdade. Trata-se de um método de 'invenção' por excelência. Ele procede *regressivamente* do condicionado à condição, propiciando a descoberta. Kant também o utiliza no âmbito teórico em seus *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können*, de 1783. Nos *Prolegomena* Kant parte da certeza acerca da matemática e da ciência natural e ascende até os princípios dos quais tanto a matemática quanto a ciência natural pode ser deduzida.

[5] Ou, ainda, de um "conhecimento moral comum pela razão" (*Gemeine sittliche Vernunft erkenntnis*).

[6] Diferentemente do método analítico, o método sintético procede *progressivamente* da condição ao condicionado. É um método propício à 'exposição' da verdade descoberta. Por essa razão Kant o

seções anteriores, isto é, que o imperativo categórico é o princípio supremo da moralidade. Dito de outra forma, a terceira seção seria o lugar apropriado para justificar o que foi assentado nas duas seções que lhe precederam. Este seria o momento de Kant responder (contra o cético) à questão: é possível um agir racional? Todavia, tal não ocorre.

Assim, a GMS é preciosa, sobretudo, por *descobrir* qual seria o princípio supremo da moralidade. Não apenas isso, diversos outros conceitos correlatos deste princípio e da filosofia moral kantiana são *descobertos* na GMS. Contudo, o cético permanece insatisfeito. Aliás, mesmo Kant parece insatisfeito. Afinal, este seria o momento de demonstrar que o homem é capaz de agir moralmente, ou, ainda, que a moralidade não é uma vã quimera. Seria, portanto, o momento de demonstrar a autonomia da vontade (*Autonomie des Willens*), isto é, que não agimos movidos unicamente por um *Triebfeder*, por móbeis meramente empíricos, mas também por um *Bewegungsgrund*. E lembremos que cada seção da GMS é dirigida a um público: a primeira seção dirige-se ao "entendimento comum" (*der gemeine Menschenverstand*) que compartilha com ele as mesmas convicções acerca do que seria certo e errado. Tal público adotaria, implicitamente, o imperativo categórico[7] em suas decisões cotidianas. A segunda seção, por outro

---

utiliza na terceira seção da GMS: ele intenta justificar perante os demais a validade objetiva da lei moral. Tal método é utilizado, também, na *Kritik der reinen Vernunft* (1781) e na KpV, sendo que nesta, apesar do uso do método sintético, Kant também mantém um elemento analítico: ele tem como pano de fundo a ideia segundo a qual o juízo moral de todos os homens reconhece a validade do imperativo categórico.

[7] Na segunda seção da GMS Kant reconhece três formulações do Imperativo categórico: 1- *Age apenas segundo uma máxima tal que possas ao mesmo tempo querer que ela se torne lei universal* (GMS. 421). Este é o único imperativo categórico. Mas dele podemos derivar todos os imperativos do dever. O primeiro imperativo que dele deriva é o da "lei da natureza": *Age como se a máxima da tua ação se devesse tornar, pela tua vontade, em lei universal da natureza* (GMS. 421). 2- *Age de tal maneira que uses a humanidade, tanto na tua pessoa como na pessoa de qualquer outro, sempre e simultaneamente como fim e nunca simplesmente como meio* (GMS. 429). Trata-se, nessa segunda formulação, da "fórmula da humanidade com fim em si". 3- *Nunca praticar uma ação senão em acordo com a máxima que se saiba poder ser uma lei universal, quer dizer, só de maneira que a vontade pela sua máxima se possa considerar a si mesma ao mesmo tempo como legisladora*

lado, é dirigida ao sujeito que adota outra perspectiva acerca do princípio da moralidade. Trata-se, aqui, de demonstrar que tal perspectiva está eivada de elementos heterônomos. Tal público adotaria imperativos hipotéticos[8] enquanto princípios morais. Além desse público, Kant tem mente o cético moral, o qual insiste no fato de que o homem só age mediante interesses particulares. Finalmente, na terceira seção o público abarca aqueles que acreditam que o homem se comporta apenas conforme leis naturais (causalidade). E, é claro, o cético também ocupa, aqui, os pensamentos de Kant.

Com efeito, o problema da terceira seção é: não podemos derivar a moralidade da pressuposição da liberdade, bem como não podemos derivar a liberdade (em sentido positivo) da pressuposição de que somos membros de um mundo inteligível. Isso nos levaria, inquestionavelmente, a um círculo vicioso.

De qualquer forma, a questão que se coloca é demonstrar que a razão pura

---

universal (GMS. 440). Trata-se, aqui, da "fórmula da autonomia". Segundo a interpretação de H. J. Paton (1967) em seu estudo seminal, temos, na GMS, cinco formulações do imperativo categórico, a saber: 1- "Age apenas segundo uma máxima tal que possas ao mesmo tempo querer que ela se torne lei universal" – *fórmula da lei universal*. 2- "Age como se a máxima da tua ação se devesse tornar, pela tua vontade, em lei universal da natureza" – *fórmula da lei da natureza*. 3- "Age de tal maneira que uses a humanidade, tanto na tua pessoa como na pessoa de qualquer outro, sempre e simultaneamente como fim e nunca simplesmente como meio" – *fórmula do fim em si.* 4- "Nunca praticar uma ação senão em acordo com a máxima que se saiba poder ser uma lei universal, quer dizer, só de maneira que a vontade pela sua máxima se possa considerar a si mesma ao mesmo tempo como legisladora universal" – *fórmula da autonomia*. 5- "Age segundo máximas de um membro universalmente legislador em ordem a um reino dos fins" – *fórmula do reino dos fins*.

[8] O Imperativo categórico, como visto na nota anterior, ordena de maneira incondicional ("deves fazer X"). Trata-se de um imperativo apodítico: vale necessária e universalmente. É o único imperativo moral. Ele é necessário a seres que não possuem uma "vontade santa" (compreendida como identificação necessária entre querer e dever). Afinal, o homem é um ser racional finito, ou seja, um "ser sensível dotado de razão", o qual pode 1) ou agir conforme a razão 2) ou agir conforme a determinações empíricas. Os imperativos hipotéticos, por seu turno, ordenam de forma condicional ("*Se* queres Y, *então* deves fazer X": Se devo fazer X porque quero Y, então Y é condição de determinação da vontade, e X é necessário tendo em vista a condição imposta por Y). Os imperativos hipotéticos são regras de conduta que estabelecem os meios para que se possa alcançar um certo fim. Podem ser *problemáticos* (fins possíveis) ou *assertóricos* (fim perseguido por todos: a felicidade).

pode ser prática, ou, ainda, que a razão pode ser um móbil (*Triebfeder*) para o agir. Dito de outra forma, pode a vontade (*Wille*), que é a razão prática, determinar o arbítrio (*Willkür*)? Afinal, não somos sujeitos racionais. Somos sujeitos capazes de racionalidade[9]. Se fôssemos algo tal como um *animal rationale* agiríamos espontaneamente por dever (*aus Pflicht*), o que, obviamente, não é o caso. Se o fosse, seria desnecessária uma fundamentação da moral e, consequentemente, uma dedução do princípio da moralidade.

Entrementes, a distinção entre *phaenomenon* e *noumenon* também está aqui. Enquanto se vê como agente racional, capaz de agir segundo determinações da razão (autonomamente, como pertencente ao mundo inteligível) o homem reconhece o princípio da autonomia da vontade (*Autonomie des Willens*). Quando, por outro lado, ele se considera como pertencente tanto ao mundo inteligível quanto ao mundo sensível, ele reconhece o princípio da autonomia[10] como imperativo categórico, ou, ainda, como princípio do querer (*Prinzip des Wollens*). A proposição "eu devo" (*ich muss*) é uma proposição prática sintética *a priori*. O terceiro termo que, por assim dizer, 'liga' o "devo" à vontade de um ser imperfeitamente racional é a ideia dessa vontade como pertencente ao reino do mundo inteligível. Tal ideia é a liberdade da vontade (*Freiheit des Willens*). Sua função assemelha-se àquela das categorias na *Kritik der reinen Vernunft* (Kant 2006, doravante 'KrV'), a saber, permitir a formulação de juízos sintéticos *a priori*[11]. Dessa forma, liberdade (*Freiheit*) é a ideia da *razão* (*Idee der Vernunft*), sem a qual não haveria juízo moral algum[12]. Nesse sentido, a liberdade é uma "possibilidade lógica". Todavia, ser uma mera possibilidade lógica não é suficiente. Faz-se mister

[9] Como nos diz Kant em sua *Anthropologie* (Kant 1998, p.278, Ak.322): "Conseqüentemente, o homem como um animal dotado com capacidade de razão (*animal rationabile*) pode fazer de si um animal racional (*animal rationale*)".

[10] *Deve* reconhecê-lo: é precisamente esta a questão da dedução/justificação.

[11] Os quais, no plano teórico, nos permitem ampliar nosso conhecimento da natureza.

[12] Assim como, por exemplo, o conceito puro de causalidade (*reine Verstandesbegriff der Kausalitätt*) é um dos conceitos puro do *entendimento* (*Verstand*) sem o qual não haveria um conhecimento teórico da natureza.

que ela tenha, também, uma "realidade objetiva" (ainda que prática). O problema, aqui, é o mesmo enfrentado na KrV: garantir a realidade objetiva do conhecimento. Neste caso, do "conhecimento prático" (*praktischen Erkenntnisse*). Aliás, já no prefácio da GMS Kant usa a expressão *praktischen Erkenntnisse*:

> Assim, pois, as leis morais, com seus princípios, em todo o *conhecimento prático* diferenciam-se de tudo o mais que contenha algo de empírico; e essa diferença não só é essencial, mas também toda a filosofia moral encontra-se inteiramente assentada sobre a sua parte pura, e, quando aplicada ao homem, não recebe um mínimo que seja do conhecimento do homem (antropologia), mas fornece-lhe, na qualidade de ser racional, leis *a priori*[13].

Isso demanda, pois, uma justificativa, algo que, pelo menos nesse contexto, Kant não oferece. Ele rejeita a ideia de que a mera consistência lógica seja condição suficiente para a aceitabilidade de proposições teóricas e práticas[14]. Colocado em outros termos, a mera possibilidade lógica não garantiria a possibilidade real[15]. Por essa razão não basta podermos pensar a liberdade: ela

---

[13] GMS, Ak 389; KANT 1995, p.27.

[14] Isso explica o porquê de Kant não ficar satisfeito apenas com a lógica geral, elaborando uma lógica transcendental. Aliás, é nesse ponto que ele rompe com o dogmatismo metafísico daqueles que, como Descartes, Leibniz e Wolff, acreditavam na possibilidade de um progresso do conhecimento atrelado apenas a uma lógica geral. Assim, no plano teórico de uma razão especulativa, as condições da intuição pura garantem os critérios para a possibilidade real (realidade objetiva) das proposições da ciência (da matemática e da ciência natural). Dessa forma, poderíamos, em Kant, asserir a possibilidade (lógica) de geometrias não-euclidianas. Todavia, Kant não está interessado em sistemas meramente formais: ele está interessado no significado objetivo, na, digamos, verdade objetiva.

[15] Aqui reside a distinção entre *denken* e *erkennen*, respectivamente. Tal problema é explicitado especialmente na KrV, no momento em que Kant está tratando justamente da dedução das categorias (§15 – *Transzendentale Deduktion der reinen Verstandesbegriffe*, B130). Com efeito, ao apontar para o resultado dessa dedução (§27 – *Resultat dieser Deduktion der Verstandesbegriffe*) ele conclui: "Wir können uns keinen Gegenstand *denken*, ohne durch Kategorien; wir können keinen gedachten Gegenstand *erkennen*, ohne durch Anschauungen, die jenen Begriffen entsprechen". Para que algo seja logicamente possível basta que ele não viole o princípio de não-contradição.

deve ser real (sem envolver intuições possíveis). E, talvez exatamente do fato de ela ter que ser real sem recorrer a uma intuição – a um fato (*Tatsache*) em sentido concreto –, Kant nos conduz à sua ideia enquanto *factum der Vernunft*.

De qualquer forma, há um importante ponto em comum entre a GMS e a *Kritik der praktischen Vernunft*[16] (1788): liberdade e lei moral são ideias que envolvem reciprocidade[17]. Tal reciprocidade se mostra explicitamente no início da terceira seção. Após definir a vontade (*Wille*) como uma espécie de causalidade (pertencente aos seres vivos enquanto racionais), bem como definir a liberdade em seu sentido negativo (independência de causas estranhas ao sujeito), ele aponta para o sentido positivo da liberdade, a saber, liberdade como autonomia.

Embora o esboço de uma dedução, na GMS, não tenha sido bem sucedido, ele foi um passo na direção certa. A teoria do *Factum* servirá como uma forma de "provar" (*beweisen*) a realidade objetiva (prática) da liberdade transcendental.

Sendo assim, a GMS concluirá, enigmaticamente, sobre esse ponto, que devemos compreender a incompreensibilidade (*Unbegreiflichkeit*) da lei moral[18]. É dessa forma que a KpV nos levará em outra direção, apresentando-nos a polêmica expressão *Factum der Vernunft*[19].

Apesar de sua importância, todavia, o termo *factum*[20] aparece apenas nos

---

[16] Em Kant (1998). Doravante 'KpV'.

[17] Sobre esse tema ler capítulo 11, "The reciprocity thesis", de Allison (1995, p. 201).

[18] "Und so begreifen wir zwar nicht die praktische unbedingte Notwendigkeit des moralischen Imperativs, wir begreifen aber doch seine *Unbegreiflichkeit*, welches alles ist, was billigermaßen von einer Philosophie, die bis zur Grenze der menschlichen Vernunft in Prinzipien strebt, gefodert werden kann" (GMS. Ak. 463).

[19] De qualquer forma, embora Kant tenha falado em "dedução" na terceira seção da *GMS*, na segunda *Kritik* ele se afasta desse termo por razões óbvias: não há dedução da lei moral/liberdade. Daí ele recorrer à figura do *Factum*. Mas o *Factum*, embora seja apenas um *Factum der Vernunft*, "prova" por si mesmo a liberdade (afinal, a lei é a *ratio cognoscendi* da liberdade). Além disso, a doutrina do *factum* não aparece na GMS.

[20] Kant utiliza a forma latina *factum* (a grafia germanizada *Faktum* aparecerá apenas em edições póstumas da KpV). Com isso ele pretende evitar que surja alguma confusão com o termo *Tatsache*, o qual significa *fato* (por exemplo, na acepção que esse termo tem na Crítica da Razão Pura – *Kritik der reinen Vernunft*, 1781), mas não fato no sentido em que a liberdade é um "fato".

seguintes momentos:

> [1] A razão prática obtém agora por si mesma, e sem ter acertado um compromisso com a razão especulativa, realidade para um objeto supra-sensível da categoria de causalidade, a saber, da *liberdade* (embora, como conceito prático, também só para o uso prático), portanto confirma mediante um *factum* o que lá meramente podia ser *pensado*. (KpV. A 9, p.15).
>
> [2] Pode-se denominar a consciência desta lei fundamental um *factum* da razão, porque não se pode sutilmente inferi-lo de dados antecedentes da razão, por exemplo, da consciência da liberdade (pois esta consciência não nos é dada previamente), mas porque ela se impõe por si mesma a nós como uma proposição sintética *a priori*, que não é fundada sobre nenhuma intuição, seja pura ou empírica [...]. Contudo, para considerar esta lei como inequivocamente *dada*, precisa-se observar que ela não é nenhum fato empírico, mas o único *factum* da razão pura, que deste modo se proclama como originariamente legislativa (*sic volo, sic jubeo*[21]). (KpV. A 56, p. 107)
>
> [3] Esta Analítica demonstra que a razão pura pode ser prática – isto é, pode determinar-se por si a vontade independentemente de um todo empírico – e isto na verdade mediante um *factum*, no qual a razão pura deveras se prova em nós praticamente, a saber, a autonomia na proposição fundamental da moralidade, pela qual ela determina a vontade ao ato. Ela mostra ao mesmo tempo que este *factum* vincula-se indissoluvelmente à consciência da liberdade da vontade, antes, é idêntico a ela; e mediante a qual a vontade de um ente racional, que como pertencente ao mundo sensorial se reconhece, do mesmo modo que outras causas eficientes, como necessariamente submetido às leis da causalidade, contudo no domínio prático, por outro lado, a saber, enquanto ente em si mesmo é ao mesmo tempo consciente de sua existência determinável em uma ordem inteligível das coisas.(KpV. A 72, p. 139)

[21] Kant, leitor e admirador de diversos autores latinos, cita o poeta romano Juvenal (60-127 d.C): *Hoc volo, sic iubeo, sit pro ratione voluntas* – "É isto que eu quero, é assim que ordeno: por razão baste a minha vontade".

[4] Contrariamente a lei moral fornece, ainda que nenhuma *perspectiva*, contudo um *factum* absolutamente inexplicável a partir de todos os dados do mundo sensorial e de todo o âmbito de nosso uso teórico da razão. Esse *factum* oferece indícios de um mundo inteligível puro e até o *determina positivamente*, permitindo-nos conhecer algo dele, a saber, uma lei. (KpV. A 74, p. 143)

[5] Também a lei moral é dada quase como um *factum* da razão pura, do qual somos conscientes *a priori* e que é apodicticamente certo, na suposição de que também na experiência não se podia descobrir nenhum exemplo em que ela fosse exatamente seguida. (KpV. A 81, p. 157)

[6] A realidade objetiva de uma vontade pura ou, o que é a mesma coisa, de uma razão prática pura, é dada *a priori* na lei moral como que mediante um *factum*; pois é assim que se pode denominar uma determinação da vontade que é inevitável, embora não dependa de princípios empíricos. (KpV. A 96, p. 187)

[7] Mas que a razão pura, sem a mistura de nenhum fundamento determinante empírico, também seja por si só prática, tinha que poder demonstrar-se a partir do *uso prático mais comum da razão*, na medida em que se abonava a proposição fundamental prática suprema como uma proposição que toda a razão humana natural, enquanto inteiramente *a priori* e sem dependência de nenhum dado sensível, reconhece como a lei suprema de sua vontade. Primeiro se tinha que estabelecer e justificar essa proposição fundamental segundo a pureza de sua origem, mesmo no *juízo dessa razão comum*, ainda antes que a ciência pudesse tomá-la em mãos para fazer uso dela, como se fosse um *factum* que antecede toda a argumentação sutil sobre sua possibilidade e sobre todas as consequências que se desejassem tirar dela. (KpV. A 163, p. 321)

[8] Ora, tudo dependeu meramente de que esse *poder* fosse transformado em um *ser*, isto é, de que se pudesse provar em um caso efetivo, como que através de um *factum*, que certas ações, quer elas sejam efetivas ou

> também só ordenadas, isto é, praticamente necessárias, pressupõem uma tal causalidade (a intelectual, sensivelmente incondicionada) (KpV. A 187, p.369).

Conforme Lewis White Beck (1963), em seu estudo fundamental da *Crítica da razão prática*, estas passagens contêm seis distintas caracterizações do *factum*: 1) "consciência da lei moral"; 2) "consciência da liberdade da vontade"; 3) "a lei"; 4) "autonomia quanto ao princípio da moralidade"; 5) "uma inevitável determinação da vontade pela mera concepção de lei"; 6) "o caso de uma ação pressupondo uma causalidade incondicional". Além disso, Kant, em diversos momentos, identifica tal *factum* com a "liberdade", com a "lei prática da liberdade" e com o "imperativo categórico"[22].

De qualquer maneira, a doutrina do *factum,* apresentada na KpV, pode ser considerada como uma consequência da ideia de que uma dedução do princípio moral, seja mediante especulação teórica, seja mediante dados obtidos pelos sentidos, está fadada ao fracasso. Assim, na KpV o princípio moral é apresentado enquanto uma certa *Faktizität*, a qual seria, por seu turno, auto evidente (não demandando, como ocorre no plano teórico, qualquer tipo de contrapartida empírica).

Já no prefácio da segunda *Kritik* encontramos seu escopo: "demonstrar que há uma razão prática pura" (KpV. Ak. 2)[23]. É o momento, pois, de, dado o princípio supremo da moralidade, fundamentá-lo. Desta feita, como Kant chega à doutrina do *factum*?

Tal questão exige algumas considerações introdutórias. Em primeiro lugar, ela exige que estabeleçamos a estrutura do livro primeiro da "Elementarlehre der reinen praktischen Vernunft": *Die Analytik der reinen praktischen Vernunft*[24].

---

[22] Na *Kritik der Urteilskraft* (§ 91, Ak. 468) ele aparece com "liberdade"; Na *Die Metaphysik der Sitten* (*Erster Teil: Metaphysische Anfangsgründe der Rechtslehre*, Ak. 252) ele aparece como "lei da liberdade" e (*Zweiter Teil: Metaphysische Anfangsgründe der Tugendlehre*, Ak. 378) como "lei em nós"; no *Opus Postumum* (Ak. 21) ele aparece como "imperativo categórico".

[23] Ou seja: pode a razão determinar a vontade?

[24] Na "Analítica" a liberdade é demonstrada, ao passo que na "Dialética" ela é postulada. Na primeira

Em linhas gerais, as primeiras seções da *Analytik* estabelecem definições que expressam princípios da razão, os quais são, por sua vez, expressos em teoremas. Após os três primeiros teoremas, Kant coloca dois problemas (§§ 1 a 6). O primeiro problema (*Aufgabe I*, §5) concerne ao caráter da vontade, a qual deve ser determinada pela mera forma da lei[25]. Aqui temos sua concepção de autonomia: "uma vontade, à qual unicamente a simples forma legislativa da máxima pode servir de lei, é uma vontade livre (*freier Wille*)". O segundo problema (*Aufgabe II*, § 6) envolve encontrar a única lei que pode ser ligada a uma vontade livre. Tal lei só poderá ser aquela que não é oriunda da experiência (de elementos estranhos à razão). Tal lei é manifesta no imperativo categórico: *Handle so, daß die Maxime deines Willens jederzeit zugleich als Prinzip einer allgemeinen Gesetzgebung gelten könne*. Esta é a lei da razão prática pura.

Com efeito, após estabelecer a lei moral (Cf. *Grundgesetz der reinen praktischen Vernunft*, §7) Kant faz, na anotação (*Anmerkung*) seguinte, uma importante asserção:

> Pode-se denominar a consciência desta lei fundamental um *factum* da razão, porque não se pode sutilmente inferi-la de dados antecedentes da razão, por exemplo, da consciência da liberdade (pois esta consciência não nos é dada previamente), mas porque ela se impõe por si mesma a nós como uma proposição sintética a priori, que não é fundada sobre nenhuma intuição, seja pura ou empírica… (KpV. Ak.56)

---

parte ela aparece como "autonomia" (*Autonomie*), na segunda parte ela é apresentada como autocracia (*Autokratie*), isto é, como condição de possibilidade da virtude e, consequentemente, do *summum bonum*, sendo este artigo de fé (ainda que prática). Por essa razão nos interessa, aqui, especificamente a "Analítica".

[25] "So ist die Vorstellung derselben als Bestimmungsgrund des Willens von allen Bestimmungsgründen der Begebenheiten in der Natur nach dem Gesetze der Kausalität unterschieden, weil bei diesen die bestimmenden Gründe selbst Erscheinungen sein müssen. Wenn aber auch kein anderer Bestimmungsgrund des Willens für diesen zum Gesetz dienen kann, als bloß jene allgemeine gesetzgebende Form: so muß ein solcher Wille als gänzlich unabhängig von dem Naturgesetz der Erscheinungen, nämlich dem Gesetze der Kausalität, beziehungsweise auf einander, gedacht warden" (KpV. Ak. 51).

Essa é, com efeito, a primeira afirmação relacionando o *factum* com a consciência da lei moral. Sua incondicionalidade estabelece que ele não pode ser oriundo de dados anteriores. Portanto, embora Kant utilize o termo *Deduktion* ao falar da "dedução das proposições fundamentais da razão prática pura", não se trata, aqui, de uma dedução. Tal parte demonstra, aliás, por que não é possível tal dedução. Dessa forma, a *Analítica* da segunda *Kritik* é deveras diferente da *Analítica* da primeira *Kritik:*

> Ora, se comparamos com isso a parte analítica da crítica da razão especulativa pura, mostra-se um notável contraste entre ambas. Nesta não são proposições fundamentais mas uma *intuição* sensível pura (espaço e tempo) o primeiro dado que tornou possível um conhecimento *a priori* e, em verdade, somente para objetos dos sentidos. – Proposições fundamentais sintéticas a partir de simples conceitos e sem intuição eram impossíveis, muito antes, só podiam ocorrer em referência àquela intuição que era sensível, por conseguinte também só em referência a objetos da experiência possível, porque unicamente os conceitos do entendimento vinculados a essa intuição tornam possível aquele conhecimento que chamamos experiência[26].

Portanto, enquanto a primeira *Kritik* visa ao conhecimento das propriedades das coisas (na medida em que são dadas à intuição), a segunda *Kritik* constitui ela mesma seu objeto (este não é colocado à razão via dados oriundos de uma experiência). Assim, logo em seguida nos diz Kant:

> Contrariamente a lei moral fornece, ainda que nenhuma *perspectiva*, contudo um *factum* absolutamente inexplicável a partir de todos os dados do mundo sensorial e de todo o âmbito de nosso uso teórico da razão. Esse *factum* oferece indícios de um mundo inteligível puro e até o *determina positivamente*, permitindo-nos conhecer algo dele, a saber, uma lei. Essa lei deve propiciar ao mundo sensorial, enquanto *natureza sensível* (no que concerne aos entes racionais), a forma de um mundo inteligível, isto é, de

[26] KANT, 2003, p.141.

> uma *natureza supra-sensível*, sem com isso romper seu mecanismo. Ora, a natureza é, no sentido mais universal, a existência das coisas sob leis. A natureza sensível de entes racionais em geral é a existência das mesmas sob leis empiricamente condicionadas, por conseguinte é *heteronomia* para a razão. A natureza supra-sensível dos mesmos entes é, ao contrário, a sua existência segundo leis que são independentes de toda a condição empírica, que, por conseguinte, pertencem à *autonomia* da razão pura[27].

Com efeito, Kant recorre à lei moral mesma como forma de obter sua, por assim dizer, condição de possibilidade. Em seus termos:

> Mas algo diverso e inteiramente paradoxal substitui esta inutilmente procurada dedução do princípio moral, a saber, que ele mesmo serve, inversamente, como princípio da dedução de uma imperscrutável faculdade que nenhuma experiência tinha de provar, mas que a razão especulativa (para encontrar entre as suas ideias cosmológicas, segundo sua causalidade, o incondicionado e assim não se contradizer a si mesma) tinha de admitir pelo menos como possível, ou seja, a da liberdade, da qual a lei moral, que não necessita ela mesma nenhum fundamento que a justifique, prova não apenas a possibilidade mas a efetividade em entes que reconhecem essa lei como obrigatória para eles[28].

Assim, ao invés de chegar à lei moral mediante uma dedução, a lei moral mesma serve como princípio para que possamos "provar" a liberdade enquanto uma causalidade da razão prática pura. Tal relação (entre lei moral e liberdade) Kant a expressa de várias formas. Aquela que talvez possa ser reconhecida como a mais explícita ocorre quando ele afirma, em uma célebre nota[29]:

> Quero apenas lembrar que a liberdade é sem dúvida a *ratio essendi* da lei moral, mas que a lei moral é a *ratio cognoscendi* da liberdade. Pois se a lei moral não fosse pensada antes claramente em nossa razão, jamais nos

[27] Ibidem, p.143.
[28] Ibidem, p.159.
[29] Ibidem, p.7.

> consideraríamos autorizados a admitir algo como a liberdade (ainda que esta não se contradiga). Mas se não existisse liberdade alguma, a lei moral não seria de modo algum encontrável em nós[30].

Em outras palavras, nossa consciência da lei moral garante tanto a lei mesma quanto a liberdade. Aliás, desde a terceira antinomia da primeira *Kritik* tínhamos assegurada a *possibilidade* da liberdade (ainda que lá sua *realidade* não estivesse "provada"). Aqui (KpV), sua realidade é garantida pela consciência da lei moral, a qual aponta para a autonomia da vontade. A ideia de liberdade, na medida em que ela é "provada" por uma lei apodítica da razão prática pura, é a "pedra angular" de toda a estrutura da razão pura, mesmo da razão em seu uso teórico[31]. Ambas, liberdade e lei moral, implicam-se mutuamente, sendo que a liberdade (sua realidade objetiva) é, aqui, "provada": *a possibilidade dos mesmos é provada pelo fato de que a liberdade efetivamente existe; pois esta ideia manifesta-se pela lei moral*.

Assim, fica claro que a doutrina do *factum* indica uma unidade da razão, bem como implica uma concepção normativa de sujeito que não possui fundamentação empírica. Trata-se, como ele indica em sua *Anthropologie*, de uma análise do homem tal como este deve, à luz da razão, ser[32]. Quanto a tal "unidade", mais adiante, na *Dialektik der reinen praktischen Vernunft*, Kant vai além

---

[30] KpV, Ak.5.

[31] "Ora o conceito de liberdade, na medida em que sua realidade é provada por uma lei apodíctica da razão prática, constitui o fecho de abóbada de todo o edifício de um sistema da razão pura, mesmo da razão especulativa, e todos os demais conceitos (os de Deus e de imortalidade), que permanecem sem sustentação nesta <última> como simples ideias, seguem-se agora a ele e obtêm com ele e através dele consistência e realidade objetiva, isto é, a possibilidade dos mesmos é provada pelo fato que a liberdade efetivamente existe; pois esta ideia manifesta-se pela lei moral" (KANT, 2003, p.5).

[32] "Uma doutrina sistemática contendo nosso conhecimento do homem (antropologia) pode ser dada ou de um ponto de vista fisiológico ou de um ponto de vista pragmático. O conhecimento fisiológico do homem pretende investigar o que a natureza fez do homem, ao passo que o conhecimento pragmático do homem pretende investigar o que o homem faz, pode, ou deve fazer de si como ser livre atuante" (KANT, 1998, p.29). Uma ideia que já está presente, a propósito, no *Émile* (1762), de Rousseau.

e nos fala do primado da razão prática sobre a teórica[33], o que significa afirmar que tal unidade é garantida pela razão prática pura. Dito de outra forma, o cético moral será, também, um cético no âmbito epistemológico. Contudo, diferentemente do que ocorre no plano teórico, no qual se faz uso de uma dedução para justificar, diante do cético, a realidade objetiva das categorias, as quais estariam justificadas, portanto, para um sujeito "esclarecido" (pelo menos do ponto de vista teórico, isto é, conhecedor da mecânica newtoniana, da lógica aristotélica, da geometria euclidiana, etc.), na KpV a capacidade de agir consoante a lei moral jaz mesmo no mais comum dos sujeitos. Afinal, o que Kant intenta é demonstrar que a lei moral se mostra como *factum* mesmo ao sujeito ordinário e não cultivado. A base da moral kantiana permanece, na KpV, a *populären sittlichen Weltweisheit*. Ele não tem por escopo "criar" uma nova moral[34].

Voltando ao *factum*, este possui uma "facticidade" *sui generis*. Em primeiro lugar, ele não é um fato empírico. Uma esclarecedora definição para fato empírico (*Tatsache – res facti*) nós a encontramos na *Crítica da faculdade do juízo* (*Kritik der Urteilskraft*, 1790, § 91; doravante 'KdU'):

> Os objetos para os conceitos, cuja realidade objetiva pode ser demonstrada (quer seja através da razão pura, quer da experiência, e no primeiro caso a partir de dados teóricos ou práticos daquela, mas em qualquer dos casos mediante uma intuição que lhes corresponda) são fatos – *Tatsachen* – (*res facti*)[35].

Dessa forma, fatos (*Tatsachen*) empíricos há inúmeros. *Factum*[36] há apenas

---

[33] Cf. "Von dem Primat der reinen praktischen Vernunft in ihrer Verbindung mit der spekulativen" (Ak. 215).

[34] Veja-se, por exemplo, a reação de Kant diante de um *Rezensent* (Gottlob August Tittel) que escreveu desabonando a KpV por esta não trazer um novo princípio. Ora, não era o intento de Kant apresentar um "princípio novo da moralidade", mas apenas uma "nova fórmula". Afinal: "Mas quem é que queria introduzir também uma nova proposição fundamental de toda a moralidade e como que inventá-la pela primeira vez?".

[35] KdU. Rio de Janeiro: Forense Universitária, 2005, p. 308.

[36] Mesmo o uso que Kant faz do termo *factum* em vez de *Tatsache* indica, segundo vemos, a

um. Ele é, como assevera Kant, o único fato da razão (*factum der Vernunft*):

> Contudo, para considerar esta lei como inequivocamente **dada**, precisa-se observar que ela não é nenhum fato empírico, mas o único *factum* da razão pura, que deste modo se proclama como originariamente legislativa (*sic volo, sic jubeo*).

Ele é um fato *sui generis* no sentido de que ele deve ser concebido unicamente no âmbito da razão prática pura. Apesar de sua *Unbegreiflichkeit*, ele possui realidade, uma realidade que é assegurada no plano prático apenas. Poderíamos chamá-lo, portanto, de fato imanente: ele é oriundo da razão (eu seu uso prático) mesma. Daí a impossibilidade de demonstrá-lo a partir de dados externos à razão mesma. Mas a "voz da razão" (*Stimme der Vernunft*), não obstante, é evidente mesmo para o entendimento mais comum[37]. Ele se mostra como um imperativo inegável[38]. Para ilustrarmos esse ponto usemos um exemplo utilizado pelo próprio Kant:

> Perguntai-lhe, porém, se, no caso em que seu governante sob ameaça da mesma inadiada pena de morte lhe exigisse prestar um falso testemunho contra um homem honrado, que ele sob pretextos especiosos gostaria de arruinar, se ele então, por maior que possa ser seu amor à vida, considera possível vencê-lo. Se ele o faria ou não, talvez ele não se atreva a assegurá-lo; mas que isso lhe seja possível, tem que admiti-lo sem hesitação. Portanto ele julga que pode algo pelo fato de ter a consciência de que o deve, e reconhece em si a liberdade, que do contrário, sem a lei moral ter-lhe-ia permanecido desconhecida (KpV. Ak. 54, p.103).

Ora, sem a consciência da lei moral o homem não poderia reconhecer a liberdade, isto é, sua autonomia. E a autonomia evoca outro conceito basilar à

---

necessidade de evitar confusão entre um fato empírico e um fato não empírico.

[37] "[…] e se a voz da razão em relação à vontade não fosse tão clara, tão intransferível, tão perceptível mesmo ao homem mais comum, ele arruinaria completamente a moralidade" (KpV. Ak. 62. p. 119).

[38] Ainda que decidamos não agir segundo sua ordem.

moral kantiana: o conceito de respeito (*Achtung*). Este conceito, diferentemente do *factum*, já aparece na GMS. Em uma valiosa nota de rodapé da primeira seção da GMS, Kant coloca a ideia de respeito nos seguintes termos:

> Poderiam objetar-me que eu, por trás da palavra respeito, busco apenas refúgio num sentimento obscuro, em vez de dar informação clara sobre esta questão por meio de um conceito da razão. Porém embora o respeito seja um sentimento, não é um sentimento *recebido* por influência; é, pelo contrário um sentimento que *se produz por si mesmo* através dum conceito da razão, e assim é especificamente distinto de todos os sentimentos do primeiro gênero que se podem reportar à inclinação ou ao medo. Aquilo que eu reconheço imediatamente como lei para mim, reconheço-o com um sentimento de respeito que não significa senão a consciência da *subordinação* da minha vontade a uma lei, sem intervenção de outras influências sobre a minha sensibilidade. A determinação imediata da vontade pela lei e a consciência desta determinação é que se chama *respeito*, de modo que se deve ver o *efeito* da lei sobre o sujeito e não a sua *causa*. O respeito é propriamente a representação de um valor que causa dano ao meu amor próprio. É portanto alguma coisa que não pode ser considerada como objeto nem da inclinação nem do temor, embora tenha algo de análogo com ambos simultaneamente. O *objeto* de respeito é portanto simplesmente a *lei*, quero dizer aquela lei que nos impomos *à nós mesmos*, e no entanto como necessária em si. Como lei que é, estamos-lhes subordinados, sem termos que consultar o amor próprio; mas como lei que nós nos impomos a nós mesmos, é ela uma conseqüência da nossa vontade e tem, de um lado, analogia com o temor, e, do outro, com a inclinação (GSM, Ak, 401).

Com efeito, *Achtung* é um sentimento, mas não um sentimento empiricamente condicionado. Temos, aqui, uma importante distinção: sentimento (*Gefühl*) e sensação (*Empfindungen*). O segundo (*Empfindungen*) é patológico, sendo, pois, provocado por objetos externos a nós. O primeiro (*Gefühl*), em contrapartida, é o efeito da *representação* de um objeto. E nesse momento parece valioso apontar para aquela passagem da GMS em que Kant nos diz: “tudo na

natureza age segundo leis. Só um ser racional tem a capacidade de agir segundo a representação das leis, isto é, segundo princípios" (GSM, Ak, 412). Assim, a determinação imediata da vontade pela lei, bem como a consciência dessa determinação, é chamada, por Kant, de respeito. Este constitui, pois, um efeito da lei (sua consciência) sobre o sujeito. Na GMS respeito (*Achtung*) tem apenas algumas linhas em uma nota de rodapé. Na KpV, entretanto, ele tem lugar de destaque, muito provavelmente em virtude de ele dar suporte ao conceito de *factum* e, dessa forma, sedimentar a metafísica da moral almejada por Kant.

Portanto, se na terceira seção da GMS Kant intentou deduzir a lei moral a partir da ideia de liberdade, na KpV ele nega uma tal dedução[39]. Temos, aqui, a intervenção do conceito de respeito, este sentimento (*Gefühl*) peculiar despertado em nós pela "consciência da lei moral" (uma das 'definições' de *factum*). Este conceito é plenamente desenvolvido na KpV em virtude de ele estar ligado à ideia de *factum* e de justificação da lei moral. Assim, a relação entre *factum* e *Achtung* é essencial à KpV. Tal união nos permite fundamentar que a razão pode ser prática.

Assim, se a lei moral é auto legisladora (não se fundamentando sobre elementos sensíveis[40]), todos aqueles que estiverem cônscios de que estão sujeitos a ela terão um móbil (*Triebfeder*) para obedecê-la: o respeito, o qual é "despertado" em nós pela consciência da lei. Uma consciência presente nos sujeitos e que deve ser "evocada" socraticamente. Nesse sentido, a tarefa do *moralische Lehrer* é fundamentalmente socrática[41]. A questão é "trazer à lume" a lei moral ínsita em nossos juízos cotidianos.

---

[39] Lembremos: na KpV Kant é enfático ao afirmar que a lei moral é *ratio cognoscendi* da liberdade, enquanto esta é a *ratio essendi* daquela. Portanto, não há "círculo" algum na KpV. Aqui a lei moral não é estabelecida a partir da liberdade: a liberdade é "conhecida" pela consciência que temos da lei moral (esta é a *ratio cognoscendi* daquela).

[40] Daí Kant rejeitar duramente a *Populärphilosophie*, isto é, a ideia de que devemos recorrer à experiência para dela tomar exemplos úteis à moral (uma ideia presente em autores tais quais Garve e Mendelssohn).

[41] Como, aliás, o próprio Kant indica em sua *Die Metaphysik der Sitten* (Zweiter Teil: Metaphysische Anfangsgründe der Tugendlehre. Ak. 376), ao falar em "catequizar socraticamente".

> Assim, no conhecimento moral da razão humana vulgar, chegamos nós a alcançar o seu princípio, princípio esse que a razão vulgar em verdade não concebe abstratamente numa forma geral, mas mantém sempre realmente diante dos olhos e de que se serve como padrão dos seus juízos. Seria fácil mostrar aqui como ela, com essa bússola na mão, sabe perfeitamente distinguir, em todos os casos que se apresentem, o que é bom e o que é mau, o que é conforme ao dever ou o que é contrário a ele. Basta, sem que com isso ensinemos nada de novo, que chamemos a sua atenção, como fez Sócrates, para o seu próprio princípio, e que não é preciso nem ciência nem filosofia para que ela saiba o que há a fazer para ser honrado e bom, mais ainda, para ser sábio e virtuoso (GMS. Ak. 403-404).

Na GMS (Ak. 402) Kant deixa claro, após enunciar o imperativo categórico, que com ele está "de acordo a comum razão humana nos seus juízos práticos e tem sempre diante dos olhos este princípio".

Uma dedução, *stricto sensu*, exigiria, pois, um conhecimento teórico prévio. Tal não ocorre no âmbito prático, onde o senso comum (*Menschenvernunft*) utiliza, ainda que sem o saber, tal princípio. Essa ideia, já presente na GMS, retorna na KpV, onde ele assere (referindo-se ao *gemeinste Verstand*):

> A regra da faculdade de julgar sob leis da razão prática pura é esta: pergunta a ti mesmo se poderias de bom grado considerar a ação, que te propões, como possível mediante a tua vontade, se ela devesse ocorrer segundo uma lei da natureza da qual tu mesmo fosses uma parte. Segundo essa regra, efetivamente, qualquer um ajuíza se as ações são moralmente boas ou más. Assim se diz: se *cada um* se permitisse enganar onde crê alcançar o seu proveito, ou se considerasse autorizado a abreviar sua vida tão logo o acometa um completo tédio dela, ou se olhasse com inteira indiferença para a necessidade de outros, e se tu também pertencesses a uma tal ordem de coisas, como irias estar de bom grado nela com o assentimento de tua vontade? Ora cada um sabe bem que, se ele se permite secretamente a fraude justamente porque nem todos também a pratiquem ou se ele sem ser notado é desumano, nem por isso todos também o seriam imediatamente contra ele; por isso esta comparação da

> máxima de suas ações com uma lei natural universal não é tampouco o fundamento determinante de sua vontade. Mas esta lei é, contudo, um *tipo* do ajuizamento daquela máxima segundo princípios morais. Se a máxima da ação não é constituída de modo tal que resista à prova na forma de uma lei natural em geral, ela é moralmente impossível. *Mesmo o entendimento mais comum julga desse modo*; pois a lei da natureza é encontrada sempre como fundamento de todos os seus juízos mais comuns, até dos juízos de experiência. Portanto ele a tem sempre à mão […][42].

Tal passagem encontra-se precisamente na seção intitulada "Da típica da faculdade de julgar prática pura", a qual seria o análogo, em filosofia prática, da "Doutrina do esquematismo dos conceitos puros do entendimento", dado que nela Kant trata de uma espécie de "esquematismo" da Razão Prática pura (de sua 'faculdade de julgar'). Dito em outros termos, na *Typik* Kant trata daquilo que muitas vezes chama-se, inapropriadamente, de "dedução", a qual recai, como momento derradeiro, na doutrina do *factum*. Aliás, pouco antes, ao apresentar a "Tábua das categorias da liberdade relativamente aos conceitos de bom e mau", sobre sua justificação (das "categorias da liberdade"), Kant é de uma tranquila (e inquietante) brevidade: "Não acrescento aqui mais nada para a elucidação da presente tábua, porque ela é por si suficientemente compreensível"[43].

Dessa forma, é impossível, bem como desnecessária, uma dedução da lei moral. Ela já está dada. Tal "dadidade" da lei moral é manifesta, como *factum*, sobretudo, no sentimento do respeito à lei. Dessa forma, de uma mera notá de rodapé na GMS, a ideia de respeito passa a ter um papel deveras importante na KpV: aqui ele implica a consciência, *a priori*, do imperativo moral (de sua validade objetiva). Ele possui um fundamento racional por ser "produzido" pela razão mesma. E a partir desse momento ele se faz "sentir". O respeito é um motivo racional que produz duas sensações sobre nossa sensibilidade, a saber, a humilhação ou dor e a elevação. Assim, se num primeiro momento nossa

[42] KpV. Ak. 123. p. 241, grifos nossos.
[43] KpV. Ak. 116, p.230.

sensibilidade é humilhada por reconhecer, a partir da consciência da lei, que se guiava por elementos heterônomos, ela em seguida é elevada por reconhecer que o sujeito é livre, isso é, autônomo (capaz de encontrar em si a lei). Portanto, pelo respeito tal lei se mostra de forma inegável aos sentidos.

E é nesse ínterim que vemos a importância do sentimento moral. A propósito, ainda quanto a este, uma importante contribuição para o debate nós a encontramos em Henry Allison e sua *incorporation thesis*[44], a qual introduziu na discussão o papel do sentimento (*feeling*, *Gefühl*). Tal tese defende que, para que ajamos de acordo com a lei moral, devemos inserir em nossas máximas incentivos não racionais. Afinal, como vimos acima, uma teoria do agir moral, da motivação moral, demanda uma abordagem do sentimento, especialmente se considerarmos que não somos santos, mas sujeitos sensíveis que podem agir de acordo com a razão. Seria, afinal de contas, um absurdo pensar que a razão nos exige algo inexequível. Se ela nos exige algo, então é ao menos possível, para nós, seguir os seus ditames, a já referida "voz da razão".

Mas a questão é: que móbeis são legítimos à luz da teoria moral kantiana, à luz da autonomia? Aliás, o receio de cair na armadilha empirista, de conceder demasiadamente ao sentimento, levou muitos comentadores a subestimarem o papel do sentimento quando tratavam da motivação moral em Kant.

De qualquer forma, a tese de Allison oferece, certamente, uma nova perspectiva para o debate. Ela pelo menos nos faz refletir sobre a necessidade de uma influência da razão sobre a sensibilidade, para que possamos agir, para que a razão se converta em ação.

Assim, para Kant somos seres finitos dotados de razão. Isso demanda móbeis para nossa ação, para que nos coloquemos em movimento. Como mencionei ao início desse texto, saber o que é moral não significa agir a partir do que é moral. Precisamos de um incentivo ao agir, um incentivo para agir. E, dado que somos sujeitos sensíveis dotados de razão, faz-se necessária uma influência

[44] Tal tese é desenvolvida de modo seminal por Allison (1990). Muitos comentadores consagrados acolheram, com entusiasmo, a tese de Allison, como Christine Korsgaard, Allen Wood e Paul Guyer.

da razão sobre a sensibilidade. O sentimento deve, então, exercer alguma influência sobre o desejo, o qual é entendido como a "faculdade do mesmo ente de ser mediante suas representações, causa da efetividade dos objetos destas representações"[45]. Uma definição similar aparecerá na Metafísica dos Costumes (doravante 'MS'): "A faculdade de desejar é a faculdade de ser, por intermédio das suas próprias representações, causa dos objetos dessas representações"[46].

Logo, para impelir um agente a agir, é preciso que a faculdade de desejar seja, digamos, "tocada". Neste caso, ela precisa ser tocada pela razão. Ainda que isso seja algo difícil, ainda assim é possível. Isso porque Kant deixa claro que nem sempre a razão determina a ação, isto é, nem sempre (pelo menos não necessariamente) o agir é autônomo. Os móbeis, condições subjetivas, "não coincidem sempre" com as condições objetivas[47]. Mas do fato de que eles não "coincidem sempre" com o motivo, depreende-se a possibilidade de que eles, ainda que raramente, coincidam. Logo, o agir racional não é impossível.

Mas o ponto é que a determinação subjetiva Kant a denomina móbil (*Triebfeder*) para a ação. Isso porque o motivo deve converter-se em móbil. De nada ele vale se ele permanece apenas um motivo. Para nos levar à ação, ele deve tornar-se móbil. Esse é o problema que Kant nos apresenta na KpV, especialmente no terceiro capítulo da "Analítica da razão prática pura", intitulado, precisamente, "Dos móbeis da razão prática pura". Na divisão da MS temos uma interessante caracterização desse "móbil": ele "liga subjetivamente com a representação da lei o fundamento da determinação do arbítrio para a realização dessa ação"[48]. E isso vale, como ele assevera nesse mesmo contexto, para "toda a legislação" [...], prescreva ela ações internas ou externas". Tais móbeis, aliás, valem para seres sensíveis finitos, dado que somos levados (na maioria das vezes) a agir por móbeis que não são oriundos da ideia de dever. E o único móbil moral é, como afirma Kant, o já mencionado respeito pela lei moral. Isso é descrito

[45] KpV, A 9, p.29, nota.
[46] MS, 210, p.15.
[47] Cf. GMS, 413, p.51.
[48] MS 218, p.26.

por Kant na seguinte passagem da KpV:

> Em primeiro lugar, a lei moral determina objetiva e imediatamente a vontade no juízo da razão; mas a liberdade, cuja causalidade é determinável simplesmente pela lei, consiste precisamente em que ela limita todas as inclinações, por conseguinte a estima da própria pessoa, à condição do cumprimento de sua lei pura. Ora, essa limitação promove um efeito sobre o sentimento e produz uma sensação de desprazer, que pode ser conhecida *a priori* a partir da lei moral[49].

Isso nos conduz, com efeito, a uma notável definição de sentimento moral, a qual Kant apresenta na MS, a denominando de "receptividade do livre arbítrio para ser movido pela razão pura prática (e pela sua lei)"[50].

Tal "receptividade" é necessária. Afinal, reconhecer o que deve ser feito não nos leva, pelo menos não necessariamente, a agir. E nós somos, conforme Kant, levados tanto à ação pela lei (em que o motivo converte-se em móbil) quanto pelo interesse particular (movidos por móbeis empíricos). No texto *Religião nos limites da simples razão* (1793; doravante 'Rel') Kant torna isso significativamente claro. Nessa obra Kant nos fala em uma "disposição para o bom" e uma "propensão – *propensio* – para o mal" na natureza humana. Um ser racional finito toma, pode tomar, o respeito à lei ou o interesse particular como móbil para a ação. Ainda no texto da Rel (23, p.36) Kant nos diz: "O espírito da lei moral consiste em que ela só seja suficiente como móbil". Mas isso não significa ignorar nossa sensibilidade, como se fossemos seres angelicais, capazes de agir unicamente a partir da ideia de dever. Como coloca Kant na Rel (33, p.42), o homem "depende também, em virtude da sua disposição natural igualmente inocente, de móbeis da sensibilidade e acolhê-los outrossim na sua máxima (de acordo com o princípio subjetivo do amor de si)"[51]. Ou, ainda, "ele acolhe de modo natural ambas as coisas na sua máxima" (Rel. 33, p.42).

---

[49] KpV, 139, p. 273.
[50] MS, 400, p.312.
[51] Essa é a passagem central para compreendermos a tese de Allison, sua *Incorporation Thesis*.

São dois os móbeis nesse contexto, o amor de si e o respeito à lei. Logo, "a diferença de se o homem é bom ou mal deve residir, não na diferença dos móbeis, que ele acolhe na sua máxima (não na sua matéria), mas na subordinação (forma da máxima): de qual dos dois móbiles ele transforma na condição do outro" (Rel. 34, p.42). Logo, trata-se da primazia de um em relação ao outro. Ambos estão envolvidos. A questão é: ou colocamos a lei (e o respeito por ela) como móbil, ou colocamos o amor de si como móbil. Mas de uma forma ou de outra, temos que ser "tocados" na sensibilidade.

Com efeito, na já referida MS Kant fala em um "prazer prático"[52], o qual está ligado ao prazer. Nem todo prazer está ligado ao desejo, como fica claro na primeira parte da *Crítica da faculdade do juízo* (*Kritik der Urteilskraft*, 1790, doravante 'KU'). No entanto, o prazer prático envolve desejo, sendo este a faculdade de ser, por intermédio das suas próprias representações, causa dos objetos dessas representações (MS 211, p.15). Temos, aqui, uma suscetibilidade, *Empfänglichkeit*, ou, ainda, *Gefühl* (MS 211, p.16, nota). Nessa mesma nota esclarecedora, Kant sustenta que esse sentimento é uma suscetibilidade diante da representação. Ele contém um efeito da representação, "seja esta sensível ou intelectual", pertencendo, pois, à sensibilidade.

Ora, essa ideia de suscetibilidade é esclarecedora. Isso porque suscetibilidade, ou, ainda, o sentimento de suscetibilidade, envolve certa dependência com relação àquilo a que somos suscetíveis. O sentimento não é algo de imediato, algo a que se quer imediatamente, mas algo que se quer como reação a uma dada representação, seja esta sensível ou intelectual: tal suscetibilidade à representação é, na definição de Kant, "sentimento o qual contem um efeito da representação (seja esta sensível ou intelectual) sobre o sujeito e pertence à sensibilidade, ainda que a representação possa pertencer ao entendimento ou à razão" (MS 212, p.16). Em suma, um sentimento sempre brota

[52] Tal prazer (*Lust*) Rudolf Eisler o identifica, em seu Kant-Lexikon, com 'sentimento', 'sentimento moral' e 'respeito'. Em seu verbete, lemos: "Praktische *Lust* s. Gefühl, Moralisches Gefühl, Achtung" (Eisler 1994).

de uma suscetibilidade, seja a uma representação da sensação, seja a uma representação da razão. Ele tem objeto, ainda que não um objeto da sensação, a qual sempre se refere exclusivamente ao sujeito, sem valor objetivo, portanto. Seu objeto pode, todavia, ser oriundo da razão.

O sentimento é, pois, uma expressão sensível (dado relacionar-se com a sensibilidade) mediante representação. Trata-se de um sentimento télico, e não cego. Ele se relaciona com o desejo, e por essa razão, como vimos anteriormente, a consciência da lei causa dor: ela, a dor, é um afastamento do objeto exterior do desejo condicionado heteronomamente.

Com efeito, o prazer estético é, sem dúvida alguma, desinteressado. Mas o prazer prático, mesmo no âmbito do sentimento moral, tem um interesse. Vejamos, novamente, a MS: o prazer prático é um "prazer que está necessariamente conectado com o desejo (do objeto, cuja representação afeta assim o sentimento)".

Nesse sentido, o sentimento determina a faculdade de desejar e produz uma inclinação e um desejo (daí seu jaez télico), o qual constitui um prazer prático. E com isso ele se torna um móbil para o agente, despertando, a sua existência, um interesse. Nos termos de Kant,

> toda a determinação do arbítrio vai desde a representação da possível ação até a ação, mediante o sentimento de prazer e desparazer, ao assumir um interesse nela ou no seu efeito; nesse caso, o estado estético (a afecção do sentido interno) é ou um sentimento patológico ou um sentimento moral. O primeiro é aquele sentimento que antecede a representação da lei, o último, aquele que só pode surgir na sua sequência (MS 399, p.312).

Assim, o sentimento moral surge da representação da lei. Tal representação engendra um interesse, um interesse da razão, da liberdade[53]. Na KpV (141, p.277) vemos que uma máxima "somente é autenticamente moral se depende do mero interesse que se toma pela observância da lei". E esses três conceitos, como assevera Kant nesse mesmo contexto, móbil, interesse e máxima, aplicam-se unicamente a seres finitos (sensíveis dotados de razão), dado estes possuírem

[53] Sobre esse tema, ver estudo seminal do Professor Valério Rohden (1981).

"uma carência de ser impelido por algo à atividade".

Não obstante, uma conexão entre arbítrio e interesse nós já a encontramos na GMS (459, p.96). Nesse contexto Kant está preocupado com o "interesse que o homem possa tomar pelas leis morais". E o homem

> toma realmente interesse por elas, cujo fundamento em nós, é o que chamamos de sentimento moral, sentimento que alguns têm falsamente apresentado como padrão do nosso juízo moral, quando é certo que ele deve ser considerado antes como o efeito subjetivo que a lei exerce sobre a vontade e do qual só a razão fornece os princípios objetivos (GMS, 460, p.96-97).

Com efeito, as passagens acima nos permitem compreender em que sentido não estamos tratando, quando falamos de sentimento moral, de ações meramente fundadas sobre o desejo, mas de um agir racional que toma como *mola propulsora* o sentimento no plano da faculdade de desejar, do prazer prático. E tal desejo pode ser moralmente bom ou moralmente mau. Isso dependerá do seu objeto. O objeto do bom é "um objeto necessário da faculdade do desejo" (KpV, 101, p.197). Afinla de contas, "a ética oferece ainda uma matéria (um objeto do arbítrio livre), um fim da razão pura" (MS, 380, p.284). Ora, tal fim desperta, também, um interesse em nós. E, quando tomamos um interesse pelo objeto, tal objeto deixa de ser um mero objeto do desejo. Ele se torna um fim. Mas, diferentemente do que ocorre no prazer pelo agradável, aqui, no plano do sentimento moral, não temos um conhecimento do objeto. Nesse caso, "o elemento subjetivo da nossa representação não pode tornar-se elemento de conhecimento, porque contem unicamente a relação da representação com o sujeito e não contem nada de útil para o conhecimento do objeto" (MS, 211, p.16).

Resumindo, dir-se-ia, então, que o sentimento é uma suscetibilidade à representação. Ele permite que estejamos cientes acerca de como a representação do objeto pode nos afetar de uma perspectiva subjetiva. E este "estarmos cientes" envolve não apenas o mero conhecimento de nosso estado, mas refere-se, sobretudo, ao agir, à práxis. Ele nos permite saber por que devemos, enquanto indivíduos, agir moralmente, tomando a lei, portanto, como

móbil. Ao tocar em nossa sensibilidade, ele deixa de ser meramente formal e ganha substancialidade, força propulsora.

Referências

ALLISON, H. *Kant's theory of freedom*. Cambridge: Cambridge University Press, 1980.

ALMEIDA, G. A. "Kant e o 'fato da razão': 'cognitivismo' ou 'decisionismo' moral?". In: *Studia Kantiana* 1.1, p.53-81, 1999.

_____. "Liberdade e moralidade segundo Kant". In:*Analytica* 2.1, p.175-202, 1997.

BECK, L. W. *A commentary on Kant's Critique of practical reason*. Chicago: The University of Chicago Press, 1960.

_____. *Essays on Kant and Hume*. New Haven: Yale University Press, 1978.

_____. *Studies in the philosophy of Kant*. New York: The Bobbs-Merrill Company, INC, 1965.

EISLER, R. *Kant-Lexikon*. Hildesheim: Olms, 1994.

FÖRSTER, E. (Ed.). *Kant's Transcendental Deductions*. Stanford: Stanford University Press, 1989.

GUYER, P. *Kant on freedom, law, and happiness*. Cambridge: Cambridge University Press, 2000.

_____. *Kant and the experience of freedom.* Cambridge: Cambridge University Press, 1996.

HAMM, C. V. "Princípios, motivos e móbeis da vontade na filosofia prática kantiana". In: Napoli, R. B.; Rossato, N. D. e Fabri, M. (orgs.), *Ética e justiça*. Santa Maria: Pallotti, 2003, p.67-82.

HENRICH, D. *Der Begriff der sittlichen Einsicht und Kants Lehre vom Faktum der Vernunft*. Hrsg: Gerold Prauss, Kiepenheuer & Witsch Köln, 1973.

KANT, I. *Werke in sechs Bänden. Band IV: Schriften zur Ethik und Religionsphilosophie*. Darmstadt: Insel Verlag, 1998.

_____. *Anthropologie in pragmatischer Hinsicht*. Stuttgart: Reclam, 1998b.

_____.*Crítica da razão prática*. Edição bilíngue. São Paulo: Martins Fontes, 2003.

_____. *Crítica da faculdade do juízo*. Lisboa: Imprensa Nacional – Casa da Moeda, 1992.

_____. *A religião nos limites da simples razão*. Lisboa: Edições 70, 1992.

_____. *Fundamentação da metafísica dos costumes*. Porto: Porto Editora, 1995.

_____. *A metafísica dos costumes*. Lisboa: Fundação Calouste Gulbenkian, 2005.

MUNZEL, G. F. *Kant's conception of moral character*. Chicago: The University of Chicago Press, 1999.

PHILONENKO, A. *Études kantiennes*. Paris: Vrin, 1982.

PATON, H. J. *The categorical imperative*. New York: Harper, 1967.

ROHDEN, V. *Interesse da razão e liberdade*. São Paulo: Ática, 1981.

15

MOTIVAÇÃO E VALOR EM HUME E KANT

André Klaudat

Introdução

Em suas filosofias morais, em particular na parte ou aspecto *fundacional*, David Hume e Immanuel Kant dão lugar especial à *motivação moral*. Hume registra o fato de que a moral é *prática*, e afirma que sem essa "vantagem" ele nunca teria se voltado ao tratamento filosófico do assunto. Kant, por sua vez, faz o *motivo do dever* ser a porta de entrada de sua filosofia moral. Ambos têm boas razões para tal opção. Em primeiro lugar, historicamente, racionalistas como Samuel Clarke procuravam esclarecer a natureza e o conteúdo das leis morais sem preocupação com o *porquê* de o homem dever se orientar por elas. A agenda filosófica passava a conter um tópico incontornável.

Em segundo lugar, para além da preocupação com o ceticismo moral, se nos centrarmos, de um ponto de vista sistemático, em formas de obrigação moral, a pergunta sobre por que estamos moralmente constrangidos, sobre qual é o *rationale* para essas formas de obrigação, parece ser, definitivamente, uma boa pergunta. Tanto Hume quanto Kant têm olhos para essa pergunta.

Em terceiro lugar – e aqui parece haver uma diferença importante nas posições dos dois filósofos – enquanto que Kant sustenta que atentar para o motivo do dever é o caminho para a absolutamente necessária e inédita explicação da natureza e *justificação* da obrigação moral (a obrigação *incondicional*), o que permitirá o esclarecimento dos fundamentos da moral; Hume, por seu turno, faz da "praticalidade" da moral uma espécie de condição de adequação material para a *explicação* do fenômeno moral, se não estivermos explicando o que de fato nos move, não estaremos explicando a moralidade. Mas quanto a essas razões para

Hume e Kant se voltarem para a motivação moral, o cuidado é recomendável se seguirmos certos intérpretes desses filósofos. Quanto a Kant, não obstante a aparência em contrário, uma ética dos deveres alicerçada *deontologicamente não* seria de fato a posição kantiana. Quanto a Hume, se é o caso que deveres e formas normativas de determinação da vontade não ocupam um lugar central em sua análise, ele não resume seu interesse na moralidade a um esforço meramente explicativo.

Meu objetivo aqui é, primeiramente, tematizar as concepções da motivação moral dos dois filósofos. Minha convicção é que em relação a essa, Hume e Kant não se distanciam quanto a um elemento central dessa motivação, a saber, o elemento *negativo* da independência dessa motivação das inclinações e dos interesses *pessoais atuais* dos agentes. Prospectivamente, parecer-me-á que as posições de Hume e Kant se distinguem quanto ao que é o sustentáculo de suas teorias morais, a concepção de valor, daquilo que ao final constituirá o *rationale* da motivação moral. Quanto a essa concepção, um modo de se chegar a ela, e de perceber a sua centralidade, é se colocar a pergunta que põe a dificuldade num ponto: a moralidade é constituída, fundamentalmente, de algo que antes *entendemos*, ou antes *sentimos*?[1]

Mesmo que, tentativamente, encontre uma diferença irreconciliável, nisso que é o conceito fundamental da ética, entre Hume e Kant, penso que, quanto à defesa das suas concepções éticas, num ponto elas *não* se distinguem: a concepção de valor ser o *rationale acessível*, se não ao julgamento moral ordinário (ou à "distinção moral" particular), então à deliberação moral (ou à avaliação do porquê do sentimento moral dever ser observado). Nesse sentido, meu objetivo principal aqui não é adjudicar as filosofias morais de Hume e Kant, mas procurar

[1] Cf. David Wiggins: "We feel *bound*. But why *reason-bound*? Why not say we feel bound by our moral nature, i.e. bound by those moral sentiments without which (we have concluded, if we feel bound by obligation) we should not recognize ourselves? (Bound then by reason of those sentiments, one may want to say)" ("Categorical requirements: Kant and Hume on the idea of duty". In: Hursthouse, Lawrence e Quinn (eds.), *Virtues and reasons, Philippa Foot and moral theory*, Clarendon Press, 1995, p.310. Cf. também p. 323-325, n.19).

averiguar, tentando me colocar nos ombros deles, quais são as nossas opções.

1. Motivação e obrigação moral

A referência de Hume ao que veio a ser chamado de "praticalidade" da moral é inequivocamente ao seu aspecto motivacional *enquanto moralidade*. A dita "vantagem" diz respeito ao que nos *afeta* sobremaneira (cf. T 455-6[2]), que, por isso, concluímos não pode ser uma "quimera". O ponto de partida de Hume é que a moralidade nos move como *moralidade*:

> Se a moralidade não tivesse naturalmente qualquer influência sobre as paixões e ações humanas, seria em vão fazer tanto esforço por inculcá-la; e nada seria mais infrutífero que a multidão de regras e preceitos com os quais todos os moralistas abundam. [...] E isso é confirmado pela experiência comum, que nos informa que os homens são muitas vezes governados pelos seus deveres e são desestimulados [*deter'd*] para certas ações pela opinião de injustiça, e impelidos a outras por aquela de obrigação (T 457).

Que Hume, nesse contexto, insista que a moralidade tem influência *naturalmente* não parece querer dizer mais que é isso que constatamos na experiência comum. Mas o que é dito ser assim influente são nossos *deveres* e *obrigações morais.* A moralidade, consequentemente, é um "princípio ativo" (T 457) em relação a nós seres humanos.

Mas se nos perguntarmos, seguindo Hume (T 473), se os "princípios" dos quais a moralidade é "derivada" (quanto ao seu conteúdo) são *naturais*, uma pergunta sobre *origens*, a resposta dele é que nossas noções das virtudes podem se subdividir em dois grupos: temos o sentido de que algumas virtudes são *artificiais*, enquanto que outras são *naturais* (T 475). O que Hume nos diz sobre essa distinção – entre natureza e artifício – pode ser uma clara indicação quanto à

[2] *A treatise of human nature*, p.455-456 (Oxford University Press, Selby-Bigge/Nidditch, 1978): assim as referências no texto.

sofisticação da sua posição sobre a "fundação da moralidade"[3].

Embora Hume seja deflacionário quanto à exuberância *intelectual* de planos e projetos de homens, ele concede propriedade à nossa representação deles como "livres e inteiramente de nossa lavra" [*entirely our own*] (T 474), desse modo vendo-os como *artifícios* a serem opostos "a outros princípios da natureza" (T 474). É nessa acepção somente que o sentido de algumas virtudes é não-natural: pois dizem respeito a planos e projetos frutos de inventividade humana.

Mas arroubos intelectualistas são frustrados pela afirmação de que, de certo ponto de vista, esses planos e projetos "são princípios tão necessários na sua operação quanto calor e frio, úmido e seco" (T 474). Na medida em que há a concessão quanto à propriedade de se falar de livres e nossos em relação a eles, então a analogia com a termodinâmica não pode dizer respeito à inevitabilidade com que, p. ex., objetos e substâncias expandem quando aquecidos, mas à *naturalidade* das operações das quais nós somos capazes[4]. A moralidade é natural ao ser (1) baseada em princípios gerais naturais, (2) oposta a milagres, e (3) oposta ao que é raro e não usual, mas não no sentido de ser (4) toda ela sem a inventividade humana para a criação de artifícios.

Voltemo-nos, então, à naturalidade das nossas operações quanto à obrigação moral *em geral* (sem distinção entre as virtudes naturais e artificiais) segundo Hume. Tomemos o caso do *dever* paterno do *cuidado* com os filhos (NB: não o "dever" de afeição com os nossos filhos, cf. T 478, abaixo), o qual se assenta sobre um sentimento natural de humanidade:

---

[3] "…Essa afecção de humanidade pode não ser estimada em geral como tão forte como a vaidade ou ambição, mas sendo comum a todos os homens, ela sozinha pode ser a fundação da moralidade, ou de qualquer sistema de censura ou elogio" (*Enquiry concerning the principles of morals*, Oxford University Press, Selby-Bigge/Nidditch, 1975, p.222 = EPM 222).

[4] "O interesse no qual a justiça está fundada é o maior imaginável e se estende para todos os tempos e lugares. Ele não pode concebivelmente [*possibly*] ser servido por nenhuma outra invenção. Ela é óbvia e se revela na primeira formação da sociedade. Todas essas causas tornam as regras da justiça firmes e imutáveis; ao menos, tão imutáveis quanto a natureza humana. E se estivessem fundadas em instintos originais, poderiam elas ter uma estabilidade maior?" (T 620).

> Nós censuramos um pai por negligenciar seu filho. Por quê? Porque mostra uma falta de afeição natural, que é o dever de todo pai. Não fosse a afeição natural um dever, o cuidado com os filhos não poderia ser um dever; e seria impossível que nós tivéssemos o dever em vista na atenção que damos à nossa prole. Nesse caso, todos os homens supõem um motivo para a ação distinto do senso de dever (T 478).

Com isso temos (1) que o senso de dever, de obrigação, é *consequente* a uma motivação natural:

> Um pai sabe que é seu dever cuidar do seu filho: mas ele também tem uma inclinação natural a isso. E se nenhuma criatura humana tivesse essa inclinação, ninguém poderia encontrar-se sob tal obrigação (T 519).

Com isso, por sua vez, temos (2) que a motivação natural nesse caso é uma *inclinação natural*.

E quanto à natureza da obrigação, ou dever, em geral, temos:

> Toda moralidade depende de nossos sentimentos: e quando qualquer ação, ou qualidade da mente, nos agrada *segundo uma certa maneira*, dizemos que ela é virtuosa; e quando o descuido ou a não realização dela nos desagrada *segundo uma mesma maneira*, nós dizemos que estamos sob a obrigação de realizá-la (T 517).

O quadro, então, parece dever ser o seguinte: (1) homens, comumente, têm afeição por seus filhos, têm uma inclinação natural por eles; (2) em função desse fato, o indivíduo que não tem essa afeição – que garantiria o cuidado – estará *obrigado* ao cuidado de seu filho por causa de uma *avaliação* negativa dessa carência: todos, inclusive o indivíduo em questão, se *desagradam moralmente* com essa situação.

O que é fundamental compreender aqui é que a *obrigação moral* da virtude natural do cuidado com os filhos *não* se assenta na *ocorrência* efetiva de um desprazer consigo mesmo (individual) ante a *comparação* com a situação de outros, como se Hume estivesse sustentando que só há obrigação para o sujeito

quando ele sente um desprazer efetivo (quando há esse sentimento nele)[5]. Ao contrário, a obrigação, o dever em questão, se constitui do modo que faz (como *exigência*) certamente em função da resposta natural dos homens aos seus filhos, mas a exigência se apresenta como tal independentemente do que ele de fato sente, *porque* outros, e ele próprio eventualmente, sentem um desprazer de um *ponto de vista comum* (T 591), do ponto de vista de um *espectador judicioso* (T 581), um desprazer de caráter *moral*.

E para que não se pense agora que Hume procede em círculo – justificando o dever moral pela moralidade do ponto de vista comum – é crucial reconhecer sua *agenda*. O interesse primário de Hume é *explicar a fonte* das "distinções morais" que distinguem o bem do mal em caracteres, ou seja, qual é a fonte da moralidade em relação a seres como nós, seres naturais (T 458-9). A investigação dele é sobre a "origem da moralidade" (T 575). Hume começa sua investigação como um "psicólogo moral", ele está interessado em primeiro lugar na nossa *psicologia moral*. David Wiggins coloca esse ponto sobre a "gênese da moralidade" da seguinte maneira:

> Sua [de Hume] primeira questão não é a questão da justificação do dever (essa questão pode surgir posteriormente), mas a questão de como seres naturais como nós podem ter alcançado um ponto a partir do qual podem mesmo pensar os pensamentos refinados e complexos que nós pensamos sobre a moralidade e suas exigências (Op. Cit., p. 300).

É em função de reconhecermos essa agenda que podemos ver que a análise de Hume do *dever* moral nos mostra essa obrigação como *secundária* em relação à constituição de um ponto de vista moral quanto às nossas interações. A *explicação* de Hume dessa constituição não vai nos ocupar aqui. Basta lembrar que o princípio é a afeição de humanidade (que faz com que nada humano nos seja indiferente), cujo fundamento se encontra no mecanismo da simpatia, que nos faz sentir uma espécie de satisfação (ou insatisfação) especial quando corrigida quanto aos nossos interesses particulares e alçada a nossos "pontos de vista

[5] Cf. Wiggins op. Cit., p.309, n. 9.

firmes e gerais" (T 581-2). É frente a esse pano de fundo que a obrigação moral, o *dever moral*, pode se apresentar, também em Hume – e é isto que é importante que se note – como uma exigência *categórica*, ou seja, independente das *inclinações atuais* do sujeito (Cf. Wiggins op. cit. Passim.). Nesse sentido, haverá "beleza moral" no comportamento naturalmente afetuoso e cuidadoso de um pai em relação ao seu filho, mas também haverá tal beleza no comportamento de cuidado *por dever* de um pai em relação ao seu filho, e não precisamos nos ocupar agora com a questão de Bernard Williams sobre qual tem mais.

O ponto *lógico* em questão é que o fato da fundação/origem de um dever ser *empírica*, de ele ter uma fundação *natural* vista como *a posteriori* (no sentido de se assentar em nossa natureza humana), *não* impede que a exigência moral em questão em seu conteúdo seja apresentada como *categórica*, sem assento nas inclinações *atuais* do sujeito. Hume afirma que "nosso senso de dever sempre segue o curso comum e natural de nossas paixões" (T 484), mas não obstante trata-se de um senso de *dever*. Se não fosse assim, empiristas em geral sobre os fundamentos da moral (como Harry Frankfurt) estariam logicamente impossibilitados de explicar/justificar as exigências morais *categóricas*.

O que esse ponto mostra, e isso é crucial, é que o motivo do dever na análise de Hume tem a característica exigida por kantianos: rastrear a moralidade, ter a "intencionalidade certa". Podemos ver isso na seguinte altercação.

O kantiano concede que o problema com o altruísta natural do exemplo de Kant *não* é que a sua ação careça de conteúdo moral por não ser regular, confiável ou até mesmo imparcial. Ela é: a ação beneficente é confiavelmente realizada quando da necessidade dos outros[6]. Segundo kantianos, o problema está na simpatia do altruísta natural não municiar o agente com um interesse na correção ou exigência da ação. A ilustração é a seguinte:

> Suponha-se que eu vejo alguém com dificuldades, tarde da noite, com uma carga pesada junto à porta dos fundos do *Museum of Fine Arts*. Por causa

[6] Cf. B. HERMAN. "On the value of acting on the motive of duty", p.4. In: *The Practice of moral judgement*, Harvard University Press, 1993.

> do meu temperamento simpático eu sinto uma inclinação imediata a ajudá-lo. Não precisamos seguir o exemplo para entender seu propósito: a classe de ações que se seguem da inclinação a ajudar os outros não é um subconjunto da classe de ações corretas ou exigidas por dever [*dutiful*][7].

Mas é claro que *esse* sujeito *não* é um sujeito virtuoso para Hume. Para ele, ajudar alguém nessas circunstâncias em Boston é certamente *burrice*, é como ele disse no 4º Apêndice ao *Enquiry Concerning the Principles of Morals* ("*Of Some Verbal Disputes*"): "Quem alguma vez diria, exceto por ironia, que tal sujeito era um homem de grande virtude, mas um emérito mentecapto [*egregious blockhead*]"[8].

Portanto, para Hume a ação de um agente virtuoso, dotado da virtude natural da beneficência, precisa ser a ação aprovada *moralmente* a partir da perspectiva de um espectador judicioso, de um ponto de vista comum, e não a ação para a qual – o alvo de Kant – o agente tem somente uma inclinação *imediata* natural. Racheal Cohon coloca o ponto destemidamente: para se agir virtuosamente ou bem para Hume, "não é suficiente fazer a coisa certa; nós precisamos fazê-lo pelas razões certas"[9]. Mas a "intencionalidade certa" para uma humeana está ligada a um motivo *natural* determinado: "Eu não manifesto minha benevolência ao ajudar um sem-teto se eu o ajudo em função do desejo de melhorar a minha reputação. O que eu faço é certo, mas não benevolente. Uma ação que manifesta a virtude da benevolência é motivada, ao menos em parte, pela preocupação como o bem-estar da outra pessoa. Essa é a razão certa" (Loc. Cit.).

Isso encerra minha abordagem da motivação moral quando da obrigação moral. Ambos os filósofos reconhecem a possibilidade de uma motivação atrelada à moralidade que é independente, por certo, das inclinações *atuais* dos agentes.

---

[7] "On the value of acting on the motive of duty", p.4-5.

[8] Cf. EPM p. 177 (Oxford University Press, 1975). A expressão crítica, em português, é do José Oscar de Almeida Marques na edição da Unicamp em 1995.

[9] "Hume on promises and the peculiar act of the mind". In: *Journal of the History of Philosophy* 44.1, p.27, 2006.

2. Motivação e valor

Mas e quanto à motivação independente das inclinações *simpliciter*? A análise de Kant da obrigação moral apresenta o seguinte ponto como a objeção principal a uma concepção como a humena dos fundamentos da moral:

> O que […] é derivado da particular predisposição natural da humanidade, o que é derivado de certos sentimentos e pendor e até mesmo, eventualmente, de uma direção particular que seria própria da razão humana e não teria de valer necessariamente para a vontade de todo ser racional – [tudo isso] pode, é verdade, vir a dar uma máxima para nós, mas não uma lei; um princípio subjetivo segundo o qual temos o pendor e inclinação a poder agir, mas não um princípio objetivo, segundo o qual estaríamos orientados a agir ainda que a isso se opusesse todo o nosso pendor, inclinação e constituição natural[10].

O que Kant parece ter em mente aqui não é a determinação da vontade, a motivação, sem qualquer, digamos, *desejar* [*Begehren*] da faculdade de desejar [*Begehrensvermögen*], mas sua determinação por desejos empíricos [*Begierde*] ou em sentido limitado (pré-determinados pelo prazer/desprazer). Qualquer *princípio* moral que se assentasse sobre essa base de desejos seria impróprio para a moralidade *a la* Kant (um princípio *material*, de acordo também com a segunda *Crítica*). Hume não parece poder concordar com isso. Para Hume, a moralidade *está naturalizada* no seguinte sentido. Vejamos.

Quando Hume nos revela do que depende o *mérito* das ações em geral, ele advoga um *naturalismo cum realismo* quanto aos primeiros motivos das mesmas:

> […] Todas as ações virtuosas derivam seu mérito somente de motivos virtuosos e são consideradas meramente como signos desses motivos. Desse princípio eu concluo que o primeiro motivo virtuoso, que confere mérito a qualquer ação, não pode nunca ser uma consideração pela virtude daquela ação, mas precisa ser algum outro motivo ou princípio natural.

[10] G 4:425. *Grundlegung zur Metaphysik der Sitten* (AA 04), p. 425 do volume da Akademie-Ausgabe. A tradução que utilizo é a de Guido de Almeida para a Barcarolla/Discurso, 2009.

> Supor que a mera consideração pela virtude de uma ação possa ser o primeiro motivo que produz a ação e a torna virtuosa é raciocinar em círculo. Antes que possamos ter tal consideração, a ação precisa ser realmente virtuosa: e essa virtude precisa ser derivada de algum motivo virtuoso e consequentemente o motivo virtuoso precisa ser diferente da consideração pela virtude da ação. Um motivo virtuoso é exigido para tornar a ação virtuosa. Uma ação precisa ser virtuosa antes que nós possamos ter uma consideração pela sua virtude. Algum motivo virtuoso, portanto, precisa ser antecedente àquela consideração (T 478).

O comprometimento de Hume com primeiros motivos naturais – isto que não podem ser senão *desejos empíricos* – está claro na seguinte passagem:

> Nenhuma ação pode ser exigida de nós como nosso dever, a menos que haja implantado na natureza humana alguma paixão ou motivo movente [*actuating*] capaz de produzir a ação. Esse motivo não pode ser o senso de dever. Um senso de dever supõe uma obrigação antecedente: e onde uma ação não é exigida por qualquer paixão natural, ela não pode ser exigida por qualquer obrigação natural… (T 518).

Aí está a máxima ciceroniana indubitável: "*Que nenhuma ação pode ser virtuosa, ou moralmente boa, a menos que exista na natureza humana algum motivo para produzi-la distinto do senso de sua moralidade*" (T 479).

Humeanos insistem que o propósito dessa máxima *não* é impedir que a moralidade se erga para além do seu ponto de partida, é antes registrar que, como fenômeno *natural*, a moralidade precisa de um ponto de partida *natural*[11]. O que, de qualquer modo, seria justamente indicado pelo uso da máxima no contexto da análise das virtudes *artificiais*. Tome-se como exemplo a 3ª Lei da Natureza quanto à justiça: a obrigatoriedade da promessa. Não obstante aparências em contrário – "[…] é evidente que nós não temos um motivo que nos conduz ao cumprimento das promessas distinto do sendo de dever" (T 518) –, a sofisticada posição de Hume sobre esse tema não pretende implicar a suspensão da máxima.

[11] Cf. WIGGINS, Op. Cit., p.312ss.

Esta "expansão da moralidade" talvez não esteja presente na constituição do senso do *dever moral natural* (por oposição ao dever ligado a uma virtude artificial) de cuidado com os filhos, mas fica *patente* na assim chamada (de modo algo desorientador) "obrigação natural" quanto à justiça, essa que se assenta no *interesse* (T 498), que por fim alcança uma forma elaborada na obrigatoriedade da promessa: "[…] Atrevo-me a concluir que promessas são invenções humanas, fundadas nas necessidades e interesses da sociedade" (T 519). Aqui estamos no contexto de um "comércio auto-interessado" (T 521) que reconhece a importância de um "concerto", de um "esquema de ações calculado para o benefício comum" – e homens "concordam em ser fiéis [*true*] às suas palavras" (T 522).

É nesse ponto que a *complexidade* da posição de Hume se revela, pois é aqui que podemos ver a mencionada *expansão* da concepção do senso de dever. Esse senso do dever é ligado de fato, por Hume, a um auto-interesse *esclarecido* (*enlightened* é o termo de A. Baier[12]). O auto-interesse é esclarecido porque se vê que sua manutenção está calculadamente vinculada ao "benefício comum". Ele está refletido mutuamente nos indivíduos. Passa-se de um mero "seu interesse" a um "seu interesse" *também reconhecido no outro como o mesmo*, e por isso merecedor de uma desenvolvida forma de consideração. O que Hume afirma é o seguinte:

> […] Quando cada indivíduo percebe o mesmo senso de interesse em todos os seus semelhantes [*fellows*] [de interesse no cumprimento de promessas], ele imediatamente realiza a sua parte em qualquer contrato por ter-lhe sido assegurado que eles não falharão nas suas [partes]. Todos eles, em concerto, entram num esquema de ações calculado para o benefício comum, e concordam em ser fiéis [*true*] às suas palavras; nada mais é exigido para se formar esse concerto ou convenção senão que cada um tenha o senso de interesse no fiel cumprimento de engajamentos, e expresse esse senso aos outros membros da sociedade. Isso imediatamente causa aquele interesse a operar sobre eles; e o interesse é

[12] Cf. "Nature and artifice, equity and justice", p.72. In: *The cautious jealous virtue, Hume on justice*, Harvard University Press, 2010.

> a *primeira* obrigação para a realização de promessas (T 522-3).

Para que reconheçamos que esse interesse (esclarecido) pode ser o motivo primeiro para o cumprimento das promessas, e ser a motivação *aprovada* para os atos de cumprimento delas, basta não confundirmos – como ensina A. Baier – essa *motivação aprovada* com as nossas *razões para a sua aprovação* (Cf. Idem). Essas razões olham para além do indivíduo, para – via o *esclarecimento* do auto-interesse – o *benefício comum*, no sentido de "de todos e de cada um", e não no sentido de ser o benefício de todos considerados em conjunto, talvez na expectativa de que ele seja discretamente distribuído a todos.

Parece poder ser desse modo que o senso de dever no caso do cumprimento da promessa não implica uma limitação da máxima ciceroniana, antes um reconhecimento de que o que está em questão não é o motivo *natural à la* virtudes naturais, mas uma "paixão ou motivo eficientes": o auto-interesse esclarecido – que *não* é esclarecido *naturalmente*. Desse modo, a expansão da compreensão do motivo do dever passa a incluir o senso de auto-interesse *esclarecido*, que para Hume *não* é natural num sentido, pois é *artificial* por estar vinculado a *convenções* humanas, mas que não deixa, mesmo assim, de ser tão *natural* quanto essas convenções, no sentido de possíveis e efetivas para a inventividade humana.

Mas será que é só esse o papel – o explicativo – da máxima ciceroniana de Hume? Voltemo-nos para Kant e ao propósito de sua objeção para chegarmos a uma resposta a essa pergunta.

A exigência kantiana de que a obrigação moral se assente em princípios *objetivos* tem seu *rationale* na sua teoria de *valor*: na sua visão da natureza e conteúdo do valor. Ou seja, as afirmações kantianas quanto à natureza não qualificada da obrigação moral – trata-se de uma "necessidade absoluta" (G 4: 389), sem referência a "um elemento empírico qualquer" (*idem*) – encontram sua justificação na sua concepção de "valor absoluto":

> Os princípios práticos são *formais*, quando abstraem de todos os fins subjetivos; mas são *materiais*, quando tomam por fundamento os fins

> subjetivos, por conseguinte certas molas propulsoras. Os fins que um ser racional se propõe a seu bel-prazer como *efeitos* de sua ação (fins materiais) são, sem exceção, relativos apenas: pois é tão-somente sua relação com uma faculdade apetitiva de índole particular do sujeito que lhes dá o valor, o qual <valor>, por isso <mesmo>, não pode fornecer princípios universais para todos os seres racionais e tampouco válidos e necessários para todo querer, isto é, leis práticas. Por isso, todos esses fins relativos são tão-somente o fundamento de imperativos hipotéticos. Supondo, porém, que haveria algo *cuja existência* tenha em si mesma um valor absoluto – o que, enquanto *fim em si mesmo*, poderia ser um fundamento de leis determinadas -, então encontrar-se-ia nele e tão-somente nele o fundamento de um possível imperativo categórico, isto é, <de uma> lei prática (G 4: 427-8).

A tese kantiana, é claro, é que "o homem – e de modo geral todo ser racional – *existe* como fim em si mesmo" (G 4: 428), ou seja, "pessoas, porque sua natureza já os assinala como fins em si mesmos", são "*fins objetivos*, isto é, coisas cuja existência é em si mesmo fim", e por isso têm "*valor absoluto*" (G 4: 428).

Extremo cuidado é recomendável aqui. O argumento de Kant não constitui um argumento por gênero e diferença específica, um *especieísmo* que dota as pessoas de valor por terem *humanidade* compreendida como a capacidade de se darem fins por razões em geral. Esse tipo de interpretação conduz a perguntas sobre o que ou quem é uma pessoa, como se Kant tivesse introduzido sua concepção de valor absoluto somente quando da introdução da noção de fim em si mesmo como base para a obrigatoriedade absoluta do Imperativo Categórico. Essa é a linha de interpretação da defesa do valor ou dignidade das *pessoas humanas*.

Um olhar mais cuidadoso ao texto revela algo diferente. É fato que Kant afirma que "a natureza racional existe como fim em si" e que isso é o *fundamento* (*Grund*) do "*princípio objetivo* da vontade" (*supra cit*), mas a referência à "natureza racional" aqui parece ser ao modo de *representacionalmente* todo ser racional se conceber: "todo ser racional [não só o homem] representa sua existência <desse modo>, em consequência de precisamente o mesmo fundamento racional, que

também vale para mim" (G 4: 429). Isso que aqui *vale* deve, portanto, ser "ao mesmo tempo um princípio *objetivo*, do qual, enquanto fundamento prático supremo, todas as leis da vontade têm de poder ser derivadas" (*idem*).

O ponto fundamental é o seguinte: na filosofia kantiana, a litania do valor e dignidade humanos é de proveito duvidoso, pois as noções *kantianas* de valor, de dignidade e de respeito entram em cena "como resultado, e não no lugar, do trabalho do formalismo"[13]. É crucial compreender que em Kant "dizer que uma coisa ou pessoa tem valor é considerar algo sobre o que ela é como uma fonte de orientação para como nós devemos agir em relação a ela" (Idem). Desse ponto de vista, a deliberação moral é antes pautada *axiologicamente* (do que teoricamente nas características distintivas da espécie humana) numa concepção do valor do *querer bem* como *raciocinar bem*, a partir de um princípio formal e de premissas que encontramos quando somos guiados pelos dois fins obrigatórios da filosofia moral kantiana[14]. Assim, devemos procurar sempre verificar quais são os modos concretos de interação entre nós que minimizam "nossas vulnerabilidades como raciocinadores [*reasoners*] práticos"[15].

O apoio teórico dessa concepção da deliberação pautada axiologicamente (ao invés de teoricamente) está justamente na concepção de valor e na sua *defesa*. *Um elemento* dessa defesa me interessa ressaltar aqui, pois a pretensão kantiana é que ele é exclusividade da posição de Kant: a *acessibilidade* relevante à deliberação moral (o *rationale* moral acessível também à avaliação moral). O curso da argumentação segue as seguintes etapas, o elemento em questão aparecerá na última.

1. Os princípios racionais da agência apresentados no procedimento do Imperativo Categórico constituem-se num *padrão* de justificação *completa* quanto à avaliação de nossas escolhas de fins e de ações para tais fins. Isso pode ser verificado na apresentação da razão moral para a interdição da falsa promessa, o

[13] HERMAN. "Embracing Kant's formalism". In: *Kantian Review* 16.1, p.57, 2011.
[14] Cf. HERMAN. "Reasoning to obligation". In: *Inquiry* 49.1, p.45, 2006.
[15] HERMAN, "Embracing Kant's formalism", p. 53.

ataque a um princípio constitutivo da agência racional: não está justificada a máxima que pretende se utilizar *causalmente* da separação da vontade de outro, que é uma faculdade de razões que vão até o fim, razões que podem justificar completamente suas avaliações do que é bom.

2. A acepção em que razões vão até o fim na vontade separada de agentes exige que essa justificação por razões com essa característica tenha seu padrão num princípio dessa vontade, num princípio que essa vontade autolegisla. Isso é a autotomia da vontade.

3. A característica da vontade racional, a de ser capaz de autonomia – da natureza racional ser um fim em sim mesmo – fornece um padrão *objetivo* para a avaliação do que é bom, esse padrão – ele próprio – constitui-se na concepção de um critério do que tem valor *absoluto*.

4. Consequentemente, qualquer avaliação de que algo é bom que seja dependente de *desejos* do agente, sejam ocorrentes, sejam desejos sem os quais o agente não se reconheceria como tal, portanto, até mesmo desejos endossados reflexivamente, é uma avaliação *relativa* ao agente, e por isso sem valor absoluto. Somente uma avaliação que pode se socorrer de princípios válidos universalmente pode ser uma avaliação objetiva que tem seu *rationale* num valor absoluto.

5. Uma avaliação que desse modo é remetida a uma fonte de valor absoluto na própria vontade racional é uma avaliação que pode, exatamente por isso, ser motivacional para o agente. É nesse sentido que o motivo moral pode ser *transparente* ao agente quanto às suas origens[16]. E a consequente estabilidade motivacional do motivo do dever, não só a expressão da boa vontade em situação de adversidade, não será propriamente a da firmeza ante à tentação, mas a de uma confiança quanto ao valor da moralidade conferindo uma abertura para a continuidade do desenvolvimento moral em resposta às mais variadas demandas[17]. Desse modo, a *acessibilidade* ao valor é garantia da possibilidade da

[16] Cf. HERMAN, "Making room for character", In: *Moral literacy*, Harvard University Press, 2007 p.4-5.
[17] Op. Cit., p. 5.

deliberação moral válida[18].

A acessibilidade em questão é bloqueada, segundo a posição kantiana, numa filosofia como a de Harry Frankfurt. A avaliação de Barbara Herman dessa filosofia pode servir como ilustração aqui:

> O que Frankfurt mostrou é que a possibilidade de razões incondicionadas não requer, como se poderia ter pensado, um valor externo, ou qualquer coisa fora da órbita do auto-erguer-se [*bootstraping*]. Elas podem surgir da natureza da nossa sensitividade [*sensitivity*] e capacidade para amar – o modo que valorizamos alguns fins últimos. Essas razões incondicionais não são mediadas pelo apego a qualquer outro fim; sua força motivacional está assegurada diretamente com o reconhecimento da aplicação de uma razão. E como o fim em que se apóia a razão é um que eu não posso abandonar, ele funciona, do ponto de vista da volição, como um valor incondicionado. Mas ainda há uma condição. Existe o valor-para-nós de fins incondicionados, paradoxicalmente condicionados pela nossa necessidade deles – a saber, se quisermos vir a ser pessoas. Quando não abandono meus fins últimos, não é por causa do *seu* valor; é porque eles vieram a desempenhar um papel co-constitutivo na minha identidade como pessoa. O trabalho é feito pela inércia. É uma inércia de tipo especial, certamente: pois o que eu não posso abandonar é a mim mesmo[19].

É essa *inércia* que temos em Hume? Não, de acordo com Annette Baier.

Segundo essa humeana, um propósito importante da máxima ciceroniana em Hume é evitar o *desvirtuamento* da motivação moral nos seguintes termos. Na medida em que o senso da moralidade é o *sentir* de uma satisfação/insatisfação na contemplação do caráter, ou seja, o sentimento moral é uma espécie de *prazer*/*desprazer*, pode-se facilmente incorrer no erro de conceber a motivação

---

[18] "…The derivation of action from the constitutive principle of volition provides a noncontingent connection between value and reasons. Failures to be moved by a reason is a sign of a deliberative error, not a lack of subjective connection to the reason's value" (Herman, "Bootstraping", p.172. In: *Moral literacy*, Harvard University Press, 2007).

[19] Herman, "Bootstraping", p.163.

moral como o desejar desses estados de satisfação com as características de caráter das pessoas. Mas, o que é fundamental para Hume, nisso seguindo Shaftesbury, é que a satisfação moral, essa forma de prazer, não seja ela própria o motivo para as ações que são moralmente aprovadas[20]. O desejo de assegurar a aprovação não pode ser o motivo à ação que é aprovado moralmente pelo espectador judicioso, julgando da perspectiva comum, de Hume. Esse erro é cometido por Adam Smith, segundo Baier[21]. O homem em Smith deseja naturalmente não só ser amado, mas ser amável, não só não ser censurado, mas não ser censurável. O que ele deseja, em Smith, é fundamentalmente essa forma derivada de satisfação que é facultada pela consciência, este *homem dentro de nós* [*the man within*], que é uma espécie de mão invisível de um mercado moral que garante que virtudes, excelências humanas, sejam produzidas conforme nossas necessidades.

Nesse sentido, o sentimento moral, ao assentar-se num *prazer* com características pessoais *reais*, não permite que a satisfação da aprovação se transforme de modo autóctone no *primeiro* motivo aprovado moralmente.

Mas esse elemento *hedônico*, resguardado na máxima ciceroniana, e inextinguível no sentimento moral, traz consigo como consequência que as virtudes sejam de fato características humanas especiais presentes numa larga medida em nós. Se não, por certo, virtudes não seriam o que nos *agrada* de um ponto de vista *comum*. Esse elemento, com o fato apontado, também pauta a *reflexividade* do juízo moral ou do senso de moralidade, o que se constitui num *endosso crítico* (não devido à mera reflexão mecânica) da sua avaliação de traços de caráter como virtudes, ou seja, como *realmente* virtudes[22].

A exigência humeana, esta que lhe faculta pretensões de normatividade, é, ao final do *Tratado*, que "não só a virtude deve ser aprovada, mas também o sendo

---

[20] Cf. A. BAIER, "Moral sentiments, and the difference they make", *Aristotelian Society*, Supplementary Volume (1995), p.17.

[21] Cf. Op. Cit. p.18-21.

[22] Cf. Geoffrey Sayre-McCord, "On why Hume's 'general goint of view' isn't ideal – and shouldn't be", passim (*Social Philosophy & Policy* 11.1, 1994).

de virtude"[23]. De fato, a visão de Hume é que as "distinções morais" – também estas com as quais Hume se compromete, como as das assim chamadas "virtudes monásticas" – são derivadas de uma origem *nobre* relacionada à *generosidade* e *capacidade* da natureza humana. O senso da moral ganha força, segundo Hume, em relação às suas avaliações quando "ao refletir sobre si mesmo, aprova aqueles princípios dos quais é derivado, e não encontra nada no seu surgimento e origem senão o que é grandioso [*great*] e bom". Fundar o senso da moral numa "simpatia ampla [*extensive*] com a humanidade" confere vantagens à operação desse senso: um "sólido argumento" como o de Hume nos mostra que temos razões para aprovar o senso de virtude e "os princípios dos quais é derivado", de modo que "nada é apresentado em qualquer lado, senão o que é louvável e bom".

Nesse sentido, a "explicação" humeana nos oferece o *rationale* acessível para que as nossas "distinções morais" sejam feitas com segurança e estejam embasadas. A versão de Hume do sentimento moral e suas distinções avaliativas confere à lista humeana das virtudes um *rationale* que serve de parâmetro para que entendamos nossas avaliações e façamos novas com a convicção de que estão justificadas. Esta é, então, a *transparência*, quanto à sua origem, do senso moral, segundo Baier:

> Nós precisamos, tanto em Descartes quanto em Locke, ser conscientes de nossas naturezas cognitiva e passional antes que possamos discernir quais normas deveriam nos guiar, mas o giro [*turn*] reflexivo, se ele é puramente cognitivo, não é suficiente para nos oferecer a compreensão moral [*to yield moral insight*]. Para que a reflexividade nos propicie resultados normativos, o que é voltado para dentro tem que ser por sua natureza avaliativo, tendo o bem – não meramente o verdadeiro – como seu objeto[24].

Portanto, a reivindicação de acessibilidade ao *rationale* da moralidade não é exclusividade kantiana[25]. Mas o *valor* nas posições de Hume e Kant, os termos nos

---

[23] T 619. Também para as citações na continuação.

[24] Op. Cit., p.29.

[25] "Esse entendimento de si, Hume acredita, torna nosso caráter moral mais sólido e nos revela a

quais o *rationale* é articulado, parece ser distinto. Em Hume, o *prazer* com o que é humano e o melhor em nós; em Kant, a *natureza racional* como fonte de interações que nos tornam merecedores de nossa dignidade.

---

felicidade e dignidade da virtude. Ele afirma que isso é uma vantagem da sua visão sobre aquelas de Clarke e Hutcheson, pois em suas doutrinas as adequações [*fitnenesses*] das coisas, ou o senso moral, permanecem opacos e não tem qualquer conexão inteligível com o que é de importância central na vida humana" (*John Rawls: Lectures on the history of moral philosophy*, Ed. Barbara Herman, Harvard University Press, 2000).

16

A FILOSOFIA MORAL DE PAUL GRICE

Frank Thomas Sautter

Introdução

Duas obras póstumas de Paul Grice – *The Conception of Value* (Grice 1991) e *Aspects of Reason* (Grice 2001) – contêm uma parcela substantiva da sua filosofia moral. Três temas se destacam:

1- Comparação entre razões teóricas (Grice as denomina "razões aléticas") e razões práticas, um tema dominante nos primórdios da filosofia moral de tradição analítica, e que recebe de Grice um toque kantiano. Esse tema ocupa parte da obra *Aspects of Reason*.

2- Defesa severa de uma abordagem teleológica aos problemas da filosofia moral e a centralidade da noção de *eudaimonia* para a mesma. Esse tema complementa a obra *Aspects of Reason*.

3- Esboço de um programa original construtivista, *pace* Carnap, de uma metafísica do valor. Esse tema ocupa inteiramente a obra *The Conception of Value*.

Tratarei, aqui, do primeiro tema – a análise proposicional *pari passu* de sentenças teóricas e sentenças práticas, na 6ª seção, mas antes disso farei uma crítica a certas formas que esse tipo de análise pode assumir, na 4ª seção, e uma breve exposição de análises similares por parte de precursores, na 5ª seção. Também examinarei brevemente os dois temas anteriormente mencionados: na 2ª seção a abordagem teleológica orientada para a consumação da felicidade, e na 3ª seção a original metafísica do valor proposta por Grice.

1. Causas finais e felicidade

Grice atribui às causas finais um papel indispensável nos fundamentos da ética[1], e, por conseguinte, a felicidade, por ser entendida como a posse e a operação de um sistema de fins para a vida[2], é uma noção central de sua filosofia moral. A discussão sobre a felicidade, na forma de aceitabilidades prudenciais, também ocorre como culminação de seu projeto de análise *pari passu* dos discursos teórico e prático, como será visto na 6ª seção.

Apesar dessa importância, a discussão sobre a felicidade é, em grande medida, embrionária. Seguindo a discussão em torno da noção aristotélica de *eudaimonia*, cuja terminologia para as suas diferentes interpretações foi fixada por Hardie (1965), Grice discute em certo pormenor a interpretação da felicidade como um fim dominante – um único fim sobressaindo-se como o bem – e a interpretação da felicidade como um fim inclusivo – um fim combinando ou incluindo dois ou mais bens[3], e decide-se por uma interpretação inclusiva[4].

Grice, entretanto, não propõe fins particulares; antes, ele fornece características que todo e qualquer sistema *pessoal* de fins deveria satisfazer para ser estável, ou seja, para poder servir eficazmente como guia para a vida. Ele enumera as seguintes características[5]:

1. Exequibilidade, ou seja, aquele que se guia pelo sistema de fins deve estar equipado, por natureza ou por treinamento, com as competências necessárias à realização dos fins do sistema.

2. Autonomia, ou seja, a realização dos fins do sistema deve depender o mínimo possível de intervenções externas.

3. Compatibilidade mútua dos fins do sistema; e não apenas isso, harmonia

[1] 2001, p.43.
[2] 2001, p.131.
[3] GRICE 2001, p.119ss.
[4] GRICE 2001, p.128.
[5] 2001, p.132-133.

desses fins.

4. Abrangência, ou seja, a capacidade dos fins para guiar a deliberação sobre toda e qualquer questão prática.
5. Suporte mútuo dos fins do sistema.
6. Simplicidade do sistema de fins.
7. Aprazibilidade dos fins do sistema, para que se possa lidar adequadamente com o problema da incontinência (*akrasia*).
8. Possibilidade de desenvolvimento maximal dos talentos naturais.

A ideia mais interessante de Grice a respeito da noção de felicidade é a aplicação a ela da noção econômica de utilidade marginal. Define-se a utilidade marginal de um bem ou serviço como a utilidade ganha (ou perdida) por um aumento (ou diminuição) no consumo do bem ou serviço. À noção de utilidade marginal está associada a seguinte "Lei da Diminuição de Utilidade Marginal": o primeiro consumo do bem ou serviço tem mais utilidade que os consumos subsequentes e assim por diante. Grice aplica uma lei análoga à interpretação inclusiva da felicidade: possivelmente há um núcleo duro de fins que todo e qualquer sistema pessoal de fins deveria comportar, mas novos fins – mudança de qualidade – ou um grau maior de realização de fins – mudança de quantidade – tendem a ser evanescentes[6].

2. A transubstanciação do homem em pessoa

Na obra *The Conception of Value*, Grice desenvolve uma concepção de valor, entendido em sentido amplo, e não apenas valor moral[7], segundo a qual valores são realidades construídas pelos humanos e, portanto, não têm realidade própria[8]. Essa defesa – uma investigação metafísica – diz muito acerca da

[6] 2001, p.117.
[7] GRICE 1991, p.25.
[8] 1991, p.37.

compreensão de Grice sobre a teorização filosófica, porque a teorização metafísica é, segundo ele, a teorização primária em Filosofia[9]. Grice segue um programa construtivista que obedece a três preceitos[10]:

> 1- Reconhecimento de uma coleção de pontos de partida metafísicos, ou seja, aquilo que é metafisicamente primário.
>
> 2- Utilização de uma coleção de rotinas ou procedimentos de construção, mediante os quais itens não primários são construídos com base em itens *mais* primários.
>
> 3- Fornecimento de uma motivação teórica para proceder de um dado estágio ao estágio seguinte.

Vejamos, na prática, como essa proposta funciona para a noção de valor.

Um ponto de partida metafísico é a noção de propriedade essencial. Grice adota uma caracterização *standard* de essência, ou coleção das propriedades essenciais, de uma sorte ou tipo de coisa, a saber, a essência é a coleção das propriedades definitórias da sorte ou tipo de coisa; propriedades essenciais, por oposição a propriedades acidentais, estão relacionadas às condições de identidade do sorte ou tipo de coisa; e uma coisa não pode perder uma propriedade essencial, sem deixar de existir[11].

É central para a proposta de Grice a distinção entre propriedade essencial e propriedade necessária: pode haver, e há, acidentes necessários, ou seja, propriedades lógica ou metafisicamente necessárias, dada a presença de propriedades essenciais, mas que não são, elas mesmas, propriedades essenciais. Grice também admite propriedades essenciais contingentes; indivíduos com degenerações em relação aos de sua espécie, por exemplo, vinculam-se a ela por ancestralidade[12]. Tais propriedades, contudo, não são importantes no presente

---

[9] 1991, p.71.
[10] 1991, p.70.
[11] GRICE, 1991, p.79.
[12] 1991, p.80.

contexto.

Também é central para a proposta de Grice que a essência de criaturas vivas inclua finalidades[13].

O sucesso da proposta de Grice depende, ainda, da possibilidade de diferentes particionamentos de uma mesma coleção de propriedades em propriedades essenciais e propriedades acidentais[14]. Sem essa possibilidade, a operação central da proposta dele, a saber, a transubstanciação do tipo de substância homem no tipo de substância pessoa também não é possível[15].

A motivação teórica fornecida por Grice tem um caráter evolucionista: a transformação do tipo de substância homem no tipo de substância pessoa consiste na "essencialização" da racionalidade, ou seja, suas aceitabilidades, em particular, e suas atitudes, em geral, passam a ser assentadas em razões, em sentido amplo, vinculadas a valores de tipo apropriado[16]; a "essencialização" da racionalidade é ditada pela necessidade de sobrevivência, ou seja, a necessidade de adaptação a um ambiente complexo e sujeito a variações, onde os instintos já não são suficientes[17]. O resultado é uma pessoa, um tipo não biológico[18].

O processo todo envolve uma "Projeção Humeana", ou seja, uma operação mediante a qual um modo de pensamento de uma criatura viva deixa de ser um atributo do pensamento e passa a ser um atributo da própria coisa pensada[19]. Assim, para utilizar o próprio exemplo de Grice, o modo de pensar "pensar-como-valoroso" a respeito de algo passa a ser, ao fim e ao cabo da operação, um atributo da própria coisa: ela é valorosa[20].

---

[13] 1991, p.81.
[14] 1991, p.81.
[15] 1991, p.82.
[16] 1991, p.82.
[17] GRICE, 1991, p.83.
[18] GRICE, 1991, p.84.
[19] GRICE, 1991, p.88.
[20] GRICE, 1991, p.88.

3. Contra o sintaticismo

> É preciso mostrar que a lógica é aplicável ao âmbito normativo antes de começar a explicar como ela é aplicável ao âmbito normativo. O método moderno consiste em supor, sem discussão, que ela é aplicável ao âmbito normativo e, então, distrair sua atenção disso (a única questão real) ao diligentemente explicar como essa aplicação se dá[21].

Entendo por "sintaticismo" o *modus operandi* segundo o qual se assume que a semelhança sintática ou estrutural de operações sob exame com as operações lógicas é condição suficiente para garantir a logicidade dessas operações *sub judice*. Sustento que esse erro é frequente, inclusive na filosofia moral. Em rigor, a lógica deôntica encontra-se viciada na origem pelo erro. O trabalho inaugural "Deontic Logic" de G. H. von Wright (1951), é marcado por sintaticismo. Nele, von Wright afirma o seguinte:

> À execução ou não execução de uma determinada ação (por um agente) denominaremos valores de execução (para aquele agente). Uma ação será denominada função de execução de determinadas outras ações, se o seu valor de execução para qualquer dado agente depende exclusivamente dos valores de execução daquelas outras ações para o mesmo agente. […] O conceito de função de execução é estritamente análogo ao conceito de uma função de verdade na lógica proposicional. […] Funções de execução específicas podem ser definidas em estrita correspondência com funções de verdade específicas[22].

A proposta de Von Wright soa inocente: ele *assume* uma semelhança entre

---

[21] Adaptação de uma passagem de Clive Staples Lewis (1970, p.273), na qual ele descreve o vício do bulverismo, ou seja, a atitude conforme a qual se assume que o oponente está errado, e, então, explica-se o erro dele: "…você precisa mostrar que um homem está errado antes de começar a explicar por que ele está errado. O método moderno consiste em supor, sem discussão, que ele está errado e, então, distrair sua atenção disso (a única questão real) ao diligentemente explicar como ele se tornou tão tolo".

[22] VON WRIGHT, 1951, p.2.

valor de execução – o tema da analogia, ou seja, seu componente mais estranho – e valor de verdade – o foro da analogia, ou seja, seu componente mais familiar, e, então, define "negação de uma ação", "conjunção de ações" etc. O resultado é trágico! Vejamos um exemplo típico:

> A ação implicação, e.g., de duas ações dadas é uma ação que é executada por um agente se, e somente se, não é o caso que a primeira ação é executada e a segunda ação não é executada pelo agente em questão[23].

A proposta é infeliz, porque assume a dependência das ações umas em relação às outras. Contudo, não é fato consumado que a execução de uma ação dependa da execução de outras ações; é suficiente considerar o caso das ações *voluntárias*. O máximo que se pode afirmar, sem presumir demasiado, é que uma ação *deve se seguir* de outra ou outras, ou que uma ação não *deve se seguir* de outra ou outras. Afirmar que uma ação *se segue* de outra ou outras, ou que uma ação não *se segue* de outra ou outras, depende de argumento adicional.

Numa autobiografia de seu envolvimento com a Lógica Deôntica, von Wright (1999) reconhece o papel da analogia para a emergência da Lógica Deôntica como disciplina acadêmica: primeiro, ele descreve a descoberta da analogia entre os modos aléticos (modos da verdade) e os quantificadores (modos da existência); depois, a analogia desses com os modos deônticos (modos da obrigação)[24]. Mais adiante, no mesmo trabalho, ele faz uma autocrítica e uma crítica à pesquisa corrente em Lógica Deôntica, centrada numa prática qualificável de sintaticismo:

> No meu primeiro trabalho procedi como se do mero fato de que se pode construir um cálculo formal com axiomas que soam plausíveis fosse o necessário para satisfazer as demandas da lógica. E penso ser correto afirmar que essa atitude ainda subjaz implicitamente, em boa medida, o trabalho realizado em lógica deôntica. Sua natureza problemática, porém, permanece até hoje um espinho em minha carne lógica, se se pode usar

[23] VON WRIGHT, 1951, p.2
[24] VON WRIGHT, 1999, p.28.

essa metáfora[25].

A ampla maioria dos filósofos e lógicos, no mínimo, é suspeita do mesmo tipo de prática.

4. Os primórdios da filosofia analítica da moral

A história da análise filosófica de sentenças é longa e intrincada, por isso limitar-me-ei a contá-la nos seus episódios relevantes para a presente discussão.

O *locus classicus* dela, na tradição analítica, é a obra *A linguagem da moral* do filósofo inglês Richard Mervyn Hare (1952). Nela, Hare distingue dois componentes: a frástica e a nêustica. A diferença entre eles pode ser melhor apreciada ao examinarmos um exemplo utilizado pelo próprio Hare. Considere as seguintes sentenças:

- Você vai fechar a porta.
- Feche a porta.

A primeira é uma sentença indicativa, a segunda é uma sentença imperativa. Ambas tratam do mesmo assunto, e isso é a frástica, mas tratam dele de modo distinto, e isso é a nêustica.

À luz dessa análise, as sentenças podem ser reescritas do seguinte modo:

- Você fechar a porta no tempo t, sim: "Você fechar a porta no tempo t" é a frástica comum às duas sentenças e "sim" é a nêustica comum a todas as sentenças indicativas.
- Você fechar a porta no tempo t, por favor: "Você fechar a porta no tempo t" é a frástica comum às duas sentenças e "por favor" é a nêustica comum a todas as sentenças imperativas.

A análise de Hare em *A linguagem da moral* é a mais conhecida, mas não a primeira e nem mesmo a melhor.

---

[25] VON WRIGHT, 1999, p. 31.

Em um trabalho de 1949 Hare faz a distinção, denominando "descritor" à frástica e "dictor" à nêustica. Antes dele, o filósofo danês Jörgen Jörgensen, em 1938 faz a distinção, criticada por Hare, entre um fator indicativo e um fator imperativo. A crítica de Hare a Jörgensen, cuja solução é incorporada à sua própria análise, diz respeito ao modo como Jörgensen entende os fatores indicativo e imperativo: sentenças indicativas têm *somente* um fator indicativo, enquanto que sentenças imperativas têm um fator indicativo *e* um fator imperativo. Isso sugere que as sentenças imperativas são derivadas e, de algum modo, inferiores às sentenças indicativas. Na análise de Hare sentenças indicativas e sentenças imperativas têm, *ambas*, uma frástica *e* uma nêustica.

Hare, em um trabalho de 1989, insatisfeito com a distinção proposta em *A linguagem da moral*, amplia essa distinção, subdividindo a "antiga" nêustica em um sinal de modo ou trópico, um sinal de subscrição ou ("nova") nêustica, e um sinal de completude ou clístico. Esse último, o clístico, praticamente não é empregado na literatura; os dois outros, o trópico e a nêustica, têm correspondentes na análise de Grice.

A "antiga" nêustica[26] acumulava duas funções distintas, a saber, fornecer um modo para a frástica e indicar um compromisso para com ela. O trópico[27] corresponderá à primeira dessas funções, enquanto que a "nova" nêustica corresponderá à última dessas funções. Ambos, trópico e nêustica, são relevantes de um ponto de vista lógico. A validade de argumentos é regida, em parte, pelos trópicos de seus componentes. Uma conclusão imperativa somente pode ser validamente obtida se ao menos uma premissa também for imperativa[28]; e em um argumento válido cuja conclusão é indicativa as premissas imperativas são irrelevantes. Para explicar a "nova" nêustica, Hare apela a duas analogias: o uso subscritivo da linguagem é análogo à intervenção do autor da peça protegido pela parábase do teatro, caso em que está comprometido com o que diz e faz, e o uso

[26] "Nêustica", em grego, significa consentir inclinando a cabeça.
[27] "Trópico", em grego, significa o modo gramatical.
[28] Trata-se de aplicação da Guilhotina de Hume.

não subscritivo é análogo ao preenchimento de um cheque sem, contudo, assiná-lo.

5. Análise proposicional de Paul Grice

Grice inicia sua análise comparando as seguintes sentenças:

- João deve estar recuperando sua saúde.
- João deve se filiar aos Alcoólicos Anônimos.

Ele as decompõe em um operador de racionalidade, um operador de modo e um radical.

O operador de racionalidade, *Acc*, lê-se "é aceitável que" e, ao que tudo indica, corresponde à "nova" nêustica de Hare, ou seja, esse operador indica o comprometimento do falante com o que é dito.

O próprio Grice esclarece que o radical *r* tem um uso idêntico ao da frástica de Hare, ou seja, ele indica um conteúdo, um tema.

O operador de modo da primeira sentença (João deve estar recuperando sua saúde) é ├ , enquanto que o operador de modo da segunda sentença (João deve se filiar aos Alcoólicos Anônimos) é ! e, ao que tudo indica, corresponde ao trópico de Hare. A escolha do símbolo ├ não é feliz, porque esse símbolo, introduzido por Frege, é um duplo símbolo: ele se compõe do traço horizontal, que indica conteúdo, e do traço vertical, que indica assentimento. Na literatura mais recente, o correspondente a ├ é !-, e não !.

O emprego dos operadores e do radical resulta na seguinte análise das sentenças anteriormente mencionadas:

- *Acc* + ├ + *r*, para a primeira sentença.
- *Acc* + ! + *r*, para a segunda sentença.

Grice acrescenta que cada símbolo cai sob o escopo de seu predecessor: o radical no escopo do operador de modo, e este no escopo do operador de racionalidade. Essas considerações também estão presentes nos ensaios tardios de Hare, nos quais há um operador adicional – o clístico, sem que Hare consiga se

decidir por uma ordem unívoca dos elementos da análise. Aparentemente, alguns exemplos sugerem uma ordem, e outros exemplos sugerem outra ordem.

Grice chama a atenção para a seguinte questão, que ele considera não decidida: há alguma sentença em alguma linguagem natural tal que uma parte dela é uma contraparte superficial de um radical? Frege, antes dele, já havia sugerido uma resposta negativa à questão, mas Grice menciona apenas Wittgenstein ao tratar do assunto[29]. Para além da questão fatual, uma questão mais interessante, pelo menos da perspectiva filosófica, é a determinação do impacto que uma resposta afirmativa ou negativa tem sobre a legitimidade da análise empreendida, se é que tem algum impacto. Para Frege, a resposta negativa implica, simplesmente, o uso necessário de uma linguagem artificial para os propósitos da análise filosófica; essa situação, Frege a descreveu esplendidamente ao compará-la a um barco que utiliza o vento para navegar contra o próprio vento.

Grice admite duas espécies básicas de aceitação[30]: a aceitação judicativa, abreviadamente J-aceitação, e a aceitação volitiva, abreviadamente V-aceitação. A V-aceitação, por sua vez, subdivide-se em um modo (inferior) intencional e um modo (inferior) imperativo[31]. O modo (superior) de J-aceitação, por analogia, subdivide-se também em dois modos inferiores. A Tabela 1 abaixo fornece o esquema para cada um desses quatro modos (inferiores):

[29] Aparentemente, a referência é §23 de *Philosophical investigations*, porque Grice fala das observações de Wittgenstein sobre os radicais na química.

[30] Grice sugere que essas duas espécies de aceitação devam ser entendidas como modos superiores. A sugestão é razoável, porque o modo está no escopo do operador de racionalidade (aceitabilidade).

[31] A sentença "Não falharei" é um exemplo do próprio Grice para o primeiro modo, enquanto que "Seja mais gentil com ele" é um exemplo dele para o segundo modo.

Suplemento	Diferencial	Operador *o*	Modo
Ø	julgar	$\vdash_A$	Judicativo "intencional"
quer que *H*	julgar	$\vdash_B$	Judicativo "imperativo"
Ø	quer	$!_A$	Volitivo intencional
quer que *H*	quer	$!_B$	Volitivo imperativo

Tabela 1: O proferidor *U* profere para o ouvinte *H* a sentença da forma *o* + *r* se *U* quer que *H* julgue que *U* <Suplemento> <Diferencial> *r*.[32]

A tabela original de Grice contém, além desses quatro modos (inferiores), quatro outros modos (inferiores) relacionados às sentenças interrogativas[33], e faz sugestões das modificações necessárias para obter modos (inferiores) relativos às interrogações nominais[34].

Essa análise conjunta, por si só, não resolve o dilema entre a alternativa kantiana de uma única razão e a alternativa de uma razão especificamente prática, relacionada analogicamente a razões não práticas. Grice ataca a questão mediante a discussão do problema da aparente relatividade das modalidades práticas e o aparente caráter absoluto das modalidades aléticas[35]. A solução preferida de Grice consiste em negar que as modalidades aléticas sejam

[32] A tabela foi simplificada em relação à tabela apresentada pelo próprio Grice para esses quatro modos (inferiores), sem prejuízo ao seu conteúdo.

[33] Sentenças interrogativas são interrogações cuja resposta é um sim ou um não; por exemplo, "Buenos Aires é a capital do Brasil?".

[34] Interrogações nominais são aquelas cuja resposta pode ser distinta de um sim ou um não; por exemplo, "Qual é a capital do Brasil?". A distinção entre sentenças interrogativas e interrogações nominais já era conhecida dos estoicos, e também foi utilizada por Frege, nas *Logische Untersuchungen*, para uma análise proposicional!

[35] GRICE, 2001, p.57.

absolutas: uma teoria ou sistema precede suas operações[36]; nesse sentido, *toda* operação, inclusive as modalidades, é relativa à teoria ou sistema do qual participa. As extensões de teorias ou sistemas são as produtoras das especializações de operadores, e não o inverso[37]. Grice menciona especificamente a seguinte sequência exótica de extensões de teorias: a teoria lógica é pressuposta pela teoria metafísica, essa é pressuposta pela teoria física, e essa é pressuposta pela teoria ictiológica[38]. Grice não o diz em *Aspects of reason*, mas, tendo em conta o que ele diz e faz em *The conception of value*, essas extensões de teorias (camadas) não são arbitrárias[39], mas respondem por uma demanda; esse procedimento é concorde com a sua concepção de um caráter construtivista da teorização filosófica.

A concepção de Grice culmina com uma caracterização bastante geral de raciocínio:

> Raciocínio $=_{def.}$ faculdade da ampliação de nossas admissões pela aplicação de formas de transição, de um conjunto de admissões a uma admissão ulterior, de tal modo a assegurar a transmissão de valor das premissas à conclusão, desde que tal valor recaia sobre as premissas[40].

"Verdade" e "bondade" são, ambos, assumidos como valores que constituem casos especiais de "satisfatoriedade"[41].

Para concluir essa exposição, apenas introdutória, da proposta de Grice, indicarei os pontos principais de seu tratamento do difícil problema das condicionais mistas, ou seja, aquelas em que há um modo judicativo e também um modo volitivo na mesma sentença, e o problema relacionado de inferência de aceitabilidades volitivas a partir de aceitabilidades judicativas. Soluções a ambos

---

[36] GRICE, 2001, p.61.
[37] GRICE, 2001, p. 61.
[38] "Ictiológico" é relativo a peixes! Grice era irreverente (ver CHAPMAN, 2008. p.6).
[39] GRICE, 1991, p.71.
[40] GRICE, 2001, p.87.
[41] GRICE, 2001, p.88.

os problemas serão exemplificadas pela apresentação e explicação de um esquema geral de justificação dos imperativos técnicos kantianos.

Antes, contudo, é preciso esclarecer a importância das condicionais mistas.

O modelo mais comum de inferência legal é o silogismo jurídico. Ele é basicamente um procedimento de subsunção sob leis. Contudo, à diferença das ciências empíricas, essas leis são normas. A leitura usual de um silogismo jurídico é aquela segundo a qual ele tem uma norma geral e hipotética como premissa maior, um juízo singular e categórico – o caso jurídico – como premissa menor, e uma norma singular e categórica – a solução jurídica – como conclusão. A inferência propriamente dita consiste em subsumir o caso jurídico sob a norma geral. Deixemos de lado a questão sobre as possíveis diferenças entre a deliberação jurídica e a deliberação moral, assumindo que a disputa sobre a autonomia do Direito em relação à Moral, e da Moral em relação ao Direito, dá-se no terreno das normas e não das inferências a partir de normas. Nesse caso, esse modelo de deliberação jurídica pode ser adaptado para uma explicação da deliberação moral. Contudo, esse modelo geral envolve dificuldades em relação a cada um de seus componentes:

1- Se a conclusão é uma norma, é necessária uma norma entre as premissas? E se ela é uma norma, isso significa que juízes, e agentes morais, criam normas?

2- A premissa menor – o caso jurídico ou moral – é, de fato, um juízo, ou ele seria já normativamente matizado?

3- A premissa maior – a norma do Direito ou da Moral –, por ser aquela sob a qual um juízo é subsumido para produzir uma norma, necessita ser um híbrido de juízo e norma? Um juízo geral e categórico responde pelo antecedente, e uma norma geral responde pelo consequente?

Essa última questão – o problema das condicionais mistas – é, talvez, o mais difícil que o modelo do silogismo jurídico, em particular, e do silogismo prático, em geral, precisa resolver, porque ele toca o próprio núcleo de interesse do

modelo, qual seja, a possibilidade de apresentação da deliberação prática como uma análoga da explicação científica[42]. Grice, porém, não encontra problemas para analisar sentenças que, a exemplo de condicionais mistas, têm um componente alético e um componente prático. A solução, indisponível para Frege e para o primeiro Hare, mas disponível para o último Hare, consiste em utilizar a distinção entre o operador de racionalidade, ou aceitabilidade, e o operador de modo[43]. Uma sentença tem um único operador de racionalidade, o operador de racionalidade, mas pode ter múltiplos operadores de modo. Grice fornece o seguinte exemplo: "para preservar uma compleição jovem, se se tem uma pele relativamente insensível, deve-se esfregar a face com manteiga de amendoim antes de ir dormir"[44].

A análise dessa sentença, em que E é "preservar uma compleição jovem", F é "ter uma pele relativamente insensível" e G "esfregar a face com manteiga de amendoim antes de ir dormir", é a seguinte: Dever (!E, ├ F; !G), na qual ";" separa antecedente e consequente. Ao modo de Kant, Grice deriva a "verdade" dessa sentença da verdade da seguinte sentença: "Deveria ser que, dado que se preserva uma compleição jovem e que se tem uma pele relativamente insensível, esfrega-se a face com manteiga de amendoim antes de ir dormir"[45], ou, em símbolos, Dever (├ F, ├ G; ├ E).

Frege, por não fazer a distinção no ato de asserção entre a aceitabilidade, neutra com respeito ao modo, e o modo, não pode admitir múltiplos atos, por exemplo múltiplos atos de asserção em uma única sentença, pois isso exigiria uma multiplicação inaceitável dos modos de inferência. Por contornar essa dificuldade,

---

[42] Por outro lado, levar a analogia muito a sério pode ser desastroso. Kelsen, na "Teoria Geral das Normas", acusa Kant de não prestar a devida atenção às diferenças entre a causalidade e a imputabilidade, embora, em teoria, Kant reconhecesse que elas são diferentes. Por exemplo, uma sequência causal não tem início e não tem fim, mas uma sequência de imputação tem tanto início como fim.

[43] Nêustica e trópico, respectivamente, na terminologia do último Hare.

[44] GRICE, 2001, p.91.

[45] GRICE, 2001, p.91.

ao admitir múltiplos e, inclusive, heterogêneos componentes vinculados a uma única aceitabilidade, a solução de Grice e do último Hare é interessante. Contudo, isso não parece ser suficiente para legitimar o tipo de análise realizada.

Essa análise comparativa de uma sentença mista com uma sentença puramente alética nos levou do plano definicional para o plano inferencial. Concluirei a seção com a análise griceana da derivação kantiana de imperativos técnicos[46].

A derivação kantiana principia com uma sentença analítica: Aquele que quer um fim, digamos A, (à medida que a razão decide sua conduta) também quer os meios indispensáveis a ele, digamos B. Grice analisa essa sentença da seguinte forma: $!A \ \& \ \vdash A \supset (B \supset A); !B$. Segundo Grice, essa sentença responde a uma "lei psicológica fundamental à medida que dela se derivam os comportamentos do querer e do julgar"[47].

Dessa deriva-se a sentença analítica da forma: $\vdash A \supset (B \supset A); !A \supset !B$. Aparentemente, do passo anterior para esse, foi utilizada uma espécie de lei de exportação da Lógica Proposicional Clássica.

Dessa deriva-se, por sua vez, a sentença analítica da forma: $\vdash A \supset (B \supset A); !A \supset$ *tem de* ("must") $!B$, que é a análise do imperativo técnico.

Considerações finais

O projeto griceano de análise filosófica da racionalidade, teórica e prática, é ambicioso. Ele quer, a exemplo de Aristóteles e Kant, a quem menciona a esse respeito, derivar consequências filosóficas da própria ideia de um ser racional[48]. Isso, a meu ver, não está muito distante daquilo que ocorre nos denominados "argumentos ontológicos": uma superposição indevida dos planos definicional e inferencial.

Para fazer justiça a Grice, ele está ciente que a proposta envolve, em

[46] 2001, p.93-94.
[47] 2001, p.94.
[48] GRICE, 2001, p.4.

primeiro lugar, a definição precisa de ser racional, e em segundo lugar, a determinação do tipo de procedimento que conta como uma derivação[49]. Ele se limita, na obra, a arriscar uma solução para a primeira tarefa.

No plano definicional, ele defende a Tese da Equivocidade, a saber, a tese segundo a qual certas expressões modais comuns aos discursos alético e prático são unívocos ao longo da barreira prática/alética, ou, caso sejam multívocos, sua multivocidade ocorre igualmente em ambos os lados da barreira[50]. Ele prefere a primeira alternativa, e a análise de sentenças apresentada na seção anterior é um dos modos de defendê-la. A Tese da Equivocidade, Grice a associa à tese de Kant segundo a qual há uma única faculdade da Razão[51].

No plano inferencial, ele propõe duas alternativas de defesa das inferências práticas: ou elas são equivalentes a inferências aléticas, ou as aceitabilidades práticas são inferíveis de aceitabilidades aléticas[52]. Grice prefere a segunda alternativa, e a defende com um estudo de caso: a derivação de imperativos técnicos kantianos.

Na execução de seu projeto Grice tem a oferecer antes um esboço, ainda que promissor, do que algo acabado. E, importante, desde o início ele não se deixa seduzir pelo que denominei "sintaticismo". Oxalá os lógicos e filósofos pudessem evitar sempre esse erro!

Referências

CHAPMAN, S. *Paul Grice: Philosopher and linguist*. Hampshire: Palgrave Macmillan, 2008.

FREGE, G. *Logical investigations*. Oxford: Blackwell, 1975 [1918-1923].

GRICE, P. *The conception of value*. Oxford: Oxford University Press, 1991.

[49] GRICE, 2001, p.4.
[50] GRICE, 2001, p.90.
[51] GRICE, 2001, p.44.
[52] GRICE, 2001, p.90.

_____. *Aspects of reason.* Oxford: Oxford University Press, 2001.

HARDIE, W.F.R. "The final good in Aristotle's ethics". In: *Philosophy* 40.154: p.277-295, 1965.

HARE, R. M. "Imperative sentences". In: *Mind,* 58, 229, p.21-39, 1949.

_____. *The language of morals.* Oxford: Oxford University Press, 1952.

_____. "Some sub-atomic particles of logic". In: *Mind* 98, 389, p.23-37, 1989.

JÖRGENSEN, J. "Imperatives and logic". In: *Erkenntnis* 7.1, p.288-296, 1937-1938.

KELSEN, H. *Allgemeine Theorie der Normen*. Wien: Manz, 1979.

LEWIS, C. S. *God in the dock: Essays on theology and ethics*. Grand Rapids: William B. Eerdmans, 1970.

VON WRIGHT, G. H. "Deontic logic". In: *Mind* 60, 237, p.1-15, 1951.

_____. "Deontic logic: A personal view". In: *Ratio Juris* 12.1, p.26-38, 1999.

WITTGENSTEIN, L. *Philosophical investigations*. Revised 4th edition. Oxford: Wiley-Blackwell, 2009, [1953].

Editora
UFPel

DISSERTATIO
FILOSOFIA