

Normatividade moral

Leonardo de Mello Ribeiro

Rogério A. Picoli

Vitor Sommovilla

(Orgs.)



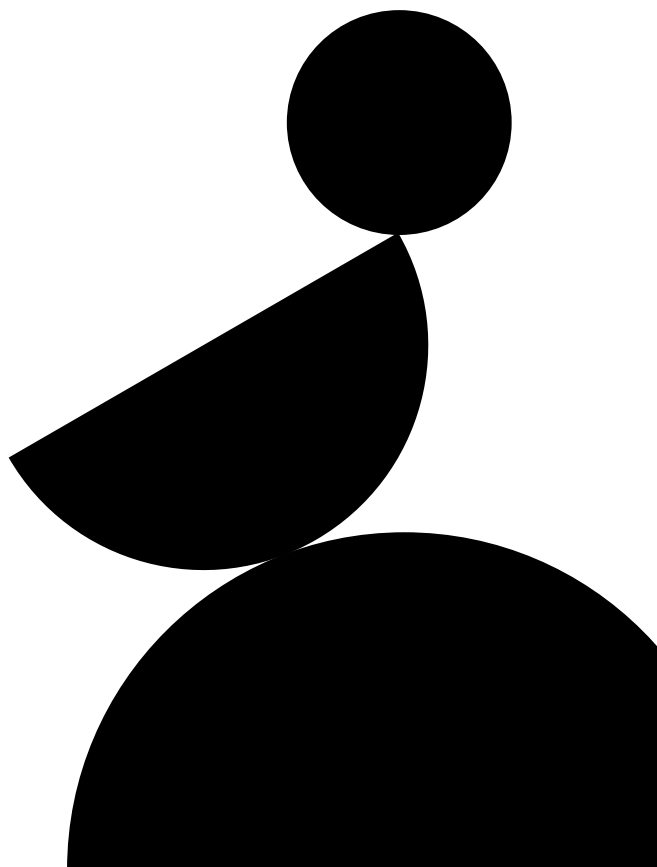
Normatividade moral

Leonardo de Mello Ribeiro

Rogério A. Picoli

Vitor Sommavilla

(Orgs.)



UNIVERSIDADE FEDERAL DE MINAS GERAIS

Reitora: Sandra Regina Goulart Almeida

Vice-Reitor: Alessandro Fernandes Moreira

FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS

Diretora: Thais Porlan de Oliveira

Vice-Diretor: Rogério Duarte do Pateo

DEPARTAMENTO DE FILOSOFIA

Chefe: Leonardo de Mello Ribeiro

Subchefe: Tadeu Mazzola Verza

PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

Coordenador: Ernesto Perini Frizzera da Mota Santos

Subcoordenador: Rogério Antônio Lopes

EDITORA PPGFIL-UFMG

Tadeu M. Verza

André J. Abath

Abílio A. Rodrigues Filho

Henrique B. Barreto

CONSELHO EDITORIAL

Daniel Pucciarelli

Miriam Campolina Diniz Peixoto

Rogério Antônio Lopes

Verlaine Freitas

CONSELHO CIENTÍFICO

Alessandro Pinzani (UFSC)

Anderson Bogéa da Silva (UFPR)

André Leclerc (UNB)

André Nascimento Pontes (UFAM)

Antônio José Ferreira Bento (Universidade da Beira Interior)

Beatriz Cecilia Bossi López (UCM)
Catarina Belo (Universidade Americana no Cairo)
Gisele Amaral dos Santos (UFRN)
Jacques Poulain (Université Paris VIII)
José Leonardo Annunziato Ruivo (UEMA)
Marco Antônio Caron Ruffino (UNICAMP)
Marcus Sacrini Ayres Ferraz (USP)
Maria Aparecida Paiva Montenegro (UFCE)
Maurício Pagotto Marsola (UNIFESP)
Sara Juliana Pozzer da Silveira (UFMT)
Vinicius Berlendis de Figueiredo[†] (UFPR)

<https://www.edppgfil.fafich.ufmg.br>

<https://www.editorappgfilufmg.com>

Avenida Presidente Antônio Carlos, 6627

FAFICH, sala 4047, 4º Piso

Pampulha, Belo Horizonte - MG.

CEP: 31270-901

Dados internacionais de catalogação na publicação (CIP)

Normatividade moral [recurso eletrônico] / Leonardo de Mello Ribeiro, Rogério A. Picoli, Vitor Somavilla (orgs.) - Belo Horizonte: PPGFIL-UFMG, 2025.

N851

1 recurso online (394 p.): pdf

Inclui bibliografias.

ISBN: 978-65-01-77098-7

DOI: 10.5281/zenodo.17478416

1. Normatividade (Ética). 2. Ética. 3. Filosofia. I. Ribeiro, Leonardo de Mello. II. Picoli, Rogério A. III. Barros, Vitor Somavilla de Souza. IV. Título.

CDD: 10.44

CDU: 017

Elaborada por Vilma Carvalho de Souza - CRB 6/1390

Copyright © Os autores

Capa e projeto gráfico: Ami Comunicação & Design

Revisão: Ana Teresa Campos Souza

Diagramação: Cecília Moreira Mafra

Finalização: Tadeu M. Verza



PPGFIL-UFMG, 2025

Esta obra foi selecionada pelo Conselho Editorial do Selo PPGFIL/UFMG após avaliação por pareceristas *ad hoc*.

O acesso e a leitura deste livro estão condicionados ao aceite dos termos de uso do Selo do PPGFIL/UFMG, disponíveis em: [https://
https://www.edppgfil.fafich.ufmg.br](https://www.edppgfil.fafich.ufmg.br)

Sumário

Introdução: normatividade moral, sua complexidade e seus desafios	9
Leonardo de Mello Ribeiro, Rogério A. Picoli e Vitor Somnavilla	
 PARTE 1: TEORIAS MORAIS E NORMATIVIDADE	
A convergência ética normativa da humanidade	45
Cinara Nahra	
 Repensando a teoria <i>normativa</i> tríplice	67
Darlei Dall'Agnol	
 Hibridismo normativo	99
Denis Coitinho	
 PARTE 2: QUESTÕES METAÉTICAS E NORMATIVIDADE	
Razão prática e normatividade moral em Rainer Forst	131
Lucas M. Dalsotto e Felipe Taufer	
 Construtivismo normativo modesto: esboço de um projeto	161
Rafael Vogelmann	

Normatividade imperiosa da moral e culpabilização: uma leitura inferencialista 189

Rogério A. Picoli

Moralidade como solo normativo comum: primeiros passos de uma teoria 225

Vitor Sommavilla

PARTE 3: MOTIVAÇÃO MORAL E NORMATIVIDADE

Razões no realismo moral naturalista: uma interpretação alternativa 255

Adelino Ferreira

A imaginação e o agir moral 281

Claudia Passos-Ferreira

A moralidade da ansiedade: problemas e perspectivas 309

Flavio Williges

PARTE 4: JUSTIFICAÇÃO MORAL E NORMATIVIDADE

Quem tem competência ética? 335

Leonardo de Mello Ribeiro

O problema da explicação e da justificação nas teorias evolucionárias da moralidade 371

Matheus de Mesquita Silveira

Introdução: normatividade moral, sua complexidade e seus desafios

Leonardo de Mello Ribeiro

Rogério A. Picoli

Vitor Somavilla

Como diz Judith Jarvis Thomson (2008, p. 1), “nosso pensamento é rico no que se costuma chamar de normatividade”. Além de normatividade moral, prudencial, social, política, jurídica, de etiqueta – domínios interpretados como paradigmas de linguagem normativa –, falamos igualmente de normatividade lógica, semântica, epistêmica, psicológica, estética, convencional (de vários tipos). Também usamos uma gama extensa e plural de conceitos normativos. Chamamos de normativos não apenas enunciados que usam conceitos *deontológicos* (como “dever”, “certo” e suas respectivas negações, extensões e variações), mas também conceitos *valorativos*

(como “bom” e suas respectivas negações, extensões e variações).¹ Dizemos, por exemplo, “A deve cumprir suas promessas”, “A deve pagar seus impostos”, “No xadrez, é errado mover a torre na diagonal”, “A não deve acreditar que está vendo um unicórnio”, “Se você acredita que p e acredita que, se p , então q , você deve acreditar que q ”, “Em português, não está certo usar crase antes de verbos”, “A é um excelente cozinheiro”, “Essa faca é muito boa”, “A é um péssimo amigo”, “Pollock é um ótimo pintor”.

Nas últimas décadas, na literatura filosófica, “normatividade” adquiriu o status de termo técnico, e seu estudo vem fomentando o que podemos chamar, seguindo Finlay (2010), de “uma significativa indústria” da produção acadêmica.² Esse dado histórico pode dar a impressão de que estamos diante de um mesmo fenômeno, ainda que com múltiplas ramificações. Mas um primeiro desafio para uma investigação sobre o tema é justamente saber se há um significado comum do conceito de normatividade que reúne todos os usos nos distintos domínios anteriormente mencionados. Se estamos dispostos a admitir que há, por exemplo, normatividade lógica, semântica, epistêmica, psicológica, estética, parecemos ter também que admitir que normatividade se estende para além de uma dimensão prática (contrariamente à impressão que o foco em domínios

¹ Thomson (2008, p. 2), seguindo Wiggins, chama os primeiros de “diretivos” (*directives*) e os segundos de “avaliativos” (*evaluatives*).

² Finlay (2010) fornece dados empíricos que indicam que o termo se consolidou sobretudo a partir da década de oitenta na produção filosófica. Ainda de acordo com Finlay, há evidência de que sua origem na história recente de uso acadêmico está vinculada ao contexto jurídico. A despeito de o termo poder ser relativamente novo no contexto filosófico, as questões que ele abarca não parecem ser novas. Entretanto, a consolidação do termo pode ser evidência de um esforço por uma melhor compreensão e sistematização dessas questões.

paradigmaticamente interpretados como normativos – moral, prudência, política, direito, etiqueta etc. – poderia gerar). Um caminho adotado por muitos autores/as contemporâneos/as identifica, em enunciados sobre *razões*, a característica definidora de enunciados normativos. Assim, todos os enunciados normativos seriam, em algum sentido, variações de enunciados do tipo “[...] é uma razão para X” – em que “X” pode ser uma ação, atitude, crença, juízo, emoção, estado de coisas etc.³ Há, porém, ao menos duas considerações que lançam dúvidas sobre essa proposta.

Primeiro, é possível alegar que existem certos enunciados sobre “demandas de normatividade” (por exemplo, sobre a coerência interna às atitudes ou crenças de um indivíduo) que são distintos de enunciados sobre razões. Pode ser o caso que um indivíduo deva fazer algo (como, por exemplo, uma exigência de racionalidade

³ Para os propósitos desta Introdução, vários refinamentos necessários para um entendimento completo sobre enunciados de razão são dispensáveis. Por exemplo, nas últimas décadas, tornou-se lugar-comum tomar *atos* ou *proposições verdadeiras* como razões – embora autores como Scanlon (1998, p. 17) refiram-se a razões como “considerações que contam a favor de algo”. Ao contrário do que essa formulação pode sugerir à primeira vista, ela não é incompatível com a tradicional teoria humiana das razões (para ação), como argumenta Schroeder (2007) em sua defesa do seu “hipoteticismo” (*hypotheticalism*), segundo o qual todas as razões normativas para ação são *explicadas* por estados mentais como desejos. Além disso, há também o problema de saber se a relação em questão é diádica (entre uma razão e aquilo para o que é uma razão) ou triádica (entre uma razão, aquilo para o que é uma razão e um agente/indivíduo em particular). A distinção tradicional na literatura moral entre razões neutras ao agente (*agent-neutral reasons*) e relativas ao agente (*agent-relative reasons*) é um desdobramento dessa questão. Por fim, há ainda a pergunta sobre o que ocupa o lugar “daquilo para o que uma razão é uma razão”. Circunscrevendo o problema ao domínio da moralidade, autores/as divergem, basicamente, sobre uma razão ser uma razão para uma ação ou atitude ou ser uma razão para que um estado de coisas seja o caso (incluindo a possibilidade de que o estado de coisas seja do tipo: “que um agente A faça X”). Nada do que se segue nesta Introdução será afetado pela ausência de resposta a todas essas questões.

instrumental, interna às suas intenções e crenças) ou que ele deva acreditar em algo (como, por exemplo, uma exigência de racionalidade epistêmica diante da evidência que tem disponível), mas que ele não tenha, respectivamente, razão *propriamente* (“objetiva” ou “tudo o mais considerado”) para fazer aquele mesmo algo (porque seria, por exemplo, moralmente incorreto) e não tenha razão *propriamente* (“objetiva” ou “tudo o mais considerado”) para acreditar naquele mesmo algo (porque seria, por exemplo, falso). Se essa for uma descrição correta, então existiria mais de um sentido de normatividade, pois haveria demandas racionais (normativas) sobre o pensamento ou ação que contrastariam e poderiam ser superadas pelo que o indivíduo teria em última instância razão (normativa) para fazer ou acreditar.⁴

Segundo, e talvez mais decisivo, mesmo que todos os usos de vocabulário normativo vinculados a todos aqueles distintos domínios acima listados sejam interpretados como enunciados de *razões*, não é claro que as razões daqueles distintos domínios possuam algo que as unifique em torno de um significado unívoco de normatividade. Citando apenas uma contraposição possível, uma razão moral pode ser de natureza radicalmente distinta de uma razão epistêmica e não haver nada que as vincule, exceto um uso comum e vago do conceito de normatividade. Por exemplo, uma razão moral pode ser essencialmente prática, enquanto uma razão epistêmica pode não ter nenhum elemento prático em sua natureza. Se essa for uma descrição correta, então haveria mais de um *tipo* de normatividade.

⁴ Ver Broome, 1999, para uma distinção nesses termos. Recentemente, Broome (2020) alterou sua nomenclatura, distinguindo “normatividade” (como “responsividade a razões”) de “racionalidade” (como “uma propriedade superveniente das relações entre estados mentais de um indivíduo”).

Para além da questão sobre se há um significado geral comum para “normatividade”, parecemos também enfrentar dificuldades mesmo circunscrevendo a questão ao domínio da moralidade, que é o tema central deste livro. Em geral, partimos do pressuposto de que a moralidade é normativa. Como diz Copp:

A ideia de que afirmações morais são normativas é mais bem interpretada como algo dado. Um filósofo pode alegar não saber o que significa dizer que afirmações morais são normativas, querendo com isso dizer que ele não possui uma teoria ou análise da normatividade, ou pode negar uma teoria específica da normatividade. Mas não vejo nenhum fundamento para negar que afirmações morais sejam normativas. (Copp, 1995, p. 11)⁵

A despeito disso, possuímos pouca clareza (e parecemos divergir de modo significativo) sobre o que se pode substancialmente extrair daquele “dado”. Atenção a essa dificuldade sobre o que derivar do pressuposto de que o domínio da moralidade é normativo é algo raramente explicitado na literatura contemporânea sobre ética e metaética. Decerto há disponível nessa literatura um extenso leque de propostas sobre a natureza e o conteúdo da normatividade moral. Mas o grande desafio é saber se aquele “dado” do qual partimos é suficiente para justificar expectativas de que seja possível unificar o fenômeno em torno de um conceito unívoco. Vejamos alguns temas clássicos e suas dificuldades associadas.

⁵ Copp (2024, p. 1) diz ainda que: “todos os principais problemas da metaética – problemas de semântica, epistemologia, metafísica e psicologia filosófica – remontam ao problema da normatividade. Eles surgem da forma que surgem porque a ética é normativa”.

Uma forma canônica de caracterizar normatividade no domínio moral se dá através da contraposição entre descrição/prescrição ou fato/valor.⁶ Enunciados prescritivos seriam normativos porque teriam uma relação direta com algum tipo de exigência ou recomendação de realização de uma ação, enquanto enunciados descritivos apenas descreveriam ou representariam uma realidade factual (empírica/natural ou não empírica/supranatural) sem, no entanto, implicar qualquer tipo de relação com a realização de uma ação.⁷ A ideia principal por trás dessas dicotomias é a de que as funções linguísticas ou semânticas de tais termos ou conceitos seriam distintas e, assim, a natureza da prescrição e do valor, por um lado, e a das propriedades e fatos reais (empíricos/naturais ou não empíricos/ supranaturais), por outro, seriam também distintas. A insistência na necessidade de

⁶ Essas duas dicotomias não precisam ser interpretadas como inteiramente coincidentes, mas um refinamento desse ponto iria além dos propósitos desta Introdução. A despeito disso, ver nota 6 abaixo para considerações que apontam na direção de distinções entre essas duas dicotomias.

⁷ Essas dicotomias análogas são defendidas de forma veemente por Hare (1952) e toda a tradição não-cognitivista. Para Hare, em particular, juízos morais e valorativos em geral são, em essência, prescritivos ou recomendatórios (no caso moral, mais especificamente, prescritivos universais). Cumprem, assim, uma função essencialmente prática e não descritiva ou representacional de quaisquer propriedades ou fatos (embora Hare admita que certos enunciados morais possam ser primariamente prescritivos e secundariamente descritivos). Mackie (1977) concorda com metade dessa história e, como resultado, ocupa um lugar peculiar nesse debate. Ele concorda com Hare e os não-cognitivistas que enunciados morais contêm um elemento prescritivo ou normativo, que não pode ser analisado/reduzido em termos de propriedades descritivas não morais. Mas, por outro lado, Mackie rejeita a tese de Hare e dos não-cognitivistas de que enunciados/juízos morais não cumprem primariamente uma função representacional/cognitiva e expressam proposições sobre propriedades e/ou fatos intrinsecamente normativos. É justamente por não encontrar razão para acreditar que existam tais propriedades e/ou fatos que Mackie defende sua versão de teoria do erro aplicada ao discurso moral.

reconhecimento de tais dicotomias é central à tradição não-cognitivista em metaética.⁸

Um problema com tais dicotomias é que elas nunca foram consensuais. Subjetivistas, relativistas, realistas naturalistas (de diferentes estirpes) e pragmatistas morais as negam. Para todos esses, embora por diferentes razões, descrição e prescrição, fato e valor, não são mutuamente excludentes, seja por combinarem-se, serem redutíveis ou identificáveis um com o outro, serem indissociáveis ou interdependentes. Assim, enunciados (morais) valorativos/prescritivos seriam, em algum sentido a ser especificado, também factuais/descritivos. Encontramos na literatura sobre o tema uma

⁸ Também, em certo sentido, a Moore (1903) e aos intuicionistas em geral. Mas a questão torna-se complexa porque há um sentido em que Moore e os intuicionistas admitem que existem proposições morais verdadeiras (autoevidentes) e fatos morais, embora o conteúdo de tais proposições e a natureza desses fatos sejam *sui generis*. Assim, embora eles concordem com não-cognitivistas que não é possível extrair uma conclusão moral de um conjunto de premissas não morais e que um termo moral não pode ser analisado ou reduzido a um (conjunto de) termo(s) não moral(is), eles não precisam endossar inteiramente a dicotomia fato/valor – exceto se por “fato” compreender-se apenas fatos empíricos/naturais ou não empíricos/supranaturais. De modo similar, eles também não concordariam inteiramente com a interpretação dada por não-cognitivistas à dicotomia prescrição/descrição. Embora seja verdade que, para Moore e os intuicionistas, propriedades morais sejam intrinsecamente normativas e que propriedades morais não sejam descritivas de propriedades empíricas/naturais ou não empíricas/supranaturais, eles pensam que termos morais são, em algum sentido, *representacionais* – Moore é explícito ao dizer que “bom” denota uma propriedade simples, indefinível, *sui generis*, assim como intuicionistas tendem a defender que “dever” ou “razão” também denota uma propriedade ou relação indefinível, *sui generis* e “primitiva”. Ver Shafer-Landau, 2003, para uma defesa de um realismo de propriedades morais *sui generis* normativas, e Larmore, 2021, para uma defesa de um realismo de razões como relações *sui generis* normativas e primitivas.) Moore, em particular, parece enfrentar a dificuldade adicional de explicar em que sentido “bom” é normativo, já que pressupõe a prioridade deste conceito sobre o conceito de dever. Como Frankena (1942) argumenta, é uma questão significativa para Moore saber por que devemos maximizar aquilo que é bom.

variedade de versões dessa relação: por exemplo, enunciados (morais) valorativos/prescritivos seriam também factuais/descritivos das preferências de um indivíduo; das preferências, interesses e necessidades de seres humanos em geral ou de uma sociedade/cultura em particular; dos interesses, necessidades, capacidades e funcionamentos da natureza humana (tornando indissociáveis ou interdependentes valor/prescrição e fato/descrição).⁹ Todas essas formas de rejeição daquelas dicotomias, e as respectivas propostas alternativas de sua superação, são evidência da dificuldade de se conferir univocidade ao conceito de normatividade no domínio da moralidade.

Essa variação na forma de se pensar a natureza da prescrição e do valor também se reflete nas várias maneiras de se pensar a *origem* e o *fundamento* da normatividade moral. Para alguns, a investigação sobre a origem e o fundamento da normatividade moral é uma investigação empírica (em grande medida, em terceira pessoa, compatível com algum tipo de investigação científica) sobre disposições e necessidades humanas (biológicas, psicológicas e sociais). Para outros, há algo irredutivelmente dado pela perspectiva de primeira pessoa no fundamento da normatividade moral (seja na forma de uma intuição, de uma atitude prescritiva, de uma emoção, do exercício da escolha livre racional,¹⁰ da observância de procedimentos universais de racionalidade prática etc.). Ainda outros identificam o fundamento da moralidade em algum tipo de acordo

⁹ Nesse rol, apesar de diferenças significativas entre suas propostas, podemos elencar tanto naturalistas reducionistas como Railton quanto neoaristotélicos (distintos) como Thomson e McDowell, e pragmatistas como Putnam.

¹⁰ A ideia de uma escolha livre racional pode ser interpretada como introduzindo ainda outro significado de “razão”: o nome da faculdade humana do raciocínio e/ou a faculdade da origem e fundamento das exigências de racionalidade *per se*.

resultante de uma vinculação recíproca, seja baseada em reivindicações e/ou expectativas comuns, seja baseada na maximização do autointeresse racional, seja baseada em convenções e práticas sociais histórica e culturalmente situadas. Essas distintas vias de identificação da origem e do fundamento da normatividade moral terão também, como resultado, distintas implicações nas respostas a questões acerca da sua *natureza*: categoricidade (ou instrumentalidade), autoridade (intrínseca ou extrínseca), supremacia (ou possibilidade de superação por outros valores e normas), universalidade (ou relatividade).

Por fim, aquela que parece ser a dimensão mais crítica da normatividade moral, por dizer respeito a sua própria aplicação prática, consiste justamente em identificar o seu *conteúdo* adequado ou verdadeiro. Historicamente, diferentes tradições de teorias normativas de primeira ordem (como a deontologia kantiana, o utilitarismo e/ou consequencialismo, o contratualismo e a ética de virtudes) fornecem respostas conflitantes sobre tal conteúdo. Em muitos momentos desse debate, é possível ter a impressão de que há algum tipo de dissonância entre os debatedores sobre o que pressupõem ser parte do conteúdo normativo da moralidade. Assim, certas tentativas de utilitaristas de responder a alguns dos desafios de deontologistas – por exemplo, que o utilitarismo é excessivamente exigente ou não é capaz de acomodar restrições implicadas por direitos individuais – nunca parecem ser inteiramente satisfatórias ao tipo de expectativa de justificação moral baseada em pressupostos deontológicos.

De modo similar, no que diz respeito aos métodos de justificação de crenças morais, intuicionistas (epistêmicos) morais nunca se satisfazem inteiramente com as respostas aos seus desafios, mesmo as

mais sofisticadas, fornecidas por adeptos de alguma versão de coerentismo moral (incluindo o método do equilíbrio reflexivo). Parte desse resultado é, obviamente, trivial: trata-se de visões concorrentes sobre o conteúdo da normatividade moral e sobre sua justificação. Também nada se segue diretamente em relação à impossibilidade de convergência de opiniões. Mas a persistência de várias disputas representa um desafio para uma investigação dessa dimensão da normatividade moral e alimenta dúvidas céticas.

Há, na produção filosófica recente sobre o tema, um direcionamento em particular que tenta reconfigurar tais disputas e impasses historicamente persistentes. Trata-se de várias iniciativas de esforço para compatibilizar aquelas diferentes tradições normativas morais (entre, por exemplo, a deontologia kantiana, o utilitarismo e/ou consequencialismo, o contratualismo e a ética de virtudes), concentrando-se em suas aproximações e fornecendo algum tipo de revisionismo conciliatório sobre seus conflitos. O resultado são versões de teorias morais pluralistas (sobre valores e/ou deveres morais) que se aproximam do espírito do intuicionismo (normativo) pluralista da primeira metade do século XX, de autores como Ross e Prichard, no sentido de reconhecerem que uma ampla gama de considerações morais – como justiça, beneficência, não maleficência, fidelidade, reparação, autonomia, reciprocidade, desenvolvimento pessoal etc. – são todas relevantes em si mesmas, podendo entrar em conflito em certos contextos. Algum tipo de critério, procedimento ou princípio mais básico para compatibilizar tais deveres e/ou valores e superar seus conflitos é parte do que tais teorias recentes almejam fornecer (e, nesse sentido específico, distanciam-se daqueles autores intuicionistas).

De modo similar, no que diz respeito à questão metodológica, não são poucas as iniciativas que tentam fornecer alguma compatibilização entre intuicionismo (epistêmico) moral e equilíbrio reflexivo, entre fundacionalismo e coerentismo, produzindo modelos “híbridos”. Mesmo autores (por exemplo, Scanlon, 2014; DePaul, 2006) que reconhecem como inevitável aderir, em assuntos morais, a alguma metodologia coerentista de justificação de crenças morais (como o método do equilíbrio reflexivo) alegam que um modelo completo e adequado de justificação moral deve também incluir algum critério de hierarquização e reconhecimento de superioridade epistêmica de certas crenças morais (ponderadas) que terminam desempenhando um tipo de papel fundacional.

Mas, apesar dessas ambições “teorizadoras” e “sistematizadoras” típicas da tradição moral filosófica, pode-se afirmar com certa segurança que seu contraponto crítico – ceticismos de várias estirpes sobre a plausibilidade da própria ideia de teoria moral, sobre as pretensões filosóficas universalistas, sobre até mesmo a possibilidade de sistematização abstrata do pensamento moral – nunca saiu de cena e permanece como desafio à questão intrinsecamente normativa sobre o que devemos moralmente fazer e sobre o que é moralmente bom, e como justificar nossas convicções a esse respeito.

Este livro explora esses múltiplos sentidos da normatividade moral. Em sua diversidade de abordagens, fornece tanto visões otimistas quanto pessimistas sobre a possibilidade de responder a várias das questões acima descritas. Assim, o objetivo central do livro não é tentar conferir uma unidade ao conceito de normatividade moral, nem fornecer um mapeamento completo do tema, mas fazer jus à complexidade do fenômeno. Essa complexidade está representada em quatro partes temáticas que dividem o livro: (1)

Teorias morais e normatividade; (2) Questões metaéticas e normatividade; (3) Motivação moral e normatividade; (4) Justificação moral e normatividade. Passamos, na sequência, a uma breve descrição dos conteúdos dessas quatro partes.

Na Parte 1, “Teorias morais e normatividade”, embora Cinara Nahra, Darlei Dall’Agnol e Denis Coitinho forneçam contribuições distintas sobre como responder a questões normativas morais básicas de primeira ordem, todas são versões “híbridas” das teorias morais tradicionais filosóficas, tentando conciliar compromissos centrais às tradições kantiana, contratualista, utilitarista e/ou consequencialista e da ética de virtudes.

No capítulo “A convergência ética normativa da humanidade”, Cinara Nahra defende a ideia de que existe um conjunto de dez “subprincípios” que formam o núcleo normativo da moralidade humana, que devem funcionar como guia normativo para as ações, sob o risco de a própria existência da humanidade, seus patamares de bem-estar e seus marcos civilizatórios serem colocados em xeque. Tais subprincípios devem valer não apenas para ações individuais, mas também para empresas e grandes corporações, assim como para o Estado. Nahra conjectura que a maioria dos seres humanos, independentemente de “cultura, posição social, raça ou gênero”, endossaria esse núcleo normativo na condução dos seus afazeres e relações. A despeito disso, sua observância, obviamente, não é garantida e, assim, Nahra chama atenção para sua relevância ainda mais notável em um mundo no qual “o lucro e o interesse próprio imediato parecem ser colocados acima de tudo”. Características específicas do panorama contemporâneo, que envolvem fatores múltiplos como a formação de poderosos oligopólios, a crescente desinformação em ambientes globais digitais, os imprevisíveis usos da

inteligência artificial, potencializam os riscos de dominação por empresas privadas e grandes corporações, algo que se reflete no controle e manipulação de várias dimensões da vida ordinária dos indivíduos comuns.

Nahra chega a tal lista de dez subprincípios fazendo eco a uma ideia pouco explorada na tradição filosófica recente e, nesse sentido, sua tese assume um caráter original e desafiador. Assim como Hare (1981), Nahra identifica aproximações significativas entre a deontologia kantiana e o consequencialismo utilitarista. Mas, enquanto Hare desenvolve essa ideia focando essencialmente na dimensão metaética, Nahra concentra-se na dimensão estritamente normativa. Apesar das distintas justificações que os modelos deontológico kantiano e consequencialista utilitarista possam fornecer, eles concordam substancialmente no que diz respeito aos conteúdos de suas moralidades. Nesse sentido, Nahra termina aproximando-se do pluralismo dos deveres *prima facie* de Ross (1930), pois, assim como Nahra, Ross também tomou como inquestionável, por exemplo, que justiça, fidelidade, beneficência e não maleficência constituam igualmente um núcleo normativo intuitivo da moralidade, ainda que os deveres associados possam entrar em conflito em situações particulares. Os dez subprincípios de Nahra, que são derivados dos compromissos centrais do imperativo categórico, do princípio da utilidade, do princípio do dano (de Mill) e da regra de ouro, envolvem desde o respeito às pessoas, o compromisso com a verdade e a honestidade, passando pela promoção da felicidade e a minimização do sofrimento alheio, até a preservação do planeta e abstenção da crueldade com outros seres vivos. A esses subprincípios, Nahra dá o nome sugestivo de “deontoutilitaristas”.

No capítulo “Repensando a teoria *normativa* tríplice”, Darlei Dall’Agnol faz alguns questionamentos sobre o modo como Derek Parfit encaminha o trabalho de compatibilização das teorias normativas em *On What Matters*. Dall’Agnol propõe um aperfeiçoamento da *Teoria Tríplice* abrindo espaço para elementos da ética aretaica e, em seguida, o foco recai sobre o refinamento da ética aretaica a ser incorporada àquele tipo de teoria. Ao final, o autor aponta algumas das vantagens da Teoria Tríplice reformulada com base na ética aretaica.

Ao tratar da deontologia kantiana, Parfit desconsidera um aspecto central: a fórmula da humanidade, segundo a qual há um valor absoluto e incondicional associado à dignidade da pessoa humana. A negligência em relação a tal aspecto tornaria a teoria de Parfit problemática e incompatível com uma “ética realmente kantiana”. Esse é o mote para Dall’Agnol passar em revista o tipo de compatibilidade que pode haver entre a ética “realmente” kantiana e o consequencialismo. Não seria possível abandonar o respeito pelas pessoas em favor do consequencialismo e seguir sendo kantiano, mas, ele acredita, as duas perspectivas podem ser coadunadas se a teoria adotar o princípio de que respeitar os direitos básicos das pessoas é algo otimizador.

No contratualismo de Scanlon, endossado por Parfit, um princípio é moralmente válido se ninguém puder rejeitá-lo de maneira razoável. Dall’Agnol sugere pensar a razoabilidade em termos aretaicos, como uma qualidade de caráter. Essa interpretação da razoabilidade é o que poderia conferir à teoria de Parfit a especificidade de uma teoria “realmente tríplice”: deontológica, consequencialista e aretaica. A razoabilidade de se aderir a princípios otimizadores universalmente aceitáveis é o que aproxima a

perspectiva contratualista da ética de virtudes. Então, defende Dall'Agnol, “podemos postular um princípio ético primeiro (PEP) incorporando o critério formal de correção das ações de uma ética *aretaica*”. As virtudes: são critérios de correção moral, não estabelecem prioridade entre o bom e o correto, reconhecem o valor intrínseco dos traços de caráter e são ancoradas em princípios. Uma ética aretaica é uma teoria da ação valiosa. Nisso ela se aproxima tanto da deontologia quanto de um consequencialismo de regras. O critério de correção fornecido pela ética aretaica é algo como a pessoa virtuosa é um juiz justo e julga imparcialmente. O critério tem relação com o agente e suas motivações, mas não é dependente deles. Em síntese, o autor propõe uma ética aretaica não relativista invocando o valor intrínseco das virtudes e a ideia de que certas qualidades de caráter são necessárias: para a promoção da boa vida e para que nos tornemos cidadãos razoáveis.

Em “Hibridismo normativo”, Denis Coitinho propõe-se a desenvolver uma teoria normativa híbrida. A chave para a construção da teoria híbrida é a análise da distinção entre o domínio da moralidade pública e da moralidade privada, especificamente, da assimetria quanto à intensidade relativa da censura em casos de ofensas à moralidade pública. O hibridismo proposto busca dar conta dessa assimetria em situações complexas, aquelas em que somos chamados a considerar e conciliar aspectos de ambas as dimensões. Tendo em vista as limitações e dificuldades das teorias monistas aplicadas isoladamente (utilitarismo, kantismo, contratualismo etc.), a melhor saída para tais situações seria uma teoria que pudesse combinar uma ética de virtudes e uma ética neocontratualista. Assim, a motivação e a justificação desse tipo de teoria é o fato de que problemas complexos demandam a mobilização de critérios

normativos diversos, por exemplo, o desafio de alocação de leitos de UTIs durante a pandemia da covid-19: o critério utilitarista da maior eficiência falharia ao priorizar pacientes com maiores chances em detrimento de idosos, crianças e pessoas com comorbidades. O critério kantiano falharia porque atender ao princípio da humanidade (nenhum agente deve ser tomado como meio) poderia resultar num aumento indesejável de mortes. Um problema análogo seria observado ao se apelar aos critérios de uma ética de virtudes.

O desafio da assimetria entre as exigências da moralidade pública e da privada, ilustrado pela diferença na intensidade da censura, parece ser devido ao fato, argumenta Coitinho, de que, no caso da moralidade privada, a esfera da autoridade é principalmente pessoal, já, no âmbito da moralidade pública, a esfera da autoridade moral é principalmente coletiva. Na dimensão privada, a autoridade normativa é em primeira pessoa, envolve certa disposição do agente e nesse sentido ela mostra-se mais desafiadora a teorias morais monistas, que tendem a impor demandas de raciocínio moral aos agentes. Por sua vez, a moralidade pública parece operar noutra direção, como uma esfera sobretudo de direitos e obrigações, mais do que de traços de caráter, portanto, menos relativista. Nela, a obrigação é acrescida de forte pressão social e cobrança em favor das normas endossadas coletivamente. O neocontratualismo seria mais adequado nessa dimensão porque, por essa teoria, os deveres são estabelecidos a partir da correlação de direitos exigidos pela vida comum “a partir da própria aceitabilidade dos envolvidos” e com base na racionalidade e razoabilidade.

O próximo desafio é especificar o procedimento apropriado para a seleção das virtudes e dos princípios que compatibilize as dimensões pública e privada. Essa teoria híbrida neocontratualista das virtudes

tem de compatibilizar os critérios de razoabilidade e de justiça das teorias neocontratualistas que se sobressaem na dimensão pública com virtudes da dimensão privada que sejam compatíveis com uma perspectiva liberal pluralista, tais como: prudência, autonomia, humildade etc. Os critérios de moralidade pública poderiam ser, por exemplo, aqueles resultantes da articulação entre duas teorias neocontratualistas: a teoria de Rawls, que ofereceria os princípios de justiça e de equidade da sociedade, e a teoria de Scanlon, que ofereceria o procedimento para a escolha de um critério de certo e errado que teria em conta justiça, razoabilidade, direitos e tolerância dos agentes. Os critérios da dimensão privada seriam oferecidos por uma teoria das virtudes eudaimonista, isto é, orientada para a promoção da felicidade pessoal e coletiva e para o florescimento humano; uma teoria que tende a responsabilizar o agente com base no seu caráter. Já os critérios de moralidade da ponte entre a dimensão privada e a pública da moralidade seriam determinados pela aplicação de um procedimento de escolha de virtudes privadas coerentes com as virtudes e os princípios demandados pela moralidade pública. A ideia central do procedimento é que a escolha seja coletiva, devendo ser justificada pela razão pública e assumindo-se que os agentes estejam de posse de todas as informações relevantes. Uma virtude privado-pública será aprovada se ela não puder ser “razoavelmente rejeitada pelos envolvidos na decisão”, o que significa que se trata de uma virtude sensível ao contexto.

Em última instância, Coitinho procura defender certa coerência entre essas teorias. A despeito de uma apelar a princípios e a outra a virtudes, o autor defende que os seus conteúdos normativos apresentam alguma semelhança. A razoabilidade dos agentes aponta para as virtudes da humildade, tolerância, coerência, prudência e

reflexividade, além da justiça e a da equidade, e ambas as teorias ancoram a normatividade na autonomia dos agentes.

Passando à Parte 2 do livro, “Questões metaéticas e normatividade”, Lucas M. Dalsotto e Felipe Taufer, Rafael Vogelmann, Rogério A. Picoli e Vitor Somavilla fornecem distintas incursões em questões metaéticas gerais sobre a fundamentação do discurso moral e a natureza do juízo moral. Um aspecto comum às abordagens desses autores é uma tentativa de derivar conclusões normativas a partir de certas premissas metaéticas: sobre a estrutura da razão prática, sobre a natureza do agente humano e da interação entre agentes desse tipo e sobre práticas discursivas de uso do vocabulário moral.

Em “Razão prática e normatividade moral em Rainer Forst”, Lucas M. Dalsotto e Felipe Taufer examinam os potenciais e carências da teoria de Forst sobre os fundamentos da moralidade, exposta em *O direito à justificação* (publicação original em 2007). Segundo Forst, contextos morais exigem que agentes racionais não apenas adotem meios adequados para perseguir seus objetivos, mas também e crucialmente que eles sejam capazes de justificar intersubjetivamente as suas ações. Para que ações sejam consideradas justificadas em contextos morais é preciso que não existam razões intersubjetivamente compartilhadas que possam rejeitá-las. Subjaz a essa abordagem a noção de um direito moral à justificação, que faz com que reivindicações morais devam ter validade recíproca (i.e., aplicável igualmente ao sujeito e ao objeto da reivindicação) e geral (i.e., aplicável a todos os agentes afetados pela ação).

Dalsotto e Taufer exploram o modo como Forst defende sua posição, com destaque para seu simultâneo compromisso com as seguintes duas ideias: i) razões morais possuem uma conexão intrínseca (ou “interna”) com a motivação para a ação; e ii) a natureza

das razões morais comporta elementos “externos” aos estados mentais e motivacionais dos agentes morais. O restante do texto põe à prova o modo como Forst busca conciliar as teses de que razões morais intrinsecamente motivam, mas são colhidas externamente aos agentes e, portanto, podem franquear um tipo de objetividade. Ademais, sublinha-se o esforço de evitar recair em uma forma de metafísica que simplesmente assumiria algum tipo de fato da razão nos moldes kantianos. Embora reconheçam o potencial do trabalho de Forst, os autores propõem suplementar a teoria com aspectos das abordagens de John McDowell e Robert Brandom. A noção de segunda natureza, tal como explorada por McDowell, daria a Forst um acesso à objetividade da moral sem necessidade de uma metafísica extravagante. Por sua vez, as reflexões de Brandom acerca do espaço das razões tornariam mais robusta a caracterização feita por Forst dos compromissos práticos que subjazem a todo uso da razão prática em contextos morais.

Em “Construtivismo normativo modesto: esboço de um projeto”, Rafael Vogelmann apresenta uma variante original de construtivismo normativo que não pretende ser capaz de estabelecer diretrizes morais válidas para qualquer agente racional concebível meramente a partir da forma da razão prática, como almejam os proponentes do construtivismo kantiano, mas também não se satisfaz com a constatação de que nenhum compromisso normativo substantivo pode ser extraído de uma caracterização relevante de agentes racionais, como sustentam os construtivistas humanos. Construtivistas kantianos (e.g., Christine Korsgaard) e humanos (e.g., Sharon Street) divergem sobre a possibilidade de se derivarem princípios morais de uma caracterização geral e bastante minimalista da razão prática, que abarcaria qualquer agente racional concebível.

Diante desse impasse, passou-se a explorar na literatura recente um projeto mais modesto, mas não menos importante: a tentativa de extrair compromissos normativos substantivos de uma caracterização mais robusta (e, portanto, menos generalizável) de agente racional. O capítulo de Vogelmann explora uma forma de levar adiante esse projeto.

Em particular, Vogelmann concebe agentes morais como agentes dotados de razão prática que avaliam outros agentes em termos de vícios e virtudes. Dessa caracterização de agente moral, argumenta-se que compromissos morais significativos podem ser inferidos. Embora esses compromissos morais não sejam derivados de uma concepção geral de razão prática, como no projeto ambicioso de Korsgaard, o autor sustenta que o resultado é importante, na medida em que nós, seres humanos reais, somos agentes do tipo caracterizado pelo seu projeto. Desse modo, por mais que os compromissos substantivos fundamentados pela teoria possam não valer para todo e qualquer agente racional, eles teriam validade assegurada para agentes como nós. De forma simplificada, Vogelmann estabelece uma conexão íntima entre valores, razões e virtudes que lhe permite excluir ideais de virtude que não estejam comprometidos com uma forma de reciprocidade, a exemplo do egoísmo racional. Um passo crucial nessa direção é a ideia de que se importar com algo ou alguém implica considerar vicioso o caráter de quem prejudica esse algo ou alguém. Nesse sentido, por exemplo, um egoísta que se aproveitasse dos outros para obter seus fins e julgasse virtuosos todos os agentes que assim se comportassem incorreria em uma contradição, visto que, ao mesmo tempo, julgaria virtuoso o comportamento de outro egoísta, por manifestar o ideal egoísta de virtude, e o julgaria vicioso, por agir de forma prejudicial aos seus fins (do primeiro egoísta). Para evitar

inconsistências como essa, seria preciso limitar significativamente o escopo de virtudes aceitáveis e, assim, aceitar um conjunto mínimo e fixo de razões morais.

Em “Normatividade imperiosa da moral e culpabilização: uma leitura inferencialista”, Rogério A. Picoli pretende estender a concepção inferencialista das práticas proposta por Robert Brandom para as práticas morais. Mais especificamente, para três práticas morais acusatórias: censura moral (*moral blame*), responsabilização moral (*attribution of moral responsibility*) e culpabilização (*attribution of guilt*). Picoli quer defender que o ferramental fornecido pela abordagem inferencialista das práticas possibilita um refinamento da caracterização das normas constitutivas destas. A caracterização e a distinção entre essas práticas podem elucidar como surgem os aspectos fenomenológicos próprios da normatividade imperiosa (*authoritative*) da moral.

Argumenta-se que o inferencialismo interpretado como uma teoria pragmática das “normas de uso-primeiro” pode oferecer uma compreensão das características fenomenológicas da normatividade imperiosa. Para tanto, é necessário fornecer uma caracterização da pragmática inferencialista e da semântica normativa propostas por Brandom. Assim, o desafio é esclarecer como, para Brandom, conteúdos conceituais e proposicionais emergem de práticas de atribuição de compromissos e autorizações monitoradas por um *scorekeeping* deôntico. O autor destaca que a leitura inferencialista das práticas, ao privilegiar as normas que as constituem e governam os usos, consegue explicar não só a correção conceitual, mas também o significado prático das nossas ações.

Tendo reunido o ferramental inferencialista, Picoli propõe-se a aplicar esse arcabouço a três práticas morais concretas: censura,

responsabilização e culpabilização. Após mostrar como cada uma delas reorganiza o espaço de razões, o autor volta a atenção para a culpabilização. Na culpabilização, há um rebaixamento do status deôntico do agente faltoso, um fechamento de suas possibilidades de justificação e a percepção da suspensão da derrotabilidade da norma violada. Disso decorre a “autoridade” irresistível da norma moral, que assume na prática de culpabilização o aspecto categórico. Tal aspecto é explicado justamente pelas normas constitutivas da culpabilização. O argumento apresentado por Picoli é que os aspectos fenomenológicos da normatividade imperiosa da moral podem ser compreendidos como decorrentes da configuração específica da prática de culpabilização, incluindo: o rebaixamento do status deôntico do agente culpabilizado, as restrições, as limitações, as interdições e o próprio fechamento do espaço de razões. Assim, defende o autor, a extensão da abordagem inferencialista das práticas, proposta por Brandom, uma espécie de “constitutivismo de práticas inferencialista”, oferece uma resposta promissora para a elucidação das características distintivas da normatividade imperiosa da moral.

No capítulo “Moralidade como solo normativo comum: primeiros passos de uma teoria”, Vitor Sommariva reflete sobre o conceito geral de moralidade. A obra de autores que concebem a moralidade como um fenômeno essencialmente moderno é discutida, seja daqueles que problematizam a moral moderna, entre outras razões, por sua conexão com a tradição cristã (e.g., Anscombe e Williams), seja daqueles que adotam o projeto moral moderno propondo alguns refinamentos (e.g., Darwall). Considerando a obra desses autores, mas julgando suas abordagens em última instância insatisfatórias, Sommariva esboça uma concepção alternativa de moral que simultaneamente leva a sério a dimensão de segunda

pessoa destacada por Darwall, mas evita seu compromisso conceitual com o universalismo, e reconhece os problemas da tradição moral moderna tal como foi criticada por Anscombe e Williams, mas sem fazer uma associação necessária entre o fenômeno moral e esse entendimento moderno. O resultado é uma concepção de moralidade livremente inspirada na ideia de solo comum, introduzida por Stalnaker. Em linhas gerais, a proposta consiste em argumentar que a moralidade é composta por um conjunto de pressuposições e expectativas que agentes morais em interação possuem uns sobre os valores e práticas dos outros. A moralidade, nessa visão, constitui um solo normativo comum, que fundamenta e justifica as práticas de responsabilização moral.

Um desafio para a abordagem é dar conta dos diversos cenários em que não há um pleno compartilhamento de valores e normas entre os agentes ou quando as expectativas e pressuposições desses mesmos agentes são divergentes. A parte final do capítulo busca elucidar a posição com recurso a alguns expedientes kantianos manifestos em partes da obra não moral de Kant. Dito brevemente, argumenta-se que algo como a comunidade moral de todos os agentes racionais vislumbrada por Kant é uma possibilidade limite, mas que agentes epistêmicos racionais ajustam suas expectativas de compartilhamento de valores e normas à luz de informações que obtêm sobre o contexto e as práticas e discursos dos seus interlocutores. Com isso, frequentemente será o caso que comunidades morais serão mais restritas que o ideal kantiano e, conseqüentemente, que o discurso moral terá força sobre um conjunto mais delimitado de interlocutores do que a totalidade dos seres racionais.

Avançando à Parte 3, “Motivação moral e normatividade”, Adelino Ferreira, Claudia Passos-Ferreira e Flavio Williges fornecem três variações em torno do multifacetado tema da natureza da motivação para agir e sua relação com a normatividade moral. Enquanto Ferreira está interessado em responder especificamente à questão sobre a racionalidade da moralidade, Passos-Ferreira e Williges, a partir de perguntas distintas, exploram o papel de sentimentos e emoções na sua relação com juízos morais.

Em “Razões no realismo moral naturalista: uma interpretação alternativa”, Adelino Ferreira defende uma releitura do modo como teorias naturalistas, como a dos realistas de Cornell, lidam com o problema do internismo vs. externismo. Os realistas de Cornell sustentam que propriedades morais dependem de propriedades naturais, embora não sejam redutíveis a estas. Realistas naturalistas não reducionistas precisam oferecer uma explicação para o problema da relação entre a dimensão prática da moralidade, a conexão entre a moralidade e as razões e disposições e atitudes do agente para agir; desafio que alimentou o clássico debate entre internismo e externismo na metaética. Realistas não reducionistas, como David Brink e Richard Boyd, são considerados externistas de motivação, isto é, eles negam que haja uma conexão necessária entre um julgamento moral sincero e estar motivado a agir de acordo com ele; a motivação apropriada dependeria de outros fatores adicionais e contingentes. Eles são considerados, também, externistas de razão, isto é, eles entendem que razões morais são objetivamente válidas, e independentes de estados mentais dos agentes. Contrariando a leitura tradicional que identifica Brink e os realistas de Cornell como defensores de uma versão forte do externismo, Ferreira quer mostrar que, no que diz respeito à relação entre razões para agir e a moralidade, uma posição como a de Brink

pode ser compatível com uma posição intermediária, um internismo (ou externismo) híbrido.

Como aponta Ferreira, determinar se as demandas morais fornecem ou não razões exige considerar questões substantivas e não questões meramente formais. O argumento de Brink é o de que a racionalidade das razões morais não pode ser derivada *a priori* nem depender exclusivamente de estados subjetivos. O ponto levantado por Ferreira é o de que Brink não afasta de forma decisiva a tese internista de que “a moralidade fornece razões para ação de forma necessária” e que “o cenário final de sua descrição do fenômeno moral inclui ao menos uma concessão a uma tese menos forte de internismo, ou, como chamaremos aqui, uma tese híbrida sobre razões e moralidade”. Então, poderíamos falar de uma relação necessária entre moralidade e razões para ação, mas tal relação só poderia ser especificada *a posteriori*. De acordo com Ferreira, uma investigação empírica e uma compreensão adequada da nossa natureza e de como o mundo funciona poderia dar conta de oferecer um tipo de razão justificatória para o agente, como as do egoísmo racional, ou seja, aquelas que seriam razões para sujeitos mesmo que eles não as reconhecessem. Se o hibridismo for plausível, as razões seguiriam sendo razões para o agente ainda que ele não as reconhecesse ou as negasse e o amoralista seria irracional, não por não se sentir motivado a agir, mas por não ser sensível a razões. Nesse sentido, Brink seguiria sendo fortemente externista de motivação, no entanto, quanto às razões, a sua posição seria compatível com uma versão enfraquecida do internismo.

Ferreira reconhece que o ponto desafiador é especificar de onde vem o “necessário” da relação entre moralidade e razões justificadoras. Ele sustenta que uma teoria centrada em propriedades de

florescimento, como a proposta por outro realista de Cornell, Richard Boyd, poderia fornecer um caminho; embora não sem desafios, a começar pelo componente teleológico da teoria. Desse modo, a despeito de a conexão entre moralidade e motivação ser defendida como uma conexão não necessária, à luz do realismo naturalista de Brink ainda é plausível considerar razões para agir necessárias no domínio moral.

No capítulo, “A imaginação e o agir moral”, Claudia Passos-Ferreira leva adiante a tradição sentimentalista, com ênfase particular no papel para a moralidade de imaginar os sentimentos dos outros. Com recurso tanto a estudos científicos quanto à argumentação filosófica, a autora enfrenta o desafio posto ao sentimentalismo por casos como o de pessoas com autismo e crianças pequenas. Para a tradição sentimentalista em questão, as capacidades de imaginação afetiva e de adoção da perspectiva do outro são essenciais para a produção de julgamentos morais genuínos. Autores como Heide Maibom e Shaun Nichols argumentam que portadores de autismo e crianças pequenas são capazes de desaprovação moral genuína, mas ainda não dispõem de ferramentas para a imaginação afetiva e a adoção da perspectiva do outro. Além desses casos, também as reações viscerais em contextos morais parecem pôr em xeque a tese sentimentalista de que a imaginação e a adoção de perspectivas são essenciais para a moralização.

Em resposta a esses desafios, Passos-Ferreira aceita a solução sentimentalista histórica proposta por Antti Kauppinen, segundo a qual as atitudes morais de portadores de autismo e as reações viscerais com conteúdo moral dependem assimetricamente de processos que envolvem a imaginação e a adoção de perspectiva e, portanto, podem ser explicadas em termos sentimentalistas. Ela propõe uma explicação

alternativa sentimentalista para o caso das crianças pequenas. Alguns estudos recentes apontam para a capacidade de crianças adotarem a perspectiva de vítimas e manifestarem preocupação. A autora argumenta que essa capacidade pode ser vista como precursora do processo sentimentalista de desaprovação moral com recurso à imaginação afetiva da perspectiva do outro. Dessa maneira, combinando a abordagem de Kauppinen com elementos originais, Passos-Ferreira oferece uma defesa robusta do sentimentalismo diante dos desafios da literatura filosófica e científica recente.

Em “A moralidade da ansiedade: problemas e perspectivas”, Flavio Williges explora o tema da normatividade moral da perspectiva de sua relação com as emoções, em particular, a da ansiedade. A abordagem se insere na corrente recente em filosofia das emoções e em psicologia moral que propõe reavaliar o papel tradicionalmente conferido às emoções ditas “negativas”, tais como medo, raiva, desprezo e ansiedade. Alguns autores contemporâneos defendem a tese de que emoções negativas merecem ser cultivadas, seja por possuírem algum tipo de valor instrumental, seja mesmo por possuírem valor intrínseco. Assim, elas poderiam passar a ser reconhecidas como desempenhando um papel central no desenvolvimento ou exercício de virtudes morais. Se a raiva for capaz de deflagrar pensamentos, por exemplo, relacionados a exigências de justiça ou for um tipo de reação emocional concomitante a juízos morais de justiça, tais autores argumentam que ela poderia desempenhar um papel epistêmico instrumental no pensamento moral ou mesmo ser parcialmente constitutiva de certos juízos morais. Se essa tese for plausível, ela implica que há um padrão normativo associado a tais emoções negativas: teríamos razões para cultivá-las no contexto moral. A tese que Williges discute consiste no

mesmo tipo de padrão teórico explicativo aplicado à emoção da ansiedade.

Concentrando-se na tese defendida por Charlie Kurth, Williges a apresenta criticamente ressaltando tanto seus aspectos positivos quanto suas limitações. A ideia básica é que a ansiedade, como uma emoção que envolve a percepção de uma ameaça a algo de importância para o sujeito, possuiria um valor epistêmico instrumental por ser capaz de auxiliar na identificação de uma situação moral relevante e no processo de deliberação moral. Além disso, a ansiedade poderia mesmo ser parte constitutiva da virtude da preocupação moral e, assim, possuir valor intrínseco. Williges concorda com críticos de Kurth que a emoção da ansiedade não parece possuir valor moral intrínseco por ser excessivamente instável e difusa. Além disso, seu valor instrumental também parece depender de recursos interpretativos aplicados ao contexto particular da situação moral relevante. A despeito dessa “ambiguidade moral”, Williges defende, por outro lado, que a emoção da ansiedade pode, em certos contextos específicos, desempenhar um papel central no comportamento moral. Isso ocorre, segundo o autor, especificamente sob a pressuposição de uma ética do cuidado. A ansiedade parece ser parte do tipo de reação característica da dimensão relacional da atitude de cuidado, na medida em que uma ameaça ao bem de alguém que nos importa é internalizada como uma forma de ameaça a nós mesmos.

Por fim, a última parte do livro, “Justificação moral e normatividade”, traz as contribuições de Leonardo de Mello Ribeiro e Matheus de Mesquita Silveira, ambos interessados em responder sobretudo a questões de epistemologia moral. Enquanto o primeiro questiona a plausibilidade do conceito de competência ética e explora

as implicações epistemológicas disso, o segundo investiga relações possíveis entre explicação e justificação moral, tendo como foco explicações evolucionárias do fenômeno da moralidade.

Em “Quem tem competência ética?”, Leonardo de Mello Ribeiro pergunta em que sentido é possível falar de competência, perícia e autoridade éticas. Traçando uma comparação com o discurso científico, Ribeiro argumenta que falar de competência, perícia e autoridade nas ciências empíricas (e mesmo na matemática) é algo rotineiro em nossas práticas epistêmicas. Além disso, entendemos como justificadas nossas crenças formadas com base no testemunho de autoridades competentes naquelas áreas. Mas, na ética, ao contrário, parecemos resistir à ideia de que estejamos justificados em adquirir crenças por testemunho de autoridades competentes. Ribeiro esclarece que sua discussão não é propriamente sobre o papel do testemunho como fonte de justificação epistêmica, mas sobre uma questão mais básica: como identificar competência, perícia e autoridade na ética. A resposta de Ribeiro a essa questão é, em grande medida, cética.

Ribeiro defende que, nas ciências, possuímos critérios objetivos e independentes (como capacidade de previsão e métodos compartilhados de testagem de hipóteses) que não parecem estar disponíveis no domínio da ética. Na ética, enfrentamos profundas discordâncias sobre o que conta como evidência e mesmo como método adequado de justificação de crenças. Parece ser inescapável, nesse domínio, julgar uma crença como correta (ou incorreta) a partir de nossas próprias perspectivas éticas, dos próprios padrões normativos ou valorativos que endossamos. Ribeiro alega que não somos capazes de isolar (totalmente) uma crença ética e testá-la através de critérios inteiramente objetivos. Uma vez que, na ética,

discordâncias não parecem ser passíveis de superação em nível substancial, padecemos, nesse domínio, da dificuldade de validar algum senso de objetividade que justifique a noção de conhecimento aplicada à ética. O autor ilustra essa tese discutindo três possíveis respostas ao problema, todas problemáticas: a tese do progresso ético, o intuicionismo e o equilíbrio reflexivo. Para ele, nenhuma dessas respostas é capaz de superar o desafio cético posto. Ribeiro, então, encerra sua discussão alegando que, se há algum sentido possível de identificação de critérios de competência, perícia e autoridade éticas, esse sentido deve inescapavelmente fazer referência a uma perspectiva *local* e *contingente* de uma “cultura ética”. Inspirado em Bernard Williams, Ribeiro argumenta que, se uma cultura ética comum for capaz de produzir alguma convergência substancial de valores, práticas e crenças éticas, critérios de identificação de competência, perícia e autoridade éticas internos a tal cultura poderiam estar disponíveis. Esse seria um sentido “deflacionado” de tais critérios, que poderia justificar, em algum nível, a formação de certas crenças éticas por testemunho dentro de uma cultura ética. A despeito disso, o autor termina defendendo que, mesmo que haja tal sentido deflacionado, o tipo de “conhecimento ético” que está pressuposto em sua base deve, ainda assim, ser distinto do tipo de conhecimento que associamos às ciências. Isso porque, além do caráter local do “conhecimento ético”, existe uma dimensão insuperável de contestabilidade no universo de discurso da ética.

No capítulo “O problema da explicação e da justificação nas teorias evolucionárias da moralidade”, Matheus de Mesquita Silveira investiga as pretensões de justificação normativa da moralidade a partir das teorias biológicas evolucionárias, em especial, aquelas de origem darwinista que ganharam proeminência ao longo do século XX

e servem de palco para um acalorado debate contemporâneo. Como Silveira esclarece, há duas perguntas a serem respondidas sobre a moralidade com base em uma metodologia informada pelo discurso científico da biologia evolucionária: (i) de onde vem a moralidade?; (ii) como fatos relacionados à evolução podem conferir normatividade à moralidade? Silveira argumenta que essas são questões de natureza distinta, mas que são facilmente confundidas na literatura (e mesmo em outras discussões filosóficas). A primeira pergunta é de natureza explicativa, enquanto a segunda é de natureza justificatória. Confundir essas duas questões pode tanto levar a equívocos interpretativos sobre a contribuição da teoria evolucionária para a compreensão do fenômeno moral quanto, na pior das hipóteses, levar a uma argumentação falaciosa. Assim, Silveira sugere que a teoria evolucionária pode contribuir para a explicação do surgimento do fenômeno moral, ainda que seja pouco plausível que ela seja capaz de justificar diretamente crenças morais e o conteúdo da moralidade tal como endossado atualmente por seres humanos.

Silveira dedica boas páginas para convencer o leitor da necessidade de esclarecimento metodológico sobre o que está em jogo em uma investigação evolucionária sobre a moralidade. Parte central do argumento do autor depende, assim, de uma compreensão precisa da diferença entre explicação e justificação. Embora Silveira não exclua a possibilidade de que explicações possam se tornar justificações, ele alega que, para tal, a mera explicação de um fenômeno natural (em termos científicos) não é suficiente. É necessária a inclusão de alguma premissa adicional que seja capaz de conferir um papel justificatório a uma explicação. No contexto da moralidade, Silveira alega que esse ponto é ainda mais básico do que aquele salientado por Moore (1903) com seu argumento da falácia

naturalista. A dificuldade central para uma “ética evolucionária” não é tanto em ser capaz de mostrar como definições naturalistas de termos morais podem ser fornecidas (o ponto de Moore) quanto em ser capaz de mostrar como uma explicação científica de um fenômeno pode também assumir um caráter normativo de justificação –no caso, de crenças morais e práticas humanas. Silveira ilustra o ponto através de uma reflexão crítica sobre a tradição do método genealógico de explicação e justificação de crenças morais, dado que explicações evolucionárias da moralidade podem ser interpretadas como fazendo uso desse mesmo método. Apoiando-se em parte na literatura contemporânea nietzschiana, o autor propõe que a tentativa de justificar (como verdadeira ou falsa) uma crença moral diretamente através de uma história genealógica de seu surgimento tem como resultado incorrer em uma forma de falácia genética. Por outro lado, se o método genealógico for utilizado apenas com propósitos explicativos, a teoria evolucionária pode fornecer contribuições relevantes para explicar o surgimento da moralidade e, indiretamente, questionar a confiabilidade de certas crenças morais.

Ao oferecer esse conjunto original de textos, esperamos que tanto o/a leitor/a já familiarizado/a com a temática quanto o/a iniciante possa se beneficiar da pluralidade de abordagens, da riqueza de detalhes e do percurso por uma extensa e atualizada bibliografia. Almejamos também ter conseguido demonstrar ao longo dessas páginas a relevância do tema e por que ele merece ser estudado por todos aqueles que tenham interesse em compreender o fenômeno moral e suas implicações em outros domínios afins teóricos e práticos. Por fim, esperamos que os textos aqui reunidos possam contribuir

para fomentar o debate e ampliar o desenvolvimento de pesquisas nessa temática no Brasil.¹¹

Referências

- Broome, J. (1999) Normative Requirements. *Ratio*, v. XII, n. 4, pp. 398-419.
- Broome, J. (2020) Rationality Versus Normativity. *Australasian Philosophical Review*, v. 4, n. 4, pp. 293-311.
- Copp, D. (1995) *Morality, Normativity, and Society*. Oxford: Oxford University Press.
- Copp, D. (2024) *Ethical Naturalism and the Problem of Normativity*. Oxford: Oxford University Press.
- DePaul, M. (2006) Intuitions in Moral Inquiry. In: Copp, D. (Org.) *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, pp. 595-623.
- Finlay, S. (2010) Recent Work on Normativity. *Analysis*, v. 70, n. 2, pp. 331-346.
- Frankena, W. F. ([1942] 1976) Obligation and Value in the Ethics of G. E. Moore. In: Goodpaster, K. E. (Org.) *Perspectives on Morality: Essays by William F. Frankena*. Notre Dame, Indiana: University of Notre Dame Press, pp. 12-23.
- Hare, R. M. (1952) *The Language of Morals*. Oxford: Clarendon Press.
- Hare, R. M. (1981) *Moral Thinking*. Oxford: Oxford University Press.

¹¹ Este livro é resultado do projeto de pesquisa Realismo e Antirrealismo: Integrando as Evidências mais Recentes ao Debate acerca da Normatividade Moral (vigente de 2022 a 2025), que contou com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, obtido através da Chamada CNPq/MCTI/FNDCT Nº 18/2021 – UNIVERSAL, processo Nº 422774/2021-9.

- Larmore, C. (2021) *Morality and Metaphysics*. Cambridge: Cambridge University Press.
- Mackie, J. (1977) *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin Books.
- Moore, G. E. (1903) *Principia Ethica*. Cambridge: Cambridge University Press.
- Ross, W. D. (1930) *The Right and the Good*. Oxford: Clarendon Press.
- Scanlon, T. M. (1998) *What We Owe to Each Other*. Cambridge, Mass.: Harvard University Press.
- Scanlon, T. M. (2014) *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Schroeder, M. (2007) *Slaves of the Passions*. Oxford: Oxford University Press.
- Shafer-Landau, R. (2003) *Moral Realism: A Defence*. Oxford: Oxford University Press.
- Thomson, J. J. (2008) *Normativity*. Chicago, Illinois: Open Court Publishing.

Parte 1

Teorias morais e normatividade

A convergência ética normativa da humanidade¹

Cinara Nahra²

Deontologia e utilitarismo são frequentemente vistos como teorias opostas e, de fato, divergem em relação a várias questões e problemas em ética normativa, prescrevendo diferentes cursos de ação e levando muitas vezes a resultados distintos e divergentes no campo da ética aplicada. No entanto, quando olhamos atentamente para ambas, provavelmente ficaríamos surpresos ao perceber que as duas teorias convergem em muitos pontos, formando um pequeno, mas muito importante subconjunto, que é comum ao imperativo categórico, de Kant, ao princípio do dano (*harm principle*) e ao princípio da maior felicidade, de Mill, bem como à regra de ouro (que esteve presente em muitas tradições culturais e até religiosas

¹ O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES) – Código de Financiamento 001. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

² Professora do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal do Rio Grande do Norte. Bolsista de produtividade do CNPq.

diferentes ao longo das eras). Que subconjunto comum seria esse? Sugiro que ele é constituído por 10 subprincípios que devem formar ou fazer parte do núcleo básico da moralidade humana. Sugiro também que a grande maioria dos seres humanos, independentemente de sua cultura, posição social, raça ou gênero, os reconheceria como subprincípios que deveríamos seguir, embora nem todas as pessoas no mundo efetivamente os sigam. O principal objetivo aqui é mostrar que esses subprincípios propostos são, na verdade, guias normativos práticos para a humanidade, e desviar-se deles pode levar a humanidade a um caminho que pode levar ao colapso moral da nossa civilização, ou até mesmo ao fim da espécie humana e/ou o fim da humanidade em nós.

1. De volta aos princípios básicos: a convergência ética normativa da humanidade

Todos nós sabemos que a humanidade é diversa. Somos agora 8 bilhões de pessoas distribuídas em pouco mais de 200 países. Em todo o mundo, há muitas culturas, muitas línguas e dialetos, hábitos e costumes diferentes. Com toda essa diversidade, é possível encontrar algum tipo de convergência moral, algum ponto em comum entre nós? Acredito que é factível encontrar uma convergência ética normativa mínima no diálogo de preceitos utilitários e deontológicos, gerando um subconjunto de princípios (subprincípios) que todos nós devemos respeitar e que, de fato, a grande maioria de nós, independentemente de nossas diferentes culturas e crenças, já os reconhece. Esses subprincípios são: a) prevenir a extinção da humanidade e o colapso da civilização; b) prevenir a extinção da vida na Terra e no universo; c) prevenir a destruição do planeta; d) abster-se de crueldade para com outros humanos e outros seres vivos; e)

abster-se de discriminação com base em raça, sexo, gênero, orientação sexual ou diferenças econômicas, sociais ou religiosas; f) não prejudicar e não matar pessoas inocentes; g) respeitar os outros; h) ser verdadeiro e honesto (no sentido mais amplo, especialmente evitando engano, manipulação e falsidade); i) contribuir para a felicidade dos outros sempre que possível; j) ajudar a minimizar o sofrimento na Terra e no universo. Argumento aqui que esses subprincípios são comuns ao utilitarismo e à deontologia e fazem parte de um subconjunto de princípios que formam um mínimo ético normativo comum da humanidade.

2. Por que devemos evitar a extinção da humanidade, o colapso da civilização, a destruição do planeta e a extinção da vida na Terra e no universo

Se há algo que pode ser dito sobre os humanos é que, com raríssimas exceções, nós queremos viver. Até agora, ao longo da História, a maioria absoluta dos humanos (passados ou presentes) queria, ou ainda quer, continuar existindo neste planeta que tem sido nosso lar por milhares de anos. Diante da inevitabilidade de nossa morte (já que ainda somos mortais), não queremos que o mundo se acabe quando morrermos.³ Queremos que nossos semelhantes vivam; a ideia de que o mundo pode acabar é aterrorizante para nós e tem sido para todos os nossos predecessores.

³ Aqui não importa se o mundo é real ou é apenas uma simulação, o fato é que as pessoas querem permanecer nele. Não importa também se elas ainda estão vivas ou já morreram. O fundamental é que a absoluta maioria das pessoas, mesmo as que já morreram, enquanto estavam vivas não queriam morrer e não queriam que todos no mundo morressem junto com elas, naquilo que se convencionou chamar de “fim do mundo”. Há aqui um compromisso importante de todos nós com a continuidade da nossa espécie e do mundo em que vivemos.

Queremos evitar a extinção da humanidade, e o senso comum aqui coincide com a filosofia por duas razões primordiais. A principal razão utilitarista contra a extinção é que, mesmo que todos no mundo estejam sofrendo terrivelmente, o fato de que ainda estejamos vivos nos dá esperança de que o sofrimento possa um dia acabar e que possamos ser felizes, o que nunca poderia acontecer se todos morressem, já que a precondição para alguém ser feliz é estar vivo. Se quisermos maximizar a felicidade, também precisamos preservar a vida, porque não há felicidade sem vida.⁴

A principal razão deontológica kantiana contra a extinção é que a vida, especialmente a vida racional, tem que ser preservada para que um universo moral seja criado. O bem supremo para Kant é construir um mundo no qual a moralidade seja maximizada e, idealmente, todos ajam moralmente. Os humanos, como seres racionais, são aqueles que podem fazer com que tal mundo ideal venha a existir na Terra. Se a humanidade for extinta, não há possibilidade de que esse mundo moral seja construído e não há mais esperança de que o

⁴ Observe que Crisp (2021) escreve que “A questão de se a extinção seria boa ou ruim no geral é obviamente muito importante, especialmente diante de potenciais eventos catastróficos na virada da história. Mas esse questionamento também é muito difícil de responder. Em última análise, não estou afirmando que a extinção seria boa, apenas que, já que pode ser, devemos dedicar muito mais atenção a pensar sobre o seu valor do que temos feito até agora.” Neste capítulo, estou levando a sério a preocupação de Crisp sobre o valor da extinção e fornecendo uma resposta à sua pergunta. Minha resposta é a de que a extinção é ruim por razões utilitárias e deontológicas. A sua malignidade gera um dos subprincípios mais importantes que a humanidade deve seguir: prevenir a extinção da humanidade. Aqui não desenvolvo a discussão proposta por Crisp sobre o possível valor da extinção, mas farei isso em outro momento, examinando cenários nos quais a humanidade está sob sofrimento excruciante ou se torna “diabólica” e sugerindo que, mesmo nesses casos, há boas razões para evitar a extinção.

mundo perfeito, ou seja aquele no qual os que agem moralmente sejam felizes também, seja criado.⁵

Combinando então ambos, deontologia e utilitarismo, vemos que evitar a extinção da humanidade é provavelmente uma das convergências mais importantes das duas teorias. Mas seria apenas a extinção que queremos evitar? E quanto aos cenários pós-Armagedom, em que muito poucas pessoas sobrevivem e, ainda assim, em condições muito precárias, com toda a tecnologia e cultura construídas em milhares de anos sendo destruídas? O que dizer de um mundo no qual todo o conhecimento desapareceu, em que os livros foram todos queimados, toda a eletricidade acabou, a internet e os telefones não mais funcionam, no qual não há mais água potável nos canos e nenhum meio de transporte além do uso puro e simples de nossas próprias pernas? Novamente, a maioria absoluta de nós certamente não gostaria de viver em tal cenário.⁶

Não queremos que nossa civilização entre em colapso assim como também não queremos que nosso lar seja destruído. No

⁵ Em Kant, o bem supremo tem dois significados: o bem supremo como supremo e como perfeito. No primeiro sentido (supremo) podemos, como espécie, nos aprimorar, agindo motivados pelo imperativo categórico. No segundo sentido do bem supremo (*consummatum* ou *perfectissimum*), ele requer que a moral seja seguida pela felicidade, ainda que esta seja algo que só podemos esperar porque a felicidade é algo que não depende da nossa vontade. Sobre isto, ver Nahra, 2013.

⁶ Também não gostaríamos de viver em um mundo onde há uma guerra (existente ou potencial) de todos os homens contra todos os homens, o que Hobbes (1996) descreveu como um estado de natureza, uma condição na qual não há lugar para a indústria porque o fruto dela é incerto, nem nenhuma cultura da terra, nenhuma navegação, nenhuma construção, nenhuma arte, nenhum letramento, nenhuma sociedade, e o que é pior de tudo, há o medo contínuo e o perigo de morte violenta, e a vida do homem, solitária, pobre, desagradável, brutal e curta. Estamos aqui, então, diante de outro consenso: nós, humanos, não queremos viver em tal estado, em um mundo que não seja civilizado.

momento, nós, humanos, ainda estamos vivendo na casa onde nós, *Homo sapiens*, nascemos, este lindo planeta azul chamado Terra. Provavelmente no futuro, se a humanidade sobreviver, poderemos nos mudar para outros planetas similares ao redor do mundo. Neste preciso momento do tempo, no entanto, o conhecimento e a tecnologia necessários para isso não estão disponíveis, e isso nos daria, pelo menos, uma razão utilitarista para evitar, por todos os meios, a destruição de nosso planeta, nossa única casa. Se não formos bem-sucedidos nessa tarefa, não há escapatória, morreremos no dia em que o planeta for destruído, e, junto a nós, toda a nossa civilização e todas as espécies que coabitam conosco. Isso, além de nos dar uma forte razão para não destruir nosso planeta, levanta outra questão crucial: importa-nos que toda a vida na Terra e no universo seja extinta?

Se toda a vida no universo for destruída, isso obviamente nos incluiria, o que é outra razão utilitarista que nos sugere que a extinção da vida no universo diz respeito a nós também. Mas e quanto a destruir toda a vida no universo, exceto a vida humana? Seria isso moralmente aceitável? Se aplicarmos a regra de ouro em sua forma negativa, a saber, “Não faça aos outros o que você não gostaria que fizessem a você”, ficaria imediatamente claro que, se não queremos que nossa espécie seja extinta, não deveríamos querer também a extinção de nenhuma outra. A exceção a isso é quando outra espécie ameaça nossa existência. Nesse caso teríamos razões utilitaristas para destruir aquela que nos ameaça (o que está em questão aqui são casos como o de vírus, por exemplo, no qual há controvérsia sobre se estes são ou não seres vivos, mas, mesmo supondo que sejam uma forma básica de vida, certamente ninguém em sã consciência argumentaria que não

deveríamos aniquilar o coronavírus, por exemplo, por causa da regra de ouro).⁷

A conclusão aqui é a de que, se temos razões para impedir a extinção de outras formas de vida na Terra e no universo, temos então uma razão adicional para impedir a destruição do planeta, já que a Terra é não apenas nosso próprio lar, mas também o único lar dessa infinidade de espécies que ninguém quer que sejam extintas. Destruindo o planeta, estaríamos condenando à extinção não apenas a nós mesmos, mas também a todas essas espécies que vivem na mesma casa que nós.

3. Por que não devemos prejudicar ou matar pessoas inocentes e contribuir, sempre que possível, para a felicidade dos outros, e para minimizar o sofrimento no mundo

Agora podemos dizer que nós, humanos, queremos evitar a extinção da humanidade, que nosso amado planeta continue existindo e que a vida se expanda no universo. Mas, além disso, que tipo de humanos e humanidade almejamos? O utilitarismo já nos ensinou que o sofrimento deve ser minimizado e a felicidade maximizada. Um mundo desejável é certamente aquele em que há a maior felicidade e o menor sofrimento possível para o maior número.

⁷ Sidgwick (1981, p. 498), na chamada passagem de distinção, afirma “Estou preocupado com a qualidade da minha existência como indivíduo em um sentido, fundamentalmente importante, no qual eu não estou preocupado com a qualidade da existência de outros indivíduos.” Esta é conhecida como a pressuposição na qual o egoísmo em Sidgwick é baseado. Assumo aqui que, embora as consequências dessa pressuposição possam ser complicadas do ponto de vista individual, uma vez que podem servir para justificar o egoísmo além do que é razoável, elas são perfeitamente aceitáveis quando pensamos sobre a consideração especial que poderíamos dar à nossa espécie quando há ameaças externas à nossa sobrevivência.

O mundo ideal seria aquele em que todas as pessoas fossem felizes e ninguém sofresse. Então, extirpar o sofrimento e tudo o que o produz, como miséria, dor, morte e doenças, deve estar na agenda da humanidade. Mas, na prática, que compromissos isso implicaria?

O mais importante, acredito, é não matar pessoas inocentes. Quando alguém que vive sua vida em paz e não prejudica ninguém é assassinada por outro (outra pessoa, um grupo ou um país), é como se a humanidade em cada um de nós também estivesse sendo morta. Nesse ponto, o imperativo categórico de Kant, em sua fórmula de humanidade, poderia ser trazido à nossa atenção. Ninguém deve ser usado apenas como um meio para nossos fins, quaisquer que estes últimos sejam. Matar alguém que está prejudicando os outros (por exemplo, matar alguém que está torturando uma criança quando esta é a única maneira de impedi-lo de torturar) é permissível porque o torturador aqui é tudo menos inocente, e matá-lo é a única maneira de barrar a tortura. Mas acabar com a vida de alguém inocente e que não está prejudicando ninguém é tratar essa pessoa como uma mera coisa, não como um ser humano.

Esse compromisso deve trazer uma erradicação completa das guerras e atos terroristas em nossa civilização. Os atos terroristas, em que as pessoas são feridas e mortas, como forma de chamar a atenção para as causas políticas, sejam elas quais forem, devem ser abolidos no mundo. Os princípios da humanidade certamente excluiriam também as guerras como tivemos até agora na História, em que pessoas inocentes são alvejadas ou, mesmo que não sejam alvos, morrem

como consequência dos atos brutais da guerra.⁸ Matar e massacrar inocentes não é compatível com o processo civilizatório.

Causar dano a outras pessoas também é algo que nenhum ser humano, ou corporação, deveria fazer. Ninguém gosta de ser prejudicado, e aqui, mais uma vez, poderíamos aplicar a milenar regra de ouro “Não faça aos outros o que você não gostaria que fizessem a você”. Este é um requisito óbvio: se você não quer ser prejudicado, por que prejudicaria os outros? Existem, é claro, muitos mal-entendidos sobre a regra de ouro. Algumas pessoas afirmam, por exemplo, que, com base nessa regra, um juiz nunca deve condenar ninguém, pois não gostaria de ser condenado. Nesse caso, no entanto, o imperativo categórico resolve o aparente problema com o requisito de imparcialidade universal: “Se eu quero algo para todos, então devo querer para mim”, ou seja: “Se eu quero que todos que praticam certos crimes (assassinato, por exemplo) sejam punidos, eu deveria querer que eu mesmo (se eu praticar esse crime) fosse punido também.” De acordo com Kant, estamos sempre prontos para tentar permitir exceções na universalidade da lei a nosso favor, mas, quando olhamos do ponto de vista universal, vemos que isto é injustificável.

Podemos chegar à mesma conclusão (não devemos causar dano os outros) aplicando diretamente o princípio do dano (*harm principle*), de Mill, segundo o qual “O único propósito pelo qual o poder pode ser exercido corretamente sobre qualquer membro de uma comunidade civilizada contra sua vontade é o de evitar danos aos

⁸ Não discutirei aqui em maiores detalhes a “doutrina do duplo efeito”, mas me alinho àqueles que a contestam. Se sabemos que civis serão mortos em um bombardeio, ainda que não tenhamos a intenção de matá-los e sim apenas de atingir o alvo, é hipócrita alegar que a ação é permissível porque a “intenção” não é matar os civis. Se as consequências do ato são previsíveis, se sabemos que inocentes serão mortos, embora não seja a intenção, somos responsáveis por essas mortes.

outros". Que o dano aos outros deve ser evitado é uma exigência do utilitarismo que coincide com o bom senso. De alguma forma, as pessoas sabem que causar dano aos outros é errado, mas, por distintas razões, elas insistem nessa prática. Prevenir danos aos outros, no entanto, também é um requisito do princípio da utilidade, de Mill, que estabelece que a promoção da felicidade é o critério final para definir o que é moral e o que não é, ou seja, a sociedade moral ideal é aquela em que todos são felizes e todos estão livres do sofrimento. Se as pessoas (e as corporações) se abstivessem de causar danos aos outros, estariam contribuindo para minimizar o sofrimento no mundo e, assim, para alcançar o ideal utilitarista de eliminar forças objetivas de sofrimento como doenças, miséria e outras.⁹

A exigência de minimizar o sofrimento dos outros é apresentada por Kant de forma diferente. Kant menciona o dever indireto de benevolência que temos uns para com os outros. Não precisamos ajudar a todos da mesma forma, mas temos o dever de auxiliar os necessitados.¹⁰ Ajudar os necessitados, no entanto, é uma forma de minimizar o sofrimento no mundo. O ponto importante aqui é que, embora não tenhamos uma coincidência entre utilitarismo e deontologia sobre o quanto devemos continuar tentando assistir os

⁹ Mill (2001) enumera grandes fontes de sofrimento físico e mental na vida, como indigência, doença e a indelicadeza, inutilidade ou perda prematura de objetos de afeto.

¹⁰ A abordagem kantiana sobre o dever de benevolência não é muito bem compreendida pela recepção contemporânea de Kant, mas na verdade o que está em jogo aqui é sua distinção entre agir por causa do dever e agir de acordo com o dever. De acordo com o filósofo, atos de benevolência que são realizados tendo como motivação a lei moral são os únicos que têm valor moral. Isso não denota que ações benevolentes motivadas por interesses pessoais, por exemplo, não devam ser praticadas, significa apenas que elas não têm valor moral, são ações de acordo com o dever, mas não por dever.

outros, o quão longe devemos ir na ajuda, ou sobre as razões para ajudar, há, no entanto, a concordância de que os necessitados devem ser ajudados.

Mais uma vez, temos uma coincidência entre as exigências da deontologia e do utilitarismo: diminuir o sofrimento no mundo é um ideal da humanidade e a humanidade claramente estaria melhor se diminuíssemos o dano que os humanos causam uns aos outros e ao planeta. No que diz respeito à contribuição para a felicidade dos outros, se olharmos mais atentamente, há também uma coincidência entre a deontologia de Kant e o utilitarismo. Como sabemos, o utilitarismo requer a maximização da felicidade, o que não é o caso da deontologia de Kant, que não tem esse requisito. No entanto, Kant afirma claramente que a promoção da felicidade dos outros assim como o desenvolvimento da nossa própria perfeição são os dois fins que são ao mesmo tempo deveres.¹¹ Assim, não é o caso de haver um grande choque entre utilitarismo e deontologia neste ponto. De fato, Kant não propõe a maximização da felicidade, mas, se pensarmos em uma convergência ética mínima entre deontologia e utilitarismo, a promoção da felicidade dos outros está presente em ambas as teorias,

¹¹ Kant (2017, MS 6: 385-8) define “a própria perfeição” e a “felicidade dos outros” como os dois “fins que são ao mesmo tempo deveres”. Kant pergunta: quais são os fins que são ao mesmo tempo deveres? E responde: Esses fins são minha perfeição pessoal e a felicidade dos outros. Ele enfatiza que não podemos inverter isso e fazer da minha felicidade pessoal e da perfeição dos outros os fins em si mesmos que seriam deveres. Observe que esses dois fins, que são ao mesmo tempo deveres, relacionam-se o primeiro (minha perfeição pessoal) ao dever de desenvolver nossos talentos e o segundo (a felicidade dos outros) ao dever de ajudar os outros. Além disso, como observa Kleingeld (2016), eu promovo a felicidade dos outros, e outros promovem a minha. A felicidade de cada um é boa aos olhos dos outros e vice-versa. Isso significa que os agentes virtuosos neste mundo visam coletivamente a felicidade de todos.

sendo um requisito mínimo sobre o tratamento que devemos dispensar uns aos outros.

4. Por que devemos abster-nos de crueldade contra outros seres humanos e outros seres vivos e de discriminação com base em idade, raça, sexo, gênero, orientação sexual ou diferenças econômicas, sociais ou religiosas

Os utilitaristas, por muito tempo, têm sustentado, corretamente, que minimizar o sofrimento é um requisito para todos os seres sencientes (todos os seres que são capazes de sofrer dor) não apenas para os seres racionais. Então, se levamos a sério o cumprimento desse requisito, o dever óbvio que devemos ter para com os outros seres, além de evitar sua extinção e impedir a destruição de seu lar, o planeta Terra, certamente deve ser o de nos abstermos de agir cruelmente com outros seres vivos, especialmente aqueles que coabitam a Terra conosco, os animais. A crueldade é provavelmente a pior forma de sofrimento. Certamente não queremos isso para nós mesmos e, portanto, não devemos ser cruéis com os outros, incluindo diferentes espécies.

Como é amplamente conhecido, para Kant, temos deveres apenas para com os seres racionais, mas, por outro lado, ele também afirma que não devemos ser cruéis com os animais e sugere que alguns de nossos deveres para com os animais são indiretamente deveres perfeitos para conosco como animais e seres morais que somos. Segundo Kant, quando tratamos os animais sem amor, é como se estivéssemos diminuindo nossa humanidade, como se fôssemos

desprovidos de sentimentos, o que configura uma violação indireta da humanidade em nossa pessoa.¹²

Portanto, mesmo que a extensão e as fontes dos compromissos e obrigações que devemos ter para com os animais sejam distintas na deontologia de Kant e no utilitarismo, há um terreno comum no que diz respeito a nosso comportamento em relação a eles: a proibição absoluta de ser cruel com os animais, de fazê-los sofrer ou de maltratá-los ou até mesmo torturá-los.

Também não devemos ser cruéis com nossos semelhantes e, novamente, essa proibição é comum entre a deontologia e o utilitarismo. Em Kant, essa proibição é clara quando olhamos para a fórmula da humanidade “Trate a humanidade, seja em sua própria pessoa ou na pessoa de qualquer outro, sempre ao mesmo tempo como um fim, nunca apenas como um meio”. No utilitarismo, uma vez que a crueldade é uma das piores formas de infligir sofrimento, ela certamente também não é permitida.

A crueldade pode ser praticada de várias formas e pode envolver (e frequentemente envolve) a imposição de sofrimento mental. Práticas como bullying, racismo, misoginia, homofobia, ageísmo (preconceito de idade), aporofobia (preconceito com os pobres), gordofobia (preconceito com os obesos), capacitismo (preconceito contra pessoas com deficiência) e preconceito contra os indivíduos por causa de suas crenças religiosas (quando são inofensivas) também

¹² De acordo com Denis (2000, p. 417), “Nossa consideração do relato de Kant sobre nossas relações com os animais mostrou que, embora a teoria moral de Kant não tenha deveres para com os animais, os deveres em relação aos animais são bem fundamentados e são deveres genuínos. Existem analogias, baseadas em nossa animalidade compartilhada, entre seres humanos racionais e seres animais não racionais. Por causa dessas analogias, as maneiras como tratamos os animais refletem e afetam atitudes e sentimentos moralmente importantes.”

são formas de crueldade porque fazem os sujeitos sofrerem mentalmente (e frequentemente fisicamente também) de uma forma que poderia ser facilmente evitável, se as pessoas parassem de agir de modo preconceituoso ou discriminatório. Do ponto de vista utilitarista, rebaixar ou discriminar os indivíduos por causa de sua idade, cor, sexo etc. é prejudicial para todas as pessoas que são discriminadas ou tratadas de modo preconceituoso, e isso pode lhes causar enormes danos mentais. Se a discriminação se tornar física, o dano é ainda mais evidente, e o utilitarismo condenaria todos esses comportamentos.

Do ponto de vista kantiano, nenhum desses comportamentos discriminatórios resiste à fórmula da humanidade.¹³ Se temos que tratar as pessoas como fins em si mesmas, temos que tratar todos dessa maneira, independentemente de suas diferenças, e, quando discriminamos e temos preconceito em relação aos indivíduos por causa dessas características, estamos rebaixando-os em sua humanidade. As pessoas têm dignidade e a dignidade não tem preço. A dignidade é um valor incondicional e incomparável. Quando as discriminamos, estamos tratando-as como coisas, não como pessoas, não como seres racionais. Sob todos os aspectos não as estamos respeitando.

Aqui, novamente, podemos encontrar um terreno ético comum mínimo entre a deontologia kantiana e o utilitarismo. Não devemos ser cruéis com os animais, não devemos ser cruéis com as pessoas, e

¹³ Mesmo que o próprio Kant tivesse algumas visões controversas sobre homossexualidade, mulheres e raça, provavelmente porque, como todos os filósofos, somos no final das contas, infelizmente, influenciados pelos preconceitos da época em que vivemos, o importante a ser notado é que, quando aplicamos a fórmula da humanidade do imperativo categórico, a discriminação e o preconceito não se sustentam porque implicam tratar alguns grupos de humanos não como um fim.

isso tem muitas implicações do ponto de vista do tratamento que devemos, e especialmente não devemos, dar aos outros. A prescrição da regra de ouro neste caso também é clara: se não queremos ser tratados com crueldade, não devemos tratar os outros com crueldade, se não queremos ser discriminados, não devemos discriminar os outros.

5. Respeite os outros, seja verdadeiro e honesto (no sentido mais amplo, especialmente evitando engano, manipulação e falsidade)

O respeito, para Kant, é um sentimento, um sentimento único porque é produzido pela lei moral. É porque somos conscientes da lei moral que esse sentimento de respeito por ela é criado e nos move a agir moralmente. O respeito está, então, intrinsecamente ligado à moralidade, e não tem nenhuma relação com agir por nossos interesses pessoais, familiares ou de grupo, mas tem a ver com agir de acordo com o que pode ser universalizado. Respeito pelos outros significa tratá-los como fins em si mesmos e não como meios para nossos fins. Em outras palavras, significa tratar os seres humanos com dignidade e consideração.

Mill, por outro lado, não menciona a expressão respeito pelas pessoas.¹⁴ No entanto, quando olhamos atentamente para o princípio do dano (*harm principle*), fica claro que o respeito pelas pessoas é derivado deste, uma vez que a proibição de causar dano aos outros é provavelmente a principal e mais fundamental forma de respeitar os outros. Por outro lado, quando olhamos para o princípio da utilidade, a conclusão é a mesma: se quisermos minimizar o sofrimento, o respeito pelos outros segue-se imediatamente, pois, quando as pessoas

¹⁴ Ver Loizidis, 2017.

são desrespeitadas, há um dano mental sendo causado (e muitas formas de desrespeito incluem também danos físicos). Então, ambos os princípios, o princípio do dano e o princípio da utilidade, exigem respeito pelas pessoas. É claro que sempre poderíamos perguntar se, por outro lado, podemos sacrificar o respeito por alguém para alcançar a maior felicidade e o mínimo sofrimento para um número maior de indivíduos (por exemplo: podemos deixar algumas pessoas inocentes morrerem para salvar milhões de outras pessoas inocentes?), mas novamente esses raros casos são a exceção e não a regra e devem ser discutidos caso a caso. O ponto importante aqui é que o princípio da utilidade torna imprescindível o respeito pelos outros e o respeito pelos outros pode ser derivado do princípio da utilidade. Existe uma pequena zona cinzenta na qual o princípio utilitarista poderia colidir com a exigência de respeito por algumas pessoas, mas isso não tornaria a maioria absoluta dos casos menos clara.¹⁵

Para resumir, o respeito pelos outros está ligado à moralidade, à felicidade e à sobrevivência de nossa espécie e de nossa civilização.¹⁶

¹⁵ Na visão utilitarista, é discutível se, em certas condições, poderíamos abrir exceções a essa regra para minimizar o sofrimento e maximizar a felicidade, como se pudéssemos torturar (sob chantagem) uma pessoa inocente para salvar milhões de outras pessoas. No entanto, deixando de lado os casos limítrofes, a proibição de ser cruel com nossos semelhantes permanece no utilitarismo tanto quanto na deontologia.

¹⁶ Dias (1995) defende a ideia da Moral do Respeito Universal, proposta por Ernst Tugendhat com base na fórmula universal kantiana da humanidade. Respeito, em sua opinião, no entanto, seria reconhecer todo ser humano como tendo direitos. Para Dias, o respeito por cada indivíduo supõe reconhecer suas necessidades básicas, e ela posteriormente desenvolveu sua teoria em vários livros e artigos. Sou muito simpática à proposta de Dias, mas a ideia de respeito que adoto aqui é menos inclusiva do que a aventada por ela. A razão para isso é que o que estou tentando encontrar aqui são pontos de convergência entre utilitarismo e deontologia compondo um mínimo ético a ser acordado por todos, indivíduos, empresas e Estado.

Uma sociedade na qual as pessoas são totalmente egoístas e não se importam com, ou não respeitam, ninguém é uma sociedade que mais cedo ou mais tarde será destruída. Mais do que isso, uma sociedade em que não respeitamos os outros é uma sociedade que perdeu o sentido do humano e, se isso acontecer, nossa civilização estará à beira do abismo, o lugar no qual a moralidade e a humanidade colapsam.

Finalmente, o subprincípio que pede às pessoas que sejam verdadeiras e honestas é provavelmente um dos mais importantes para a sobrevivência da humanidade e de nossa civilização, hoje. Como disse Aristóteles (1986, 1011b25), definindo a verdade: “Dizer do que é que não é, ou do que não é que é, é dizer o falso, enquanto dizer do que é que é, e do que não é que não é, é dizer o verdadeiro.” Façamos agora o seguinte experimento de pensamento: imaginemos que todos afirmem o que é falso de uma forma que, a partir de agora, todos afirmem (e acreditem) que o que é não é, e que o que não é, é. De repente, perderíamos qualquer correspondência com a realidade. Todo o conhecimento acumulado da humanidade seria perdido, nossa História, nossa Ciência, nossa Filosofia e nossos valores teriam desaparecido. O que poderia resultar disso? Um mundo em que os seres humanos são substituídos por um bando de zumbis esquizofrênicos e perigosos que perderam toda a sua conexão com o mundo em que vivem e criaram sua própria realidade, que tem muito pouco ou nada a ver com a Civilização que nós levamos milhares de anos para construir.

Essa experiência de pensamento, que poderia ter parecido inverossímil no início do século, é, infelizmente, uma ameaça que agora enfrentamos nas redes sociais, dada a quantidade de informação falsa e desinformação proposital que aí tem sido propagada. A disseminação desses comportamentos e de suas consequências pode

ser catastrófica e, novamente, temos muito a aprender com os princípios utilitaristas e deontológicos. Kant alertou sobre os perigos de mentir e fazer falsas promessas, mostrando por que essas práticas não podem ser universalizadas: pois produzem uma contradição na Razão e até impossibilitam promessas, criando um mundo em que ninguém mais acredita em ninguém e no qual até mesmo coisas simples como contratos perdem a razão de ser, dado que sabidamente não serão cumpridos, destruindo as bases de nossa sociedade. Quanto ao utilitarismo, Mill, embora admita mentiras em certas ocasiões, para o bem maior (como salvar vidas, por exemplo), adverte, no entanto, que devemos ser muito cautelosos para evitar o enfraquecimento da confiança na veracidade.¹⁷

Por que a confiança na veracidade é tão importante para a humanidade? Porque a verdade é a verdadeira pedra filosofal. É da verdade que provém todo o conhecimento no e do mundo. A falsidade pode estar em toda parte, há muitas maneiras de estar em erro, mas apenas uma maneira de estar em correspondência com o mundo, ou seja, dizendo do que é que é, e dizendo do que não é que não é.¹⁸ Há uma primazia ontológica e epistemológica da verdade sobre a falsidade e a principal razão para isso é que buscar a verdade e alcançar a verdade é o principal fator que conduziu a humanidade a sobreviver e construir a civilização que criamos. O conhecimento entre gerações nunca teria sido transmitido se todos tivéssemos passado conhecimento falso e informações falsas para as gerações

¹⁷ Mill (2001) adverte que a exceção não deve se estender além da necessidade e deve ter o menor efeito possível no enfraquecimento da confiança na veracidade, devendo, se possível, ter seus limites definidos.

¹⁸ Por exemplo: no que diz respeito ao *Opus majus*, dizer que Roger Bacon escreveu esse livro é a única maneira de dizer a verdade, mas dizer que foi escrito por Hegel, John Lennon ou Madonna são apenas três das infinitas maneiras de dizer o falso.

futuras.¹⁹ Porém, se continuarmos a propagar massivamente a desinformação (tanto sem intenção quanto proposital, ou seja, *misinformation* e *disinformation*, respectivamente), é precisamente isso que acontecerá no futuro próximo.²⁰

Uma das principais razões pelas quais somos agora a espécie que somos e evoluímos para sermos racionais e até mesmo capazes de construir inteligência artificial é porque nós, humanos, como um todo, como espécie, temos sido, em geral, comprometidos com a verdade.

Se não nos apegarmos à veracidade e à honestidade e permitirmos que o oposto seja universalizado, aniquilaremos a razão, como alertou Kant, e estaremos à beira de destruir a moralidade e a humanidade, colapsando nossa civilização tal como a conhecemos.

6. Conclusões

Duas conclusões principais podem ser tiradas com base em tudo o que foi dito aqui. A primeira é que, mesmo que o utilitarismo e a deontologia possam prescrever uma extensão distinta de atitudes que devemos ter em relação aos outros (sejam elas quais forem), há uma clara zona de convergência entre ambos e entre a regra de ouro.

Deixando de lado os casos limítrofes e as diferenças na forma como os deontologistas e utilitaristas justificam suas perspectivas, há, para além das divergências em seus pontos de vista em muitos aspectos, uma clara convergência sobre práticas e atitudes que nós, indivíduos, corporações e Estados, nunca deveríamos adotar e

¹⁹ Veja, por exemplo, o gerenciamento do fogo. Se cada geração não tivesse passado para as outras gerações as informações corretas sobre como criar fogo, cada geração teria que aprender como fazê-lo a partir do zero, e nosso progresso tecnológico teria sido muito mais lento. Certamente, se fosse esse o caso, nossa civilização não seria o que é hoje.

²⁰ Ver Nagra, 2013.

algumas que devemos adotar sempre. Essa convergência pode ser resumida, inicialmente, nos 10 subprincípios aqui propostos. Esses 10 subprincípios deontoutilitaristas fazem parte do núcleo dos subprincípios morais da humanidade, derivados desse diálogo entre ambas as teorias e a milenar regra de ouro.

A segunda conclusão é que, quando olhamos para a forma como essas exigências se desenvolvem, estamos agora confrontados não só com obrigações individuais, mas também com as obrigações e deveres das empresas e do Estado. Em uma era em que o lucro e o interesse próprio imediato parecem ser colocados acima de tudo e de todos, esses subprincípios aparecem como guias para conduzir não apenas indivíduos, mas também gestores públicos, CEOs e acionistas de empresas em uma espécie de retorno aos fundamentos primeiros e aos consensos básicos no campo normativo. O mundo hoje é dominado por empresas privadas que detêm um poder colossal e que têm controle sobre muitos aspectos da vida dos cidadãos comuns. Esses cidadãos são completamente impotentes diante dessas empresas, perdendo sua autonomia e sendo obrigados a entregar seus dados pessoais a qualquer momento que acessem a internet. Esses cidadãos, eu, você que está lendo este artigo, todos nós, estamos sendo controlados e manipulados sobre as informações as quais teremos ou não acesso e quanto devemos pagar por qualquer serviço e produtos, já que os grandes oligopólios dominam o que deveria ser definido pela concorrência e pelo livre mercado ou por regulações estatais. Mas estamos, hoje, no pior dos mundos possíveis: não há nem a livre concorrência nem regulações do Estado. Estamos, literalmente, à mercê dos grandes oligopólios e das big tech.

Essas empresas e corporações deveriam ter obrigações e deveres a serem cumpridos, e muitos deles saltam aos olhos quando miramos

esses subprincípios da humanidade e os usamos efetivamente como guias para estabelecer políticas públicas, especialmente agora que a Inteligência Artificial começou a ser massivamente usada. Tanto no setor público quanto no setor privado, em todos os níveis de responsabilidade, os detentores de poder deveriam estar sujeitos a regulamentos, e esses regulamentos deveriam ser, em última instância, derivados de princípios normativos da humanidade que podem ser obtidos (é minha sugestão final aqui) a partir da convergência entre prescrições da deontologia, do utilitarismo e da milenar regra de ouro.

Referências

- Crisp, R. (2021) Would Extinction Be So Bad? Disponível em: <https://www.newstatesman.com/ideas/agora/2021/08/would-extinction-be-so-bad/>. Acesso em: 8 de outubro de 2014.
- Denis, L. (2000) Kant's Conception of Duties Regarding Animals: Reconstruction and Reconsideration. *History of Philosophy Quarterly*, v. 17, n. 4, pp. 405-423.
- Hobbes, T. *Leviathan* ([1651] 1996): *The Matter, the Form and the Power of a Commonwealth, Ecclesiastical and Civil*. 1. ed. Oxford: Oxford University Press.
- Kant, I. (2017) *The Metaphysics of Morals*. Trad. Mary Gregor. Cambridge: Cambridge University Press.
- Kleingeld, P. (2016) Kant on “Good”, the Good, and the Duty to Promote the Highest Good. In: Höwing, T. (Org.) *The Highest Good in Kant's Philosophy*. Berlin: De Gruyter, pp. 33-49.
- Loizidis, A. (2017) John Stuart Mill: Individuality, Dignity, and Respect for Persons. In: Giorgini, G.; Irrera, E. (Orgs.) *The Roots of Respect: A Historic-Philosophical Itinerary*. Berlin: De Gruyter, pp. 187-206.

Mill, J. S. (2001) *Utilitarianism*. Oxford: Oxford University Press.

Nahra, C. (2013) The Deontoutilitarianist Manifest. *Ethic@*, v. 12, pp. 169-192.

Sidgwick, H. (1981) *The Methods of Ethics*. Indianapolis: Hackett Publishing Company.

Repensando a teoria *normativa* tríplice¹

Darlei Dall’Agnol²

1. Introdução

O que torna um ato humano, tal como, por exemplo, não manter promessas ou não ajudar pessoas em extrema necessidade, algo moralmente errado? Alguma propriedade intrínseca ao próprio ato? Algum motivo torpe que o engendrou? Os maus resultados que se seguiram? Essas questões são do domínio da ética normativa. Um dos problemas éticos atuais é que há diferentes teorias que respondem de maneira diversa e, algumas vezes, até mesmo antagônicas a tais questões. No livro *Three Methods of Ethics*, a kantiana Marcia Baron, o

¹ Trabalho primeiramente apresentado em João Pessoa, em 22 de março de 2024, na International Conference *The Normativity of Morality*, realizada na Universidade Federal da Paraíba (UFPB). Agradeço aos colegas Leonardo Ribeiro e Vitor Somavilla pelo convite e aos participantes pelas questões e comentários. O trabalho foi, posteriormente, apresentado num seminário de doutorandos coordenado pelo professor Jean-Christophe Merle na Universität Vechta (Alemanha), a quem agradeço o convite e a proveitosa discussão. Não por último, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio ao meu projeto “Bioética Clínica e Filosofia da Medicina”.

² Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal de Santa Catarina. Bolsista de produtividade em pesquisa do CNPq.

consequencialista Philip Pettit e o eticista de virtudes Michael Slote apresentam a divergência normativa desta maneira:

Nos últimos anos, três formas de pensar sobre a moralidade passaram a dominar em grande parte o panorama do debate ético. Esses três são o consequencialismo, que enfatiza os bons resultados como base para avaliar as ações humanas; a ética kantiana, que se concentra nos ideais de lei universal e no respeito pelos outros como base da moralidade; e a ética da virtude que vê as questões morais do ponto de vista do agente moral com caráter ou motivos virtuosos. (Baron; Pettit; Slote, 1998, p. 1)

Depois de exporem e defenderem as suas respectivas teorias normativas, cada autor(a) objeta ao oponente e, finalmente, Baron, Pettit e Slote terminam o livro em profundo desacordo ético e moral, tendo como implicação metaética o relativismo. De fato, soluções divergentes, por exemplo, ao problema trolley, parecem demonstrar que o desacordo moral é insuperável.

Há autores, entretanto, que trabalham para mostrar possíveis congruências normativas e superar o relativismo ético. Algumas teorias híbridas já foram propostas no século passado (por exemplo, o utilitarismo kantiano de Richard Hare), mas hoje um enfoque normativo misto que tem dominado o debate é a Teoria Tríplice (TT) elaborada por Derek Parfit em *On What Matters* (2011). Na primeira parte deste trabalho, irei então expor tal enfoque normativo bem como apresentar alguns questionamentos sobre o modo como Parfit acredita ter avançado no debate normativo superando divergências. Ele argumenta que algumas formas de deontologia são compatíveis

com uma versão do consequencialismo e que seus defensores estão subindo a montanha por diferentes lados. Muitos autores discordem de Parfit, mas, mesmo assim, recomendam uma metodologia eirênica (Crisp, 2020, p. 269). Na segunda parte, proporei, então, algumas reformulações visando aperfeiçoar a TT a partir da chamada “ética de virtudes” ou, ao menos, uma de suas variantes que caracterizarei como ética *aretaica* (do grego *areté* = excelência ou virtude). Na terceira parte, abordarei a questão “O que é uma ética aretaica?”, discutindo como ela pode ser incorporada na TT. Finalmente, nas considerações finais, apresentarei algumas vantagens teóricas de uma Teoria Tríplice Reformulada (TTR).

A situação atual da ética normativa é incômoda, pois não contribui para a superação das divergências e, portanto, também não ajuda a resolver os graves problemas morais que enfrentamos. Estamos presenciando ameaças existenciais (de novas guerras nucleares, crises climáticas com eventos extremos constantes como as inundações no Rio Grande do Sul, novas tecnologias, em especial a Inteligência Artificial com potencial de ameaçar a vida como um todo etc.). Sem querer apresentar soluções definitivas, penso que devemos, enquanto filósofos, buscar uma ética normativa capaz de auxiliar na construção de um mundo melhor, começando por superar as nossas diferenças teóricas.

2. A versão de Parfit da TT

A proposta de uma teoria mista ou híbrida não é, em si, uma novidade. Por exemplo, o enfoque bioético conhecido como “princípioalismo” (Beauchamp; Childress, 2018) compatibiliza *à la* Ross princípios deonticos (as normas *prima facie* válidas do respeito pelas pessoas e da justiça como equidade) com princípios, também válidos

à primeira vista, da não maleficência e da beneficência. Um problema desse enfoque é que ele pode engendrar aparentes dilemas morais que precisam, eventualmente, ser dissolvidos apelando para uma espécie de metaprincípio. Já mencionamos também o utilitarismo kantiano de Richard Hare, para quem a linguagem moral tem uma *forma* kantiana (a prescritividade, a universalizabilidade e a sobreposição normativa dos juízos morais) e um *conteúdo* utilitarista (a satisfação de preferências). Um problema dessa teoria é que preferências não razoáveis ou até mesmo irracionais precisam ser evitadas, mas como? O decisionismo de Hare parece não dar conta de casos simples como o do nazista que, descobrindo-se judeu, prefere ir para campos de concentração. Há que se mencionar também o projeto, menos conhecido, do wittgensteiniano Jeremy Wisniewski denominado “Ética Clarificatória”: a ética de Kant e a de Mill seriam *compatíveis* se lermos o imperativo categórico como um esclarecimento da lógica da racionalidade moral e o princípio da utilidade como clarificação dos elementos constituintes do florescimento humano e animal em geral. Essa perspectiva não será explorada aqui.

Já famoso pelas suas contribuições para a ética apresentadas em *Reasons and persons* (por exemplo, para a ética populacional), o filósofo britânico Derek Parfit elaborou outra obra seminal na qual também defende a congruência entre algumas teorias normativas. Combinando uma ética kantiana com o consequencialismo de regras (todos devem seguir princípios otimizadores porque eles são os únicos que podem ser desejados como leis universais) e com um tipo específico de contratualismo, a saber, o elaborado por Scanlon (todos devem seguir princípios que nenhuma pessoa razoável rejeitaria), Parfit propõe a TT. A primeira formulação foi feita nestes termos:

TT: um ato é errado justamente quando tais atos são desautorizados por algum princípio que é

(1) um dos princípios cujas leis universais fariam as coisas correrem melhor,

(2) um dos únicos princípios cujas leis universais todos poderiam desejar racionalmente, e

(3) um princípio que ninguém poderia razoavelmente rejeitar. (Parfit, 2011, pp. 413-414)

É relativamente fácil perceber a presença de um princípio consequencialista (no item (1)) e outro kantiano (2), embora seja um pouco mais difícil notar a tese contratualista (3), dadas as peculiaridades do enfoque normativo de Tim Scanlon.

Alguns esclarecimentos adicionais se fazem necessários para compreender adequadamente o projeto de Parfit de unificação dos três princípios num único metaprincípio. Iniciemos pelo primeiro (1). Nem toda teoria consequencialista está sendo levada em conta, por exemplo o egoísmo ético e o utilitarismo-de-ato foram deixados para trás. O primeiro porque a busca da realização do autointeresse apenas, ou seja, a não preocupação com as outras pessoas, como talvez o Dilema do Prisioneiro mostre, é subotimizadora. Somente a colaboração mútua é otimizadora. O segundo, porque o utilitarismo-de-ato hedonista parece permitir ações que geralmente consideramos injustas (por exemplo, tirar a vida de uma pessoa inocente visitando um hospital para transplantar órgãos para vários pacientes) e nem sempre um(a) agente, por ser falível e não conseguir antever todas as consequências dos seus atos, pode estar certo de que quebrar determinada regra seja otimizador. Além disso, nem tudo o que importa é redutível ao prazer. Há uma pluralidade de portadores de valor em si, como o conhecimento, as virtudes, a contemplação da

beleza etc. permitindo diferentes estilos de vida valiosos de se viver. Portanto, o consequencialismo precisa estar baseado em *regras* que maximizem o pluralismo axiológico. Esse ponto será retomado na segunda parte.

Também é necessário esclarecer que a adoção do princípio da ética kantiana realizada por Parfit (vide acima (2)) está baseada apenas na fórmula *geral* do imperativo categórico, a saber, na seguinte metarregra: “*age apenas segundo a máxima (regra subjetiva) pela qual possas, ao mesmo tempo, querer que ela se torne uma lei universal*” (cf. Kant, 4: 421). Há uma discussão enorme entre comentadores kantianos sobre como melhor interpretar tal imperativo e, acima de tudo, suas implicações práticas. Alguns chegam a duvidar se ele é realmente capaz de orientar ações dada a sua universalidade vazia. Depois de questionar, refazer, tentar extrair consequências de tal fórmula geral, Parfit chega à conclusão de que ela pode funcionar. Todavia, sua reformulação adequada é crucial. Nesse sentido, Parfit (2011, p. 340) prefere a interpretação de Scanlon segundo a qual é errado agir a partir de alguma máxima a não ser que *todos* possam racionalmente querer que seja verdade que todos acreditem que tais atos são moralmente permitidos. Essa reformulação seria compatível com as “fórmulas da autonomia” e do “reino do fins”.

O autor de *On What Matters* rejeita, todavia, o principal conteúdo da ética de Kant: a ideia de um valor absoluto, incondicional etc. das pessoas (sua dignidade intrínseca talvez baseada no princípio metafísico *imago dei*), ou seja, a chamada “fórmula da humanidade” ou do “fim em si”, já que não somos, necessariamente, os únicos seres racionais. Tal fórmula diz que devemos agir sempre considerando pessoas como fins em si e nunca como *meros* meios. Nos capítulos 9 e 10, Parfit recusa o que chama “princípio do mero meio” afirmando que

ele comete vários erros, em especial que, nem sempre que agimos erradamente, tratamos pessoas dessa maneira e que, nem sempre quando estamos tratando uma pessoa como mero meio, nós estamos agindo incorretamente. Este último ponto leva a uma rejeição quase completa da ética kantiana e é, portanto, altamente questionável. Segundo Parfit (2011, p. 221), entretanto, quando causamos pequenos danos a uma pessoa (por exemplo, quebramos um dedo dela numa situação em que esse ato permite que salvemos a vida de alguém), contra a sua vontade, estaremos tratando-a como mero meio, mas isso seria, *pace* Kant, correto do ponto de vista moral. Como veremos, essa tese torna a proposta da TT de Parfit incompatível com uma ética realmente kantiana baseada no respeito pelas pessoas, sendo necessário repensá-la aqui.

É preciso, agora, discutir até que ponto a ética kantiana é compatível com o consequencialismo e se o próprio Kant o rejeitou inteiramente. Se lembrarmos da crítica de Mill, então parece haver algum espaço para debate:

[...] ele [Kant] falha de modo quase grotesco em mostrar que haveria qualquer contradição, qualquer impossibilidade lógica, para não dizer física, na adoção por todos os seres racionais de regras de conduta escandalosamente imorais. *Tudo o que ele consegue mostrar é que as consequências de sua adoção universal seriam tais que ninguém iria escolhê-las.* (Mill, 1987, p. 52; grifos no original)

Autores como MacIntyre (1985, p. 46) concordam e citam como exemplo máximas que passariam pelo crivo do imperativo categórico, tais como “persiga todas as pessoas que mantêm falsas crenças

religiosas”. Talvez kantianos discordem e defensores de uma ética aretaica não julguem correto fazê-lo. Seja como for, no sentido *estrito*, entretanto, Kant não foi consequencialista, pois basta lembrar da sua exigência de seguir o *dever pelo dever* (o conceito de vontade boa tal como ilustrado na interdição absoluta de mentir mesmo por amor à humanidade ironizado por Benjamin Constant).

Num sentido *amplo*, todavia, ao testar as máximas (regras subjetivas) de ação, ele se pergunta sobre *as consequências* positivas ou negativas de segui-las para transformá-las em leis práticas. Por isso, as éticas neokantianas atuais (a teoria da justiça como equidade de Rawls, a ética da discussão de Apel e Habermas e a ética do respeito universal de Tugendhat) não são irresponsáveis, isto é, não ignoram o que se segue da adoção de certos princípios. Podemos, todavia, deixar a vontade boa de lado ou deflacioná-la considerando o seu rigorismo também criticado por Hegel, mas não o conceito de respeito pelas pessoas. Por isso, é importante frisar que Parfit não pode simplesmente abandonar esse aspecto da ética de Kant, pois já não faria sentido algum falar de ética *kantiana*. Como veremos, parece possível compatibilizar esse aspecto material da ética kantiana com uma ética consequencialista, *pois respeitar os direitos básicos das pessoas pode ser considerado otimizador*. Uma teoria consequencialista que pudesse ser realmente adjetivada como “kantiana” poderia ter essa formulação como princípio primeiro e não receberia as objeções endereçadas à TT de Parfit, por exemplo por Herman (2011, p. 83), por abandonar a ideia do respeito pelas pessoas como limitador da maximização da felicidade, em especial, se for hedonisticamente concebida.

Finalmente, é preciso observar que a compatibilização da teoria consequencialista com outras éticas normativas (item 3 acima) diz

respeito apenas a outro enfoque também considerado deontológico, mas de uma formulação bastante específica e pouco comum para uma teoria contratualista, ao menos, se pensarmos na filosofia política clássica (Hobbes, Locke e Rousseau) ou tomarmos a de Rawls como modelo. O que pode implicar a exigência que sigamos princípios que ninguém poderia *razoavelmente* rejeitar? A definição e a exemplificação feitas pelo próprio Scanlon em *What We Owe to Each Other* parecem apontar na direção de um certo tipo de sensibilidade moral. Por exemplo, ao discutir um suposto direito humano à água, Scanlon argumenta que, se, numa pequena vila qualquer, alguém é dono da única fonte e poderia, racionalmente falando, querer toda a água para si, julgaríamos isso como algo não razoável. Em outros termos, considerando uma concepção instrumental de racionalidade (cálculo de meios para fins), não é irracional querer toda a água para si. Todavia, isso que à primeira vista parece melhor é, como vimos acima, subótimo, ou seja, seria melhor cooperar. As outras pessoas têm uma reivindicação razoável à partilha da água. Para colocar nos próprios termos de Scanlon: entre a ideia de irracionalidade e o ideal do que é mais racional acreditar ou fazer, “Eu sugiro que julgamentos sobre o que é ou não razoável fazer ou pensar são relativos a um corpo específico de informações e um conjunto de razões que podem ser menos do que completos.” (Scanlon, 1998, p. 32) A teoria normativa de Scanlon parece, enfim, aproximar-se do particularismo, e torna-se crucial ver até que ponto é possível ser compatibilizada com o consequencialismo que pode possuir uma concepção instrumental de racionalidade. O que causa estranheza, todavia, é que o razoável é visto como estando entre o racional e o irracional, mas então como fazer convergir aquilo que é otimizador com aquilo que é razoável num único princípio? Essa é uma dificuldade para o projeto de Parfit.

Além disso, a questão da definição da “razoabilidade” é problemática, pois é bastante genérico o que entendemos por ela. Enfim, o que é razoável? Trata-se de uma qualidade de caráter? De uma virtude? Um modo de agir resultante de determinadas qualidades de caráter (por exemplo, coragem, temperança, justiça e sabedoria)? Se esse for o caso, não seria possível elaborar uma teoria *realmente* tríplice que englobasse a terceira mais importante teoria normativa contemporânea ao lado do consequencialismo e do deontologismo, a saber, a ética de virtudes mencionada no início deste trabalho? Esse ponto será o tema principal da próxima seção.

Em 2015, antes da publicação do terceiro volume de *On What Matters* (2017), tive a oportunidade de discutir essa questão com o próprio Parfit. Comecei perguntando-lhe, dado que estava escrevendo o terceiro volume para a obra, se uma verdadeira TT não deveria incluir a ética de virtudes ao invés de *uma* consequencialista e *duas* deontológicas. No final, Parfit concordou que, para algumas variantes de ética de virtudes, seria possível compatibilizar o critério de correção ou incorreção de um ato com os princípios acima elencados de outras teorias supostamente rivais. Por isso, eu aguardei ansiosamente a publicação do novo volume que, embora não aprofunde a sugestão e se dedique mais a problemas metaéticos, reconhece explicitamente esse ponto. Parfit escreveu:

Esta visão [Consequencialismo de Motivos] é parecida com algumas formas sistemáticas de *Ética de virtudes* que fazem reivindicações similares sobre disposições ou traços de caráter que melhor promovem o florescimento humano. (2017, p. 418)

Embora não se assuma ali o critério de correção das ações dessa terceira corrente importante da ética contemporânea, penso que podemos, com algumas qualificações, fazê-lo para poder avançar no debate normativo. Esse ponto será desenvolvido na terceira parte. Podemos então, finalmente, passar à discussão sobre como reformular a TT compatibilizando um consequencialismo kantiano com uma ética *aretaica*.

3. Reformando a TT

Antes de propor uma reformulação da Teoria *Normativa* Tríplice (TT) de Parfit (ele defendeu também, no terceiro volume de *On What Matters*, uma Teoria *Metaética* Tríplice, que não discutiremos aqui), convém esclarecer que várias teorias éticas normativas foram recusadas. Além do egoísmo ético e do utilitarismo-de-ato principalmente o clássico de base hedonista, outros enfoques como a moralidade de senso comum, o intuicionismo, a moral da vontade boa de Kant e, mais importante para os propósitos do presente artigo, as versões que chamaremos aqui de “tradicionalistas” da ética de virtudes (por exemplo, a de MacIntyre ou, entre as orientais, o confucionismo) são incompatíveis com um consequencialismo de regras. Tais teorias de virtudes são, no fundo, relativísticas, pois estão baseadas em diferentes modelos, exemplares de pessoas boas, que formam tradições morais incompatíveis entre si. O próprio MacIntyre reconhece esse ponto ao citar mais de meia dúzia de tradições de justiça e racionalidade prática. Teríamos, por conseguinte, que discutir os *critérios* de escolha dos modelos de ação (as supostas pessoas virtuosas, por exemplo, Jesus ou Buda, Confúcio ou Maomé). O que se recusa também de qualquer versão tradicionalista de ética de virtudes é certa interpretação do caráter formal da correção de ações.

Em outros termos, para uma TT, apenas no sentido *prático*, mas não no sentido epistêmico ou ontológico, o critério de correção das ações morais de uma ética *aretaica* pode ser incorporado em direção a uma unificação. Essa tese será esclarecida e desenvolvida na terceira parte deste trabalho.

O ponto de partida para uma reformulação adequada da TT pode ser uma espécie de “fenomenologia ética”, ou seja, uma descrição dos ingredientes constitutivos da vida moral e, acima de tudo, de suas respectivas funções. Parece claro que há sentimentos morais positivos e negativos: simpatia, empatia, culpa, ressentimento, vergonha etc. Também não há dúvida de que as ações geram hábitos e costumes e, eventualmente, tradições. Hábitos formam traços de caráter, ou seja, se são considerados bons, são louvados e encorajados como virtudes. Nem toda prática ou tradição é, entretanto, geradora de atos corretos. Há também, na vida moral, ingredientes mais claramente normativos, ou seja, regras e princípios. Os princípios são guias gerais ou universais de conduta enquanto regras especificam atos particulares. Eles formam uma espécie de “rede de normas”, cada uma desempenhando uma função. São os elementos mais objetivos, pois, por exemplo, sentimentos são elusivos. É a partir de uma rede de normas, com padrões normativos objetivos, que é possível caracterizar mais claramente ações como sendo corretas ou não. Também é bom salientar que princípios desempenham uma *função justificatória*, ou seja, legitimam as regras que, por sua vez, *especificam* ações, eliminando seu caráter abstrato e genérico. Se existir, então, algum princípio ético primeiro (PEP), ele pode ter uma status *apriorístico*, fulcral. Esse ponto será retomado a seguir. É crucial, finalmente, salientar que não basta, na moral, agir em conformidade apenas com normas, mas é necessário internalizá-las e *seguí-las*, isto é, cumprir o

ato que é ordenado, seja bom ou correto em si. Portanto, qualquer princípio normativo primeiro precisa começar com “Siga normas...”.

Para complementar o metaprincípio, é importante evitar reducionismos ao caracterizar o valor de um ato. Parece improvável que elementos pré-ato (digamos, no T_{-1} , por exemplo, intenções, deliberações ou escolhas) sejam suficientes para determinar a correção da ação. Por isso, uma visão como a intuicionista clássica que sustentava que existem atos corretos *sejam lá quais forem as circunstâncias* parece ferir a descrição mais elementar da responsabilidade humana. Esse é um problema também da ética de Kant, e a defesa de um *fiat iustitia, et pereat mundus* é simplesmente insustentável. Precisamos responder pelo que se segue das nossas ações. Além disso, a depender do conhecimento, declarado até mesmo impossível das intenções das ações, jamais poderíamos caracterizá-las como corretas. Por outro lado, um tipo de consequencialismo radical, ou seja, que sustenta que o valor de um ato (T_0) é *somente* determinado pelos seus resultados, também parece ferir a caracterização de certas ações como sendo *prima facie* erradas, por exemplo, cometer homicídio. Por conseguinte, não é apenas aquilo que ocorre no T_{+1} e o que se segue indefinidamente que conta para estabelecer a correção das ações. A depender do conhecimento de *todos* os resultados das nossas ações para caracterizá-las como corretas, talvez jamais possamos fazê-lo. É claro que somos responsáveis até mesmo por consequências não antevistas, mas será que podemos enumerar todas as previsíveis? A discussão sobre a *natureza do ato* também é relevante assim como a atitude do agente. Por conseguinte, não existem atos corretos *sejam lá quais forem as intenções*. Podemos invocar aqui o “princípio das totalidades orgânicas”, de G. E. Moore, que, segundo Parfit (2011, p. 465), é uma contribuição original do *Principia ethica*. O

princípio diz que “O valor de um todo não deve ser assumido como simplesmente a soma dos valores de suas partes”, e, então, deve-se considerar *motivos + ato + resultados* uma totalidade orgânica cujo valor não é a mera soma de suas partes e que, cada parte, pode contribuir diferentemente para o valor da ação *como um todo*. Por exemplo, na distinção jurídica entre ato doloso ou culposo, a intenção pesa diferentemente bem como o efetivo resultado da ação. Parece óbvio, fenomenologicamente falando, que há, por um lado, ações aparentemente boas em si (ajudar alguém), mas que podem trazer más consequências, por exemplo, condicioná-las a continuar nessa condição e, por outro lado, ações aparentemente incorretas (punir alguém), mas que trazem bons resultados. Tanto num caso quanto no outro pode haver várias intenções, mas nenhuma sozinha pode determinar o valor do ato.

Aqui podemos, então, tentar introduzir um elemento de avaliação *prima facie* dos atos como sendo valiosos (bons ou corretos) em si, mas cuja valoração completa dependerá de outros fatores, por exemplo, intenção e resultados. Neste ponto, podemos invocar vários filósofos clássicos que não rejeitariam essa tese inicial. Por exemplo, para Aristóteles (EN 1107a10), há ações (adultério, roubo e homicídios) que são más em si mesmas. Também há, diga-se de passagem, sentimentos maus em si, como a inveja, o ciúme e a falta de vergonha. Por outro lado, atos resultantes das excelências morais são valiosos em si mesmos, por exemplo, um ato de justiça distributiva. Um consequencialista *sofisticado* como Mill também não rejeitaria avaliar algumas ações como boas e más em si. Em *The Subjection of Women*, ele escreveu: “[...] O princípio que regula as relações sociais existentes entre os dois sexos – a subordinação jurídica de um sexo ao outro – é errado em si mesmo, [...]” (In: Mill, 1994, p. 119). Além do prazer, então,

há, para Mill, outros portadores de valor ou desvalor intrínsecos. Como veremos mais adiante, as virtudes para ele também podem ser valiosas e desejadas em si e não somente como agregados de um todo. Claramente, Kant considerava todas as leis morais como sendo imperativos não hipotéticos, ou seja, prescrições que ordenavam atos como bons ou maus em si mesmos. Deflacionando o caráter categórico, podemos dizer que ações morais são corretas em si. Finalmente, a distinção de Wittgenstein, em *Lecture on Ethics* (1965, p. 5), entre usos relativos ou absolutos de “bom” ou “correto” aponta nessa direção: pode não fazer diferença para alguém ser um bom ou mau jogador de futebol, mas não é moralmente permitido inventar uma calúnia (“*preposterous lie*”): *devemos* nos comportar melhor. Portanto, um PEP pode começar com “Siga normas que prescrevem ações (ou omissões) *prima facie* valiosas em si [...]”.

Valendo-nos da TT original de Parfit, podemos, agora, introduzir a qualificação a partir de um consequencialismo-de-regras kantiano, a saber, que tudo o que importa é o quão bem as coisas vão. Cabe perguntar, claro, “para quem?”. Ora, uma vez que o egoísmo ético já foi recusado, podemos inserir uma qualificação formal do tipo “otimizadora”, mas com uma explicitação seguinte, a saber, que seja “universalmente desejável ou aceitável”. Nesse sentido, não há razões *a priori* para defender o hedonismo ou qualquer outra teoria monista para todas as pessoas e, na verdade, se deixarmos de lado a ideia da superioridade de uma forma de vida sobre outras abandonando dessa maneira a tese aristotélica da *ergon* do ser humano, temos, portanto, como resultado o já mencionado pluralismo axiológico e, a partir deste, o político, o religioso etc. Em outros termos, dado que há vários portadores de valor (prazer, conhecimento, beleza, amizade, virtudes etc.), cada pessoa pode criar um projeto de vida buscando alguns

desses bens valiosos em si enquanto outras pessoas dedicam suas vidas a outros projetos. É necessário, então, o respeito pelas pessoas, que tem prioridade lógica, para possibilitar a coexistência pacífica e a cooperação mutuamente vantajosa, sendo essas governadas por uma democracia agonística. Tais projetos de vida serão razoáveis e excluem a possibilidade de uma “nação de demônios,” isto é, indivíduos que se tratam sem respeito (como meros objetos), como partes de uma comunidade moral. É exatamente a exigência do respeito que, entre outras coisas, torna possível um pluralismo razoável de fato (mudando um pouco a expressão de Rawls “*the fact of reasonable pluralism*”). Conforme veremos mais adiante, essa é a função de uma moral comumente partilhável. Seja como for, o PEP pode ser incrementado dessa maneira: “Siga normas que prescrevem ações aparentemente valiosas em si mesmas, mas que também sejam otimizadoras, universalmente aceitáveis e ...”, sendo necessário complementá-lo agora com uma ética *aretaica*.

Lembremos que Parfit utilizou originalmente o contratualismo de Scanlon com a noção de razoabilidade. Ora, dizer que uma pessoa razoável não rejeitaria princípios otimizadores no sentido que seriam universalmente aceitáveis parece introduzir um elemento *aretaico* no primeiro princípio. Facilita, em certo sentido, aproximar o contratualismo da ética de virtudes (Coitinho, 2016), embora o primeiro ofereça uma fundamentação fraca da moralidade. Apelos à razoabilidade são conhecidos em diversos âmbitos da vida humana. Por exemplo, no direito procura-se, nos processos judiciais em que se julgam crimes, mostrar evidências “*beyond all reasonable doubt*”. Não entrarei, entretanto, na discussão sobre a natureza de tais evidências, mas assumirei que há provas suficientes de crimes a partir de evidências robustas. Na medicina, havia o padrão tradicional da

tomada de decisões, paternalisticamente determinado pelo profissional da saúde, mas hoje há o apelo àquilo que é saudável para uma pessoa razoável (*the reasonable person*). Novamente, não discutirei os pormenores, mas assumirei que esse padrão objetivo existe. Isso não significa, todavia, que ele não possa variar (por exemplo, o que conta como risco razoável numa cirurgia para um paciente: 1% chance de óbito ou 10% ou até mesmo 50%?), depende de cada paciente. Por isso, talvez seja necessária uma conversação *real* entre as pessoas envolvidas para uma tomada de decisão e não uma projeção do que seria idealmente “razoável” a partir de uma perspectiva particular.

Um ponto central, então, é entender melhor como usamos o conceito “razoável”, por exemplo quando dizemos para alguém “Seja razoável!”. A expressão parece referir-se a uma qualidade de caráter ou pressupor até mesmo um conjunto de virtudes. Esse parece ter sido o pensamento de Rawls (1993, p. 58s): “Em vez de definir o razoável diretamente, específico dois dos seus aspectos básicos como virtudes das pessoas: (i) propor e respeitar princípios de cooperação (justiça); (ii) ter disposição para reconhecer o peso do julgamento.” A razoabilidade seria, então, um aspecto das virtudes das pessoas. Claramente, aqui, podemos interpretar a razoabilidade como um “subproduto” de se ter determinadas virtudes, ou seja, ela seria superveniente às excelências como, somente para ilustrar, as quatro *aretai* primárias ou também chamadas “cardeais”: coragem, temperança, justiça e sabedoria. Esse ponto será retomado a seguir. Seja como for, é crucial também discutir o estatuto epistêmico da razoabilidade. Nesse sentido, não é necessário abraçar um fundacionalismo do tipo forte, cartesiano, pois, como escreveu Wittgenstein: “Mas o que os homens consideram razoável ou não

muda com o tempo. Em certos períodos, os homens consideram razoável o que, em outros períodos, não consideram e vice-versa.” (OC, §336) Se podemos, dessa maneira, assumir que ser razoável é ter um traço de caráter superveniente às virtudes básicas, então podemos postular um PEP incorporando o critério formal de correção das ações de uma ética *aretaica*. Esse aspecto da TT será desenvolvido na próxima seção.

4. O que é uma ética *aretaica*?

Já é hora, então, de estabelecermos melhor o que podemos considerar o lugar de uma ética *aretaica* numa TTR. Para começar, é importante lembrar que há uma variedade grande de éticas de virtudes. Como Darwall mostra (1998), uma ética de virtudes pode ser moral ou não moral; direta ou indireta etc. Por exemplo, a versão tradicionalista de MacIntyre, pretensamente inspirada em Aristóteles, sustenta que virtudes são qualidades necessárias para atingir os bens internos a certas *práticas* que são justificadas por narrativas no interior de *tradições*. Essa visão, todavia, conforme vimos anteriormente, tem implicações relativistas e não parece ter sido a expressa na *Ethica nicomachea*, que defende, contra o *ethos* grego vigente à época, uma concepção de bem humano supremo (*eudaimonia* perfeita) composta prioritariamente pela vida dedicada à teorização filosófico-científica, embora reconheça que, para outros estilos de vida, há também a possibilidade de atingir um tipo de felicidade. Claro que o argumento aristotélico está baseado numa biologia metafísica e precisa ser adaptado, hoje, depois da teoria evolucionista, mas a versão de MacIntyre não é a mais promissora por levar, conforme já vimos, ao relativismo ético. Por isso, nesta seção, vamos apresentar uma variante de uma ética de virtudes que chamarei

“*aretaica*”, fortemente inspirada em Aristóteles, diferenciando-a dessas concepções.

Para compreender o que significamos por “ética aretaica”, convém apresentar alguns componentes básicos: i) um critério de correção das ações morais; 2) a tese da não prioridade entre “bom” e “correto”; iii) o valor intrínseco das qualidades de caráter; iv) a fundamentação das virtudes em princípios. Antes de discuti-los, convém definir a “virtude” ou “excelência moral” como sendo uma propriedade de caráter adquirida por bons hábitos, ou seja, pela repetição contínua de atos louváveis. Também é importante salientar, *ab initio*, que uma ética aretaica é uma *teoria* da ação valiosa (boa ou correta) e não uma mera ética de virtudes não preocupada de forma central com essa caracterização. Como notou acertadamente Crisp (2015, p. 258), uma ética de virtudes como teoria da ação correta aproxima-se a uma forma de deontologia (e, eu acrescentaria, do consequencialismo-de-regras).

Iniciando, então, pelo primeiro ingrediente, é crucial lembrar que, geralmente, uma ética de virtudes sustenta que um ato é correto se e somente se é o tipo de ato que seria performado por um(a) agente virtuoso(a) (Statman, 1997, p. 7; Crisp, 2008, p. 33). Em outros termos, um modelo de pessoa virtuosa é considerado o próprio critério daquilo que se deve fazer. Todavia, há diferentes maneiras de entender essa tese. Não é claro que Aristóteles a tenha sustentado, mesmo que tenha incluído o *phronimos* na própria definição de excelência moral. Há uma ambiguidade: o certo é fazer o que a pessoa com sabedoria prática faz ou devemos agir com sabedoria prática seguindo a regra correta? Há passagens da *Ethica nicomachea* que apontam para a segunda alternativa, por exemplo, a seguinte: “Não faz diferença se foi uma pessoa boa que defraudou uma pessoa má ou uma pessoa má que

defraudou uma pessoa boa, nem se foi uma pessoa boa ou má que cometeu adultério. A lei considera apenas a diferença causada pela lesão e trata as partes como iguais” (Aristóteles, EN 1132a3-5). Em outros termos, a pessoa virtuosa segue a regra correta (*orthos logon*), ou seja, um juiz justo julga imparcialmente. Essa é uma ideia comum na filosofia prática aristotélica e pode ser aceita pelo consequencialismo de regras milliano e tranquilamente por uma ética kantiana e pode, por conseguinte, ser incorporada na ética aretaica. Vemos, aqui, claramente a necessidade de embasar as virtudes em princípios sólidos evitando a crítica feita por Kant no início da *Fundamentação*, a saber, que as virtudes não são incondicionalmente louváveis, pois a coragem de um homicida torna-o mais abominável. Voltaremos a esse ponto a seguir quando enunciarmos um PEP.

É por isso que não podemos simplesmente dizer que o correto é fazer aquilo que uma pessoa boa faria nas circunstâncias sem outras qualificações. Ao contrário, ser virtuoso, por exemplo, ser um juiz justo é julgar imparcialmente. Se quisermos reservar um lugar privilegiado às virtudes em relação às ações, essa tese não pode significar uma ética verdadeiramente aristotélica, como bem apontou Irwin (2007, p. 198), nem no sentido ontológico (a correção dependeria de ter sido produzida por uma pessoa virtuosa) nem epistêmico (somente a pessoa virtuosa sabe o que é o correto), mas apenas no sentido prático (nós, provavelmente, faremos a coisa certa se cultivamos as virtudes). Em outros termos, não se trata, na ética aretaica aqui defendida, de assumir um critério de correção moral qualificado pelo(a) agente, mas apenas relacionado com ele(a), ou seja, assumindo a exigência de que o(a) agente tenha a motivação virtuosa. Esse ponto é compatível com a vontade boa kantiana, obviamente deflacionada, e com um consequencialismo de motivos,

tornando a TT *reformulada* uma opção realmente viável para superar as divergências atuais.

Um segundo ponto importante é metaético. Geralmente, uma ética de virtudes defende que “bom” tem primazia sobre “correto”, diferenciando-se de uma ética puramente deontológica. O utilitarismo segue, em geral, esse preceito teórico também definindo o correto em função do bom. Na verdade, uma ética aretaica não precisa assumir essa tese metaética. Vou sustentar, então, em termos de pressuposição lógico-semântica, a tese da não prioridade: nem “bom” nem “correto” serão considerados termos primitivos, mas ambos são definíveis a partir de “valor”, ou seja, no sentido moral, daquilo que é intrinsecamente valioso. Esse ponto permite superar a divisão entre teorias deontológicas e teleológicas. Em termos de epistemologia moral, sustento que o saber-que (x é errado) é uma condição necessária, porém não suficiente para agir bem. É necessário também *saber-como* agir de modo correto, por exemplo, valorando intrinsecamente. Chamo essa epistemologia moral de “cognitivismo prático”, e ela pode ser mais bem ilustrada em atitudes como a de cuidado: não apenas é necessário o sentimento moral da simpatia, e atos para preservar e incrementar o bem-estar de alguém, mas ele deve ser praticado *pelo benefício de quem está sendo cuidado*. Da mesma maneira, *saber-como* respeitar implica não apenas reconhecer a personalidade de alguém, mas efetivamente cumprir as obrigações em função dos direitos da pessoa *por ela mesma*. Essas duas atitudes são as pedras fundacionais de *uma* Moralidade Comumente Partilhável (MCP), capaz de garantir a coexistência pacífica e a cooperação mutuamente vantajosa entre agentes. Aliás, é bom que se tenha claro que a elaboração tem um eventual PEP a partir de uma TT que será apresentado como justificando essa moral comum (MCP) e não a

suposta moral a-histórica, perene etc. Na MCP, o respeito tem prioridade lógica (não haverá coexistência pacífica se não começarmos por ele); o cuidado possui prioridade ontológica, ou seja, reconhecemos que devemos o nosso ser ao cuidado como filósofas feministas mostraram recentemente (Held, 2006, p. 3). Finalmente, em termos ontológicos, assumirei uma perspectiva não naturalista sem postular entidades estranhas, mas apenas não sucumbirei ao factualismo. Definirei “pessoa” como agente capaz de seguir normas que não representam estados de coisas, *stricto sensu*, naturais. Em outros termos, trata-se de um fenômeno prescritivo e não descritivo ou empírico.

Um terceiro elemento central de uma ética aretaica é que virtudes são desejáveis por si, pois têm valor intrínseco, mas também contribuem para o bem maior tanto pessoalmente quanto para o bem comum. Dito de outro modo: as virtudes são desejáveis por serem boas para o agente, mas também em nome da realização de seu projeto de vida (sua concepção de vida boa) e também como capazes de contribuir para realizar o bem público, ou seja, as instituições necessárias para a coexistência pacífica e as múltiplas formas de cooperação (não apenas econômica, mas cultural, lúdica etc.). Nesse sentido, é crucial distinguir uma ética de virtudes relativista como a de MacIntyre de uma ética aretaica baseada em excelências não relativistas. Como sustenta Martha Nussbaum (1988), há qualidades de caráter necessárias para realizar quaisquer projetos de vida. Podemos acrescentar, e não há contradição nisso, que também são necessárias enquanto virtudes cívicas, ou seja, para sermos cidadãos razoáveis numa MCP capaz de legitimar a forma de governo democrática etc. Por isso, se um PEP é capaz de justificar normas intermediárias em diferentes domínios (por exemplo, na política, o Estado de Direito; na

bioética, o consentimento informado etc.), ele também poderá justificar normas que exigem o cultivo de virtudes primárias e deixar à discricionabilidade dos agentes, de suas funções específicas ou projetos particulares de vida, outras qualidades. Por exemplo, para aderentes da moral cristã, além das quatro virtudes primárias, poderíamos acrescentar a fé, esperança e caridade; para budistas, outras virtudes etc. Algumas excelências poderiam ser requeridas para funções específicas, tal como um bom pai terá que ter paciência etc. Essa qualidade não é requerida de todos os agentes, pois nem todos têm filhos. Aqui, entretanto, não deve restar dúvidas quanto à chamada “Tese Aristotélica” em relação às quatro virtudes básicas: uma vida virtuosa é a melhor para os seres humanos (Darwall, 2004, p. 75). Essa tese é crucial para uma ética aretaica, sendo as excelências valiosas por si mesmas, embora possamos desejá-las também em nome de um bem maior pessoal ou público. Esse ponto será complementado a seguir.

Para fins de ilustração, vamos caracterizar, brevemente, as quatro virtudes primárias constitutivas de uma ética aretaica. A coragem diz respeito às nossas emoções, em especial, ao medo, mas principalmente à força moral de fazer o que é correto. A temperança lida com nossos desejos e sentimentos e é, basicamente, uma capacidade de moderação necessária para *seguir* regras. A justiça, como vimos acima, diz respeito às noções como a da imparcialidade, à da reciprocidade etc. e, fundamentalmente, à equidade. A sabedoria, que nos define enquanto espécie, é um saber-como aplicar o conhecimento para agir corretamente e viver bem. Tais virtudes foram defendidas desde Platão, passando pelos estoicos (e Cícero), via Agostinho, até a ética contemporânea em autores como Peter Geach etc. Parece, então, adequado dizer que quem possui essas virtudes

intrinsecamente valiosas (mesmo que nada delas se seguisse, ainda assim as desejaríamos) será uma pessoa virtuosa, ou seja, tanto realizará seu projeto de vida quanto será um(a) cidadão(a) capaz de contribuir para o bem comum. Por isso, elas são necessárias para a construção de uma MCP.

O quarto elemento importante de uma ética aretaica é embasar as excelências morais em princípios sólidos. Lembremos da crítica de Kant no início da *Fundamentação* (4: 393), a saber, que a coragem de um assassino torna seus atos ainda mais abomináveis, mas também que as virtudes são fundamentadas via imperativo categórico considerado um princípio sintético *a priori*. Acredito que não há como propor uma variante de ética de virtudes sem levar essa crítica a sério e resolver o problema que ela aponta. A presente versão de ética aretaica evita o problema embasando as excelências no PEP. A caracterização de Kant, todavia, pressupõe a necessidade incondicional, ou seja, uma epistemologia fundacionista dificilmente sustentável hoje. Conforme vimos acima, uma epistemologia falibilista é mais adequada para a compreensão da natureza de um primeiro princípio na TTR.

Finalmente, estou numa posição de poder enunciar o PEP de forma completa. Para ter clareza sobre a reformulação que estou propondo, citarei aqui o resumo da explicação sobre a incorreção das ações feita pelo autor de *On what matters*. Parfit sustenta que: “[...] um ato é errado justamente quando tais atos são desaprovados pelos princípios que são otimizadores, unicamente desejáveis universalmente e não razoavelmente rejeitáveis” (2011, p. 413). Em outros termos, um ato é errado quando não é permitido por princípios que são otimizadores (critério consequencialista), desejáveis de modo unicamente universal (elemento deontológico) e não razoavelmente

rejeitáveis (aspecto reformulável em termos aretaicos). Introduzirei, então, duas modificações principais: primeiro, transformarei a explicação da incorreção dos atos num princípio *normativo* capaz de proibi-los; em segundo lugar, introduzirei um critério *prático* de correção das ações a partir da ética aretaica apresentada acima. Podemos, assim, postular o seguinte PEP: Siga normas (*princípios, regras*) *que prescrevem ações (ou omissões)* *prima facie valiosas em si mesmas, mas que também sejam otimizadoras e universalmente aceitáveis inclusive por uma pessoa virtuosa considerando-as desde um ponto de vista imparcial*. Esse princípio justifica uma teia de regras (princípios intermediários e regras) bem como as quatro virtudes primárias acima apresentadas.

Esse princípio, permanecendo fixo como uma dobradiça, ou seja, sendo postulado como válido *a priori* (mas não necessariamente incondicional), dá sustentação justificatória para outras normas intermediárias, por exemplo para o Estado Democrático de Direito (e, portanto, é o fundamento jurídico da política) e regras em domínios específicos: somente para ilustrar, o consentimento informado na bioética. Dois esclarecimentos adicionais se fazem, agora, necessários. Considero o PEP uma proposição prática fulcral, no sentido wittgensteiniano, capaz de justificar normas intermediárias e regras específicas para nossa forma de vida (*Homo sapiens*) e, eventualmente, podendo, por nós, ser projetado para outros agentes, sejam eles animais não humanos ou agentes artificiais. Em outros termos, possui caráter apriorístico e é universalmente aplicável para todos os agentes. Ainda assim, é necessário recusar um fundacionalismo forte e compreendê-lo como falível.

O segundo esclarecimento é este: o PEP é o fundamento normativo, embora falível, de uma MCP. Uma moral comum é

condição para a coexistência pacífica e a colaboração mutuamente vantajosa entre pessoas, com crenças e valores divergentes, e pressupõe o respeito recíproco e o cuidado mútuo. Estes, por sua vez, levam a uma série de políticas públicas necessárias para a construção de uma esfera comum na qual diferentes sistemas axiológicos podem coexistir, possibilitando que as pessoas virtuosas realizem seus projetos de vida. Esse é um dos principais resultados deste trabalho e outros serão retomadas na conclusão.

Para encerrar esta seção, procurarei mostrar que uma ética aretaica integrada numa TTR não é incompatível nem com uma ética consequencialista sofisticada (de regras) nem com uma ética kantiana, embora já tenha pressuposto a fusão das duas acima. No que diz respeito a Kant, já mencionamos a sua crítica, no início da *Fundamentação*, às virtudes isoladamente consideradas ou não justificadas por princípios (por exemplo, que a coragem torna um assassino mais temível), mas precisamos lembrar também que a *ethica* para ele nada é senão uma “doutrina das virtudes”, sendo a outra parte de sua *Metafísica dos costumes* uma “doutrina do direito”. É preciso salientar que, formalmente falando, só há uma virtude para Kant, ou seja, a força de vontade no cumprimento do dever (MS, 6: 394) e, nesse sentido, poderíamos pressupô-la como capacidade para executar o PEP. Claramente, mesmo em Kant, seguir certas máximas (por exemplo, desenvolver talentos) forma bons hábitos e gera virtudes. Que Kant considere a virtude o “bem supremo” (*oberster Gut*) não significa a prescrição de um estoicismo para todos (virtuosismo), dado que ela pode ser parte constitutiva de um bem maior (*höchsted Gut*). Por outro lado, há que se recusar, numa sociedade plural e secularizada como a nossa, a tese de Kant de que a virtude é um constituinte do mais alto bem somente atingível pós-morte na sua

completude necessária com a felicidade garantida pela imortalidade da alma. Esse componente metafísico da ética de Kant é incompatível com uma TTR, embora possa ser aceito como parte de algum credo particular de uma moral específica, mas não para uma MCP. É possível, então, interpretar, como Parfit faz, a “fórmula do bem maior” desta maneira: “Todos deveriam se esforçar para promover um mundo de virtude universal e felicidade merecida” (2011, p. 408). Assim, alguns kantianos (por exemplo, Onora O’Neill) não negariam que a ética kantiana é compatível com uma aretaica, embora não o seja com uma ética de virtudes no sentido tradicionalista. Seja como for, um PEP pode justificar um princípio intermediário de virtude mostrando que é otimizador, racionalmente aceitável e razoável seguir o princípio intermediário kantiano de virtudes, a saber, “[...] seguir máximas com fins que podem, ao mesmo tempo, ser leis universais”. Em vez da perfeição própria como fim universalizável, é necessário pensar em termos de realização de projetos pessoais e, ao invés da felicidade alheia, é preciso incrementar o bem comum como sendo um que é ao mesmo tempo um dever. Como percebemos, há transformações importantes a serem feitas na ética de Kant para incorporar uma ética kantiana na TTR, em especial, a integração da maximização do respeito pelas pessoas.

Um consequencialismo (de *regras* que especifique qualidades de caráter desejáveis) também não é incompatível com uma ética aretaica. Desde o livro de Frankena chamado *Ethics*, tem-se discutido como aproximar teorias aretaicas ao utilitarismo como, por exemplo, a partir do utilitarismo valórico ou de caráter de Railton ou o utilitarismo de motivos de Adams. Talvez, a questão remonte ao próprio Mill: “Mas será que a doutrina utilitária nega que as pessoas desejam virtude...? Muito pelo contrário. Ela sustenta não somente

que a virtude deve ser desejada, mas que ela deve ser desejada desinteressadamente, por si mesma” (1987, p. 82). Essa seria, então, mais uma razão para sustentar um pluralismo axiológico a partir do consequencialismo. O próprio Mill não deixa dúvidas sobre esse ponto: “Os ingredientes da felicidade são vários, e cada um deles, desejável em si mesmo *e não meramente quando considerado como expansão de um agregado*” (Mill, 1998, p. 82, grifos nossos). Temos claramente presente aqui a ideia de que virtudes são desejáveis por si e podem incrementar o bem comum nos termos detalhados acima. Por isso, convido consequencialistas-de-regras, kantianos não ortodoxos e defensores da ética aretaica a juntar-se na elaboração de uma teoria unificada procurando o cume da montanha que parece ter a feição de um PEP.

5. Observações finais

Para encerrar este capítulo, gostaria de enfatizar duas implicações da adoção de uma ética aretaica numa TTR: ela ajuda a desenvolver duas atitudes fundamentais, a saber, a do cuidado respeitoso e a do respeito cuidadoso. Não é necessário, aqui, fazer uma análise final daquele que é um dos elementos constituintes principais de uma MCP, a saber, o respeito pelas pessoas que implica o seu reconhecimento como tal e o cumprimento das obrigações em relação a elas, valorando-as intrinsecamente, para reconhecer que Parfit não parece estar certo (2011, p. 221) ao recusar esse princípio intermediário, ou seja, que tratar pessoas como meros meios, como coisas, é moralmente errado. Autores como Tugendhat, Darwall e Dworkin estão certos em insistir na centralidade do respeito para a moralidade. Nesse sentido, podemos dizer que respeitar pessoas é otimizador,

universalmente aceitável e é o que um indivíduo virtuoso faz considerando as coisas imparcialmente.

Também não é necessário apresentar uma análise última do cuidado enquanto outra forma de valoração intrínseca de indivíduos vulneráveis simpatizando com eles e incrementando o bem-estar deles para perceber que, assim como pensadoras feministas sustentam, o cuidado não é apenas algo pessoal (cuidado de si), mas também algo político e até mesmo global. Por isso, a democracia (e lembrando do pluralismo axiológico, a agonística) é *prima facie* valiosa em si embora também se justifique pelos seus resultados. Essa, entre outras, é uma das implicações importantes da TTR.

Para terminar e não deixar uma questão polêmica aberta sem um tipo de resposta, gostaria de retomar o problema trolley, mostrando que ele não representa nem um dilema real nem leva a uma opção entre teorias normativas incompatíveis. Uma ética puramente consequencialista estabelece como obrigatório puxar a alavanca e salvar cinco vidas mesmo ao custo de matar uma pessoa inocente; uma ética deontológica ortodoxa sustenta que nada se deve fazer. O que diria uma TTR? A resposta parece ser esta: embora seja relevante levar em consideração que o caso exige não olhar para as pessoas amarradas nos trilhos para decidir, não é eliminada a autorreflexão do(a) agente do porquê encontra-se neste contexto e nem se tem alguma responsabilidade especial pela segurança do local. Por conseguinte, a coisa certa a fazer para alguém com alguma atribuição pública não é necessariamente a mesma para um indivíduo comum sem esse tipo de incumbência. Desse modo, para a primeira pessoa seria aparentemente obrigatório puxar a alavanca, mas proibido para a segunda, que pode nada fazer. Portanto, fazer o que é obrigatório para cada um é *prima facie* correto, mas também otimizador e aceitável por

todos, inclusive por agentes virtuosos considerando as coisas imparcialmente.

Referências

- Anscombe, G. E. M. (1997) Modern Moral Philosophy. In: Crisp, R.; Slote, M. (Orgs.) *Virtue Ethics*. Oxford: Oxford University Press, pp. 26-44.
- Aristotle. (1984) *The Complete Works*. Barnes, J. (Org.). Princeton: Princeton University Press.
- Baron, M.; Pettit, P.; Slote, M. (1998) *Three Methods of Ethics*. Oxford: Blackwell.
- Beauchamp, T.; Childress, J. (2018) *Principles of Biomedical Ethics*. 8. ed. Oxford: Oxford University Press.
- Crisp, R. (1997) *Mill on Utilitarianism*. London: Routledge.
- Crisp, R. (2008) *Reasons and the Good*. Oxford: Clarendon Press.
- Crisp, R. (2015) A Third Method of Ethics? *Philosophy and Phenomenological Research*, v. 90, n. 2, pp. 257-272.
- Crisp, R. (2020) Are We Climbing the Same Mountain? Moral Theories, Moral Concepts, Moral Questions. *Zemo*, v. 3, pp. 269-278.
- Crisp, R.; Slote, M. (1997) *Virtue Ethics*. Oxford: Oxford University Press.
- Darwall, S. (1998) *Philosophical Ethics*. Boulder County: Westview Press.
- Darwall, S. (2004) *Welfare and Rational Care*. Princeton: Princeton University Press.
- Geach, P. (1997) *The Virtues*. Cambridge: Cambridge University Press.
- Hare, R. (1997) *Sorting Out Ethics*. Oxford: Clarendon Press.
- Held, V. (2006) *The Ethics of Care: Personal, Political, and Global*. Oxford: Oxford University Press.

- Herman, B. (2011) A Mismatch of Methods. In: Parfit, D., *On What Matters*, v. 2. Oxford: Oxford University Press, pp. 83-115.
- Hurka, T. (2003) *Virtue, Vice, and Value*. Oxford: Oxford University Press.
- Irwin, T. (2007) *The Development of Ethics*, 3 v. Oxford: Oxford University Press.
- Kant, I. ([1785-1797] 2004) *Practical Philosophy*. Trad. e Org. M. J. Gregor. Cambridge: Cambridge University Press.
- MacIntyre, A. (1985) *After Virtue*. Indiana: Duckworth.
- Mill, J. S. ([1861] 1987) *Utilitarianism*. New York: Prometheus Books.
- Mill, J. S. ([1859] 1994) *On Liberty and Other Writings*. Cambridge: Cambridge University Press.
- Moore, G. E. ([1903] 1993) *Principia Ethica*. Baldwin, T. (Org.). Cambridge University Press.
- Nussbaum, M. (1988) Non-Relative Virtues: An Aristotelian Approach. *Midwest Studies in Philosophy*, v. 13, n. 1, pp. 32-53.
- O'Neill, O. (1989) *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.
- Parfit, D. (2011-2017) *On What Matters*, 3 v. Oxford: Oxford University Press.
- Railton, P. (2003) *Facts, Values, and Norms*. Cambridge: Cambridge University Press.
- Rawls, J. (1993) *Political Liberalism*. New York: Columbia University Press.
- Rawls, J. ([1971] 1999) *A Theory of Justice* (ed. rev.). Oxford: Oxford University Press.
- Ross, D. ([1930] 1998) *The Right and the Good*. Stratton-Lake, P. (Org.). London: Hackett.

- Scanlon, T. M. (1998) *What We Owe to Each Other*. Cambridge, Massachusetts/London, England: Belknap Press of Harvard University Press.
- Slote, M. (1992) *From Morality to Virtue*. Oxford: Oxford University Press.
- Statman, D. (Org.) (1997) *Virtue Ethics: A Critical Reader*. Edinburgh: Edinburgh University Press.
- Wisnewski, J. (2007) *Wittgenstein and Ethical Inquiry*. Dublin: Continuum.
- Wittgenstein, L. (1965) *A Lecture on Ethics*. The Philosophical Review, v. 74, n. 1, pp. 3-12.
- Wittgenstein, L. (1989) *Werkausgabe in 8 Bänden*. Frankfurt am Main: Suhrkamp.

Hibridismo normativo¹

Denis Coitinho²

1. Introdução

O objetivo deste capítulo é explicar em maior detalhe as características de uma teoria moral mista e por que precisamos de uma teoria normativa híbrida para lidarmos com um conjunto significativo de problemas práticos complexos. Após apresentar as razões da necessidade de tal teoria, o próximo passo será fazer uma distinção entre a moralidade privada e a pública. A ideia básica é que há uma assimetria normativa, de forma que censuramos menos os agentes por suas decisões erradas na esfera privada da moral, mas somos mais exigentes com o comportamento dos cidadãos no domínio público da moralidade e, por isso, é importante distinguir essas duas esferas da moral e esclarecer em que medida elas estão conectadas. Na sequência, apresento o procedimento de escolha das virtudes que são necessárias para garantir a felicidade pessoal e coletiva. Por fim, problematizo a razão da teoria mista procurar conciliar especificamente a ética de virtudes com o

¹ Publicado originalmente em uma versão ampliada como Capítulo 9 do livro Coitinho, D. *Contrato & virtudes III*. São Paulo: Edições Loyola, 2024.

² Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade do Vale do Rio dos Sinos (UNISINOS). Bolsista de produtividade em pesquisa do CNPq.

neocontratualismo e não alguma outra teoria, tal qual o utilitarismo, o kantismo etc.

Veja que problemas morais e políticos complexos exigem, muitas vezes, critérios normativos diversos para a identificação de soluções. Uma teoria moral tradicional usa em geral apenas um critério normativo, como a maximização do bem-estar, considerando o utilitarismo, a universalizabilidade e não instrumentalização, considerando o kantismo, a razoabilidade, considerando o neocontratualismo, ou mesmo algum agente virtuoso, como o prudente, tendo em mente a ética de virtudes. O ponto é que, em muitas situações, são exigidos outros critérios, fazendo-se necessário usar um padrão normativo diferente para resolver a questão, o que revelaria um problema de coerência interna da teoria. Como uma teoria utilitarista usaria um critério deontológico de liberdade sem se mostrar incoerente internamente? Ou como uma teoria deontológica, como a kantiana, poderia usar um critério de maximização do bem-estar sem recair em contradição interna? Como a ética de virtudes poderia fazer uso de um princípio da dignidade humana para assegurar os direitos humanos sem comprometer a sua própria estrutura?

2. Situando o problema da normatividade

Partimos do problema de como distribuir bens escassos, como o ocorrido na recente pandemia da covid-19. De posse de uma teoria utilitarista, se usaria um princípio agregacionista para saber como devemos agir, distribuindo algum bem, digamos leitos de UTIs em uma pandemia, de maneira que ele traga os melhores resultados, considerando todos os envolvidos. Assim, o cálculo a ser feito é identificar quais pessoas terão mais chance de sobreviver fazendo uso

de leitos de UTIs em comparação com outras que terão menos chance de sobrevivência. E isso parece implicar que terão preferência os agentes mais jovens e com menos comorbidades, sendo consequência disso, então, que indivíduos mais idosos e mais doentes morrerão em maior número, o que é, no mínimo, questionável, pois, intuitivamente, acreditamos que devemos dar preferência às pessoas mais vulneráveis, como idosos, crianças e doentes. Por sua vez, com uma teoria deontológica em mãos, como a kantiana, não se poderia fazer uso desse importante cálculo bem-estarista realizado pelo utilitarismo, uma vez que o princípio da não instrumentalização ou princípio da humanidade diz que nenhum agente pode ser tomado apenas como meio, devendo ser tomado também como fim, o que implica reconhecer a igualdade moral de todos. O problema é que, no caso em tela, isso implicará a possível morte de um número maior de pessoas, o que não é desejável. E, com uma teoria das virtudes, provavelmente se chegaria a um resultado semelhante à teoria deontológica, pois se apelaria para o critério de benevolência com os vulneráveis, isto conectado com o critério de justiça, o que não parece inteiramente adequado, considerando que acreditamos que devemos salvar o maior número de vidas possível.

É bem conhecido na literatura filosófica o problema trolley (*trolley problems*), que mostra algo similar ao que procurei apontar no exemplo acima. De forma geral, são apresentados dilemas morais para ver quais intuições os agentes apresentam, e isso com o objetivo de identificar o princípio moral correspondente. Um dos dilemas mais conhecidos questiona a correção de se sacrificar a vida de alguém para salvar um número maior de indivíduos. Quando se pergunta às pessoas o que elas devem fazer ao verem um trem desgovernado a 100 km/h que irá atropelar cinco operários no trilho à frente, sendo

possível, contudo, acionar uma alavanca para desviá-lo para um trilho alternativo onde há apenas um operário, a resposta geralmente dada é sim, deve-se desviar o trem. Por sua vez, se a pergunta é se devemos empurrar uma pessoa robusta de uma passarela para salvar os cinco operários nos trilhos do trem, a resposta geralmente dada é não. A interessante questão que exige reflexão é saber por que o princípio que parece certo no primeiro caso é tido como errado no segundo exemplo? O princípio no primeiro caso claramente aponta para a ideia de que é correto sacrificar uma pessoa para salvar cinco, isto é, de que devemos salvar o maior número possível de indivíduos, e isso porque a vida humana importa, o que nos remete ao princípio da maximização do bem-estar. Por sua vez, o princípio no segundo caso nos diz que é errado matar um ser humano inocente, não importando as consequências, princípio este que se choca com o primeiro. Seria semelhante a situação de se ver como errado matar uma pessoa saudável para transplantar seus órgãos para cinco doentes graves que precisam desse tipo de tratamento para sobreviver. Note que há um conflito dos princípios aqui. Por um lado, sabemos que é correto o esforço para salvar o número maior de pessoas. Ao mesmo tempo, sabemos que é errado matar um ser humano inocente. Então, o que isso parece significar? Uma alternativa seria dizer que se está considerando a morte no primeiro caso apenas como um efeito colateral, não a identificando com a intenção do agente. Quer dizer, apenas a intenção de matar é que seria propriamente o erro moral.³

³ Os *trolley problems* são uma série de experimentos mentais em ética e psicologia, envolvendo uma série de dilemas éticos, na forma de se é correto ou não sacrificar uma pessoa para salvar um número maior de agentes. O ponto central é fazer uso das intuições morais sobre certos casos particulares para se chegar a certos princípios éticos. Originalmente, o dilema colocado por Foot foi apresentado na figura de um

Fiz referência ao problema trolley para ressaltar meu argumento central, de que não conseguimos resolver certas questões complexas apenas fazendo uso de uma teoria tradicional, que utilizará um único critério normativo que, em geral, se opõe aos critérios das outras teorias éticas. Por exemplo, o utilitarismo diria que é correto salvar o número maior de vidas, uma vez que seu princípio moral básico é o da maximização do bem-estar e, assim, seria correto acionar a alavanca para desviar o trem e salvar os cinco operários, mesmo com o ônus da morte de um agente. Aqui temos a aceitação do princípio do sacrifício. E, por esse mesmo princípio, também seria correto empurrar o homem robusto com o fim de salvar os cinco operários, bem como seria correto fazer o transplante para salvar os cinco doentes. O problema é que, de acordo com nossos juízos morais comuns, é errado tirar a vida de um agente inocente, mesmo considerando as boas consequências. Por outro lado, o kantismo diria que é errado acionar a alavanca para desviar o trem visando salvar os cinco operários, uma vez que o princípio da universalizabilidade e, sobretudo, o da não instrumentalização proibiriam qualquer ação que considerasse uma pessoa apenas um meio para o bem das outras, trazendo por consequência a morte de cinco agentes. E isso parece problemático porque também concordamos que devemos salvar o número maior de pessoas possíveis, talvez porque pensamos que a vida humana tem valor. E mesmo que o utilitarismo e o kantismo quisessem valorizar esses juízos morais comuns, eles teriam um problema de coerência interna na própria teoria.

condutor de um trem e não a de um espectador. Sobre o tema, ver Foot, 1967 e Thomson, 1976 e 1985. Ver, também, Mikhail, 2011, pp. 78-82.

Por essa razão, seria desejável poder contar com uma teoria moral que dispusesse de vários critérios normativos, como o critério utilitarista das melhores consequências, o critério deontológico da não instrumentalização, que garante o respeito universal à dignidade humana, e até mesmo um conjunto de virtudes que são padrões normativos não da ação correta, mas da vida boa, o que parece contribuir no caso de precisarmos contar com uma compreensão mais abrangente do certo e errado. E esses critérios normativos parecem adequados e não arbitrários porque eles fazem parte da maneira com que refletimos e decidimos nossas ações, estando conectados com nossas intuições morais básicas. Veja que, nessa teoria moral mista, a todo momento, estamos chamando atenção para o valor de nossas intuições morais, uma vez que não estamos comprometidos com a utilização de fatos morais, que serviriam de fundamento último dos juízos éticos. Assim, tomamos as nossas intuições morais comuns como ponto de partida para sabermos o que devemos fazer e o que devemos uns aos outros. Porém, muitas vezes, nossas intuições morais são contraditórias entre si. Por isso elas devem ser testadas por sua coerência com um sistema coerente de crenças. E, na medida em que houver alguma contradição, elas devem ser revisadas. Assim, um dos papéis muito importantes de uma teoria moral, penso, é esclarecer essas intuições comuns e alinhá-las coerentemente, num processo permanente de ajuste mútuo entre essas intuições morais coletivas, certos princípios éticos e até mesmo as crenças factuais relevantes no caso.⁴

⁴ Intuições morais são um tipo de conhecimento direto sobre a correção ou erro de algo, diferentemente de um raciocínio inferencial. Importante frisar que não tomamos as intuições morais como sinônimos de intuições individuais. Ao contrário, tomamo-las como intuições morais coletivas ou comuns, isto é, as intuições a que

Assim, creio que seja relevante considerar nossas intuições morais comuns a respeito da diferença entre a moralidade privada e pública.

3. Moralidade privada e pública

É algo aparentemente comum em nossa vida social o fenômeno da assimetria normativa no domínio moral, de forma que, irrefletidamente, censuramos menos os agentes por suas decisões e ações que irão impactar apenas na felicidade de sua própria vida, mas censuramos mais os agentes por suas decisões e ações que terão um forte impacto na vida de todos. Por exemplo, censuramos menos um agente que é desleal a um amigo, ou que é covarde em uma dada situação que requeria coragem, ou que não tem resiliência para enfrentar as diversas pressões cotidianas do que seria o caso com um agente que assedia sexualmente colegas de trabalho ou, no limite, comete uma violência sexual com alguma delas, ou um ato racista no seu emprego. A censura moral no primeiro caso é mais branda, não implicando uma condenação ao caráter moral do agente, nem gerando algum sentimento de ressentimento ou indignação (atitudes reativas), nem mesmo gerando uma mudança de disposição em relação à felicidade da pessoa. Talvez seja apenas um julgamento de

chegamos por um processo de progresso moral, isto é, de correção da arbitrariedade nas avaliações éticas. São exemplos dessas intuições a consideração de que punir um agente inocente é errado, bem como é errado diminuir a credibilidade epistêmica de alguém em razão de seu gênero, raça ou classe, assim como é errado ser cruel com os animais. Apenas elas terão uma certa autoridade epistêmica presumida. No âmbito das intuições individuais, encontramos mais facilmente o desacordo moral, como é o caso em questões como aborto, eutanásia ou veganismo, entre outras. Sobre as intuições morais, ver McMahan, 2013, pp. 104-105, 118-119. Ver, também, Mikhail, 2011, pp. 111-117.

que a ação ou a atitude do agente é errada. Diferentemente do segundo caso, em que a censura moral é mais forte, gerando, no mínimo, um sentimento de indignação pelo ato sexista e racista, ou, no limite, podendo ocasionar tanto uma condenação do caráter quanto uma mudança disposicional.⁵

Mas o que isso parece mostrar? Penso que uma maneira de interpretar o fenômeno é reconhecer que, de forma não consciente, julgamos diferentemente os casos morais em que apenas a própria pessoa da ação é atingida dos casos em que todos os agentes (cidadãos) de uma comunidade são ou podem ser atingidos. O dano, no primeiro cenário, é pessoal, no outro, é coletivo. Por exemplo, não ser resiliente trará um dano apenas para a própria pessoa que está agindo dessa maneira, diferentemente de se ser racista, em que o dano atinge toda a sociedade. De outra forma, a ação, no primeiro caso, impactará apenas a felicidade do próprio agente, enquanto, no segundo caso, impactará a felicidade de toda comunidade. Se isso estiver ao menos parcialmente correto, creio que seria produtivo fazer uma distinção entre duas esferas da moralidade, a saber, a moralidade

⁵ Penso que a dificuldade em reconhecer essa assimetria normativa se dê pela diversidade semântica que a censura possui. Não conhecemos ainda muito bem o que seja de fato a censura e como ela funciona. Um sentido de censura muito bem conhecido é como uma atitude reativa. Strawson defende que responsabilizamos moralmente as pessoas porque temos sentimentos reativos pelos atos errados ou corretos cometidos e, assim, sentimos ressentimento, indignação, culpa, ou, alternativamente, sentimos gratidão. Nessa compreensão, responsabilidade é tomada como censurabilidade. Ver Strawson, 2008, pp. 4-22. Outra importante maneira de compreender a censura é em sua perspectiva cognitiva, de forma que ela seria um julgamento, um juízo avaliativo, sobre as ações, atitudes ou caráter dos agentes. Ver Watson, 2004, p. 265. Sobre o problema da censura, ver Coates e Tognazzini, 2013, pp. 3-26. Ver, também, McCormick, 2022, pp. 11-39.

privada e a moralidade pública, o que não é o mesmo que a distinção usual entre a esfera privada e pública da vida.⁶

Na esfera da moralidade privada, a autoridade normativa é em primeira pessoa, de forma que necessita da disposição do agente para querer ser um indivíduo melhor. Nessa dimensão, ele não é obrigado pela lei nem pelas convenções sociais. Por exemplo, doar para o Médicos sem Fronteiras ou para a Cruz Vermelha é um ato correto e que é geralmente elogiado pelas pessoas, uma vez que isso revela solidariedade com os mais vulneráveis do mundo. Não há uma lei que obrigue atos de solidariedade nem há uma convenção social de que devemos obrigatoriamente ajudar os mais vulneráveis. Essa será uma decisão do sujeito. Por isso, não doar não é um motivo para censura. Veja que aqui não há uma dimensão dos direitos, que serviria de base para o dever moral. Por exemplo, os habitantes de países pobres que geralmente recebem ajuda humanitária não têm direito a exigir essa ajuda dos cidadãos ao redor do mundo. Seria o mesmo caso de alguém que é vegano por razões morais, isto é, por considerar errado matar os animais não humanos para comer. Os animais não humanos não têm

⁶ A distinção usual entre as esferas da vida privada e pública reivindica que os indivíduos devem ter liberdade para escolher como viver, qual religião ter, qual partido político votar e como viver moralmente a sua vida, mas devem ser obrigados a seguir as leis e as regras de justiça. Nessa diferenciação, a religião e a moral fazem parte da vida privada dos agentes, enquanto a política e o direito fazem parte de sua vida pública. Essa distinção é muito comum em modelos liberais de justiça. A distinção que estou propondo é um pouco diferente. Em vez de pensar na moral como um todo em oposição à política e ao direito, separo duas partes da moral, uma privada, que terá relação direta com a escolha do agente em agir/ser de uma certa forma, e outra pública, que estará mais conectada com a política e o direito, sendo um tipo de obrigação moral perfeita, podendo ser entendida como padrões morais minimamente aceitáveis que definem a decência humana e encorajam a vida social. Sobre o tema, ver Hampshire, 1978, pp. 23-53.

o direito assegurado à vida, uma vez que é comum o consumo de carne, além de não ser ilegal consumir proteína animal.

Veja que, nessa dimensão privada da moralidade, não se fere nenhum direito que seja integralmente reconhecido pela comunidade política e, por isso, não é implicada nenhuma punição porque não se desrespeita nenhuma lei. Não é um caso para punição a não doação para o Médico sem Fronteiras ou Cruz Vermelha, do mesmo modo que não é passível de punição o comer carne. Seria o mesmo em situações como mentir, ser desleal ou trair. São atos claramente reconhecidos como errados, mas não são passíveis de punição, que é um tipo de censura, a saber, censura legal. Assim, é importante notar que, nessa esfera privada da moral, o nível de censura é mais brando do que na esfera pública da moral, podendo-se aceitar mais facilmente o desacordo moral, isto é, nossa discordância a respeito do que é certo e errado.

Por isso, parece que a descrição dessa dimensão privada da moralidade é mais adequada à ética de virtudes, pois ela é uma teoria que tem por base a crença na força da disposição do agente para ser uma pessoa melhor, isto é, virtuosa, o que implicará bem deliberar, sendo o mesmo que encontrar a mediania, quer dizer, o meio-termo entre os extremos, meio-termo não aritmético, mas relativo ao agente e às características da própria virtude, que é alcançada pela repetição das ações virtuosas, isto é, por um processo de habituação, a fim de formar o caráter virtuoso. Importante notar que uma ética principialística, tal como utilitarismo, kantismo ou mesmo o contratualismo, não parece ter muita relevância nesse domínio, pois ela exigiria do agente um tipo de raciocínio moral muito demandante,

precisamente num âmbito em que não há uma obrigação perfeita, e isso por não existirem aqui direitos reconhecidos publicamente.⁷

Já na esfera da moralidade pública, ao contrário, a autoridade normativa é em segunda pessoa, de forma que há uma exigência social para um certo tipo de comportamento. Para falar nos termos de Darwall, teríamos aqui uma “obrigação bipolar” (Darwall, 2013, p. 20), isto é, uma obrigação intersubjetiva entre os cidadãos de uma dada comunidade moral e política. Nessa dimensão, o agente é obrigado não (apenas) por sua consciência, mas tanto pelas leis quanto pelas convenções sociais. Por exemplo, torturar pessoas, ou mais especificamente, torturar soldados em uma guerra é um ato errado, e isso porque está em desacordo com as convenções sociais e políticas, bem como está em desacordo com a lei. Torturar pessoas é um ato ilegal que é passível de punição. Veja que aqui há uma dimensão dos direitos, que será uma das bases do dever moral de não torturar. Por exemplo, há leis nacionais, tais como Constituições, que condenam a tortura, bem como há leis internacionais, tal como a Declaração Universal dos Direitos Humanos, que condenam a tortura e, mais, tomam isso como um critério para justificar até mesmo uma intervenção internacional. Importante chamar atenção de que aquele que seria torturado tem um direito tanto à sua integridade como à sua

⁷ Deveres ou obrigações morais são formas moralmente requeridas de conduta. Podem ser perfeitos ou imperfeitos. Deveres perfeitos implicam um direito correlato. São aqueles deveres estritos, que não são passíveis de interpretação, sendo obrigatórios tanto moralmente como legalmente, e sua violação torna a vida intolerável. Por exemplo, não matar. Já os deveres imperfeitos, dependem do juízo valorativo (moral) de cada agente, estando abertos à liberdade de cada um. Por exemplo, a caridade. São casos de deveres perfeitos os de não matar, não roubar, não fraudar. São tipos de deveres imperfeitos: ser caridoso, solidário e amigo. Ver Rainbolt, 2000, pp. 233-234. Ver, também, Stocker, 1967, pp. 507-517.

vida, e isso assegurado pelo conjunto legal já referido. Nessa dimensão, há uma clara correlação entre razões morais e razões político-jurídicas.

Note que, nessa dimensão pública da moral, a censura é mais forte, uma vez que o ato errado irá contra um direito reconhecido publicamente e, em muitas situações, esse ato errado será equivalente a uma ação ilícita que é passível de punição, o que nada mais é do que um tipo específico de censura, isto é, uma censura feita por uma autoridade estatal com a intenção de reprová-lo e reparar de alguma maneira o dano causado. A moralidade pública, assim, pode ser vista como a moralidade comum, consistindo nas mais amplas convicções morais compartilhadas pelo grupo. Por exemplo, parece fazer parte da moralidade comum atualmente a crença de que a tortura é errada, de que a escravidão é injusta, de que a discriminação, seja por raça, gênero ou classe, deve ser condenada, de que toda a intolerância é nefasta, entre outras. Dessa forma, o nível de censura é mais forte nessa esfera pública da moral, e se espera mais fortemente o consenso normativo. Talvez isso explique, ao menos parcialmente, o porquê de a relatividade ética não ser tão pacífica aqui como o é na esfera privada da ética.⁸

Por isso, parece que a descrição desse domínio público da moral é mais adequada ao contratualismo, uma vez que ele é uma teoria

⁸ Em linhas gerais, o desacordo moral é a evidência social de que alguém cometeu um erro. Por sua vez, o consenso ético é a evidência social de que estamos corretos, isto é, de que estamos garantidos em acreditar que p. Há aqui uma justificação das decisões/ações por um certo grau de confiança nas crenças asseguradas. Na esfera privada da moral, pode ter baixo grau de confiança. Já na esfera pública da moralidade, necessita-se de maior grau de confiabilidade. E isso exige humildade epistêmica, de forma que essa humildade nos leva a não intervir na vida dos outros. Ver Nguyen, 2019, pp. 348-349.

moral e política que procura estabelecer os deveres dos agentes a partir de sua correlação com os direitos, determinando os princípios morais para a vida comum a partir da própria aceitabilidade dos envolvidos, isto é, a partir das características de agência de racionalidade e razoabilidade. Ao contrário do utilitarismo e kantismo, que estabelecem princípios morais absolutos com base, sobretudo, na racionalidade humana, com uma autoridade normativa em terceira pessoa, o contratualismo, assim como o proposto por Rawls e Scanlon, por exemplo, estabelece os princípios morais intersubjetivamente, isto é, através de um procedimento em que a razoabilidade será um critério central para a identificação do consenso moral.

Um último comentário a respeito dessa distinção apresentada deve ser feito. Defendi que devemos distinguir entre a moralidade privada e pública e isso em razão da existência de uma assimetria normativa irrefletida, que traz como consequência a atribuição menor de censura no domínio privado da moralidade e maior no seu domínio público.⁹ Agora, quero defender que devemos estabelecer uma

⁹ O domínio público da moralidade pode ser visto como centrado em critérios como o de justiça, direitos, liberdade, e o domínio privado estaria mais ligado aos critérios de lealdade, deveres e autoridade, para exemplificar. Importante destacar que em *The righteous mind*, Haidt investiga a extensão do domínio moral descritivamente, afirmando que a mente moral foi formada em bases evolucionistas, apresentando seis fundamentos básicos: cuidado (*care*), justiça (*fairness*), liberdade (*liberty*), lealdade (*loyalty*), autoridade (*authority*), santidade (*sanctity*), critérios estes que explicariam nossas preferências políticas. Liberais, por exemplo, estariam centrados nos critérios de cuidado, justiça e liberdade, e os conservadores, em todos os seis, o que incluiria a lealdade, autoridade e santidade como princípios normativos importantes para avaliação das políticas públicas. Ver Haidt, 2012, p. 114. Minha hipótese é a de que, com a complexidade da vida em sociedade, nossa mente moral evoluiu de forma a incorporar essa assimetria normativa. Sobre a teoria dos fundamentos da moral, ver Shweder *et al.*, 1997, pp. 119-169.

conexão importante entre esses dois domínios, identificando certas virtudes privadas que sejam coerentes com certas virtudes e princípios éticos públicos. Mas vejamos isso ao tratar do procedimento de escolha das virtudes.

4. Teoria contratualista das virtudes

O ponto central da teoria contratualista das virtudes é tentar compatibilizar o critério de razoabilidade e justiça das teorias neocontratualistas, que leva em conta, também, as consequências das ações, com o critério das virtudes, como a prudência, integridade, autonomia, humildade etc., mas isso dentro de um desenho liberal, isto é, que respeita o pluralismo de valores. A ideia básica é que nós podemos escolher, sob certas circunstâncias, as virtudes privado-públicas que serão fundamentais para a garantia de nossa felicidade, seja ela pessoal, seja ela coletiva, da mesma forma que podemos escolher princípios de justiça para o ordenamento público, bem como outros princípios morais para ordenar nossa vida comum. Essa proposta toma por base uma importante distinção entre o âmbito privado e público da moralidade, de maneira que, na moralidade privada, a base da decisão do agente será fornecida por uma ética de virtudes, enquanto, na moralidade pública, os cidadãos poderão contar com os critérios de razoabilidade e justiça, que levam em conta as consequências, tal como em uma teoria neocontratualista como a de T. Scanlon, que determina que a ação errada é determinada por um princípio que não pode ser razoavelmente rejeitado (Scanlon, 1998), ou como na teoria da justiça como equidade de J. Rawls, que determina que os princípios de justiça serão escolhidos em uma situação de simetria conhecida como posição original sob o véu da ignorância (Rawls, 1971).

Com isso em mente, é importante destacar que o procedimento de escolha das virtudes se dará apenas onde ocorrer a conexão entre as esferas privada e pública da moral, de forma a identificar certas virtudes privadas que sejam coerentes com certas virtudes e/ou princípios morais públicos. Esse procedimento não servirá para a escolha dos princípios ou virtudes na esfera da moralidade pública, da mesma forma que não servirá para a escolha das virtudes no âmbito privado da moral, estando circunscrito ao que podemos chamar de uma esfera privado-pública da moralidade. Para maior clareza, deixem-me esquematizar essa distinção como a que estou propondo:

(i) Moralidade pública: Nessa dimensão, pode-se utilizar a teoria neocontratualista para a escolha dos princípios morais. Ao se fazer uso da teoria de John Rawls, por exemplo, estar-se-ia equipado para a escolha dos princípios de justiça para orientar a estrutura básica da sociedade, que nos oportunizam um princípio da igual liberdade e um princípio da igualdade equitativa de oportunidade e atenção ao bem comum. E, ao se utilizar a teoria contratualista de Thomas Scanlon, por outro lado, estar-se-ia equipado para a escolha dos princípios morais mais relevantes para determinar o certo e errado de forma mais ampla, através de um procedimento de escolha que leva em conta a razoabilidade dos agentes. Aqui os critérios centrais são os de justiça, direitos, razoabilidade e tolerância, sendo selecionados a partir de um procedimento racional/razoável de escolha.¹⁰

¹⁰ No Capítulo 5 de *What we owe to each other*, Scanlon apresenta a estrutura central de seu neocontratualismo, dizendo que um ato é correto se e somente se ele puder ser justificado por princípios que os envolvidos não poderiam razoavelmente rejeitar, isto é, um ato será correto se e somente se ele for justificado aos outros. Com isso já se pode ver que a ideia de justificabilidade é tomada de duas formas, propiciando (i) a

(ii) *Moralidade privada*: Nessa dimensão, pode-se fazer uso de uma ética de virtudes, de forma a auxiliar os agentes a pautarem suas vidas a partir do critério de vida boa ou florescimento humano, que demanda o agente a ser uma pessoa melhor, isto é, a ser virtuoso nas mais diferentes situações da vida. Esse modelo ético exige que a pessoa seja corajosa, resiliente, moderada e amiga, responsabilizando moralmente o agente a partir de seu próprio caráter. Aqui os critérios normativos centrais serão todas as virtudes epistêmicas e morais que auxiliam na felicidade pessoal do agente, que, por sua vez, auxiliam na felicidade geral da comunidade, sendo escolhidas individualmente a partir do seu próprio caráter moral.¹¹

(iii) *Moralidade privado-pública*: Apenas nessa dimensão é que se usará o procedimento de escolha das virtudes, de forma a poder identificar certas virtudes privadas, como a prudência, integridade, autonomia e humildade, entre outras, que serão coerentes com certos princípios e virtudes públicos, como a justiça, que inclui a liberdade e a igualdade, bem como com a razoabilidade, a tolerância, entre outros. Por exemplo, as virtudes da coragem, da moderação, da resiliência, da lealdade não parecem coerentes com os critérios de justiça ou da razoabilidade e, por isso, não seriam escolhidas nesse domínio.

base normativa da moralidade do certo e errado e (ii) a caracterização mais geral de seu conteúdo. Ver Scanlon, 1998, p. 189.

¹¹ A ética de virtudes parece estabelecer uma ligação estreita entre as virtudes, o caráter e a responsabilidade moral, o que pode trazer como consequência uma concepção abrangente de responsabilidade. Isso pode significar que o agente se sentirá responsável por fazer tais e tais ações ou agir a partir de um certo padrão moral excelente em razão de seu próprio caráter virtuoso que foi formado por meio das diversas escolhas realizadas e as ações subsequentes. Aqui há uma forte conexão entre as virtudes e a vida boa, ou melhor, com o florescimento humano. Sobre isso, ver Hursthouse, 1997, pp. 223-226.

Por outro lado, a virtude da autonomia parece coerente com o princípio da liberdade, de forma similar que a virtude da humildade parece coerente com o princípio da tolerância. A ideia geral do procedimento é que alguém pode propor uma virtude da dimensão (ii) que seja coerente com os critérios normativos da dimensão (i), justificando a virtude proposta e sendo aprovada se não for razoavelmente rejeitada.

Com essa distinção esclarecida, passo agora a apresentar os detalhes desse procedimento de escolha das virtudes privado-públicas.

A ideia geral, como dito anteriormente, é que alguém possa propor uma virtude privada que seja coerente com os princípios e/ou virtudes morais que orientam a esfera pública da moralidade, devendo justificar sua proposta, e que os envolvidos na escolha possam aceitar ou recusar essa virtude sugerida ao grupo. A escolha, então, se caracteriza por ser coletiva, isto é, ela deve ser feita pelos cidadãos ou por representantes de cidadãos. É uma escolha coletiva e não individual, devendo ser justificada pela razão pública, e os agentes estão de posse de todas as informações relevantes sobre a realidade social e pessoal. Por exemplo, sabem como uma sociedade funciona, tanto em termos econômicos, como em termos sociais e políticos. Da mesma forma que sabem como as pessoas são. Assim, a virtude privado-pública será aprovada se ela não for razoavelmente rejeitada pelos envolvidos na decisão. Isso significa dizer que essa escolha será sensível ao contexto, de forma que cidadãos em contextos normativos diferentes, ou em contextos históricos diferentes, poderão escolher virtudes diferentes, tendo em mente que essa escolha deve assegurar tanto a felicidade pessoal como a coletiva.

A título de ilustração, amizade, coragem, lealdade e moderação não parecem candidatas apropriadas, pois essas virtudes não se mostram coerentes com os princípios e/ou virtudes da moralidade pública, como os de razoabilidade, tolerância, justiça. Por exemplo, a lealdade e amizade podem se contrapor à exigência de justiça que, em muitas circunstâncias, demanda imparcialidade, como seria o caso de alguém que, por lealdade ou amizade a um amigo, não o denuncia por um crime cometido. Por outro lado, a integridade, autonomia, prudência e humildade parecem coerentes com os princípios morais públicos já referidos anteriormente e parecem difíceis de ser rejeitados razoavelmente, isto é, com o uso público da razão, de forma que a argumentação não poderia estar fundada apenas em razões pessoais, tal como razões religiosas ou baseadas em tradições. Veja que, para Aristóteles, a virtude privada demandada pela responsabilidade política seria a astúcia (*deinotes*), que é a capacidade calculativa de produzir resultados consistentes com um dado objetivo. Já, para Adam Smith, essas virtudes seriam o autocontrole, a benevolência, a prudência em conexão com a justiça, enquanto, para Stuart Mill, elas seriam a curiosidade intelectual, modéstia e a mente aberta.¹²

As virtudes escolhidas nesse procedimento podem ser tanto epistêmicas como morais, mas, preferencialmente, deveriam ser virtudes epistêmicas, tomadas como condição de possibilidade para as

¹² Thomas Nagel, em *Ruthlessness in Public Life*, diz algo similar, defendendo a tese de que a moralidade privada e a moralidade pública não são independentes entre si, podendo compartilhar uma base comum, sem uma ser derivada da outra. Para ele, a impiedade (*ruthlessness*) não poderia existir na esfera da moralidade pública. Como na fórmula de Radbruch, em que o direito não poderia conviver com uma extrema injustiça, teríamos, aqui, então, que a política não poderia conviver com uma extrema crueldade. Ver Nagel, 1978, p. 82.

virtudes morais e para o respeito aos princípios éticos públicos identificados, e isso para assegurar o desenho liberal da proposta. Veja a virtude da prudência (*phronesis*). Ela é uma virtude epistêmica, é claro, pois é a capacidade/disposição para identificar os meios mais adequados para realizar um fim bom. Entretanto, ela é condição de possibilidade tanto para outras virtudes epistêmicas, como a curiosidade e mente aberta, por exemplo, quanto para as virtudes morais, como coragem, moderação e justiça. Importante destacar que as virtudes escolhidas pelos cidadãos e cidadãs podem ser usadas para o estabelecimento de políticas públicas, diretrizes educacionais, e mesmo como critério de priorização na alocação de recursos, mas não podem ser motivo de punição, apenas de promoção. Assim, poder-se-ia ter por diretriz educacional a promoção da autonomia dos alunos, de maneira que eles possuam a capacidade analítica de resolver problemas de forma autorregulada, mas não seria o caso de puni-los na hipótese de permanecerem heterorregulados.

Nesse procedimento em tela, as virtudes são escolhidas pela sociedade e não pelo Estado. Por isso, seu ponto de partida seria identificar a força social, isto é, ver a demanda feita pela sociedade por uma ou por um conjunto de virtudes relevantes. A título de ilustração, a integridade passou a ser exigida no mundo dos negócios e do trabalho, de forma que não basta mais atualmente ter apenas a expertise técnica em uma dada função, mas que o agente deva demonstrar coerência entre os valores assegurados e suas ações, estando conectado mais estreitamente com os valores das organizações, por exemplo, e isso para evitar casos de corrupção e assédios de todo tipo, como o sexual e moral. A humildade também passou a ser demandada em muitos campos, como o dos negócios, da política e até da ciência. Por fim, uma forma interessante de pensar

nas virtudes escolhidas é imaginar a sua eficácia social, isto é, procurar prever quais as consequências que esses padrões normativo-morais trariam para auxiliar tanto na prosperidade como na estabilidade de uma dada sociedade.

5. Conciliando a ética de virtudes com o contratualismo

Nas seções anteriores, apresentei as principais características da teoria contratualista das virtudes, inclusive distinguindo entre as esferas privada e pública da moral e detalhando o procedimento de escolha das virtudes privado-públicas, ressaltado o porquê de devermos fazer uso de uma teoria normativa híbrida, em contraposição às teorias mononormativas, tentando conectar uma ética de virtudes com uma teoria neocontratualista. Entretanto, faltou argumentar detalhadamente por que deveríamos conciliar especificamente uma ética e epistemologia das virtudes com o neocontratualismo e não harmonizar algumas outras teorias éticas, tal como o utilitarismo, o kantismo ou a ética do cuidado.¹³ Nesta seção, pretendo argumentar nessa direção, ressaltando a força das teorias propostas, bem como evidenciando a coerência entre elas.

¹³ Derek Parfit, por exemplo, em *On what matters*, propôs uma teoria mista, chamada de teoria tripla ('TT'), tentando conciliar vários critérios normativos em um único princípio moral, tais como o consequencialista, com o critério da otimização da ação, o deontológico, com o critério da desejabilidade universal e, também, o contratualista, com o critério da rejeitabilidade razoável: "TT: Um ato é errado apenas quando tais atos não são permitidos por algum princípio que seja otimizado, universalmente desejável e que não seja razoavelmente rejeitável" (Parfit, 2011, p. 413). A limitação da proposta se dá por ela ser apenas uma teoria principialística, não congregando as virtudes como critérios normativos, o que penso não ser muito eficaz na esfera privada da moral, que precisará contar com a disposição do agente para querer ser melhor. Mas ela pode ser muito eficaz na esfera pública da moralidade, penso. Teríamos que testar em vários cenários para ver se TT é superior aos princípios contratualistas de justiça e da rejeitabilidade razoável.

Começo com o neocontratualismo. A intuição basilar dessa teoria normativa é a de estipular que a melhor maneira de identificar a objetividade para os juízos morais e políticos é através da coerência entre eles e os princípios éticos que seriam escolhidos em razão de serem aceitáveis para os envolvidos. Em vez de tentar encontrar o critério normativo na ordem natural ou humana, ou na vontade ou razão divina, o contratualismo apela para a ideia de justificação pela aceitabilidade social. Quer dizer que as normas morais encontram legitimidade em sua capacidade de obter acordo, acordo esse feito sob determinadas condições apropriadas e realizado por aqueles agentes sobre os quais as normas serão aplicadas diretamente. Isso pode ser formulado da seguinte maneira:

Um ato X é correto ou justo, em certas circunstâncias C, sse ele for aprovado por uma regra R, escolhida por agentes As, que tem a característica de ser aceitável pelos envolvidos em razão de sua razoabilidade.¹⁴

A partir dessa formulação, alguns elementos centrais da teoria se destacam. Um desses elementos é o seu falibilismo. Ela não se funda em fatos morais, propondo um procedimento de construção dos

¹⁴ Sayre-McCord explica o contratualismo como uma teoria moral e política que defende que as normas éticas ou instituições políticas podem obter a legitimidade em sua habilidade de assegurar um acordo. Ver Sayre-McCord, 2013, pp. 332-333. Por sua vez, Freeman diz que o contratualismo intuitivamente apela para a ideia de que, uma vez que as normas morais impõem exigências aos agentes e restringem sua conduta, estas normas devem ser livremente aceitas pelos envolvidos, considerando seu benefício mútuo. Assim a característica central do contratualismo “[...] é a ideia de que justiça e moralidade são os princípios que podem e devem ser livremente aceitos entre iguais e sob condições definidas apropriadamente” (Freeman, 2006, p. 57). As teorias desenvolvidas por J. Rawls e T. Scanlon são exemplos de abordagens neocontratualistas, e são as teorias que estão servindo de modelo para a proposta de uma teoria mista.

princípios éticos e políticos a partir de algumas características centrais dos agentes, tais como coerência e razoabilidade. Assim, o neocontratualismo se opõe claramente ao antirrealismo, mas, ao mesmo tempo, não assume nenhuma forma tradicional de realismo, tal como o naturalismo ou intuicionismo. Isso conduz ao importante reconhecimento dos limites da razão, de forma que um comportamento desejável seria a razoabilidade, que é uma disposição de agir a partir do senso de justiça e reconhecer as limitações da razão para descobrir crenças morais verdadeiras, que seriam verificadas por sua correspondência ao mundo. Uma outra característica muito relevante é a sua epistemologia coerentista, propondo o procedimento de equilíbrio reflexivo para identificar a coerência entre juízos morais, princípios éticos e crenças científicas, sendo uma forma de justificação pessoal dos agentes. A ideia básica é que a justificação será obtida quando se encontrar um amplo sistema coerente de crenças, de forma que a reflexividade será a marca central do agente moral. Esse modelo se contrapõe tanto ao fundacionalismo quanto ao ceticismo moral e toma o conhecimento ético não como crença verdadeira justificada, mas como crença justificada em um equilíbrio reflexivo amplo, isto é, que é justificada por sua coerência com um sistema amplo de crenças, o que inclui princípios morais e crenças factuais.

Veja que, pelas características identificadas na teoria, pode-se reconhecer que o consenso é a base da justificação. Da mesma forma que estabelecemos a legitimidade política em democracias liberais contemporâneas, a saber, através do consentimento dos envolvidos – seja pelo do voto por sufrágio universal ou seja pela aprovação das leis via decisão parlamentar, por exemplo –, o neocontratualismo explica a legitimidade política com essa mesma base consensual, e toma o consenso como forma de se obter a justificação das crenças. Assim,

justificação e legitimação coincidem. E isso é importante para se entender a fonte da normatividade, que se encontra na própria capacidade reflexiva humana, ou seja, em sua autonomia, não sendo a obrigação derivada de um comando de alguém com autoridade legítima, tal como Deus ou Soberano, nem considerada como uma entidade intrinsecamente normativa ou como um fato que ela corretamente descreve. Assim, o agente pode se pautar por regras morais universais, tendo a capacidade de identificar essas razões morais e agir, ao menos moderadamente, a partir delas.¹⁵

Passo agora à análise da ética de virtudes. De forma geral, a ética de virtudes centra sua atenção no caráter do agente para a determinação do padrão de correção moral, que é o virtuoso, e isso porque se pressupõe que uma vida virtuosa auxiliará na felicidade pessoal e coletiva, o que implica um especial destaque ao processo de deliberação moral, na forma que é o próprio agente que deve encontrar a medida exata sobre o que deve ser feito, e isso levando em conta várias circunstâncias pessoais, como as emoções e crenças apropriadas, bem como considerando a própria natureza das virtudes. Isso é assim porque a ética trata de problemas que se caracterizam pela inexatidão, de forma que as questões sobre o justo e o bom são cercadas de incerteza e diversidade de opiniões. Por isso, o papel do

¹⁵ Korsgaard, em *The Sources of Normativity*, defende corretamente que é um fato incontestável da vida humana que nós temos valores e estabelecemos normas para a convivência, e que esses valores não são iguais a fatos, e que a autoridade dessas normas pode ser localizada em quatro fontes distintas, a saber, o voluntarismo, em que a obrigação é derivada de um comando, o realismo, que toma a obrigação como uma entidade intrinsecamente normativa, o endosso reflexivo, que considera a obrigação fundada na natureza humana, em especial, nos sentimentos e, por fim, o apelo à autonomia, de forma que a obrigação seria fundamentada na própria vontade do agente. Ver Korsgaard, 1996, pp. 18-20.

agente moral é pesar as razões morais e decidir sua ação a partir de uma escolha do melhor caminho a ser seguido. Isso pode ser formulado como segue:

Um ato X é correto ou justo sse ele for aprovado por um agente virtuoso, que é quem age virtuosamente, isto é, quem apresenta um traço de caráter na ação que é necessário para uma vida bem-sucedida.¹⁶

A partir dessa caracterização geral, penso que alguns elementos essenciais dessa teoria normativa se destacam. O primeiro elemento que quero chamar atenção é que ela faz uso de conceitos éticos *thick* e não de conceitos éticos *thin*. Os conceitos morais *thin*, como dever, correto ou bom, só possuem capacidade prescritiva, isto é, a capacidade de dizer como se deve agir, não tendo uma contraparte descritiva. Por sua vez, as virtudes são conceitos éticos *thick*, tais como coragem, justiça, lealdade, modéstia, entre outras, pois, além de prescreverem como se deve agir ou como um certo traço de caráter é desejável, elas possuem um importante elemento descritivo, considerando que cada virtude terá características específicas. Por exemplo, ser justo não é igual a ser corajoso, pois a primeira virtude implica ser uma disposição para dar a cada um o que é devido, enquanto a segunda é mais bem compreendida como uma disposição

¹⁶ Hursthouse, apresenta o modelo da ética de virtudes de forma similar. Para ela, uma ação é correta sse ela for o que um agente virtuoso faria em determinadas circunstâncias, sendo um agente virtuoso aquele que age virtuosamente, isto é, é aquele que tem e exercita as virtudes. E uma virtude é um traço de caráter que um ser humano precisa para florescer ou viver bem. Ver Hursthouse, 1997, p. 219. Foot, por sua vez, caracteriza a virtude por sua força corretiva, de forma a evitar alguma tentação ou deficiência de motivação e como tendo um valor moral positivo, de maneira que uma ação com valor moral positivo ou uma boa ação deve ser vista como uma ação de acordo com a virtude. Ver Foot, 2002, pp. 8-14.

para controlar o medo. A importância dessa distinção se dá porque a maioria dos autores influenciados pelo positivismo lógico defendeu que juízos morais seriam sem sentido, não podendo ser considerados verdadeiros ou falsos, pois os conceitos morais, como bom, correto, dever e obrigação, seriam apenas prescrições, sem nenhuma correspondência descritiva no mundo. E termos morais prescritivos seriam apenas uma questão de aprovação ou desaprovação subjetiva. Com a inclusão de termos morais *thick* no debate, a ética de virtudes nos oportunizou uma rota alternativa para defender a objetividade da linguagem moral.¹⁷

A segunda característica relevante da ética de virtudes é a de como ela pensa o conhecimento moral. Numa perspectiva aristotélica, a ética não tem a mesma exatidão do que a matemática, por exemplo. Ela é uma ciência, é claro, mas é prática, podendo determinar a verdade do que é justo e bom apenas de forma aproximada, o que implica a ideia de um agente moral que terá que deliberar bem para saber como viver/agir, não estando aberto para ele a possibilidade de descoberta de certos princípios éticos universais verdadeiros que lhes dirão sempre o que deve ser feito, tornando-o, assim, imune ao erro. Adicionalmente, a epistemologia das virtudes não trata o conhecimento como crença verdadeira justificada, uma vez que se pode chegar a esse alinhamento tripartido aleatoriamente, mas como

¹⁷ Na parte final do seminal artigo “Modern Moral Philosophy”, Anscombe defende que o padrão normativo das virtudes seria superior ao dos deveres, e isso porque ele nos evidencia as circunstâncias que devem ser refletidas pelo agente, apresentando uma parte descritiva, além da prescritiva. Por exemplo, as noções de “justiça” e “injustiça” seriam superiores às noções de “moralmente certo” e “moralmente errado”, pois determinam o que é justo e injusto em certas circunstâncias e, assim, saber-se-ia que é injusto tirar a propriedade de alguém sem o devido processo, bem como não pagar as dívidas e não cumprir as promessas. Ver Anscombe, 1958, pp. 15-19.

expressão de certas virtudes intelectuais dos agentes, tais como prudência, criatividade, curiosidade, humildade, entre outras. Assim, o conhecimento é mais bem entendido como uma performance bem-sucedida, sendo como acertar o alvo, o que é muito relevante em termos morais, pois a ação virtuosa é paradigmaticamente vista como aquela que encontra a mediania, que é uma espécie de alvo entre os extremos.¹⁸

Por fim, quero destacar o que creio ser o mais relevante na teoria das virtudes, a saber, que ela está alinhada com nossa linguagem cotidiana, pois, de fato, nós usamos a linguagem das virtudes em nossas vidas, e isso como uma maneira de censurar ou elogiar as pessoas por suas ações, atitudes e caráter. É corriqueiro censurarmos alguém por ser egoísta ou por cometer um certo ato cruel, da mesma forma que é usual elogiarmos alguém que é solidário ou que se preocupa com o bem-estar dos outros, demonstrando generosidade. Palavras como gratidão, justiça, solidariedade, resiliência, amizade, para exemplificar, fazem parte da forma como nós prescrevemos o comportamento moral em nossa sociedade. Inclusive a palavra caráter, na sua modalidade de bom e mau, faz parte do vocabulário comum, como uma forma de caracterizar e diferenciar as pessoas que se comportam corretamente das que se comportam moralmente de forma errada. Acredito que nenhuma teoria moral que se pretende bem-sucedida poderia subutilizar os critérios normativos das virtudes,

¹⁸ Linda Zagzebski, por exemplo, compreende o conhecimento nesses termos, como sendo um contato cognitivo com a realidade a que se chega por ações de virtudes intelectuais, que são, de fato, formas de virtudes morais, tais como a prudência, autonomia, perseverança, humildade, coragem, paciência etc. Ela define o conhecimento como um contato cognitivo com a realidade a que se chega por atos de virtude intelectual, ou como um estado de crença verdadeira, ou como um estado de crença, a que se chega por atos de virtude intelectual. Ver Zagzebski, 1996, pp. 270-271.

fazendo um uso majoritário dos princípios éticos, pois isso parece criar um afastamento da teoria em relação à realidade, o que não parece ser muito promissor. Estou pensando, sobretudo, no utilitarismo e no kantismo, como exemplo de teorias morais principialísticas que não dão a importância devida a esses critérios normativos referidos. É o caso, também, do neocontratualismo, mas de forma mais fraca, pois há a exigência de certas virtudes públicas, ao menos na teoria rawlsiana, como as de razoabilidade, tolerância, amizade cívica, entre outras. A ética de virtudes, ao contrário, já está sintonizada linguisticamente com a sociedade, o que parece preferível.

6. Considerações finais

Após ressaltar a força do neocontratualismo e da ética de virtudes, é importante, por fim, mostrar a coerência entre esses modelos morais. Em que pese essas teorias usarem critérios normativos distintos, a primeira fazendo um uso central de princípios éticos e a segunda utilizando as virtudes de forma hegemônica, pode-se reconhecer que o conteúdo normativo de ambas é muito semelhante. São centrais para essas abordagens a razoabilidade dos agentes como forma de reconhecer os limites racionais para a descoberta de princípios verdadeiros, o que já nos conduz à humildade ou tolerância como um comportamento desejável, bem como a coerência para alcançar uma boa deliberação, o que conduzirá a uma atitude de prudência e reflexividade do agente moral. Também, ambas as teorias tomam a justiça e a equidade como critérios normativos essenciais para determinar a distribuição dos bens, bem como a reparação do erro cometido (punição), isto é, para orientar as principais instituições políticas, sociais e econômicas da sociedade. Similarmente, as duas parecem se basear fortemente na autonomia

como fonte privilegiada da normatividade, seja para determinar/escolher princípios para orientação da ação, seja para determinar como o agente deve viver sua vida.

Com isso em mente, creio que é possível apostar na superioridade de uma teoria moral plurinormativa para melhor poder lidar com problemas práticos complexos, inclusive podendo contar com uma distinção entre moralidade privada e moralidade pública que possibilita enfrentar mais adequadamente o problema da assimetria normativa. É claro que resta ainda analisar se a proposta de uma teoria contratualista das virtudes não sofreria objeções relevantes, de forma a obrigar uma reformulação da teoria, ou mesmo um total abandono do esforço teórico. Mas isso já está além do escopo deste texto.

Referências

- Anscombe, G. E. M. (1958) Modern Moral Philosophy. *Philosophy: The Journal of the Royal Institute of Philosophy*, v. 33, n. 124, pp. 1-19.
- Coates, J.; Tognazzini, N. (2013) The Contours of Blame. In: Coates, J.; Tognazzini, N. (Eds.) *Blame: Its Nature and Norms*. New York: Oxford University Press, pp. 3-26.
- Darwall, S. (2013) *Morality, Authority, and Law: Essays in Second-Personal Ethics I*. Oxford: Oxford University Press.
- Foot, P. (1967) The Problem of Abortion and the Doctrine of the Double Effect. *Oxford Reviews*, v. 5. Reimp. In: Foot, P. (1978) *Virtues and Vices*. Oxford: Blackwell.
- Foot, P. (2002) Virtues and Vices. In: Foot, P. *Virtues and Vices and Other Essays in Moral Philosophy*. New York: Oxford University Press, pp. 1-18.

- Freeman, S. (2006) Moral Contractarianism as a Foundation for Interpersonal Morality. In: Dreier, J. (Ed.) *Contemporary Debates in Moral Theory*. Oxford: Blackwell, pp. 57-76.
- Haidt, J. (2012) *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York/Toronto: Pantheon Books.
- Hampshire, S. (1978) Public and Private Morality. In: Hampshire, S. (Ed.) *Public and Private Morality*. New York: Cambridge University Press, pp. 23-53.
- Hursthouse, R. (1997) Virtue Theory and Abortion. In: Crisp, R.; Slote, M. (Eds.) *Virtue Ethics*. Oxford: Oxford University Press, pp. 217-238.
- Korsgaard, C. (1996) *The Sources of Normativity*. New York: Cambridge University Press.
- McCormick, K. (2022) *The Problem of Blame: Making Sense of Moral Anger*. Cambridge: Cambridge University Press.
- McMahan, J. (2013) Moral Intuition. In: LaFollette, H.; Persson, I. (Eds.) *The Blackwell Guide to Ethical Theory*. Oxford: Wiley Blackwell.
- Mikhail, J. (2011) *Elements of Moral Cognition: Rawl's Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. New York: Cambridge University Press.
- Nagel, T. (1978) Ruthlessness in Public Life. In: Hampshire, S. (Ed.) *Public and Private Morality*. New York: Cambridge University Press, pp. 75-91.
- Nguyen, C. T. (2019) Self-Trust and Epistemic Humility. In: Wright, J. C. (Ed.) *Humility: The Virtues*. New York: Oxford University Press.
- Parfit, D. (2011) *On What Matters*, v. I e II. Oxford: Oxford University Press.
- Rainbolt, G. (2000) Perfect and Imperfect Obligations. *Philosophical Studies*, v. 98, n. 3, pp. 233-256.

- Rawls, J. (1971) *A Theory of Justice*. Original Edition. Cambridge, MA: Harvard University Press.
- Sayre-McCord, G. (2013) Contractarianism. In: LaFollette, H.; Persson, I. (Eds.) *The Blackwell Guide to Ethical Theory*. Oxford: Wiley Blackwell, pp. 332-353.
- Scanlon, T. M. (1998) *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Shweder, R. et al. (1997) The “Big Three” of Morality (Autonomy, Community, Divinity) and the “Big Three” Explanations of Suffering. In: Brandt, A.; Rozin, P. (Eds.) *Morality and Health*. New York: Routledge, pp. 119-169.
- Stocker, M. (1967) Acts, Perfect Duties, and Imperfect Duties. *The Review of Metaphysics*, v. 20, n. 3, pp. 507-517.
- Strawson, P. (2008) *Freedom and Resentment and Other Essays*. New York: Routledge.
- Thomson, J. J. (1976) Killing, Letting Die, and the Trolley Problem. *Monist*, v. 54, pp. 204-217.
- Thomson, J. J. (1985) The Trolley Problem. *Yale Law Journal*, v. 94, n. 6, pp. 1.395-1.415.
- Watson, G. (2004) *Agency and Answerability*. Oxford: Clarendon Press.
- Zagzebski, L. (1996) *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge: Cambridge University Press.

Parte 2

Questões metaéticas e normatividade

Razão prática e normatividade moral em Rainer Forst

Lucas M. Dalsotto¹
Felipe Taufer²

1. Introdução

O que se viu em filosofia política nos anos que se sucederam à publicação de *Uma teoria da justiça* (1971) foi uma onda sem precedentes de esforço intelectual em torno do problema da justiça. Vários filósofos (por exemplo, Dworkin, 1990; Nagel, 1991) se filiaram ao liberalismo igualitário de Rawls com o objetivo de aprimorá-lo e desenvolvê-lo. Mas, assim como atraiu adeptos, a teoria também foi alvo de importantes críticas. Enquanto Robert Nozick (1974) argumentou que o liberalismo igualitário de Rawls era insuficientemente liberal, Amartya Sen (1985) e Gerald Cohen (1992,

¹ Professor do departamento de filosofia, do programa de pós-graduação em filosofia e do programa de pós-graduação em direito da Universidade de Caxias do Sul (UCS).

² Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade de Caxias do Sul (UCS).

1995) o acusaram de ser insuficientemente igualitário.³ Outros teóricos (Sandel, 1982; MacIntyre, 1981) ainda se opuseram à teoria rawlsiana sob a alegação de que ela parte de uma posição ontológica *atomista* de pessoa, ignorando o fato de que os “indivíduos só poderiam crescer e se autorrealizar no interior de um contexto particular” (Gargarella, 2008, p. 144). Em *The Right to Justification*, Rainer Forst (2011) desenvolveu uma *teoria crítica da justiça* bastante original e ambiciosa ao incorporar elementos da ética do discurso habermasiana para fortalecer o que ele vê como limitante nas teses igualitárias de Rawls. Nos termos de Forst (2011, p. 5), uma concepção de justiça somente poderia ser razoavelmente justificada a partir da ideia de um *direito (moral) à justificação* cujo principal *bem* seria “o poder socialmente efetivo de demandar, questionar e oferecer justificações”. Os princípios e direitos resultantes dessa concepção renovada de justiça formam o que Forst (2011, p. 6), levando a teoria de Rawls mais além de seus limites, chamou de “construtivismo político”.

Forst discute temas que vão desde embates sobre a natureza da moralidade até teses sobre democracia, direitos humanos e justiça global. Mas um ponto do projeto que desperta particular atenção diz respeito à sua consideração da questão normativa fundamental de como um *dever de justificar* pode ele mesmo ser moralmente justificado, e à proposta de uma resposta à tal pergunta recorrendo ao conceito de *propriedades dialógicas* da razão prática. O problema que pretendemos aqui examinar é o de compreender em que medida Forst obtém sucesso na defesa da tese de que o caráter categórico da moralidade requer um fundamento incondicional sem incorrer nas

³ Note apenas que Sen e Cohen tomam caminhos diferentes para realizar essa crítica. O primeiro segue a via do utilitarismo de bem-estar e o segundo a via do marxismo analítico.

mesmas dificuldades enfrentadas por teóricos como Kant, por exemplo.⁴ A hipótese que dá suporte a este estudo pode ser dividida em uma parte positiva e outra negativa. A parte positiva é a de que o conceito de propriedades dialógicas da razão prática somente ganha plausibilidade quando combinado às noções de *segunda natureza*, de John McDowell (1995), e de *espaço das razões*, de Robert B. Brandom (1994), as quais permitem Forst defender o que gostaríamos de chamar de combinação entre um *externismo (reflexivo) de razões* e um *internismo motivacional*. Ela é positiva, pois essa combinação dá plausibilidade a sua proposta. A parte negativa é a de que tais noções apenas desempenham adequadamente seu papel sob um pano de fundo pragmatista no qual as teorias de McDowell e Brandom se desenvolvem, o que não está claro se Forst aceita. Esse lado da hipótese é negativo, pois aponta uma lacuna de explicitação na própria base do *construtivismo político* de Forst.

Para cumprirmos tal itinerário, inicialmente apresentaremos as linhas gerais do projeto de Forst e alguns de seus conceitos fundamentais (seção 2). Então, consideraremos a famosa objeção de Bernard Williams à ideia de *razões externas* e a resposta de Forst a ela (seção 3). Na sequência, analisaremos as dificuldades de tipo kantiano que Forst enfrenta ao defender o caráter categórico da moralidade e de que forma ele escapa delas apelando ao conceito de propriedades dialógicas da razão prática (seção 4). No entanto, para evidenciar a força da teoria de Forst, argumentaremos que tal conceito somente funciona adequadamente quando combinado às noções de *segunda natureza*, de McDowell (seção 5), e *espaço das razões*, de Brandom

⁴ *Grosso modo*, Kant foi acusado de endossar uma teoria metafisicamente inflacionada ao apelar a um fundamento *a priori* para a moral.

(seção 6) – as quais parecem ser por Forst (2011, pp. 279, 1110; 1111) pressupostas. Por fim, defenderemos que, embora essas noções deem robustez à proposta de Forst, o sucesso de sua teoria depende de uma adesão mais profunda e explícita ao pragmatismo daqueles autores (seção 7).

2. Do princípio (moral) à justificação

Forst (2011) parte da ideia de que a caracterização clássica de seres humanos como *animais racionais* evidencia que somos seres de *justificação*, ou seja, seres que *respondem* a razões e que podem ser *responsabilizados* por não agirem de acordo com elas. Razões não são propriedades privadas e devem ser publicamente acessíveis aos agentes. Elas constituem “um *jogo comum de dar e pedir razões*”, de modo que justificar uma ação não exige apenas avaliar se seus fins (e os meios para sua consecução) são racionalmente possíveis (Brandt, 1994, p. 171). Para Forst, justificar exige, além disso, uma verificação sobre se as razões oferecidas em suporte desses fins podem ser *razoavelmente* justificadas diante daqueles relevantemente afetados por tal ação.

É claro que a resposta à questão “O que devemos fazer?” não demanda, em todas as instâncias de ação possível, que razões morais sejam oferecidas. Se João pretende viajar para a Itália no próximo verão, então ele deve tomar posse dos meios necessários para a realização desse fim (e.g., comprar as passagens aéreas, reservar os hotéis nos quais permanecerá etc.). Esse é o *fundamento racional* da intenção de João. Caso ele tome esses meios e realize o fim desejado, sua ação pode ser caracterizada como fundamentada racionalmente. O que lhe é exigido é apenas que a razão explique a sua ação, e não que ela seja justificada intersubjetivamente. Tal ideia é mais bem

conhecida como o *princípio de racionalidade instrumental*. A grande maioria de nossas ações pode ser explicada a partir desse princípio. Porém, quando a questão “O que devemos fazer?” é feita em um contexto moral, enquanto *seres de justificação* que somos, exigimos uma justificação distinta dessa. É preciso “oferecer razões que possam justificar as ações de acordo com os critérios que são válidos no interior de um contexto moral” (Forst, 2011, p. 15). Forst chama isso de *justificação razoável*, cuja ideia básica é a de que tanto os fins quanto os meios para a realização de uma ação devem ser intersubjetivamente justificados. Uma ação *x* é razoavelmente justificada na medida em que nenhuma razão compartilhada conta contra a realização de *x*. Tal ideia pode ser identificada com o tipo de exigência que é feita pelo *imperativo categórico* kantiano, por exemplo.

O coração da proposta de Forst reside no conceito de *direito (moral) à justificação*, o qual tem sua origem no princípio fundamental da razão prática: o *princípio de justificação*. Ele demanda que as razões normativas oferecidas como solução aos problemas relevantes sejam justificadas em conformidade com as condições de validade de *reciprocidade* e de *generalidade*. Esses dois critérios constituem as razões que justificam as reivindicações morais dos agentes. De uma parte, o critério de reciprocidade afirma que ninguém pode fazer uma reivindicação moral que ele próprio negaria aos demais, isto é, não se pode reivindicar algo para si recusando-se a reconhecer que os outros tenham igual direito a fazer a mesma reivindicação. Esse critério veda o recurso a “verdades superiores” que não sejam compartilhadas ou aceitáveis para todos. De outra, o critério de generalidade diz que as razões devem ser compartilhadas por todos aqueles que sejam

relevantemente afetados por determinadas ações ou eventos (Forst, 2011).

Conforme o princípio de justificação razoável, “a característica definidora das razões que justificam reivindicações morais é a de que elas não podem ser geral e reciprocamente rejeitadas” (Forst, 2011, p. 21). Neste ponto, é importante ressaltar que Forst toma de empréstimo o princípio de *rejeitabilidade não razoável*, de Thomas Scanlon (1998), para mostrar que sua proposta tem uma vantagem explicativa em relação a uma teoria pura do consenso como a de Habermas (1984). A vantagem é evidenciada pelo fato de que os critérios de reciprocidade e de generalidade possibilitam que juízos sobre a justificação de reivindicações morais sejam feitos mesmo em casos de desacordo. Uma reivindicação moral justificada por razões gerais e recíprocas pode ser classificada como não razoavelmente rejeitável, ainda que um consenso não seja possível. Reivindicações normativas são sempre revisáveis, e sua validade depende de que nenhuma razão geral e recíproca lhes seja contraposta.

Dessa forma, razões morais que atendam aos critérios de reciprocidade e de generalidade “justificam ações que têm um caráter normativo moralmente incondicional e são, em um sentido estrito, categoricamente vinculantes” (Forst, 2011, p. 21). Razões desse tipo surgem no interior de práticas mútuas de justificação, e os agentes que as constituem são membros do *reino das razões*.⁵ Eles reconhecem a si mesmos como seres autônomos e morais que devem uns aos outros determinadas razões em defesa de suas ações. Esse processo de reconhecimento exige uma reconhecimento da capacidade humana à razão e de sua responsabilidade autônoma, o que constitui, no vocabulário

⁵ O que corresponde, em termos morais, ao conceito kantiano de *reino dos fins*.

de Forst (2011, p. 38), “um espaço de liberdade, um espaço de expectativas normativas e de razões justificatórias”.

Por isso, a teoria de Forst não está baseada em um fundamento *neutro*, “mas no princípio *moral* à justificação, isto é, no *direito moral individual* e substantivo à justificação” (2011, p. 5, grifos nossos). E é exatamente nesse ponto que reside um elemento central de sua proposta, pois mesmo teorias *pós-metafísicas* não podem se eximir do trabalho de prover um fundamento último sob o qual repousa todo o edifício teórico. O construtivismo moral e político de Forst possui um *fundamentum inconcussum* no sentido de que a capacidade humana de reconhecer, compreender e aplicar o princípio de justificação está fundada no conceito de razão prática. É a partir desse fundamento que é possível *construir* uma ideia substantiva sobre quais deveres e obrigações temos uns para com os outros, por exemplo. No entanto, recorrer à ideia de razão prática traz consigo dificuldades fundamentais: (1) ou assumir a incondicionalidade dos juízos morais e, dessa forma, correr o risco de ter que lançar mão de uma explicação transcendental do tipo do *fato da razão* kantiano; (2) ou aceitar um *internismo de razões* e, nesse caso, contradizer seus propósitos iniciais, já que o subjetivismo motivacional desse tipo de teoria recorreria a aspectos alheios aos procedimentos de justificação (e.g., “o conjunto motivacional do sujeito”, para usar a expressão de Bernard Williams (1981)).

Esse parece ser o tipo de dilema com o qual sua teoria se defronta. Na sequência, analisaremos as razões pelas quais Forst se recusa a tomar o caminho descrito em (2).

3. O externismo de razões de Forst e o desafio de Williams

Considerando as discussões contemporâneas sobre razões para ação, não encontramos dificuldades para dizer que a proposta de Forst está comprometida com o *externismo de razões*, cuja ideia básica é a de que a razão em seu uso prático “não deve apenas justificar a ação, mas também dirigir a vontade humana” (Forst, 2011, p. 18). Em geral, essa visão remonta a Kant (1997, pp. 19-20; MS Ak VI: 213), no sentido de que almeja renovar a seguinte posição: “o arbítrio humano [...] é um arbítrio tal que é certamente *afetado*, mas não *determinado*, pelos impulsos, e não é, pois, puro por si mesmo (sem uma prática adquirida da razão)”.

Uma das principais objeções ao externismo encontra-se representada na tese de Bernard Williams (2010, p. 35) de que “existem apenas razões internas à ação”. Forst (2011) dedica boa parte do primeiro capítulo de seu livro para a discussão sobre o que ele denomina de *neo-humanismo* de Williams, o qual afirma (i) que a *forma* da explicação da racionalidade de uma ação é internamente vinculada a algum elemento do conjunto motivacional do agente; (ii) que crenças falsas sobre o que é o caso ou crenças falsas sobre o que poderia satisfazer as motivações do agente não fornecem a explicação da racionalidade de sua ação; (iii) que é possível que um agente esteja errado sobre ter uma razão interna para fazer algo ou mesmo que desconheça as razões que possui para satisfazer suas motivações; e (iv) que processos deliberativos (coletivos ou individuais) podem revelar razões internas aos agentes que se engajam nesses processos (Williams, 1981). Em termos gerais, essas quatro proposições formam

o núcleo da posição de Williams, e as duas últimas são fundamentais para Forst.⁶

Williams estabeleceu uma dificuldade fundamental com a qual todo externista precisa lidar. A dificuldade é a de que “o conteúdo presente em um enunciado sobre razões externas deve ser esclarecido considerando o que significa este *passar a acreditar*” em razões para agir (Williams, 1981, p. 84). O autor alega que um *internismo de razões* é capaz de explicar esse fenômeno na medida em que processos deliberativos podem tanto subtrair elementos do conjunto motivacional de um agente, e, assim, mostrar-lhe que não possui uma razão para a ação que antes pensava ter, quanto revelar razões que o agente engajado em processos deliberativos não conhecia previamente. Mas será que esse “passar a acreditar” em uma razão para agir não “envolve essencialmente a aquisição de uma nova motivação?” (Williams, 1981, p. 84). Williams argumenta que, não fosse esse o caso, então o ônus de mostrar no que consistiria esse *processo puramente racional* recairia totalmente sobre os ombros do externista de razões.

O problema do “passar a acreditar” é descrito por Forst (2011, p. 23) como o da “conexão entre esses dois tipos de razões”, a saber, internas e externas. Apesar de distinguir entre *razões normativas* e *razões motivadoras* e, desse modo, filiar-se à tese externista ao admitir que “razões normativas nos dizem em que medida uma ação é permitida ou exigida”,⁷ Forst (2011, p. 23) diz que *esse nível de*

⁶ É relevante chamar atenção para o fato de que o uso dos termos *internismo* e *externismo* nos debates sobre a *natureza* das razões e sobre a *motivação* não é unívoco.

⁷ Sobre isso, Forst (2011, p. 24) ainda diz que: “na medida em que a tese externista não se refere primariamente à questão das razões motivadoras, mas à qualidade normativa das razões justificadoras, ela corretamente enfatiza a independência objetiva e intersubjetiva das razões justificadoras para ações particulares ou normas em

explicação (i.e., sobre a *natureza* de razões) é abandonado quando questionamos se uma pessoa tem *boas* razões para sua ação. Forst (2011, p. 23) afirma que “a tese internista [sobre a motivação – F.T. e L.D.] deve ser endossada na medida em que [...] uma ação moral deve ser praticada sobre fundamentos morais, ou seja, baseada em razões moralmente justificadas”. Esse *internismo motivacional* não é um *internismo de razões*,⁸ mas uma posição acerca da conexão entre dois tipos diferentes de razões (normativas e motivacionais). A tese de que as razões pelas quais uma ação é correta e de que as razões que motivam alguém *podem ser* as mesmas não implica a defesa da tese de que as propriedades de uma razão são constituídas apenas motivacionalmente (Korsgaard, 1986).⁹ No entanto, surge um problema: como Forst consegue compatibilizar um externismo de razões com um internismo motivacional?

Assim como Korsgaard (1986), Forst aceita as exigências que Williams faz de que o externismo de razões precisa explicar como um agente “passa a acreditar” que possui uma razão para agir. Por isso, Forst (2011, p. 24) sustenta que, dado que as pessoas são “razoáveis em um sentido prático, existe uma conexão interna entre razões e

comparação a motivos subjetivos, o que não requer explicar a ‘natureza’ das razões em termos de um realismo moral”.

⁸ Para uma análise detalhada sobre o uso dos termos *internismo* e *externismo*, ver Darwall, 1983 e 1992.

⁹ Essa distinção entre a natureza das razões e a natureza da motivação é feita por Christine Korsgaard (1986, p. 5) no intuito de mostrar que há posições que duvidam tanto de uma quanto de outra, e que não há uma conexão necessária entre as duas. Ela sustenta que a posição de Williams (1981) é a de que seu ceticismo sobre a racionalidade da motivação depende de seu ceticismo sobre a natureza das razões. No entanto, Korsgaard (1986, p. 23) pretende mostrar que, do modo como Williams argumenta, as coisas parecem ser o contrário: “como se o ponto motivacional – a exigência internista – possuísse alguma força na limitação do que pode contar como um princípio da razão prática”.

motivos”. Se os critérios externos e intersubjetivos de generalidade e de reciprocidade informam se uma razão para a ação pode ou não ser razoavelmente rejeitada, então eles também devem dizer como ela motiva alguém a agir e, por consequência, como explica a sua ação.

[D]e um lado, a perspectiva de cada pessoa deve ser capaz de fazer parte das justificativas morais e estar presente – de uma forma reflexiva correspondente – nas razões justificadas e, de outro, razões morais são, enquanto razões compartilhadas, intersubjetivamente válidas e, portanto, possuem um conteúdo objetivo. (Forst, 2011, p. 24).

Tudo indica que, para Forst (2011, p. 25), as proposições (iii) e (iv) da teoria de Williams devem ser incorporadas a um externismo de razões porque ajudam a mostrar como “as deliberações do agente podem ser crítica e inovadoramente aplicadas a certas motivações”. Entretanto, o internismo de Williams não é apenas uma teoria sobre como os processos deliberativos explicam o “passar a acreditar” de um agente e, por conseguinte, sua motivação para agir. Ele também é um internismo de razões na medida em que estende a explicação motivacional para as propriedades da razão. É exatamente por isso que Williams reconhece que, “para além dos desejos de uma pessoa, e o que quer que eles permitam ou exijam, não há um ‘espaço de razões’ que possua o seu próprio poder de validação” (Forst, 2011, p. 26). Isso porque, embora Williams defenda que os processos deliberativos revelam razões para a ação, ele acredita que um agente somente chega à conclusão de que possui uma razão para fazer algo “por meio de uma sólida rota deliberativa a partir das motivações que este já possui em seu conjunto motivacional atual” (Williams, 2010, p. 35).

A proposta de Williams permanece presa ao paradigma neo-humano de que são os desejos e outros estados conativos do aparato motivacional do agente que determinam o espaço do raciocínio. Trata-se de uma teoria sobre razões que não explica de modo adequado, especialmente em contextos morais e políticos, a diferença entre fundamento racional e justificação razoável. Isso cria duas dificuldades para os argumentos de Williams. Primeiro, que o fato de alguém já estar convencido a fazer algo não responde à questão de se essa ação pode ser moralmente justificada; e, segundo, que o processo de “passar a acreditar” e formar uma motivação para agir deve ter por base um conjunto de motivações *já existente*. Forst (2011, p. 27), então, conclui que o internismo de razões deve ser rejeitado, sob pena de ter que se negar a tese de que existe “um espaço intersubjetivo de razões justificadoras que permanece aberto *a todos* e que legitima ações tanto quanto reprovações” a partir do princípio de justificação.

Mas, se Forst apenas rejeita o internismo de razões, aceitando uma forma de internismo motivacional, o que ainda precisa ser explicado é como razões normativas podem alterar o aparato motivacional do sujeito. Forst (2011, p. 28) aceita a conclusão internista de que não há um processo puramente racional que legitime a classificação de certas condutas como irracionais. Contudo, aceitar isso significa meramente que não existem razões morais *a priori* que sejam independentes-de-agentes, o que não implica dizer que existem processos de deliberação em que razões construídas com base nos critérios de generalidade e de reciprocidade possam alterar a motivação do agente. Disso se segue que os agentes que não aceitam essas razões como guias de suas ações estariam sendo *irrazoáveis*, mas não propriamente *irracionais*.

Forst parece satisfatoriamente incorporar certas reivindicações do internismo de Williams a fim de compatibilizar um externismo de razões com um internismo motivacional. Ele também parece argumentar com sucesso acerca da insuficiência das objeções internistas contra a tese da incondicionalidade dos juízos morais, evidenciando, assim, por que (2) não deveria ser aceito. O que precisamos agora avaliar é se Forst é capaz de defender o caráter categórico da moralidade sem recorrer a argumentos transcendentais de tipo kantiano (e.g., a doutrina do *fato da razão*).

4. Um problema para o externismo de razões

Se a adesão de Forst ao internismo motivacional o ajuda a explicar como razões podem produzir uma *intenção* de agir de acordo com elas, sua aceitação do externismo evidencia como razões morais devem ser objetivamente compartilhadas (e.g., Darwall, 1983; Korsgaard, 1996). Contudo, a compatibilização entre os aspectos cognitivo e volicional da razão prática somente é possível sob o pano de fundo do seguinte pressuposto: “o reconhecimento de si e dos outros como pessoas morais razoáveis com um direito à justificação e um dever de justificar” (Forst, 2011, p. 31). Assim como o *fato da razão* kantiano serve como fundamento para a tese de que a vontade humana pode ser determinada tão só racionalmente,¹⁰ do mesmo

¹⁰ “A razão, da qual unicamente pode surgir toda a regra que contém necessidade, em verdade põe necessidade também nesse seu preceito (pois sem isso ele não seria nenhum imperativo) [...] Mas, para a sua legislação, requer-se que ela necessite pressupor simplesmente a si mesma, porque a regra só é objetiva e universalmente válida se vale independentemente de condições subjetivas e contingentes, que distinguem um ente racional de outro” (Kant, 2016, p. 35; KpV, AK V: 38-39). Para a crítica de Kant contra perspectivas que procuram fundamentar razões no aparato

modo o *reconhecimento* desempenha tal papel em Forst. Mas o externismo de razões defendido por este último não postula um fundamento independente das condições subjetivas que explicam as propriedades da determinação racional de nossa prática. Para Forst, as próprias propriedades dialógicas da razão prática são justificações razoáveis que atuam sobre a subjetividade motivacional de um agente, a qual, por sua vez, é mobilizada em contextos morais. Por isso, ele afirma que o fundamento da razão prática é a existência da “autocompreensão *prática* da moralidade” (Forst, 2011, p. 31). O que ainda carece de esclarecimento, porém, é saber se Forst obtém sucesso ao oferecer um *fundamento prático* em suporte de seu externismo de razões sem precisar se socorrer de argumentos transcendentais.

Em “A moralidade como um sistema de imperativos hipotéticos”, Philippa Foot (2022, p. 3) chamou atenção para o fato de que a filosofia moral de Kant está repleta de obscuridades, e que “poucos moralistas contemporâneos tentariam defendê-la de todas elas”. Uma das obscuridades é o estatuto especial que Kant atribui às propriedades das razões que constituem o *imperativo categórico*. Quando Kant (2016, p. 52; KpV, AK V: 55-56) defende a tese de que razões morais determinam a vontade de um agente “pela simples forma da lei”, isso pressupõe que as condições subjetivas sejam governadas inteiramente por normas constituídas por propriedades de um tipo muito especial. Vários estudiosos entenderam que Kant apelou à doutrina do *fato da razão* para explicar o caráter incondicional da moral. Todavia, contrapondo-se a essa interpretação, Onora O’Neill (2002, p. 81) tentou mostrar que tais características de regras não derivativas que

motivacional dos agentes, ver as Anotações I e II do § 3 da *Crítica da razão prática* (Kant, 2016, pp. 38-44; KpV, AK V: 41ss).

“possuem a forma da lei” não precisam ser interpretadas como se fossem propriedades de um tipo muito especial ou derivadas de um *fato* da razão prática. Isso porque, com esse conceito, Kant tem em vista a consciência comum que temos sobre a imperatividade das regras morais. Conforme O'Neill, há uma semelhança estrutural entre o modo ordinário da consciência de nossos juízos sobre a imperatividade de regras morais e a noção de autonomia. “Se a autonomia fosse uma *propriedade* das pessoas”, então, de fato, ela possuiria um estatuto muito especial. No entanto, ela é apenas uma propriedade da *forma da lei* e do caráter *não derivativo* das regras morais (O'Neill, 2002, p. 84). Mas alguém poderia ainda perguntar: se a propriedade que constitui essas razões não é uma faculdade especial das pessoas, que propriedade seria ela então? A despeito da resposta que possamos oferecer à questão, parece-nos ser prudente reconhecer que talvez essa seja uma obscuridade da filosofia moral kantiana que permanece ainda sem uma explicação satisfatória.

Esses problemas, no entanto, não são muito diferentes da busca por um *fundamento prático* que motiva o externismo de razões de Forst. Como Kant, Forst (2011, p. 32) também alega que as propriedades que constituem as razões normativas e o princípio de justificação “não podem ser derivadas de qualquer outro fundamento”. Por isso, qual seria a diferença entre um *fundamento racional* de tipo kantiano e um *fundamento prático* do tipo que Forst tem em mente¹¹?

¹¹ Em “Crítica e justificação em Rainer Forst”, Rúrion Melo (2013, pp. 25-26) emprega a expressão *fato da justificação* para se referir à ideia de que “a própria razão prática [...] se manifesta em decorrência da exigência de justificação intersubjetiva em contextos normativos”. Tal expressão parece reforçar a suspeita de que, para além de haver um paralelismo entre as propostas de Kant e Forst, o segundo pode incorrer nas mesmas dificuldades do primeiro ao tentar defender o caráter categórico e incondicional da moralidade.

Parece que, ao dizer que razões normativas não podem pressupor outra coisa que não a si mesmas, Forst incorre nos mesmos problemas que Kant quando este diz que a razão não pode pressupor nada além de si mesma. A passagem mais clara em que Forst (2011, p. 35) tenta fornecer uma saída e esclarecer o que entende por *reconhecimento* é quando afirma que a capacidade de um indivíduo ser uma pessoa moral “emerge em contextos nos quais esse indivíduo deve aos outros (geral e reciprocamente) razões justificatórias e [...] reconhece que (e como) ele é responsável [...] perante os outros, *sem* apelo a qualquer outra razão”. Forst (2011, p. 36) denomina isso de “um insight prático de segunda ordem” e diz que essa prática é “não metafísica, recursiva e imanente aos contextos de análise dos critérios de validação de razões morais”. Além disso, ele acrescenta uma descrição sobre a intersubjetividade das propriedades dessa razão ao esclarecer que o reconhecimento do direito à justificação só é possível “aos membros de um espaço de razões justificatórias compartilhadas (e comumente construídas)” (Forst, 2011, p. 38). Apesar de Forst não se comprometer com a obscura doutrina do *fato da razão*, ele não explicita *quais são* as propriedades que constituem a sua imagem renovada de razão prática.

5. As propriedades dialógicas da razão prática e a ideia de segunda natureza

Os agentes que participam do *jogo de dar e pedir* razões precisam reconhecer a si próprios e aos outros como pessoas moralmente razoáveis e portadoras de um direito à justificação. Se assim for, a questão “Por que ser moral?” não é uma que visa convencer o cético ou o amoralista a respeito do caráter vinculante da moralidade, “mas indicar em que sentido uma pessoa que raciocina praticamente tem

uma visão não apenas do ‘como’, mas também do ‘quê’ da justificação” (Forst, 2011, p. 32). De uma parte, essa questão já pressupõe a necessidade de uma razão para participar do jogo de dar e pedir razões, o que a caracteriza como imanente ao contexto pragmático da moralidade. De outra, a resposta a ela remonta à questão da fonte da autoridade acerca de quem pode racional ou razoavelmente justificar essa “primeira” razão. É sobre *o que a constitui* que o construtivismo de Forst parece enfrentar as mesmas dificuldades que Kant. Daí a afirmação de Forst (2011, p. 34) de que

o ponto central é apenas que o “fundamento” da moralidade não está na preocupação com a própria vida boa ou mesmo na preocupação em realizar uma visão do bem, mas, em vez disso, em uma preocupação com outros no interior do espaço do que é geral e reciprocamente devido [...].

A pergunta que surge é: “O que constitui esse espaço das razões?” ou “Que tipo de propriedades o explica?”. O argumento de Forst (2011, p. 37) consiste em dizer que o princípio de justificação “ficaria pairando no ar, por assim dizer, se uma reflexão recursiva desse princípio não fizesse parte de uma reflexão prática de segunda ordem sobre o dever de justificar e o direito à justificação”. Embora elimine boa parte das respostas metafísicas acerca da natureza desse espaço, o autor é pouco claro a respeito do que propriamente o constitui. Em uma breve passagem do texto, Forst (2011, p. 38) apenas diz que a determinação volitiva dos agentes que participam desse espaço remonta ao fato de que sua “capacidade para a razão se tornou uma *segunda natureza* deles”. Mas, sem uma explicação adicional, esse conceito pouco contribui para a compreensão do que seja o espaço das

razões. Por isso, a fim de esclarecer esse elemento, analisaremos os pressupostos pragmatistas desse conceito tematizado, originalmente, por John McDowell (2001).

Apesar de McDowell (1995; 2001) não ser um construtivista moral, sua tese pragmatista de fundo está comprometida com um externismo de razões. Em seu influente texto “Might There be External Reasons”, McDowell (1995) acusa Williams de um psicologismo na medida em que este adota o ponto de vista das razões internas para explicar razões que seriam externas em algum sentido relevante. O ponto de McDowell é o de que, se um internismo motivacional não consegue conceber nenhuma externalidade aos processos racionais, isso não se deve ao fato de que tais processos sejam constituídos por propriedades muito especiais como as de um *fato da razão*. Ao contrário, trata-se de um limite próprio das explicações que um internismo motivacional pode fornecer. McDowell (1995, p. 80) mostra que a externalidade das razões possui uma dignidade própria que lhe faculta um domínio de explicação particular, de modo a não ser preciso enfrentar “as dificuldades de se aceitar o tipo de transcendência” a que Kant parece recorrer. A validade externa de razões morais exige apenas a adoção do pressuposto pragmatista de que a verdade dos juízos de um determinado domínio depende somente dos padrões de correção do próprio domínio, visto que “razões morais externas não são elas próprias externas à moral” (McDowell, 1995, p. 80).

Com relação à especificidade de contextos normativos, a tese de McDowell é homóloga a de Forst (2011, p. 31) quando este reconhece o caráter autorreflexivo de cada contexto ao afirmar ter em vista uma “autocompreensão *prática* da moralidade”. A normatividade da prática humana do jogo de dar e pedir razões é autorreflexiva no

sentido de que ela envolve modos particulares de autocompreensão acerca da própria prática. A política, por exemplo, opera dentro de um esquema no qual seus conceitos não são externos à própria prática política, e a verdade de seus juízos depende apenas dos padrões de correção internos a esse domínio para tornar a prática em questão racionalmente inteligível. Isso vale para a moral e para qualquer outro âmbito prático de nossa vida social.

No caso da moral, as razões que constituem os padrões de correção e tornam os processos morais racionalmente inteligíveis são experienciadas em *segundo grau* pelos participantes da prática (McDowell, 1995). Isso significa dizer que, quando estamos em um contexto moral, não experienciamos a racionalidade de nossas ações e as orientações de valores, conceitos e razões que a fundamenta somente de uma forma causal ou psicológica. “Não concebemos nossos valores como devendo sua autenticidade ao nosso conjunto motivacional ou a qualquer outra coisa na gênese psicológica do processo de virmos a tê-los” (McDowell, 1995, p. 81). Esse processo autorreflexivo diz respeito ao fato de como assumimos o que *somos* em cada contexto normativo. Geralmente, assumimos que somos dotados de uma identidade moral e que o jogo de dar e pedir razões nesse contexto constitui a nossa própria natureza como seres racionais.

A noção de “verdade” que opera nos conceitos, razões e valores do contexto moral não requer nenhuma metafísica da razão no sentido kantiano de que haveria uma faculdade exterior constituída por propriedades muito especiais. McDowell (2001) recorda que, como seres humanos, não somos apenas *sujeitos de conhecimento*, mas também *agentes* que buscam realizar certos propósitos no mundo. Nossa forma de ser racional não diz respeito apenas a aspectos

cognitivos de nossa prática, mas também a aspectos volitivos dela. Contudo, não existe uma “vontade” exterior aos contextos particulares nos quais ela se manifesta. Há uma racionalidade no modo como nos autocompreendemos como seres dotados de vontade moral ou política, e a autorreflexão própria de cada contexto, que é constituída pelo jogo de dar e pedir razões, organiza nossa sensibilidade de um modo diferente do que quando o faz em nossas experiências causais (McDowell, 2001). Dito de outro modo, em contextos morais, a inteligibilidade não é algo que nos é imputado pelas coisas do mundo (i.e., a primeira natureza), nem constitui uma forma suprassensível de acesso a elas (i.e., metafísica). A inteligibilidade dessas razões, em vez disso, é inerente à perspectiva associada de quem compreende o “escrutínio reflexivo de uma prática moral” (McDowell, 2001, p. 190). Exatamente por isso, ela se estabelece como uma espécie de *segunda natureza* em oposição às coisas do mundo. Contudo, a objetividade de suas propriedades não é fundada independentemente da disposição racional dos seres humanos como sujeitos de conhecimento e agentes. Ela é uma forma de objetividade reflexiva: nós nos tornamos *sensíveis* e passamos a *ver* certas razões *porque* é isso que constitui aquele contexto normativo.

Se isso faz sentido, então os critérios de imanência pós-metafísicos não são violados, pois o conceito de segunda natureza de McDowell está comprometido com uma tese pragmatista a respeito do que significa ser autorreflexivo. Se há diferentes formas de sê-lo, cada contexto pragmático permite que os agentes constituam suas identidades práticas através de cada tipo específico de autorreflexão. É da *natureza* do jogo (moral) de dar e pedir razões que nos assumamos como pessoas dotadas de uma identidade moral. É aí que reside a objetividade desse contexto. Essas propriedades são

dialógicas e constituem uma espécie de segunda natureza em relação ao mundo natural porque qualquer autocompreensão prática não pode desconsiderá-las, sob pena de o agente estar negando sua própria identidade.

Nosso argumento aqui é o de que o conceito de segunda natureza de McDowell cumpre uma dupla função na teoria de Forst: (a) ajuda-o a adotar um externismo de razões sem que isso implique o abandono de seu construtivismo moral; e (b) evidencia a constituição reflexiva e o modo de experiência das propriedades dialógicas da razão prática. Todavia, ainda não está claro o que Forst (2011, p. 38) quer dizer quando afirma que somos membros de um “espaço de razões justificatórias”.

6. Tornando a proposta mais robusta: a ideia de espaço das razões

Se assumimos que somos seres de razão nesse sentido autorreflexivo, então a segunda natureza na qual uma comunidade moral se abre à prática da justificação constitui todos os seres humanos como “membros de um espaço compartilhado (e comumente construído) de razões justificatórias” (Forst, 2011, p. 38). Nele, nós: (i) consideramos a nós mesmos e aos outros como seres capazes de razão, o que significa que podemos não somente dar e pedir razões, mas também razoavelmente rejeitá-las ou aceitá-las a depender das justificativas apresentadas; (ii) somos autônomos porque experienciamos nossas ações de um modo que não é apenas causalmente determinado e porque são *nossas* as razões que podem ser fornecidas ou esperadas; e (iii) somos responsáveis porque podemos exigir uns dos outros que essas ações atendam às exigências do princípio de justificação. O que agora precisa ser esclarecido é como a proposta de Forst é capaz de explicar o *comprometimento*

prático que essas três determinações do espaço de razões expressam, e entendemos que a melhor forma de fazer isso é considerando a ideia de espaço das razões de Brandom.

A pedra de toque do projeto filosófico de Brandom (1994) reside na combinação da tese wittgensteiniana de que as normas são implícitas às práticas sociais com a tese sellarsiana de que somos seres responsivamente racionais. Segundo tal perspectiva, somos seres que seguem regras racionalmente, que fazem avaliações de adequação ou não adequação dessas regras uns em relação aos outros e que realizam ambas as ações se situando em um espaço de razões. Mas esse espaço em que somos responsivamente racionais, e no qual nossas ações e crenças são implicitamente estruturadas por normas, não possui qualquer estranheza metafísica.¹² Trata-se de um espaço intersubjetivo e socialmente determinado no qual damos e pedimos razões uns aos outros. Brandom (1994, p. 243) chama essa estrutura discursiva das razões para ação de “compromissos práticos” ou de “compromissos discursivos práticos”. E, a respeito da relação entre razão prática, ação e justificação de compromissos práticos, ele afirma o seguinte:

[a]ções são coisas que agentes fazem intencionalmente [...], e agir intencionalmente é produzir não inferencialmente uma performance que é tanto o reconhecimento de um compromisso prático (no caso de intenções na ação) quanto resultado de exercitar uma disposição diferencial confiável para responder a tal

¹² Ou, como Brandom (1994, p. 648) prefere dizer, “nós sempre estamos dentro do jogo de dar e pedir razões. Nós habitamos um espaço normativo, e é, a partir dessas práticas implicitamente normativas, que nós [...] avaliamos as adequações das aplicações de conceitos”.

reconhecimento (no caso de intenções prévias). O reconhecimento do compromisso prático pode ser pensado como as intenções com as quais a performance é produzida. (Brandom, 1994, p. 244)

Frequentemente, nossas ações são baseadas em razões sem serem intencionais. Isso ocorre porque muitas delas são estruturadas com base em razões internas, tais como quando coçamos a cabeça ou ligamos o interruptor “sem estar[mos - F.T. e L.D.] cientes dos compromissos que nos fornecem a razão que poderíamos citar” para explicar tais eventos (Brandom, 1994, p. 244). Segundo Brandom, toda intenção é uma razão, mas o contrário não é verdadeiro. O agir intencional é produzido no jogo de dar e pedir razões, e somente seres constituídos por uma segunda natureza são capazes de agir de tal forma. Quem age intencionalmente “assume um compromisso para agir que é especificamente uma responsabilidade *justificatória*” (Brandom, 1994, p. 244). As propriedades dialógicas que constituem a externalidade reflexiva das razões tornam-se evidentes quando consideramos nossa capacidade de justificar nossas ações perante uns aos outros e de exigir que os agentes reconheçam sua responsabilidade e o compromisso que têm de agir de tal e tal maneira.

No caso de Forst, os processos deliberativos realizados de acordo com o princípio de justificação são capazes de mudar as nossas intenções com base na responsabilidade justificatória que, na terminologia de Brandom, caracteriza as ações. Sempre que um pai exige de seu filho que este faça algo porque ele é seu pai, ou quando pessoas homofóbicas são repreendidas pelos juízos preconceituosos que emitem, as práticas sociais desses agentes não estão somente ocupando uma posição no espaço das razões constituído pela segunda

natureza, mas também respondendo racionalmente a certos compromissos práticos implícitos nela.¹³ Toda vez que ocupamos um lugar no espaço das razões, estamos reconhecendo os outros como seres capazes de razão, autônomos e responsáveis. A autoridade normativa dos compromissos práticos, no entanto, não se limita à sua constituição e responsividade racional em contextos pragmáticos. Ela também diz respeito ao fato de que razões para ação são *construídas* através de duas formas de reconhecimento: (a) de compromissos práticos; e (b) de podermos nos submeter a exercícios racionais que nos possibilitem reconhecer compromissos que *antes não reconhecíamos*.

Se isso faz sentido, então a tese de Brandom auxilia Forst a mostrar que seu internismo motivacional não apenas está comprometido com a tese de que razões criam e motivam desejos, mas também que estas podem ser construídas a partir do reconhecimento de compromissos práticos. Além disso, ele poderia defender que o princípio de justificação e os critérios de generalidade e reciprocidade são uma explicação de como os processos deliberativos sobre a justiça podem nos colocar em uma posição mais confiável¹⁴ para reconhecermos os compromissos que temos uns com os outros que antes não reconhecíamos. Forst (2011, p. 38) parece assumir esse pressuposto pragmático implicitamente na seguinte passagem:

¹³ Esta é uma suposição da tese pragmatista de Brandom de que nossos compromissos práticos não são totalmente autônomos em relação aos compromissos doxásticos de situações cotidianas, em que o senso comum é posto, a todo tempo, em questão. A respeito disso, ver Brandom, 1994, pp. 233-270.

¹⁴ Esse é um elemento importante da proposta na medida em que não fornece uma posição privilegiada ou externa à prática de justificação das razões que temos para agir.

O reconhecimento da liberdade humana acompanha o reconhecimento da capacidade humana para a razão e para a responsabilidade autônoma, e, nesse sentido, a autonomia é, em aspectos moralmente relevantes, vinculada não somente ao reconhecimento recíproco, mas também ao reconhecimento dos critérios de reciprocidade e generalidade. O “espaço de liberdade”, assim, não é “livre de normas”, mas um espaço de expectativas normativas e razões justificatórias.

Parece plausível afirmar que o reconhecimento dos critérios de reciprocidade e generalidade funciona como uma espécie de *disposição diferencial confiável*, a qual nos coloca em posição para reconhecermos razões para a ação que antes não reconhecíamos. Por conseguinte, o espaço das razões é um espaço de liberdade não apenas no sentido de que somos seres (i) dotados de razão e (ii) que experienciam de modo autônomo sua responsividade racional diante do mundo. (ii) ainda expressa que devemos ser capazes de conceber que as razões que passam a ser reconhecidas por essa disposição diferencial confiável são *nossas* razões. Reconhecemos que as ações que podem ser exigidas geral e reciprocamente são práticas sociais que dizem respeito à nossa segunda natureza, de modo que podemos ter expectativas e exigir ou fornecer justificativas com bases em nossas razões sobre as ações uns dos outros. Parece-nos que as ideias de *disposição diferencial confiável* e *reconhecimento de compromissos práticos* ajudam Forst a explicar o estatuto da construção de razões ao mostrar que as propriedades dialógicas da razão prática nos tornam seres responsáveis. Portanto, acreditamos que a relação entre compromissos práticos e a intersubjetividade humana fornece um

pressuposto pragmatista que pode dar maior robustez à proposta de Forst.

7. Considerações finais

Como vimos, a concepção de razão prática de Forst pressupõe um externismo de razões e um internismo motivacional. Ao aderir ao segundo, Forst oferece uma explicação convincente para contornar as críticas céticas de Williams. Mas isso apenas resolve o problema de que razões externas não precisam fornecer qualquer explicação estranha a respeito da motivação. Há ainda um risco real de Forst ter de assumir uma visão metafísica inflacionada para explicar seu externismo de razões, uma vez que ele não esclarece adequadamente os pressupostos que explicam as propriedades que constituem essas razões.

Com isso em mente, recorremos a duas proeminentes figuras da filosofia prática contemporânea a fim de esclarecer os pressupostos pragmáticos que poderiam ajudar Forst a realizar tal tarefa. De uma parte, buscamos defender que as propriedades dialógicas da razão prática não transcendem o domínio do mundo natural e formam a estrutura inteligível na qual o jogo de dar e pedir razões ocorre. Isso significa dizer que, se Forst realmente assume o pressuposto pragmatista do conceito de segunda natureza, tal como seu texto parece indicar, então o caráter autorreflexivo desse tipo de externismo o auxilia a mostrar que as propriedades dialógicas da razão prática não possuem qualquer estatuto metafísico estranho. De outra parte, procuramos explicitar o modo como essas propriedades são capazes de nos comprometer praticamente. Argumentamos que um uso mais explícito e detalhado do conceito de espaço das razões possibilitaria a Forst explicar como os critérios de razoabilidade nos colocam em uma

posição confiável, embora não privilegiada, para determinarmos racionalmente nossas ações. Aderir a essa ideia esclareceria como os comprometimentos práticos adquiridos em virtude da natureza do próprio processo deliberativo fazem a ponte entre o externismo reflexivo de razões e o internismo motivacional já endossados por Forst.

Contudo, até que ponto Forst é capaz de acomodar essas exigências sem gerar contradições internas é algo que exige uma investigação mais detalhada, a qual não poderemos realizar aqui.

Referências

- Brandt, R. (1979) Freedom and Constraints by Norms. *American Philosophical Quarterly*, v. 16, n. 3, pp. 187-96.
- Brandt, R. (1994) *Making TT Explicit*. Cambridge: Harvard University Press.
- Cohen, G. (1992) Incentives, Inequality, and Community. In: Peterson, G. (Org.) *The Tanner Lectures on Human Values*. Salt Lake City: University of Utah Press, pp. 263-329.
- Cohen, G. (1995) The Pareto Argument for Inequality. *Social Philosophy and Policy*, v. 12, n. 1, pp. 160-85.
- Darwall, S. (1983) *Impartial Reason*. New York: Cornell University Press.
- Darwall, S. (1992) Internalism and Agency. *Philosophical Perspectives*, v. 6, pp. 155-174.
- Dworkin, R. (1990) Foundations of Liberal Equality. In: Peterson, G. (Org.) *The Tanner Lecture on Human Values*. Salt Lake City: University of Utah Press, v. 11, pp. 1-119.
- Foot, P. (2022) A moralidade como um sistema de imperativos hipotéticos. *Filosofia Unisinos*, v. 23, n. 3, pp. 1-9.

- Forst, R. (2002) *Contexts of Justice: Political Philosophy Beyond Liberalism and Communitarianism*. Trad. John M. M. Farrell. Berkeley: University of California Press.
- Forst, R. (2011) *The Right to Justification: Elements of a Constructivist Theory of Justice*. Trad. Jeffrey Flynn. New York: Columbia University Press.
- Gargarella, R. (2008) *As teorias da justiça depois de Rawls: um breve manual de filosofia política*. São Paulo: Martins Fontes.
- Habermas, J. (1984) *The Theory of Communicative Action*, v. 1. Trad. Thomas McCarthy. Boston: Beacon.
- Kant, I. (1997) *Groundwork of the Metaphysics of Morals*. Trad. Mary Gregor. Cambridge: Cambridge University Press.
- Kant, I. (2016) *Crítica da razão prática*. Tradução Valério Rohden e Monique Hulshof. Petrópolis: Vozes.
- Korsgaard, C. (1986) Skepticism about Practical Reasons. *Journal of Philosophy*, v. 83, n. 1, pp. 5-25.
- Korsgaard, C. (1996) The Reasons We Can Share. In: *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, pp. 275-310.
- MacIntyre, A. (1981) *After Virtue: A Study in Moral Theory*. Notre Dame: University of Notre Dame Press.
- McDowell, J. (1995) Might There Be External Reasons. In: Harrison, R.; Altham, J. E. (Orgs.) *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press, pp. 68-85.
- McDowell, J. (2001) Two Sorts of Naturalism. In: McDowell, J. *Mind, Value, and Reality*. Cambridge, Massachusetts and London, England: Harvard University Press, pp. 165-197.

- Melo, R. (2013) Crítica e justificação em Rainer Forst. *Cadernos de Filosofia Alemã: Crítica e Modernidade*, v. 22, pp. 11-30.
- Nagel, T. (1991) *Equality and Partiality*. Oxford: Oxford University Press.
- Nozick, R. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- O'Neill, O. (2002) Autonomy and the Fact of Reason in the *Kritik der Praktischen Vernunft*. In: Höffe, O. (Org.) *Kritik der Praktischen Vernunft: Klassiker Auslegen*. Berlin: Akademie, pp. 81-97.
- Sandel, M. J. (1982) *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.
- Scanlon, T. (1998) *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Sen, A. (1985) *Commodities and Capabilities*. Amsterdam: North-Holland.
- Williams, B. (1981) Internal and External Reasons. In: Williams, B. *Moral Luck*. New York: Cambridge University Press, pp. 77-89.
- Williams, B. (2010) Internal Reasons and the Obscurity of Blame. In: Williams, B. *Making Sense of Humanity*. New York: Cambridge University Press, pp. 35-45.

Construtivismo normativo modesto: esboço de um projeto

Rafael Vogelmann¹⁵

1. Introdução: construtivismo modesto

O realismo normativo é a tese de que há verdades ou fatos sobre o que temos razão para fazer que são independentes de nós (usarei “fatos normativos” ou “enunciados normativos” para me referir a fatos ou enunciados sobre razões normativas para agir). Essas verdades são independentes de nós porque são independentes de todas as nossas crenças, atitudes ou práticas, reais ou hipotéticas. O construtivismo normativo, por outro lado, é a tese de que as verdades ou fatos normativos dependem de tais crenças, atitudes ou práticas e, como tal, dependem de nós.

Os construtivistas sustentam que as razões e o valor são “construídos” a partir da atitude de valorar. A formulação mais popular dessa tese é fornecida por Sharon Street. Ela sustenta que “a verdade de uma alegação normativa consiste no fato de que essa alegação se

¹⁵ Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal de Santa Maria.

segue, como uma questão lógica ou instrumental, do ponto de vista prático” do agente sobre qual juízo em questão se pronuncia (2012, p. 40). O ponto de vista prático do agente é constituído pelos juízos normativos com os quais ele está comprometido. E os requisitos lógicos e instrumentais em jogo são interpretados por Street como exigências que são constitutivas da própria atitude de valorizar algo e não pressupõem quaisquer valores substantivos.

Essa formulação superficial do construtivismo normativo será suficiente para nossos propósitos por enquanto. Com ela em vista, podemos distinguir entre construtivistas normativos kantianos e humianos (Street, 2012, p. 41). O que diferencia essas versões do construtivismo é sua resposta à seguinte questão: é possível derivar conclusões normativas substantivas, com conteúdo moral, da perspectiva do *raciocínio prático como tal*? A perspectiva do raciocínio prático *como tal* é a de um agente praticamente racional, abstração feita de qualquer compromisso normativo substantivo. Portanto, a perspectiva do raciocínio prático *como tal* é caracterizada simplesmente pelos requisitos procedimentais da racionalidade prática,¹⁶ isto é, pelos requisitos lógicos e instrumentais que são constitutivos da própria atitude de valorar (sejam eles quais forem). Street ocasionalmente se refere a isso como uma *caracterização formal* de um ponto de vista prático.

Os construtivistas kantianos dão uma resposta positiva a essa questão: dadas as exigências procedimentais que regem o raciocínio

¹⁶ Requisitos procedimentais são aqueles que ditam como devemos deliberar, mas não exigem o compromisso com nenhum valor em particular. Exemplos de requisitos procedimentais são a exigência de adequação entre meios e fins e a exigência de que os fins do agente sejam consistentes. Ver Parfit, 1997, pp. 101-102 e 2001, pp. 63-65, para uma apresentação da distinção entre requisitos substantivos e procedimentais de racionalidade prática.

prático, conclusões normativas substantivas e moralmente significativas decorrem de qualquer perspectiva prática, independentemente dos compromissos normativos que a constituem. Em outras palavras, há verdades normativas que podem ser construídas a partir de todas as perspectivas práticas possíveis. Os construtivistas humianos negam que conclusões morais substantivas decorram do ponto de vista do raciocínio prático como tal. De acordo com eles, a perspectiva do raciocínio prático como tal não implica nenhum compromisso normativo específico. Os compromissos normativos que se seguem da perspectiva prática de um agente são uma função de seus “pontos de partida avaliativos particulares, contingentemente dados” (Street, 2012, p. 41). O construtivista humiano aceita que pessoas que partem de compromissos normativos suficientemente diferentes podem não compartilhar nenhuma razão moral.

A principal representante da abordagem kantiana no debate atual é Christine Korsgaard. O objetivo declarado da terceira palestra de seu *The Sources of Normativity* é “mostrar que, se valorizamos algo, devemos reconhecer que temos obrigações morais” (Korsgaard, 1996, p. 96). Em seu trabalho mais recente, ela argumenta que o imperativo categórico é um princípio constitutivo da agência e, como tal, aplica-se a toda perspectiva prática, independentemente das atitudes e compromissos específicos que a constituem. Ela alega ainda que razões morais reconhecíveis podem ser derivadas do imperativo categórico como ela o entende (Korsgaard, 2009, pp. 25-27).

A versão humiana do construtivismo é defendida por Street. Ela argumenta que, da perspectiva prática como tal, nada se segue. Dela não se pode nem mesmo derivar verdades normativas egoístas (como a afirmação de que o agente tem razão para valorizar sua própria vida),

muito menos verdades morais (Street, 2012, p. 54). De acordo com o construtivismo humiano, toda verdade normativa é contingente e depende dos compromissos normativos substantivos do agente em questão.

Note que o debate aqui diz respeito ao escopo das razões morais. Os kantianos defendem uma combinação de absolutismo moral (a tese de que as exigências morais se aplicam a todos) e racionalismo moral (a tese de que todos a quem as exigências morais se aplicam têm razão para agir de acordo). Os kantianos identificam o conjunto de agentes para os quais as exigências morais são normativas com o conjunto de agentes racionais como tais. A abordagem humiana é incompatível com a conjunção de absolutismo e racionalismo, mas há duas maneiras de rejeitá-la. Um humiano pode alegar que as exigências morais se aplicam a todos (sustentando o absolutismo) e negar que todos aqueles a quem elas se aplicam têm uma razão para obedecê-las (ou seja, negar o racionalismo). Ou ele pode defender o racionalismo e negar o absolutismo, sustentando que todos a quem as exigências morais se aplicam têm razões para agir de acordo, mas que as demandas morais se aplicam apenas a um subconjunto de agentes racionais, a saber, aqueles que têm compromissos normativos substantivos que implicam preocupações morais. Em todo caso, os humianos restringem o escopo normativo da moralidade. Eles identificam o conjunto de agentes para os quais as exigências morais são normativas com o conjunto de agentes que têm compromissos normativos substantivos dos quais razões morais podem ser derivadas.

A distinção assim formulada reconhece dois possíveis fundamentos para razões normativas (Walden, 2018, p. 126). De um lado, há a razão prática *como tal*. Ou seja, a capacidade universal e

genérica de raciocínio prático compartilhada por todo agente praticamente racional concebível. E, de outro, há nossos juízos e atitudes normativas contingentes – os valores que acontece de reconhecermos.

Enquadrar o debate dessa maneira compromete o kantiano com um projeto muito ambicioso, que parece enfrentar o seguinte dilema. Se partirmos de uma concepção mínima de raciocínio prático, que inclua apenas compromissos que são constitutivos da atividade de raciocínio prático e que ninguém poderia inteligivelmente questionar, então é duvidoso que possamos derivar desse ponto de partida mínimo uma reivindicação de qualquer coisa como razões morais universais e substantivas. Como diz Schafer (2015), dado um ponto de partida minimalista, o procedimento de construção provavelmente produzirá resultados igualmente mínimos. Essa é a objeção humiana padrão ao construtivismo kantiano: ainda não se demonstrou que razões morais substantivas podem ser derivadas de um ponto de partida tão pobre. Por outro lado, se o kantiano apelar para uma concepção mais robusta de raciocínio prático, uma que incorpore compromissos que não sejam constitutivos da atividade de raciocínio prático como tal, então é mais provável que ele seja capaz de fornecer a derivação desejada de razões morais, mas ele pode ser acusado de ter abandonado o ambicioso projeto de derivar razões morais da razão prática como tal.

Embora eu acredite que a versão ambiciosa do construtivismo kantiano descrita por Street e associada ao trabalho de Korsgaard seja vítima desse dilema, isso não significa que a versão relativista e humiana do construtivismo vença a disputa. Uma versão mais modesta, mas, ainda assim, interessante, do construtivismo kantiano é possível, mesmo que os objetivos mais ambiciosos associados a essa

tese não possam ser alcançados. Considere uma teoria que parte de uma concepção dos agentes relevantes que vai além da concepção mínima e incontroversa de razão prática, mas sem caracterizá-los em termos de compromissos normativos substantivos. Se for possível mostrar que razões morais podem ser derivadas da perspectiva prática de qualquer agente desse tipo, seja lá quais forem seus compromissos normativos particulares, teremos um resultado significativo, mesmo que mais limitado do que aquele almejado por kantianos ambiciosos.

De fato, essa é uma interpretação possível da posição de Kant. De acordo com essa leitura, longe de partir de uma concepção incontroversa e mínima da razão prática, a tese de Kant sobre o domínio normativo começa com uma concepção não trivial da natureza da razão prática – uma concepção que incorpora requisitos normativos controversos (Schafer, 2015). Essa interpretação de razão prática incorpora a exigência de que os agentes, na medida em que sejam racionais, se esforcem para desenvolver um conjunto de compromissos práticos que sejam *sistematicamente justificáveis não apenas para o próprio agente, mas também para os outros*. Essa é uma exigência robusta e de forma alguma incontroversa. No entanto, mostrar que as razões morais podem ser derivadas de uma perspectiva prática assim caracterizada é um resultado importante.

Essa forma de *construtivismo kantiano modesto* ocupa um meio-termo entre o construtivismo humiano e a forma ambiciosa de construtivismo comumente associada a Kant. Ela concede aos humianos que razões morais não se seguem de todos os pontos de vista práticos concebíveis. No entanto, não concede a alegação de que o conjunto de agentes aos quais razões morais se aplicam e sobre os quais têm autoridade normativa se restringe ao conjunto de agentes que, por uma questão de fato contingente, se preocupam com fins

morais. Ela sustenta que razões morais podem ser derivadas de uma concepção mais robusta dos agentes relevantes que, embora não seja trivial, *não é* caracterizada em termos de compromissos morais substantivos. Se bem-sucedida, essa tese implica que o conjunto de agentes aos quais razões morais se aplicam é mais amplo do que o conjunto de agentes que de fato se importam com fins morais, mas não tão amplo a ponto de ser idêntico ao conjunto de todos os agentes racionais concebíveis.

A tese que defenderei é uma versão do construtivismo modesto. A concepção dos agentes relevantes da qual partirei, entretanto, não é a mesma de Kant. Concebo os agentes morais (ou seja, aqueles agentes para os quais as exigências morais são normativas) como aqueles que não são apenas dotados de razão prática (isto é, da capacidade para raciocínio prático), mas também seres que avaliam outros agentes em termos de virtudes e vícios (ou, para ser mais neutro, qualidades e defeitos de caráter). Meu objetivo é mostrar que, dadas algumas suposições razoáveis, certas verdades morais muito gerais, mas também centrais, podem ser derivadas do ponto de vista prático de qualquer agente assim caracterizado. Isso fica muito aquém da afirmação de que razões morais podem ser derivadas da razão prática como tal. Mas, se isso puder ser estabelecido, é um resultado significativo – e tão mais significativo na medida em que somos seres desse tipo.

Como o ponto de vista prático relevante aqui é o de um agente que avalia outros agentes em termos de virtudes e vícios, chamo essa tese de *construtivismo de virtudes*.

Resumidamente, o argumento a ser desenvolvido é o seguinte: (1) os enunciados normativos sobre razões que um agente está disposto a fazer refletem seus valores, e cada alegação sobre razões

normativas implica uma descrição parcial da virtude; (2) assumimos que a concepção de virtude de um agente reflete seus valores: padrões de comportamento que violam esses valores são considerados viciosos; padrões que sinergizam com esses valores são considerados virtuosos; (3) isso nos fornece um teste de coerência para padrões de valoração: um padrão de valoração que implica uma concepção de virtude, mas também implica um juízo normativo que, por sua vez, implica uma concepção incompatível de virtude, é incoerente; (4) é possível mostrar que padrões de valoração que violam exigências morais reconhecíveis (como a exigência de, pelo menos, uma forma fraca de reciprocidade) falham nesse teste. Ou seja, certas restrições morais reconhecíveis decorrem da perspectiva prática de qualquer agente praticamente racional que avalia outros agentes à luz de seus próprios valores.

2. Razões e virtude

A viabilidade da forma de construtivismo que defenderei depende da alegação de que há uma estreita conexão entre virtude e razões normativas. Isso pode parecer óbvio: se defendermos que o que devemos fazer é aquilo que temos mais razões (isto é, razões decisivas) para fazer e que devemos agir como uma pessoa virtuosa agiria, podemos chegar à conclusão de que o que temos razões decisivas para fazer é o que uma pessoa virtuosa faria se estivesse em nosso lugar.

Infelizmente, Bernard Williams mostrou que essa ideia simples é falsa. Ao considerar o que devemos fazer, uma coisa que devemos levar em conta são as maneiras pelas quais ficamos aquém da virtude. Por exemplo, um agente virtuoso exibiria temperança. Mas, se eu souber que sou intemperante e que posso ser incapaz de controlar meus apetites em determinadas circunstâncias, posso ter razões para

evitar me colocar nestas. E isso não é algo que um agente virtuoso faria ou teria razões para fazer (Williams, 1995, p. 190).

Outros exemplos do mesmo tipo são fáceis de produzir. Todos nós temos razão para melhorar a nós mesmos, mas somente porque somos imperfeitos. Uma pessoa perfeitamente virtuosa não trabalharia para melhorar a si mesma e não teria razão para fazê-lo simplesmente porque não tem defeitos de caráter para aprimorar. Ou considere a descrição de Watson (1975, p. 210) de um agente que acabou de perder um jogo amistoso de squash para um adversário irritante. Uma pessoa virtuosa demonstraria espírito esportivo e educação ao apertar calmamente a mão do adversário. O jogador de squash de Watson, no entanto, é muito irritável e, se tentasse cumprimentar o adversário, não conseguiria controlar sua raiva e acabaria batendo nele com a raquete. Por conta de sua cabeça quente, portanto, ele tem uma razão para ir embora sem cumprimentar o adversário – algo que o agente virtuoso não faria e não tem razão para fazer.

O que esses casos mostram é que razões podem ser baseadas em defeitos de caráter. E, se esse for o caso, a afirmação de que uma consideração nos fornece uma razão para agir de certa forma somente se uma pessoa virtuosa decidiria agir dessa forma à luz dessa consideração não pode ser verdadeira. Não nos é possível explicar as razões baseadas em defeitos de caráter em termos do raciocínio prático de um agente perfeitamente virtuoso que se encontra nas mesmas circunstâncias, simplesmente porque um agente perfeitamente virtuoso nunca se encontraria nessas circunstâncias.

Embora o argumento de Williams seja bem-sucedido em mostrar que não podemos conectar virtude e razões dessa maneira direta, ele não mostra que não há uma conexão interessante entre virtude e

razões. Defenderei que há, de fato, uma relação estreita entre virtude e razões normativas.¹⁷

A premissa básica do meu argumento é a afirmação de que, quando um agente não age como deveria, deve ser possível explicar por que ele não respondeu à razão que tinha para agir dessa forma, e essa explicação deve apontar para uma imperfeição no agente. Esta pode consistir no fato de o agente ignorar algum fato relevante que lhe dá razão para agir. Nesse caso, é possível dizer que ele é ignorante. Mas, se o agente não agiu como deveria, mesmo estando ciente dos fatos relevantes, então ele não é como deveria ser. A explicação da falha do agente em responder às suas razões deve mostrar que há algo errado com ele – que ele é, em certa medida, falho. É essa falha que explica o fato de ele não responder às suas razões. Em outras palavras, sempre que um agente deixa de responder a uma razão decisiva (embora esteja ciente dos fatos relevantes), deve se expor a uma crítica pessoal de certo tipo.¹⁸ Podemos formular essa premissa da seguinte forma:

(Princípio de crítica) Se um agente tem uma razão decisiva para agir, está ciente dos fatos relevantes e não responde a essa razão, então ele se expõe a alguma crítica pessoal que explica sua falha em responder a essa razão.

¹⁷ O argumento apresentado nesta e nas próximas duas seções é desenvolvido em maior detalhe no meu Vogelmann, 2024b.

¹⁸ O tipo de crítica em questão é direcionado à capacidade de raciocínio prático do agente – sua capacidade de tomar decisões e responder a razões. As críticas relacionadas a outros aspectos da personalidade do agente (como, por exemplo, a alegação de que ele é tedioso ou tem uma voz irritante) são irrelevantes.

Se ele está ciente de que tem uma razão decisiva, mas, mesmo assim, não responde a ela, ele é *acrático*. Se ele está ciente dos fatos que lhe dão razão, mas não reconhece essas razões, então se expõe a algum outro tipo de crítica pessoal. Mas que tipo?

A segunda premissa em meu argumento é a afirmação de que um agente *perfeitamente* virtuoso é irrepreensível. O agente perfeitamente virtuoso (seja lá o que isso for) é alguém que não exhibe o tipo de imperfeição que o exporia a uma crítica pessoal que poderia explicar uma falha em responder a uma razão normativa. Se alguém se expõe a essa forma de crítica pessoal, então há algum aspecto de si mesmo e de sua capacidade de pensamento prático em relação ao qual ele poderia melhorar. Portanto, ele não é *perfeitamente* virtuoso.

O princípio de crítica implica que, se um agente não se expõe a nenhuma crítica pessoal, então ele não deixa de responder a nenhuma razão decisiva. Ou seja, desde que não seja ignorante, esse agente sempre decide agir como deve agir e pelas razões apropriadas. De acordo com nossa segunda premissa, um agente perfeitamente virtuoso não se expõe a nenhuma crítica pessoal. Segue-se que um agente perfeitamente virtuoso e não ignorante sempre age como deveria agir e o faz pelas razões apropriadas.

Portanto, se um agente deve fazer algo, mas não responde à razão decisiva em jogo (embora esteja ciente dos fatos que lhe dão a razão), então ele não é perfeitamente virtuoso. Ou seja, o agente tem um caráter em alguma medida defeituoso. Isso significa que, quando uma pessoa deixa de responder a uma razão decisiva (e não ignora nenhum fato relevante), ela se expõe ao tipo de crítica que é direcionada ao seu caráter (assumindo que a alegação de que o agente é *acrático* também aponta para uma falha de caráter). Se alguém tinha uma razão para ser gentil com outra pessoa, mas não reconheceu essa razão, então é rude

ou talvez cruel, e isso explica o fato de não ter respondido à razão. Se alguém tinha uma razão decisiva para ajudar outra pessoa, mas não reconheceu essa razão, então é insensível ou egoísta. Se a pessoa tinha uma razão decisiva para perdoar alguém, mas não respondeu a essa razão, talvez ela seja vingativa. E assim por diante.

Deixando de lado, por enquanto, as razões que se baseiam em defeitos de caráter, temos que, se alguém tem uma razão decisiva para fazer A nas circunstâncias C, ele faria A em C caso fosse perfeitamente virtuoso. Supondo, por enquanto, que o mesmo ideal de virtude se aplica a todos, segue-se que todo agente que está em conformidade com o ideal de virtude faz A em C. Se alguém não faz A em C, é menos do que perfeitamente virtuoso – seu caráter é, até certo ponto, falho.

3. Um teste negativo para juízos normativos

Contudo, algumas razões são baseadas em falhas de caráter. O jogador de squash irritado de Watson tem uma razão para ir embora sem apertar a mão do adversário *porque* ele é uma pessoa irritada. Alguém que não fosse tão irritado e exibisse a virtude do espírito esportivo não teria essa razão. Portanto, sua razão se baseia no fato de que ele tem uma falha de caráter. Isso mostra que não podemos sustentar que o fato de um agente ter uma razão decisiva para fazer A nas circunstâncias C implica que isso é o que um agente perfeitamente virtuoso faria nessas circunstâncias. Tampouco podemos sustentar que o fato de que qualquer agente perfeitamente virtuoso decidiria fazer A nas circunstâncias C pela razão R implica que é isso que você deve fazer nessas circunstâncias.

Não obstante, permanece o fato de que, sempre que alguém deixa de agir de acordo com uma razão decisiva, demonstra certa falha de caráter. Se alguém tem uma razão decisiva para agir que se baseia

em um defeito de caráter e falha em responder a essa razão, isso revela uma imperfeição de caráter adicional. Se o jogador de squash do Watson não responder à sua razão para ir embora, ele demonstra ter outro defeito de caráter além de sua irascibilidade. O defeito que ele demonstra dependerá do modo segundo o qual ele falha em responder à razão. Se ele reconhecer a razão, mas, movido pela raiva, não responder a ela, então ele é *acrático* e não tem autocontrole. Se ele não reconhecer a razão, ele pode ser uma pessoa violenta ou talvez impetuosa, ou talvez lhe falte autoconhecimento em um grau reprovável. Essas falhas explicam por que ele não reconheceu a razão.

Agir de acordo com suas razões decisivas sempre mostra que a pessoa está mais próxima da virtude do que estaria se não agisse dessa forma. Não podemos, portanto, afirmar que um agente tem uma razão decisiva para agir de uma determinada maneira se não estivermos dispostos a sustentar que ele tem um defeito de caráter caso falhe em agir de acordo com essa razão. Dessa forma, as afirmações normativas sobre nossas razões restringem o espaço da virtude possível. Isso fica mais claro no caso de razões que não são baseadas em nenhum defeito de caráter, mas também é verdade no caso de razões que são baseadas em defeitos de caráter. Por falta de um termo melhor, vamos chamar as primeiras de *razões perfeitas* e as últimas de *razões imperfeitas*.¹⁹

Suponha que eu defenda que, nas circunstâncias C, eu tenha uma razão decisiva *perfeita* R para fazer A, digamos, ajudar uma pessoa

¹⁹ Essa distinção não corresponder à diferenciação entre deveres perfeitos e imperfeitos. Uma pessoa pode ter uma razão perfeita para cumprir um dever perfeito, como a razão para não matar. Pode-se ter uma razão perfeita para cumprir um dever imperfeito, como a razão para realizar um ato benevolente. Pode-se ter uma razão imperfeita para cumprir um dever imperfeito, como a razão para melhorar o próprio caráter. E pode-se ter uma razão imperfeita para cumprir um dever perfeito, como é o caso do jogador de squash irritado de Watson.

necessitada. Isso me compromete com a afirmação de que, se eu fosse perfeitamente virtuoso, eu faria A em C à luz de R, porque (i) R é uma razão que eu teria mesmo que fosse perfeitamente virtuoso (pois não se baseia em nenhum defeito de caráter) e (ii), sendo R uma razão decisiva, eu manifestaria uma falha de caráter se não respondesse a ela. Portanto, supondo que o ideal de virtude que se aplica a mim se aplica também aos outros, a alegação de que tenho uma razão decisiva perfeita para fazer A nas circunstâncias C (que podem, é claro, incluir informações sobre minhas preferências, desejos substantivos, relações pessoais etc.) implica que qualquer agente perfeitamente virtuoso que se encontrasse nas mesmas circunstâncias faria A. Se eu sustentasse que outro agente perfeitamente virtuoso poderia agir de forma diferente nas mesmas circunstâncias, isso implicaria que alguém poderia falhar em responder R sem se expor a nenhuma crítica direcionada a seu caráter. Ou seja, isso implicaria que R era, na melhor das hipóteses, uma razão suficiente, mas não decisiva, para fazer A.

Portanto, supondo que o mesmo ideal de virtude que se aplica a mim se aplica igualmente aos outros, ao afirmar que tenho uma razão *perfeita* e decisiva para agir de uma determinada forma, eu me comprometo com a afirmação de que qualquer agente perfeitamente virtuoso agiria da mesma maneira. Nesse sentido, as afirmações normativas para efeito de que temos razões perfeitas e decisivas para agir de certa forma restringem o espaço da virtude possível. Elas nos comprometem com a alegação de que um agente perfeitamente virtuoso agiria de uma maneira específica, restringindo, assim, o domínio dos possíveis modos pelos quais os agentes virtuosos podem raciocinar de forma prática e agir. Outro jeito de apresentar o mesmo ponto é dizer que toda afirmação sobre razões perfeitas pinta uma imagem parcial da virtude.

Isso significa que nosso ideal de virtude e nossos juízos normativos sobre razões perfeitas devem estar alinhados. Nossos juízos normativos sobre razões perfeitas devem estar de acordo com as possibilidades que nosso ideal de virtude reconhece. Portanto, nosso ideal de virtude impõe um limite aos juízos normativos que podemos aceitar. Se nosso ideal de virtude puder ser determinado independentemente de nossos juízos normativos, poderemos *testar* os juízos normativos contra nosso ideal de virtude. Se um juízo normativo restringir demais o espaço da virtude possível, excluindo modos de comportamento ou raciocínio prático que consideramos compatíveis com a virtude, então (supondo que não queiramos revisar nosso ideal de virtude) devemos rejeitar esse juízo. O mesmo pode ser dito se um juízo normativo não restringir suficientemente o espaço da virtude possível, admitindo nesse espaço modos de comportamento ou raciocínio prático que sejam incompatíveis com nosso ideal de virtude.

Para ilustrar o ponto, admita, por ora, nosso ideal moral de virtude e suponha que um agente perfeitamente virtuoso seja altruísta, justo, gentil, atencioso, honesto etc. Quaisquer juízos aceitáveis sobre razões decisivas perfeitas terão de ser compatíveis com esse ideal de virtude. Não se pode ter uma razão decisiva perfeita para ser cruel a fim de obter um ganho pessoal, por exemplo, pois isso implicaria que todos os agentes perfeitamente virtuosos agiriam dessa forma. Isso é muito restritivo, pois exclui do espaço da virtude possível agentes compassivos e bondosos, e muito permissivo, pois inclui no espaço da virtude possível agentes cruéis.

4. Razões imperfeitas

Juízos normativos sobre razões imperfeitas também são limitados por nosso ideal de virtude, mas não de maneira tão direta. Considere um agente que seja terrivelmente corrompido, insensível e cruel. Se o argumento da seção anterior estiver correto, então (assumindo nosso ideal moral de virtude) esse agente não pode ter uma razão perfeita para agir de forma cruel no intuito de obter ganhos pessoais. Ele poderia ter uma razão imperfeita para fazer isso? A resposta é não. Os casos em que um agente tem uma razão imperfeita são de dois tipos: (i) ele tem uma razão para se comportar de uma forma que um agente perfeitamente virtuoso não se comportaria porque se encontra em uma situação em que este último nunca se encontraria; e (ii), dados os defeitos de caráter do agente, a tentativa de agir como um agente perfeitamente virtuoso daria errado, resultando em uma ação que violaria um valor importante. Em ambos os casos, agir à luz da razão imperfeita é uma maneira de responder de forma apropriada a um valor que um agente perfeitamente virtuoso poderia compartilhar.

Valorizar algo envolve reconhecer razões para reagir de alguma forma positiva a este algo, e afirmar que algo tem valor nos compromete com a alegação de que essas razões são de fato boas razões para reagir dessa maneira (Street, 2012, pp. 43-44; Scanlon, 1998, p. 95). Por exemplo, se você valoriza a amizade, verá o fato de seu amigo precisar de ajuda como uma razão para ajudar. E, se você julga que a amizade é valiosa, então você vê esse mesmo fato como uma boa razão normativa para ajudar. Da mesma forma, se você acha que o fato de seu amigo precisar de ajuda é uma razão para ajudar, então você acredita que ele, sua amizade ou (muito provavelmente) ambos são valiosos.

Assim, sempre que um agente responde a uma razão normativa, ele está respondendo a um valor. Esse é o caso tanto com razões perfeitas quanto com razões imperfeitas. Considere o jogador de squash do Watson. Suponha que ele decida se afastar do adversário à luz da consideração de que, se ele se aproximasse, acabaria batendo nele com sua raquete. Se esse é o caso, então esse jogador valoriza a integridade física do adversário (e a valoriza mais do que as convenções sociais do esporte). Se ele de fato tem uma razão decisiva imperfeita para se afastar, então a integridade física de seu oponente é valiosa (e mais valiosa do que as convenções sociais que ele viola ao se afastar). É somente porque tentar agir como um agente perfeitamente virtuoso resultaria em uma violação desse valor que o jogador de squash tem uma razão imperfeita para agir de uma maneira diferente.

Portanto, assim como uma disposição para responder a razões perfeitas, uma disposição para responder a uma razão imperfeita é uma disposição para responder da maneira apropriada a algum valor. O que faz com que um agente tenha uma razão imperfeita para agir não é o fato de que o panorama de valores é diferente para um agente perfeitamente virtuoso e um imperfeito, mas sim o fato de que as possibilidades de ação são diferentes para o agente perfeito e para o imperfeito. Se um agente perfeitamente virtuoso estivesse na mesma situação que o jogador de squash (apenas livre de raiva), um curso de ação que respeitasse tanto o valor das convenções de espírito esportivo quanto o valor da integridade física do oponente estaria disponível para ele. O mesmo curso de ação não está disponível para o jogador de squash do Watson. E é isso que explica por que ele tem uma razão imperfeita para agir de outra forma.

Se razões imperfeitas correspondem a valores dessa forma, então elas também são restringidas por nossa ideia de virtude. Se algo tem valor apenas na medida em que temos razão para reagir positivamente a ele e agentes perfeitamente virtuosos sempre respondem corretamente a suas razões, então algo tem valor somente na medida em que valorizá-lo for compatível com a virtude perfeita. Os valores que fundamentam razões imperfeitas decisivas também devem ser compatíveis com a virtude. Eles devem ser valores que agentes perfeitamente virtuosos sustentam. Se o jogador de squash tem uma razão decisiva para ir embora, é porque o valor da integridade física de seu oponente supera o valor das convenções de espírito esportivo. Um agente perfeitamente virtuoso, portanto, valorizaria a integridade física do oponente em detrimento do espírito esportivo e reconheceria que o jogador de squash deveria, à luz desses valores, ir embora.²⁰

5. Um teste de coerência para padrões de valorização

Nossos juízos normativos sobre razões para agir refletem nossos valores. E eu argumentei que, se eu julgo que tenho uma razão decisiva perfeita R para fazer A, eu me comprometo com a afirmação de que, se eu fosse perfeitamente virtuoso, faria A em C à luz de R porque (i) R é uma razão que eu teria mesmo que não tivesse nenhuma falha de

²⁰ Essa tese pode parecer semelhante à compreensão de razões segundo o modelo de conselheiro de Smith (1995). Mas esse não é o caso. De acordo com Smith, o que temos razão para fazer é o que aconselharíamos nosso eu real a fazer se fôssemos totalmente informados e perfeitamente racionais do ponto de vista procedimental. O que eu defendo é que devemos sempre agir em resposta aos valores que um agente perfeitamente virtuoso poderia valorizar. O agente perfeitamente virtuoso em questão não é uma versão idealizada de mim mesmo, mas qualquer agente que satisfaça o ideal de virtude e seja, portanto, irrepreensível. Algo é valioso não porque nossos eus idealizados o valorizariam, mas porque valorizá-lo (e valorizá-lo até certo ponto) é compatível com o ideal de virtude.

caráter e (ii), R sendo uma razão decisiva, eu exibiria uma falha de caráter caso não respondesse. Supondo que o ideal de virtude que se aplica a mim se aplica igualmente aos outros, a afirmação de que eu tenho uma razão decisiva perfeita para fazer A em C implica que qualquer agente perfeitamente virtuoso que se encontrasse nas mesmas circunstâncias faria A. Da mesma forma, se eu julgo que qualquer agente perfeitamente virtuoso concebível faria A em C à luz de R, segue-se que, para qualquer pessoa P, R é uma razão decisiva para que P faça A em C, a menos que P tenha uma razão imperfeita suficiente para agir de forma diferente.

Suponha agora que os juízos de virtude de um agente reflitam seus valores: os padrões de comportamento que violam estes últimos são considerados viciosos; os padrões que sinergizam com esses valores ou os promovem de outra forma são considerados virtuosos. Isso parece plausível. Valorizar algo é se importar com esse algo. E, se você se importa com algo ou alguém, encontrará falhas e defeitos naqueles que o prejudicam e apreciará aqueles que o promovem ou protegem de alguma forma. O que estou supondo é simplesmente que uma pessoa avalia outros agentes à luz de seus próprios valores. O que mais poderia fornecer as bases para essa avaliação?

Esse conjunto de conexões conceituais entre valor, razão e virtude nos permite estabelecer um teste de coerência para valores. Nossos valores são refletidos tanto em nossa avaliação de outros agentes e seus traços de caráter (ou seja, em nossa concepção de virtude) quanto em nossos juízos normativos. Mas cada um de nossos juízos normativos nos compromete com uma concepção parcial de virtude. Se um padrão de valoração produzir simultaneamente uma concepção parcial de virtude e um juízo normativo que implica uma concepção conflitante de virtude, então ele é internamente

incoerente. Um padrão coerente de valoração deve produzir uma concepção de virtude e juízos normativos que resultem numa imagem compatível de virtude.

Considere, por exemplo, uma pessoa que valoriza seu bem-estar (ou sua felicidade, seus interesses, objetivos etc.) e não valoriza o bem-estar (ou a felicidade, os interesses, os objetivos...) dos outros. E suponha que ela endosse esse padrão de valoração – ela julga seu bem-estar valioso e que o bem-estar dos outros não importa. Se surgir uma oportunidade de promover seu próprio bem-estar à custa do bem-estar de outra pessoa, esse agente encontrará razões para aproveitar essa oportunidade (supondo que seus outros valores não sejam prejudicados em virtude dessa escolha). Assim, por exemplo, se ele se deparar com uma oportunidade de roubar sem ser descoberto, encontrará uma razão perfeita decisiva para fazê-lo (se não for esse o caso, então há algo mais que ele valoriza e que lhe dá razão para não roubar – como os interesses de outras pessoas, sua relação etc. –, mas estamos assumindo que não é esse o caso). Ainda supondo que essa pessoa admita que o ideal de virtude que se aplica a ela é o mesmo que se aplica a seus pares, esse juízo normativo implica que qualquer agente perfeitamente virtuoso roubaria nessas circunstâncias. Ela está comprometida com uma imagem de virtude segundo a qual os agentes virtuosos são *oportunistas*. Seus valores, entretanto, produzem uma imagem diferente da virtude. Como avaliamos outros agentes à luz de nossos valores, ela verá uma disposição para agir de forma a prejudicar seu bem-estar como um defeito de caráter, e disposições para agir de formas que concordam com seu bem-estar, como um bom traço de caráter, como uma qualidade. Suponhamos que outro agente a roube, prejudicando assim seu bem-estar. Ela está comprometida com o juízo de que esse agente agiu de acordo com a virtude. Ao mesmo tempo,

no entanto, à luz de seus valores, esse agente será percebido como um mau agente, como tendo manifestado uma falha de caráter (ou, pelo menos, como tendo agido como se tivesse uma falha de caráter). Há uma incoerência em sua perspectiva prática.

E não há como corrigir essa incoerência enquanto o agente oportunista mantiver o mesmo padrão de valoração. Enquanto ele valorizar apenas o seu bem-estar, ele continuará achando que tem uma razão decisiva para roubar. Se o agente oportunista pensasse que tem uma razão para não fazer isso, por exemplo, porque isso prejudicaria o bem-estar dos outros, então ele estaria reconhecendo que o bem-estar dos outros tem algum tipo de valor (pelo menos como um limite para suas ações). E, enquanto ele mantiver esse padrão egoísta de valoração, e enquanto avaliar outros agentes à luz de seus valores, ele continuará a ver aqueles que ofendem seu bem-estar como falhos. Não há como se esquivar aqui. O agente oportunista não pode escapar da incoerência simplesmente dizendo “tudo bem, então eu admito que a um agente que me rouba é possível ser perfeitamente virtuoso, mesmo que ele me pareça imperfeito”. O problema é que, pelo menos sob a suposição de que avaliamos outros agentes à luz de nossos valores, ele deve ver o agente como falho – essa é a avaliação que faz sentido à luz de seus valores, daquilo que importa para ele.

A incoerência aqui não é uma incoerência em juízo, mas uma incoerência interna aos valores do agente. A única maneira pela qual esse agente é capaz de evitar a incoerência é modificando seu padrão de valoração. O agente oportunista pode sustentar que seu bem-estar não tem valor ou que o bem-estar dos outros tem algum tipo de valor que lhe dá razões para não os prejudicar a fim de promover seus próprios objetivos. Se ele não valorizasse seu próprio bem-estar, o infrator não lhe pareceria falho. Mas ele também não encontraria

razões para roubar no intuito de promover seu próprio bem-estar. Se ele valoriza seu bem-estar, ele verá o infrator como falho e, então, só poderá restaurar a coerência atribuindo valor ao bem-estar dos outros (como um limite para suas próprias ações). Se esse for o caso, ele se comprometerá com a afirmação de que o valor do bem-estar dos outros lhe dá pelo menos uma razão perfeita e decisiva para não roubar e, conseqüentemente, com a alegação de que o agente virtuoso não roubaria dos outros para promover seu próprio bem-estar.

Um argumento semelhante pode ser construído com relação a outros valores. Se uma pessoa valoriza sua vida e integridade física, ela verá um agressor como vicioso. Se essa mesma pessoa não valoriza a vida e a integridade física dos outros, ela encontrará razões para atentar contra a vida ou a integridade física dos outros quando puder ganhar algo com isso. Supondo que o ideal de virtude que se aplica a ela seja o mesmo que se aplica a seus pares, isso a compromete com a afirmação de que é assim que um agente virtuoso se comportaria. Esta pessoa está comprometida com uma imagem de virtude segundo a qual os agentes virtuosos são *predatórios*. Mas isso está em contradição com sua avaliação do agressor. Ela pode corrigir essa incoerência reconhecendo que a vida e a integridade física de outras pessoas têm algum valor que lhe dá razões para não atentar contra elas. Vamos nos referir a esse tipo de valor como *valor negativo* – ele é negativo porque dá uma razão para não violar o valor, em vez de uma razão para promovê-lo ou se preocupar positivamente com ele. Se o nosso agente encontra um valor negativo na vida e na integridade física dos outros, precisa rejeitar o juízo normativo segundo o qual ele pode ter uma razão decisiva para atentar contra a vida ou a integridade física dos outros para obter lucro. Ele deve sustentar que tem uma razão perfeita e decisiva para não fazer isso. Supondo que o

mesmo ideal de virtude se aplique a todos, segue-se que um agente virtuoso não atentaria contra a vida ou a integridade física de outros para obter uma vantagem pessoal.

O mesmo entendimento vale para o valor que alguém atribui à sua liberdade, à sua virtude, às suas relações e assim por diante. Supondo que o mesmo ideal de virtude se aplique a todos, se alguém valoriza essas coisas, então deve atribuir pelo menos um valor negativo à liberdade, à virtude e às relações dos outros. Isso significa que devemos ver razões para não violar esses valores, mesmo que isso promova os valores que importam para nós.

Se isso estiver correto, então as conexões que identifiquei entre valor, razão e virtude (da perspectiva de um ser que avalia o caráter dos outros) impõem uma certa estrutura a nossa perspectiva prática. No mínimo, essa estrutura exige uma *forma fraca de reciprocidade*. Se eu encontro valor em minha vida, minha felicidade, meus objetivos, minha liberdade, minha virtude, minhas relações etc., então eu deveria pelo menos encontrar um valor negativo na vida, na felicidade, nos objetivos, na liberdade, na virtude, nas relações etc. dos outros. Caso contrário, minha perspectiva prática será incoerente.

Essa proposta permite compreender a regra de ouro como uma estratégia de reflexão prática. Quem viola a exortação “faça aos outros o que gostaria que fizessem a você”, e faz aos outros o que não gostaria que fizessem a ele, incorre em uma forma de incoerência. Não porque haja algum tipo de incoerência bruta em tratar os outros de uma forma que você não gostaria de ser tratado, mas porque os valores desse agente (ou, pelo menos, os valores que são exibidos em sua ação) implicam juízos incompatíveis sobre virtude e vício. Ele vê como virtuoso um comportamento que não é capaz de assumir como virtuoso quando é sua vítima. O mesmo pode ser dito da injunção

“pense como você se sentiria se alguém lhe fizesse o que você está planejando fazer”. Essa é uma estratégia viável de reflexão prática, não porque haja uma incoerência bruta em fazer algo que faria você se sentir mal se fosse o alvo dessa forma de comportamento, mas porque pensar em como você se sentiria é também pensar em como perceberia o agressor. Se você sentisse raiva, ressentimento, medo ou repulsa em relação ao agressor, então ele lhe pareceria imperfeito. Isso significa que, à luz de seus valores, você não conseguiria encontrar virtude no comportamento dele. E isso contradiz as consequências do juízo normativo que você defende quando afirma que tem razões para se comportar dessa maneira.

6. Conclusão

Se essa forma de construtivismo modesto puder ser defendida, então esse é um resultado significativo. Ele mostra que, para seres que avaliam o caráter dos outros à luz de seus valores, certos padrões de avaliação tipicamente imorais são internamente incoerentes.

Certamente, entretanto, essa abordagem fica aquém do que um construtivista kantiano ambicioso esperaria. Ela é duplamente modesta. Primeiro, não implica que qualquer ser dotado de razão prática esteja comprometido com valores morais. Em segundo lugar, ela não fornece uma derivação completa da moralidade nem mesmo do ponto de vista do tipo de agente com o qual o construtivismo modesto se ocupa. Não há garantia de que todo valor moral possa ser justificado da maneira descrita aqui. Alguns sim, como os valores negativos discutidos. Mas é possível muito bem haver outros valores morais que podem ser rejeitados sem incoerência. O que o construtivismo modesto permite defender, portanto, não é uma objetividade moral forte. Em vez disso, se for bem-sucedido, ele

mostra que há alguns valores ou razões morais que são passíveis de ser derivados da perspectiva de um agente que avalia os outros à luz de seus valores. É possível que esses valores sirvam como pontos fixos morais (Cuneo; Shafer-Landau, 2014) para o equilíbrio reflexivo. Mas pontos de vista morais muito diferentes podem ser desenvolvidos em torno desses pontos fixos compartilhados. Portanto, o construtivismo modesto é compatível com uma boa dose de relativismo moral, mesmo quando insiste que alguns valores morais centrais não podem ser rejeitados (por um tipo específico de agente) sob pena de inconsistência.

Além disso, o argumento aqui se baseou em várias suposições controversas. Presumi, sem argumentar, que um agente avalia seus pares à luz de seus valores de forma bastante direta. E que, qualquer que seja o ideal de virtude de uma pessoa, ela deve admitir que o mesmo ideal de virtude se aplica igualmente a ela e aos demais. Essas premissas exigem uma defesa e só podem ser efetivamente defendidas caso sejam significativamente qualificadas. Talvez de maneira surpreendente, acredito que elas possam ser defendidas apelando para uma forma de sentimentalismo sobre juízos de virtude. A discussão dessa possibilidade exige um outro artigo, mas um exemplo servirá para ilustrar o resultado de combinar a tese aqui defendida sobre a conexão entre juízos normativos e virtude com uma forma de sentimentalismo.

Considere a forma de sentimentalismo defendido por Hume. Segundo o filósofo, a virtude é a capacidade de um agente ou traço de caráter de produzir em nós aprovação (Hume, 2000; 3.3.1.3). E a aprovação consiste em um tipo particular de sentimento. Segundo Hume, quando consideramos um traço de caráter, podemos conceber como um agente dotado desse traço afeta as outras pessoas ao seu

redor. Se imaginamos que esse traço de caráter produz prazer nos indivíduos afetados, é-nos possível partilhar desse prazer por meio de uma reação empática. A aprovação, segundo Hume, consiste justamente nesse prazer resultante da empatia por aqueles afetados pelo traço de caráter. Isso significa que só é factível experimentar aprovação em resposta a traços de caráter que sejam úteis ou agradáveis aos demais e que, portanto, um agente virtuoso é sempre útil ou agradável aos outros. Isso significa que um agente virtuoso não pode, por exemplo, ser cruel ou frio. Dada a conexão entre juízos normativos e virtude defendida aqui, segue-se que ninguém é capaz de ter uma razão perfeita para agir de maneira fria ou cruel. Em geral, modos de comportamento que produzam dor e sofrimento para os demais não podem produzir aprovação e, portanto, não são compatíveis com a virtude. Se esse é o caso, a ninguém é possível ter razão perfeita para agir dessa maneira. Consequentemente, todos devem reconhecer ao menos um valor negativo no bem-estar alheio.

O sentimentalismo de Hume é problemático por várias razões. Em particular, ele traça uma conexão muito estreita entre juízos avaliativos e nossas respostas afetivas.²¹ Contudo, o ponto aqui não é defender essa versão particular de sentimentalismo, mas ilustrar o fato de que, ao contrário do que alguém, por vezes, esperaria, a combinação de uma forma modesta de construtivismo normativo com uma forma de sentimentalismo sobre a virtude pode levar não a um relativismo desenfreado, mas à defesa do tipo de ponto fixo moral que discuti aqui.

²¹ Em Vogelmann (2024a), detalho meu entendimento da relação entre juízos avaliativos e reações afetivas.

Referências

- Cuneo, T.; Shafer-Landau, R. (2014) The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies*, v. 171, n. 3, pp. 399-443.
- Hume, D. (2000) *Tratado da natureza humana*. Trad. Déborah Danowski. São Paulo: Editora Unesp.
- Korsgaard, C. (1996) *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Korsgaard, C. (2009) *Self-Constitution*. Oxford: Oxford University Press.
- Parfit, D. (1997) Reasons and Motivation. *Proceedings of the Aristotelian Society*, v. 71, pp. 99-146.
- Parfit, D. (2011) *On What Matters*, v. 1. New York: Oxford University Press.
- Scanlon, T. M. (1998) *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Schafer, K. (2015) Realism and Constructivism in Kantian Metaethics 1: Realism and Constructivism in a Kantian Context. *Philosophy Compass*, v. 10, n. 10, pp. 690-701.
- Smith, M. (1995) Internal Reasons. *Philosophy and Phenomenological Research*, v. 55, n. 1, pp. 109-131.
- Street, S. (2008) Constructivism About Reasons. In: Shafer-Landau, R. (Org.) *Oxford Studies in Metaethics*, v. 3. Oxford: Oxford University Press, pp. 207-246.
- Street, S. (2010) What Is Constructivism in Ethics and Metaethics? *Philosophy Compass*, v. 5, n. 5, pp. 363-384.
- Street, S. (2012) *Coming to Terms with Contingency: Humean Constructivism About Practical Reason*. In: Lenman, J.; Shemmer,

- Y. (Orgs.) *Constructivism in Practical Philosophy*. Oxford: Oxford University Press, pp. 40-59.
- Vogelmann, R. (2024a) Enactive Evaluative Sentimentalism. *Revista Kriterion*, v. 65, n. 157, pp. 1-26.
- Vogelmann, R. (2024b) Squeezing the Good into the Right: The Connection Between Virtue and Reason. *Filosofia Unisinos*, v. 25, n. 3, pp. 1-12. DOI: 10.4013/fsu.2024.253.10
- Walden, K. (2018) *Practical Reason Not as Such*. *Journal of Ethics and Social Philosophy*, v. 13, n. 2, pp. 125-153.
- Watson, G. (1975) Free Agency. *Journal of Philosophy*, v. 72, pp. 205-220.
- Williams, B. (1995) *Replies*. In: Altham, J.; Harrison, R. (Orgs.) *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*. Cambridge: Cambridge University Press, pp. 185-224.

Normatividade imperiosa da moral e culpabilização: uma leitura inferencialista

Rogério A. Picoli¹

1. Introdução: a normatividade imperiosa da moral

Aquilo a que nos referimos, num sentido amplo, como a normatividade prática diz respeito ao caráter orientador e sancionador das normas e a critérios que compõem os nossos diferentes sistemas de regulação de conduta. Essa normatividade está associada à força prescritiva e à autoridade que as normas exercem sobre as nossas ações, decisões, apreciações e crenças, mas também sobre aquilo em função do que tais ações, decisões, apreciações e crenças são consideradas e avaliadas como apropriadas ou inapropriadas, corretas ou incorretas. Aqui, vou assumir que normas morais são tanto critérios direcionadores, orientadores, quanto critérios de apreciação e avaliação moral.

¹ Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal de São João del-Rei.

O foco deste texto é a normatividade prática específica da moral. As considerações que se seguem estão circunscritas ao domínio da metaética em vez do domínio da metanormatividade. Pretendo examinar o aspecto da moralidade relacionado ao que devemos ou não fazer, certas obrigações e proibições morais, bem como à força, à autoridade e à exigência especiais que certas regras morais nos impõem. Vou denominar esse aspecto, que é notadamente destacado em determinadas regras morais, de “a normatividade imperiosa (*authoritative*) da moral”. De acordo com Howard e Laskowski

[...] razões para não dirigir embriagado exibem uma normatividade “robusta” ou “genuína”, enquanto razões para não mover seu peão três casas exibem uma normatividade “meramente formal” ou “genérica” [...] chamaremos isso de distinção entre normatividade imperiosa (*authoritative*) e não imperiosa (*non authoritative*). (Howard; Laskowski 2025)²

Assim, vou usar a expressão “normatividade imperiosa da moral” para me referir à força especial de certas normas morais. Em particular, à força contrastante com uma noção fraca ou ordinária de normatividade que normalmente associamos, como nas regras do xadrez, a meras recomendações, sem que estas impliquem exigência, obrigação ou força compulsoriamente vinculantes. É por conta desse aspecto compulsório de certas obrigações morais que empregamos a expressão “normatividade imperiosa (*authoritative*)”.³

² Exceto quando indicado em contrário, todas as traduções neste texto são de minha responsabilidade.

³ A expressão “normatividade robusta” tornou-se corrente graças à contribuição de David Plunkett (2019). Tendo em vista que em português o termo “robusto” remete a

Retornando à normatividade imperiosa da moral, o ponto é que algumas normas vão além da mera orientação, recomendação e direcionamento. Por exemplo, no jogo de futebol, há regras que vetam certos tipos de conduta, e a gravidade das sanções que acompanham as suas violações é indicadora de uma espécie de normatividade forte. No entanto, parece que uma categoria de proibição séria, uma categoria de autoridade especial, envolve a regra de que “o contato entre os jogadores não pode envolver agressões desleais e intencionalmente incapacitantes”. Esse é um tipo de proibição que se coloca para além dos demais.

A despeito de o inferencialismo de Robert Brandom ter sido pensado como uma teoria normativa para explicar o significado de expressões e de conceitos e, assim, ser uma abordagem orientada para a normatividade conceitual, ela parte da análise da normatividade associada às práticas. Além disso, a teoria tem repercussões para a compreensão de como os tipos de normatividade em diferentes domínios estão conectados à centralidade de certas práticas. O inferencialismo semântico, o pragmatismo normativo e o expressivismo metalinguístico juntos mostram como as normatividades epistêmica, lógica e prática são facetas de uma mesma prática discursiva. O aspecto que pretendo explorar é uma

qualidades como vigoroso, sólido, potente, resistente, penso ser mais prudente não endossar a terminologia porque este pode não ser o caso. A expressão “normatividade imperiosa (*authoritative*)” é originalmente atribuída a David Copp e foi difundida por Tristram McPherson (2018). Embora pudéssemos manter em inglês o termo *authoritative* ou adotar o neologismo autoritativo, creio que os adjetivos imperiosa e imperioso capturam suficientemente bem o tipo de ênfase que se pretende: a natureza da aparente força dominante, daí imperiosa, de certas exigências morais. Adiante, elucidaremos um sentido mais específico para a normatividade robusta associada à natureza da normatividade prática (McPherson, 2018; Plunkett, 2019).

interpretação inferencialista das práticas da moral como um possível caminho para a compreensão do caráter imperioso que conferimos à normatividade de certas regras morais. Pela expressão “compreensão do caráter” me refiro à pretensão de elucidação de por que certos aspectos se vinculam, ou por que nós os percebemos como vinculados, a determinadas regras morais que assumimos como imperiosas.

Há uma grande diversidade de teorias e de posições que procuram oferecer uma explicação para a natureza da normatividade, inclusive a normatividade imperiosa da moral. No domínio da normatividade prática, que reúne as normatividades moral, jurídica, prudencial, estética etc., têm predominado perspectivas inspiradas nos debates metaéticos (Tappolet; Voizard, 2011). Como os desafios e dificuldades dessas abordagens teóricas aparecem associados a aspectos metafisicamente mais ou menos carregados, na próxima seção, mencionarei brevemente as principais posições, a fim de, comparativamente, indicar algumas das possíveis vantagens da interpretação inferencialista das práticas morais.

Na seção 3, apresento alguns elementos em termos dos quais tem sido apresentada a normatividade imperiosa da moral. Especificamente, trato dos aspectos que Terry Horgan e Mark Timmons (2020) consideram centrais nessa caracterização que eles denominam fenomenológica. Recorrendo a Nathan Howard e Nicholas Laskowski (2025), aponto que a noção de autoridade, mais do que a noção de superioridade, captura em parte o sentido da imperiosidade da moral. Além disso, considero a sugestão de que essa autoridade pode ser elucidada por uma teoria que ofereça uma leitura das práticas morais específicas.

A fim de marcar os elementos mais importantes que distinguem uma abordagem da normatividade como a sugerida por Brandom de

outras abordagens, na seção 4, eu apresento rapidamente as posições acerca do elemento-chave da noção de normatividade da moral num *continuum*, partindo de posições realistas mais robustas em direção a posições menos comprometidas metafisicamente.

Na seção 5, apresento a teoria de Brandom com ênfase nos elementos pragmáticos, especialmente, os aspectos relacionados à ideia de práticas tal como sustentada pela abordagem inferencialista. Destaco também os componentes associados à semântica proposta por Brandom e à sua explicação sobre como o significado de sentenças e conceitos emerge e está atrelado a aspectos constitutivos das práticas.

Na seção 6, procuro estender a visão das práticas nos moldes propostos pelo inferencialismo para a distinção e análise de três práticas centrais na moralidade: a censura, a responsabilização e a culpabilização. A ideia é mostrar que, entendidas a partir da abordagem inferencialista, cada uma destas corresponde a uma configuração distinta do espaço de razões e que as regras constitutivas das práticas de culpabilização oferecem uma melhor compreensão e elucidação dos aspectos que caracterizam a imperiosidade da moral.

2. Abordagens da normatividade imperiosa da moral

Partindo das posições robustamente realistas da normatividade imperiosa da moral que apelam à existência de entidades normativas, propriedades, qualidades ou fatos, as abordagens serão dispostas num *continuum*, em função de seus graus de compromissos metafísicos, mais ou menos realistas. Dada a difusão da nomenclatura do tipo “*x-first*”, tais como, *razões-primeiro*, *adequação-primeiro* etc.,⁴ farei uso

⁴ Ver Schroeder, 2021; Bastian, 2020; Cosker-Rowland, 2023 e Mchugh e Way, 2016.

dela para realçar o elemento-chave das diferentes espécies de explicação.

Do lado do chamado realismo metafisicamente robusto, a abordagem que poderíamos denominar “propriedades normativas-primeiro” assume que existem fatos morais e que eles dizem respeito a propriedades normativas que são objetivas e independentes de desejos e de opiniões dos agentes. As propriedades normativas da moral podem ser pensadas tanto em termos não naturalistas quanto naturalistas (Leary, 2020). As abordagens do tipo *razão-primeiro* sugerem que a justificação de nossas ações ou crenças envolve inescapavelmente o apelo a razões objetivas e de validade universal; o que conta é a validade e a força dessas razões (Schroeder, 2021; Scanlon, 2014).⁵

Uma tentativa de enfraquecer os compromissos metafísicos é tentar explicar a normatividade imperiosa da moral sem assumir diretamente a existência dessas espécies de entidades, mas aceitando ou endossando certos tipos de práticas ou a validade de certas atitudes em relação às quais os nossos conceitos normativos fundamentais ganham significado (Copp; Morton, 2022). A pergunta central dessa estratégia minimalista é: o que estamos fazendo quando empregamos um dado conceito normativo da moral? Normalmente, as teorias que tentam responder a esse questionamento tratam os significados dos

⁵ Outras abordagens metafisicamente carregadas são: abordagens do tipo *valores-primeiro*. Estas assumem que a normatividade imperiosa da moral está assentada em valores que consideramos fundamentais, ou seja, aqueles que importam e são centrais e indispensáveis para uma vida humana significativa e para a vida em sociedade, como dignidade, liberdade e felicidade (Scheffler, 1994; Chang, 2010). Podemos seguir listando abordagens da *adequação-primeiro* (Howard e Rowland, 2022; Mchugh, 2016; Mchugh, 2022); abordagens do *dever-primeiro* (Zimmerman, 2008; Brown, 2011); e abordagens das *virtudes-primeiro* (Hursthouse, 1999; Sladek, 2019).

termos e declarações morais apelando a práticas, usos e funções, eventualmente, como expressões de acordos e consensos.

A estratégia minimalista nos ajuda a compreender a diversidade e a capturar algo da normatividade imperiosa da moral sem recorrer a posições metafisicamente carregadas. Assim, do lado minimalista, podemos nos referir à abordagem das *funções-primeiro*; à abordagem da *evolução-primeiro* (Ruse, 1998; Joyce, 2006; Sterelny, 2012; Tomasello, 2016); a abordagens constitutivistas que consideram aspectos *constituíntes da agência-primeiro* (Korsgaard, 2009, 2019; Haase, 2019; Velleman, 2000; Katsafanas, 2016); a abordagens das *atitudes-primeiro* (Darwall, 2006; Wallace, 1994; Smith, 2005); a abordagens da *motivação-primeiro* (Schroeder, 2007; Smith, 1994; Prinz, 2015; Nichols, 2004; Rosati, 2003, 2016). Essas abordagens oferecem uma elucidação de determinados aspectos da normatividade imperiosa da moral e a maior parte fornece modelos descritivos ou funcionais, mas não uma justificação para a autoridade normativa da moral. Uma teoria como a proposta por Brandom pode ser entendida como uma abordagem de práticas, do tipo das *normas de uso-primeiro*. Isto é, a normatividade em geral é constitutiva das práticas sociais.

O que pretendo fazer é apresentar uma tentativa de estender essa abordagem inferencialista de Brandom para algumas das práticas morais específicas. Normalmente, as explicações sobre a normatividade imperiosa da moral assumem que a moralidade é um fenômeno complexo global e que as explicações do caráter distintivo desse fenômeno fazem referência a fatos morais objetivos gerais ou a mecanismos psicológicos universais. Tento argumentar que a ênfase nas distinções e no exame da estruturação dessas práticas morais, a exemplo do que Brandom propôs para a nossa compreensão do

significado, é capaz de elucidar a maior parte dos aspectos em relação aos quais caracterizamos a normatividade imperiosa da moral. O que pretendo explorar é quais aspectos desse caráter imperioso podem ser satisfatoriamente elucidados em termos da abordagem inferencialista das *normas de uso-primeiro*.

3. O desafio do caráter imperioso da normatividade moral

Os detalhes de uma caracterização adequada da normatividade imperiosa da moral são objeto de disputa. Alguns autores entendem que definir e especificar em que consistem as características da normatividade imperiosa da moral segue sendo um mistério (Blackford, 2016). Seja como for, não faltam esforços de definição dessa imperiosidade. Horgan e Timmons (2020), por exemplo, listam um conjunto de atributos peculiares da moralidade que, eles acreditam, uma teoria satisfatória teria de dar conta. Vamos assumir que os atributos elencados a seguir são os relevantes da normatividade imperiosa da moral. Alguns dos aspectos apontados pelos autores dizem respeito aos fundamentos da normatividade, enquanto outros são relativos ao modo como a experimentamos ou ao tipo peculiar de experiência e efeitos que associamos a ela.

Sobre os aspectos relacionados aos fundamentos, o primeiro a ser destacado é o de que algumas regras morais são percebidas como imperiosas no sentido de que possuem *uma autoridade e uma superioridade*; elas nos oferecem um critério pelo qual podemos ordenar os tipos de exigências morais. Além disso, as razões sobre as quais se assentam certas regras morais não apenas exercem sobre nós uma autoridade, mas essa autoridade parece ser derivada de uma fonte *independente da mente*, isto é, independente das nossas opiniões. Trata-se de algo a que agentes e comunidades estão sujeitos

e a que eles devem responder; essa autoridade, nós a percebemos e compreendemos como *objetiva*, não como subjetiva. Além disso, quando dizemos que “agimos porque entendemos que aquela ação é nossa obrigação”, a autoridade é percebida como algo que se impõe de fora, por isso, de uma *fonte externa*.

Horgan e Timmons chamam ainda a atenção para outros três aspectos distintivos associados à “fenomenologia” da moral: primeiro, assumimos que a normatividade imperiosa é algo que temos de considerar nas nossas avaliações, que temos o dever de agir em conformidade, como algo que nos “arrasta” (*pull*) em direção a determinadas atitudes avaliativas e cursos de ação. Segundo, ainda que se decida ignorar uma razão moral, a *aderência e o peso* dessa razão persistem, frequentemente com um custo psicológico e emocional de culpa ou remorso. Terceiro, tendemos a pensar que a autoridade da normatividade moral é *motivacionalmente eficaz* no sentido de que ela, intrinsecamente, gera o tipo de motivação apropriada. Quarto, podemos acrescentar, as razões geradas por regras normativamente imperiosas são percebidas como *inescapáveis*, no sentido de que, independentemente de qualquer outro fator, estamos sujeitos à força justificadora dessas razões.

Howard e Laskowski (2025) apresentaram argumentos que enfraquecem os aspectos associados aos fundamentos da normatividade imperiosa da moral, em particular, à suposta superioridade. Os argumentos apontam que a metáfora da “superioridade” se mostra uma noção filosoficamente problemática e, no limite, supérflua em termos de efeitos práticos. Restaria, então, a ideia de uma autoridade, e esta possivelmente pode ser explicada pelas práticas, sem a necessidade de se supor um critério “superior”. Os argumentos desenvolvidos por Howard e Laskowski (2025)

problematizam e desafiam a ideia de normatividade imperiosa em ambas as direções, da superioridade e da autoridade. Para os autores, a motivação e os pressupostos que nos levam à postulação da normatividade imperiosa são muito limitados, e não é evidente que se tenha de apelar à necessidade dessa noção para lidar com deliberações práticas. Por uma questão de espaço, não temos como reproduzir os detalhes dos argumentos, mas a ideia geral é a de que, no limite, fora da visão de domínios normativos incomensuráveis, a noção de um carácter imperioso da normatividade moral mostra-se supérflua.⁶

Um argumento em particular, atribuído a McPherson e analisado por Howard e Laskowski (2025), permite-nos examinar criticamente a suposta necessidade de uma normatividade imperiosa associada à moralidade. O que esperamos da normatividade prática é que ela nos forneça um critério para o encaminhamento de uma escolha, queremos evitar que esta seja arbitrária e injustificada. Para que uma escolha favoreça uma alternativa e exclua as demais de uma forma não arbitrária, o agente precisa apresentar um critério que explique por que a alternativa selecionada tem um status especial e, também, por que ela deve estar, no espaço deliberativo, situada à frente das demais, fornecendo razões compreensíveis para a preterição dessas alternativas. O ponto decisivo aqui é o de que a não arbitrariedade exige não apenas a justificativa da escolha preferida, mas também

⁶ O conceito de superioridade de um domínio teria um aspecto paroquial, tendo origem numa concepção controversa sobre como os “sabores normativos” das razões diferem entre si e, por isso, sugerem Howard e Laskowski, a noção soa mais como “uma solução em busca de um problema”. Além disso, por depender do estabelecimento de um domínio superior e peculiar para dar conta do desafio de pesar tipos de normatividade, ou de combinar “sabores” normativos diferentes, a normatividade imperiosa passa a gerar controvérsias filosóficas que em nada contribuem para a solução do desafio da deliberação prática (Howard e Laskowski, 2025).

uma racionalização do lugar que as demais alternativas ocupam no conjunto da decisão. Escolher uma opção como “especial” só é justificável se o critério que a torna especial for, por sua vez, capaz de ordenar simultaneamente o conjunto das alternativas com alguma coerência, e tendemos a assumir que esse é o papel da autoridade que associamos à normatividade imperiosa.

Um desdobramento dessa ideia é que, se é isso que esperamos do critério, então não está claro por que temos de postular a necessidade da superioridade da normatividade imperiosa. O critério não precisa ser superior, ele só precisa estabelecer a autoridade de uma alternativa perante as demais. Essa autoridade pode ser pensada como um aspecto constitutivo da própria prática moral deliberativa, dado que estamos considerando práticas racionais de ponderação, de justificação e de decisão. Podemos assumir que tais práticas já são regidas por padrões internos e constitutivos de validade normativa. Se, como veremos, a autoridade for entendida como enraizada em regras constitutivas das práticas racionais, não é necessário supor a superioridade categorial da normatividade moral. A imperiosidade da normatividade moral pode ser defendida em termos de autoridade, não como superioridade.

A seguir pretendo mostrar que a leitura fornecida pelo inferencialismo de Brandom nos conduz a um entendimento sobre como as regras que constituem e regulam as práticas podem assegurar o caráter de autoridade da normatividade imperiosa da moral. Acredito que uma leitura inferencialista das práticas e da normatividade imperiosa da moral pode ainda elucidar satisfatoriamente os demais aspectos fundacionais e fenomenológicos, apontados por Horgan e Timmons, característicos desse tipo de normatividade, quais sejam: a objetividade, a

independência de mente, a fonte externa, a atração, a aderência e a motivação. Para tanto, além de expor os elementos gerais da teoria de Brandom, que são as práticas, a sua normatividade constitutiva, o jogo de dar e pedir razões e o *scorekeeping* deôntico, é preciso considerar, e esse é o argumento principal, que a moral envolve um conjunto de práticas muito distintas. Vou me concentrar em apontar as práticas da censura, da responsabilização e da culpabilização e pretendo mostrar que as características da normatividade imperiosa da moral podem ser entendidas como decorrentes das normas que constituem e regulam esta última, a prática moral da culpabilização.

4. O inferencialismo de Brandom

A teoria de Brandom envolve uma recusa de perspectivas representacionistas. A normatividade conceitual não depende de representações, ela emerge das responsabilidades e compromissos que assumimos ao participar da prática discursiva de fazer asserções. Entretanto, mais genericamente, Brandom defende que os diferentes tipos de normatividade, inclusive o campo mais amplo da normatividade prática, podem ser entendidos como emergindo das estruturas gerais das nossas práticas. Como destaca Brandom em *Making it explicit* (1994, p. 253): “[...] as nossas práticas discursivas têm uma estrutura implícita, a ser revelada ou explicitada, que é intrinsecamente normativa [...]”, de modo que “[...] ser racional é colocar-se no espaço de dar e pedir razões”.

Na construção dos significados, o inferencialismo de Brandom prioriza as normas que regulam os usos e que constituem as práticas discursivas. Ele parte de uma compreensão sobre o que fazemos implicitamente quando usamos conceitos nas nossas práticas discursivas, daí o lema “Primeiro, as normas de uso!” e, por isso, me

referi a essa abordagem como a das *normas de uso-primeiro*. Nessa visão, são as normas em uso nas práticas que regulam os conteúdos conceituais, proposicionais e inferenciais, mas também os conteúdos expressivos, deônticos e práticos. De acordo com Brandom, o que é distintivo sobre julgamentos e ações, atos para os quais razões são demandadas, “[...] é a maneira como eles são governados por regras” (Brandom, 1994, p. 9).

Brandom entende que aplicar uma regra corretamente não exige sempre uma outra regra, mas sim certa competência prática do agente manifestada e reconhecida dentro do jogo das práticas. Nessa visão, é a competência do agente, as suas atitudes normativas, que sancionam as performances dos atos, os desempenhos, como corretos ou incorretos, válidos ou inválidos, adequados ou inadequados (Brandom, 1994, pp. 230-231). As normas que governam as nossas ações nas práticas decorrem da intencionalidade implicada nas nossas atitudes normativas de atribuição e de reconhecimento de significados. Assim, é a interpretação pragmática das atitudes que oferece a resposta à pergunta sobre qual é a capacidade da agência que faz com que o desempenho daquele envolvido na prática seja correto ou incorreto. A resposta é a postura e a atitude avaliativa da constituição da agência que está implícita na prática. O ato de avaliar é o que permite revelar as atitudes normativas. Esse é um aspecto central da abordagem das práticas proposta por Brandom (1994, p. 8).

Brandom (1994, pp. 34-36) interpreta as atitudes avaliativas fundamentalmente com relação a sanções e ele pensa que há duas formas de se sancionar um comportamento: de um lado, a sanção pode ser externa e direta sobre o infrator e, de outro lado, ela pode ser indireta e interna, isto é, ela pode ser especificada quanto a regras que afetam o status, as autorizações e os compromissos do infrator dentro

da prática. Nesse caso: “[...] o significado normativo da transgressão é ele próprio especificado em termos normativos (do que é apropriado, do que o transgressor teria direito de fazer)” (Brandom, 1994, p. 43). Desse modo, as punições para uma violação de norma são “[...] uma alternância em outros status normativos” (Brandom, 1994, p. 43). A norma referente a um direito ou permissão é perfeitamente inteligível quando compreendida relativamente a atitudes expressas como sanções, agora apresentadas em termos normativos, ou seja, com relação a se ter ou não direito ou permissão. Nesse sentido específico, a sanção é interna ao sistema normativo.

Na visão de Brandom, a própria prática social institui normas que estendem a atribuição de atitudes normativas a todos seus participantes e implicitamente supõe normas de desempenho cujas avaliações são expressas pelas atitudes avaliativas e sancionadoras dos integrantes. O exemplo sugerido pelo próprio Brandom é o da promessa. Prometer envolve assumir e atribuir compromissos, ações que, juntas, criam, respectivamente, os polos de responsabilidade e de autoridade: ao prometer, o agente concede aos outros o direito de confiar no cumprimento e, também, a autorização da sanção em caso de descumprimento. Na versão mais sofisticada da prática, se aquele que promete falha, a comunidade o pune com a perda de reconhecimento como alguém capaz de novas promessas (Brandom, 1994, pp. 164-165). Podemos usar esse modelo, que especifica as normas constitutivas da prática da promessa, para examinar como se configuram os arranjos das normas constitutivas das demais práticas da moral e tentar identificar como elas se distinguem umas das outras.

Assumindo a tese atribuída a Sellars de que o significado dos conceitos é firmado pelas relações inferenciais que as proposições estabelecem entre si, Brandom interpreta a correção da aplicação de

um conceito em termos da adequação ou inadequação às inferências autorizadas e não autorizadas pela proposição. Assim, o significado de uma asserção é completamente dependente de sua relação com outras asserções dentro da rede inferencial que cada uma delas autoriza. As asserções conectam-se às demais por meio das inferências que elas autorizam de modo que o significado emerge da posição de cada asserção dentro dessa rede de inferências e dos compromissos assumidos pelos participantes do discurso (Brandom, 1994, p. 87). Por causa disso, é crucial, para compreender o significado de uma asserção, entender o papel inferencial que ela desempenha uma vez proferida, ou seja, qual é o conjunto das demais asserções que ela autoriza e que podem ser inferidas a partir dela.

Um ponto-chave da visão inferencialista de Brandom é a maneira como ele entende o envolver-se numa teia de relações inferenciais, o envolver-se na prática discursiva de asserir ou o participar de um jogo discursivo. Brandom argumenta que a participação no jogo das práticas implica assumir que somos agentes racionais, inseridos em sistemas normativamente regulados. Essa regulação das práticas envolve o compromisso de oferecer razões para nossas crenças e ações, bem como de demandar razões dos outros participantes (Brandom, 1994, pp. 138-139). Assim, o “jogo de dar e pedir razões” é definidor das práticas; práticas nas quais os participantes são também, num sentido constitutivo, responsáveis por justificar suas afirmações.

Para Brandom (1994, pp. 86-88), fazer uma asserção envolve marcar uma posição em um campo inferencial, que inclui todas as conexões que ela tem com outras asserções. As inferências podem ser, admite Brandom, de diferentes tipos: a lógica (e as inferências formais) expressa um tipo de inferência mais fundamental nas

práticas, que são as inferências materiais. As inferências materiais são inferências cuja validade é dependente apenas dos conteúdos específicos dos conceitos empregados nas asserções, como em “Está chovendo agora”, logo “A rua está molhada” (1994, pp. 98-101). Desse modo, a lógica tem a função de tornar explícitos, na forma de regras, compromissos que até então permaneceram implícitos nas práticas inferenciais (Brandom, 1994, p. 117) As inferências podem também ser inferências práticas que vinculam crenças, desejos ou fins a ações e decisões, e, em geral, elas envolvem inicialmente o reconhecimento de uma situação para, em seguida, conectá-la à definição de certos cursos de ação (Brandom, 1994, p. 245).⁷

O próximo passo de Brandom é explicar como o ato de asserir, de conferir conteúdo proposicional, pode ser considerado um exemplo do princípio de constituição do elemento normativo da linguagem por meio da prática discursiva. As inferências materialmente corretas são governadas pelas normas usadas implicitamente na prática discursiva, sendo estas últimas explicadas pela pragmática normativa.

Um caminho teórico é, portanto, disponibilizado
[partindo-se] daquilo que as pessoas fazem para aquilo

⁷ Para Brandom, as inferências podem ser ainda: pragmático-normativas, como aquelas que dizem respeito ao que se deve fazer para que compromissos, permissões e obrigações, dentro do jogo de dar e pedir razões, sejam atualizados quando um participante faz uma nova asserção; inferências anafóricas, que se referem ao reconhecimento de coerência conceitual envolvida no encadeamento de referências a objetos ou conceitos no discurso, em que há a necessidade de recuperar algo do que foi dito e de manter a coesão; inferências pragmáticas, as quais abrangem implicaturas, convenções de uso da linguagem e o contexto conversacional, de modo que, a partir do que alguém diz, o ouvinte pode extrair entendimentos implícitos sobre intenções, sugestões ou expectativas (Brandom, 1994, p. 158).

que elas querem dizer, de suas práticas para os conteúdos de seus estados e expressões. Dessa forma, uma teoria pragmática adequada pode fundamentar uma teoria semântica inferencialista; suas explicações sobre o que é, na prática, tratar inferências como corretas são o que, em última análise, autoriza o apelo às propriedades materiais da inferência, que podem então funcionar como primitivos semânticos. (Brandom, 1994, p. 134)

Nessa visão de Brandom, a prática discursiva de fazer asserções, entendida em termos de um jogo de dar e pedir razões, possui uma normatividade inerente que implica a atribuição de papéis aos participantes. Faz parte das regras constitutivas da prática o fato de que quem faz uma asserção se compromete com o conteúdo do que foi dito, isto é, quando um falante faz uma asserção, ele se compromete a sustentá-la ou a justificá-la se for desafiado. Esse comprometimento estabelece a responsabilidade deôntica, na qual o participante é o responsável por sua asserção. Ao mesmo tempo, é concedida aos demais participantes a autorização para que eles façam inferências a partir da asserção, questionem o conteúdo e, se for o caso, desafiem o responsável pela asserção. Esse processo e a dinâmica discursiva criam uma rede de direitos e prerrogativas e de obrigações, compromissos e responsabilidades entre os participantes da prática:

Ao produzirem afirmações, os sujeitos realizam dois tipos de ações. Primeiramente, autorizam afirmações subsequentes (e os compromissos que expressam), tanto compromissos concomitantes de sua parte (consequências inferenciais) quanto reivindicações por parte de seu público (consequências comunicacionais). Ao fazerem isso, tornam-se responsáveis no sentido de

serem passíveis de resposta por suas reivindicações. Ou seja, também assumem uma responsabilidade específica de tarefa, a saber, a responsabilidade de demonstrar que têm direito ao compromisso expresso por suas afirmações, caso esse direito seja questionado. (Brandom, 1994, p. 173)

Considerando que a validação ou o desafio do conteúdo de uma asserção altera toda a dinâmica da distribuição de papéis e o status dos compromissos de cada um dos participantes, a prática discursiva envolve um contínuo monitoramento do cumprimento das regras da prática e de uso dos conceitos, bem como a constante atualização do status dos compromissos dos participantes a cada proferimento. À medida que surgem novas asserções, objeções, retratações, um participante pode deixar de ser o questionador e passar a ser o “asseridor” de outro enunciado, ou ainda assumir um compromisso novo. Comprometer-se com o conteúdo de uma asserção e ter direitos em relação a esse conteúdo equivale a receber ou conferir um status deôntico (Brandom, 1994, pp. 153-154).

O desenrolar da prática gera mudanças nos status dos participantes, daí a importância do monitoramento e atualização a cada novo ato assertivo. Brandom (1998, pp. 141, 188) denomina esse processo de monitoramento do status dos participantes dentro das práticas de *scorekeeping* deôntico, uma forma de monitoramento recíproco e contínuo das posições dos participantes dentro do jogo.⁸ O

⁸ De acordo com Gabriel Ferreira (2025): “[...] o que Brandom chama de *scorekeeping* [...] podemos traduzir genericamente por ‘monitoramento’ – consiste exatamente na prática pela qual todos os participantes de uma dinâmica discursiva, em especial no que diz respeito ao espaço no qual razões são relevantes, ou seja, no *logical space of reasons*, mapeiam as relações inferenciais das asserções dos demais e constroem uma espécie de placar inferencial que regula e monitora a cogência dos movimentos praticados naquele espaço”.

scorekeeping deôntico envolve o controle das atribuições de compromissos, permissões e responsabilidades no desenrolar das interações discursivas. É exatamente devido à centralidade de algumas das normas com as quais os participantes se engajam, se eles de fato “estão no jogo”, que a abordagem de Brandom é considerada uma semântica normativa. Como ele mesmo diz: “[...] o significado de um ato de fala são os modos como ele altera quais compromissos e prerrogativas os participantes mutuamente atribuem aos – e reconhecem dos – demais” (Brandom, 1994, p. 190).

A capacidade de participar adequadamente do jogo de dar e pedir razões numa base inferencialista, conforme propõe Brandom, é essencial para o que ele define como racionalidade. A racionalidade envolve não apenas compreender as regras, mas também saber participar adequadamente da dinâmica do “jogo de dar e pedir razões”. Então, ela não se refere a uma capacidade de pensar corretamente e de agir em conformidade a regras. Compreender um conceito é saber como ele se “encaixa” na estrutura inferencial do discurso, isto é, quais consequências decorrem de aplicá-lo em determinadas circunstâncias e quais condições justificam seu emprego. O participante que realiza a asserção é responsável por sustentá-la, mas também por corrigi-la ou revisá-la, quando desafiado. Ser racional envolve participar ativamente das práticas e ser capaz de entender as regras que governam o uso dos conceitos e de utilizá-los de forma coerente e justificada, bem como de demonstrar sensibilidade às normas que regem tais usos (Brandom, 1994, pp. 98, 117, 231, 253).

Em síntese, o conteúdo de uma expressão ou conceito é estabelecido pelo seu papel inferencial, pelas relações de compromisso e autorização que eles estabelecem com outras

asserções e conceitos. O que confere conteúdo a um juízo não é principalmente a sua correspondência ou representação da realidade, mas o modo como ele se insere no jogo de dar e pedir razões. A prática discursiva envolve sempre a dinâmica de atribuição de status e o *scorekeeping* deôntico, o monitoramento recíproco e a atualização dos status dos participantes. A explicitação da articulação de obrigações e compromissos inferenciais implicada na prática das asserções revela, também, a estrutura normativa e a dinâmica da prática; essa compreensão pode ser estendida para as práticas em geral, inclusive para as práticas morais.

5. A imperiosidade da normatividade da moral à luz das práticas

Não temos espaço, aqui, para apresentar detalhes sobre as distinções entre três tipos de práticas centrais na moralidade. Vou oferecer nesta seção apenas uma primeira aproximação dos aspectos distintivos da censura ou crítica moral (*moral blame*), da responsabilização moral (*responsibility attribution*)⁹ e da culpabilização (*moral guilt attribution*). Pretendo mostrar alguns dos elementos centrais que caracterizam essas práticas em termos inferencialistas para evidenciar que aqueles aspectos fenomenológicos destacados que normalmente atribuímos à normatividade imperiosa de certas regras morais, embora normalmente estendidos à moralidade como um todo, ganham sentido quando compreendidos como aspectos exigidos, em particular, pela prática moral de culpabilização.

⁹ Aqui, eu emprego *responsibility attribution* para marcar a inspiração strawsoniana, e porque *moral responsabilization*, em inglês, soa como uma prática mais formal e institucionalmente estabelecida.

De modo aproximado, podemos distinguir entre a prática da censura moral e a da responsabilização da seguinte maneira: a prática da censura envolve considerações acerca de uma ação reconhecida como uma falta moral, e o que se busca é estabelecer uma compreensão de o que exatamente configura a falta e oferecer uma caracterização genérica da autoria da ação faltosa. Ela não inclui a imputação de responsabilidade a um determinado agente, embora as práticas possam se sobrepor. A autoria pressupõe um agente com certas capacidades, motivações ou disposições reconhecíveis (inabilidade, malícia, negligência etc.), enfim, pressupõe um tipo de agência a ser especificada no processo da censura; uma pessoa concreta pode encaixar-se nesse tipo de agência especificada, então, tal agência é realizável; uma determinada pessoa possivelmente instancia o agente que reúne as condições de agência estabelecidas pela autoria genérica. O foco da censura está em especificar a falta e caracterizar a autoria. Nesse sentido, a censura não envolve atribuição de autoria, e, portanto, de responsabilização, a uma dada pessoa.¹⁰ A censura é altamente exposta à derrotabilidade. Não apenas a norma assumida como a fonte da falta pode ser derrotada, a derrotabilidade pode decorrer de conflitos de normas e da especificação das competências da agência, bem como da especificação do conteúdo intencional, do conteúdo epistêmico, do controle do agente etc.

A responsabilização, por seu turno, refere-se ao processo pelo qual, tomando como base a autoria genérica estabelecida na censura,

¹⁰ Considero a relevância da atribuição como apresentada por Peter Strawson (2008), contudo procuro aqui distinguir a prática da censura da prática de atribuição de responsabilidade. Essa caracterização mais restrita da censura nos permite compreender por que ela é algo que, como apontou o próprio Strawson (2008, pp. 16-17), varia em graus e a atitude reativa associada pode ser aplacada ou agravada em função da caracterização da autoria genérica.

busca-se atribuir a autoria da falta a uma pessoa determinada. Enquanto a censura pergunta se há e que tipo de agência é capaz de produzir a ofensa, a responsabilização pergunta “essa pessoa específica, concretamente, atende ou se ajusta às exigências da autoria genérica censurável estabelecida?”. Nesse sentido, a responsabilização envolve reconhecer que o agente é inserido num contexto acusatório associado a um juízo quanto à sua adequação à autoria genérica definida. Quando um participante é colocado na condição de responsabilizável ou como possível autor da falta, atribui-se-lhe um novo conjunto de compromissos deonticos. Agora, ele precisa responder pelas ações, ele precisa oferecer razões que justifiquem ou afetem a pretensão de atribuição ou, então, razões que a invalidem, tornando o alvo imune ou inimputável.

O que importa na censura é especificar ou pôr em evidência, de uma forma aberta e sujeita à derrotabilidade, os aspectos que possivelmente tornam determinada ação um tipo de falta moral e que deve ser considerada na caracterização da agência de uma autoria genérica. A ênfase das inferências é sobre inferências conceituais, proposicionais, práticas e normativas. A caracterização da autoria genérica pode tender à saturação, mas não se assume de partida que o processo precisa chegar a termo. Na responsabilização, o caso é diferente porque há um contexto acusatório que afeta o status, direitos e prerrogativas do acusado. As partes são menos colaborativas e mais competitivas. Os compromissos com justificação e defesa são mais sérios nas suas repercussões. Além disso, o contexto é mais fechado porque se espera que o processo chegue a termo com a determinação da autoria. Ganham relevância as inferências materiais em torno de fatos, evidências, atitudes e intenções que associem a autoria genérica

ao acusado. A derrotabilidade dos juízos torna-se crítica porque ela pode interromper, encerrar e reverter a tentativa de imputação.

De modo análogo à caracterização da prática da promessa oferecida por Brandom e à caracterização das práticas da censura e da responsabilização que esquematicamente oferecemos, podemos falar da prática de atribuição de culpa (*moral guilt attribution*) ou culpabilização. Embora essa atribuição, em geral, seja algo que se segue à responsabilização, não é necessário que elas aconteçam nessa ordem.¹¹

Do ponto de vista da configuração do jogo inferencial, a culpabilização envolve um rebaixamento sério do status normativo do participante porque o alvo da culpabilização já não está mais autorizado a mobilizar certos direitos, a fazer uso de determinados conceitos e está sujeito a interdições discursivas. As inferências que ele normalmente autorizaria com suas ações e alegações podem ser invalidadas ou recusadas pelos demais participantes. Ademais, as ações da pessoa culpabilizada já não mais autorizam inferências práticas de responsabilidade; assume-se a derrotabilidade automática das justificativas do culpado.

Além disso, o direito discursivo que ele teria a expor razões, oferecer explicações, justificar ações, explicitar intenções, reivindicar reconsiderações etc. pode ser levado em conta, mas não necessariamente, ou pode até mesmo converter-se em exigências e obrigações que, se não atendidas, resultam num agravamento ainda maior do seu status. Algo análogo se dá em relação às inferências práticas que o agente culpabilizado estaria autorizado a fazer: elas são

¹¹ Além disso, alguém pode ser colocado na condição de culpado independentemente da censura ou da responsabilização, pode igualmente culpabilizar a si mesmo (*self-guilt*) ou antecipar a culpa.

convertidas em exigências e obrigações. Os seus desejos e intenções são descredenciados, já não resultam em permissões, como é o caso dos agentes não culpabilizados. O agente não pode mais omitir-se sem que autorize novas inferências restritivas contra si e ele perde a prerrogativa de iniciativa inferencial, passando de uma condição de exercício ativo da autonomia para uma condição de estar sujeito às iniciativas dos demais participantes do jogo. Nesse sentido, o processo de culpabilização fecha o espaço de razões para o agente, ele passa a ser tratado como alguém que já teve a chance de apresentar justificações e falhou. Por isso, as suas possibilidades de justificação já estão esgotadas e as suas razões já não são mais relevantes no atual estágio da prática.

Os movimentos inferenciais da pessoa culpabilizada envolvem o reconhecimento expresso da norma violada e da própria culpa. Da mesma forma, implicam o compromisso moral com a mudança de conduta, a aceitação das implicações da falta para o rebaixamento do seu status. Nesse contexto, o silêncio e a insistência na negação da falta podem ser interpretados como nova violação ou como o agravamento da falta pela recusa da autoridade da norma.¹²

A conexão mais evidente entre a culpabilização e a compreensão do caráter imperioso da normatividade moral está associada ao fato de que a configuração da prática de culpabilização pressupõe, de partida, a absolutização da norma. A norma cuja violação autoriza a

¹² Tais comportamentos e reações são importantes porque o agente culpabilizado pode vir a se beneficiar da absolvição. Esta refere-se à ação normativa dos participantes por meio da qual o campo inferencial do agente culpabilizado é parcialmente ou totalmente restaurado e suas prerrogativas, autorizações e reivindicações são revertidas ou restabelecidas. Assim, a absolvição é uma atualização profunda do status deontico como consequência do reconhecimento de alguma perda de fundamento para a sustentação da culpabilização.

culpabilização adquire, dentro da prática, o status de norma categórica e inviolável. Além disso, reconhecer a atribuição da culpa pela falta é admitir que já estão afastadas as possibilidades de justificação e de desculpa e, portanto, a derrotabilidade daquela norma. Nesse sentido, as normas da prática de culpabilização estabelecem que a eventual derrotabilidade inferencial da norma infringida, que serve de fundamento à imputação, deve ser considerada suspensa. A norma definidora da falta torna-se impermeável a exceções. Em resumo, as regras constitutivas da prática de culpabilização convertem a norma violada num imperativo normativo absoluto, e é exatamente a sua violação enquanto imperativo que autoriza a imputação de culpa e o rebaixamento do status deôntico da pessoa alvo. O agente não pode se desligar nem abdicar dos compromissos que acompanham sua posição no espaço de razões. É a própria constituição pública da agência que torna a autoridade normativa das regras morais independente da aceitação subjetiva.

Caracterizada desse modo, vemos que a prática de culpabilização coloca o agente como um devedor diante da comunidade dos participantes. O fato de o agente culpabilizado ter de admitir que ele não agiu como deveria indica que a prática oferece um critério de correção que não é idiossincrático. Essa aceitação e o correspondente reconhecimento fazem com que a norma seja percebida ou experimentada como tendo uma autoridade sobre o agente.

Essa interpretação inferencialista nos oferece um entendimento pragmático acerca da origem dos aspectos de autoridade, de independência de mente, de exterioridade, de objetividade e de inescapabilidade da normatividade imperiosa da moral. Resta ainda

examinar se ela fornece a elucidação da aderência e da motivação para a ação.

A dor autorreflexiva e a consciência de estar na condição de ter de prestar contas configuram como o reconhecimento da falta pela violação do limiar posto pela norma. O agente vivencia internamente a culpa como sofrimento e dor moral em decorrência da falta. A ativação desse mecanismo emocional é a forma como o agente experimenta diretamente a autoridade categórica da norma. A condição de ter de prestar contas deve materializar-se em atos concretos e isso reforça o vínculo do agente com a norma moral, vínculo que é experimentado como a aderência e peso da normatividade da moral.¹³

Quanto à motivação, a abordagem inferencialista não assume que a associação entre a crença ou juízo moral e a motivação correspondente seja uma relação necessária. Noutra direção, a abordagem também não precisa assumir que razões morais são externas ao agente e contingentes. Na perspectiva inferencialista, ser motivado a agir respondendo à força normativa das razões morais é uma condição constitutiva de participar da prática normativa de dar e pedir razões, de justificar-se e de responder a compromissos. Então, alguma motivação moral é necessária para responder adequadamente no campo deôntico no qual o agente se insere na condição de participante do jogo de dar e pedir razões. A motivação moral, então,

¹³ A prática da culpabilização carrega uma forte carga emocional porque o sentimento que emerge da culpa envolve uma experiência emocional desconfortável, dolorosa e, simultaneamente, autorreflexiva e orientada ao outro, que é devida à admissão de uma falta, real ou potencial, a transposição de um limiar moral, e é acompanhada de um sentimento de responsabilidade pessoal por tal transgressão. Ver Miller, 2015.

está implicada na exigência prática decorrente da sua posição deôntica no jogo normativo.

Quando examinamos de perto a prática da culpabilização, na perspectiva inferencialista, constatamos o fechamento do espaço de razões para o agente culpabilizado e o seu campo inferencial é profundamente reconfigurado. Culpabilizado, o agente perde o direito de reivindicar inferências práticas comuns e de atribuir a si intenções e motivações, antes tomadas como legítimas. O espaço de razões acessível torna-se estreito e altamente normativo, resta ao culpabilizado um conjunto bastante limitado de proposições. Essas proposições não são meramente constatações, elas estão inferencialmente vinculadas a certas ações esperadas. “Agi errado” é acompanhada da manifestação da atitude de arrependimento. “Reconheço a minha falta” envolve manifestação de uma disposição à reparação. Restam ao culpabilizado apenas proposições desse tipo como opções discursivas legítimas, então, em qualquer proferimento seu, já estão implicadas as suas correspondentes exigências práticas. É essa configuração do espaço inferencial que faz parecer que, para cada proferimento moral, há uma crença que necessariamente carrega consigo a motivação exigida.

No campo de possibilidades discursivas da prática de culpabilização, restam críveis apenas as proposições às quais, por expectativa, a motivação se associa. Se o agente culpabilizado diz “reconheço minha falta”, mas não age em conformidade, ele passa a ser considerado como alguém hipócrita ou insincero. Não há espaço estável para uma crença moral “honesta” desvinculada do compromisso de agir. A necessidade da motivação é algo pressuposto pela configuração da posição deôntica do agente. Assume-se que o culpabilizado, em relação aos seus proferimentos limitados, está

motivado a agir naquela direção. Reconhecer que há uma ligação normativa entre proposições e motivações dentro do espaço de razões que constitui o nosso jogo moral não significa assumir que a conexão psicológica entre crença moral e motivação seja verdadeira, muito menos necessária. Essa conexão necessária é uma suposição implicada na prática, noutros termos, o internismo forte é apenas aparente.

6. Considerações

Neste texto, busquei expor o problema da normatividade imperiosa da moral e o desafio de elucidar os seus aspectos fundacionais e fenomenológicos. Apresentei as perspectivas teóricas que tentam elucidar o caráter distintivo da normatividade imperiosa da moral a fim de mostrar o que separa a abordagem inferencialista das demais, tanto em relação àquelas comprometidas com um realismo robusto quanto em relação àquelas com as quais compartilha a aposta na estratégia minimalista. Apontei que a extensão do inferencialismo de Brandom para a análise das práticas morais configura uma abordagem do tipo *normas de uso-primeiro*, ou seja, as normas que regulam como atribuímos censura, responsabilidade e culpa. As tentativas de explicar o caráter imperioso da normatividade moral assumem a moralidade como um fenômeno complexo e, por isso, tendem a enfatizar seus aspectos muito gerais, mas tentei evidenciar que a abordagem inferencialista pode ser mobilizada para tratar o problema da normatividade imperiosa da moral de um modo mais sensível às distinções qualitativas entre as diferentes práticas morais.

Em seguida, explorei quais aspectos da normatividade imperiosa da moral poderiam ser mais satisfatoriamente elucidados em termos

da abordagem *normas de uso-primeiro* e com base nos recursos oferecidos pela pragmática inferencialista. Segundo Howard e Lakowski (2025), a metáfora da superioridade, ao fim, mostra-se uma noção filosoficamente problemática e, no limite, supérflua. Esse exame evidenciou a ideia de autoridade como a mais promissora, especialmente porque ela pode ser explicada pelas práticas, sem a necessidade de se supor um critério “superior”.

Na apresentação do inferencialismo, destaquei que, nessa teoria, a normatividade emerge das responsabilidades e compromissos assumidos nas práticas sociais. Mostrei que o inferencialismo prioriza as normas que regulam os usos e que são elas que determinam conteúdos conceituais, proposicionais, expressivos e práticos, pois julgamentos e ações são essencialmente governados por regras. Apontei, também, que é a competência para a atitude avaliativa do participante que sanciona atos internos à prática como corretos ou incorretos. Nesse quadro, as práticas sociais instituem compromissos e responsabilidades cuja força depende do reconhecimento contínuo; esse modelo pode ser aplicado a outras práticas morais. Participar de uma prática implica entrar num jogo de dar e oferecer razões e, ao mesmo tempo, envolve a ativação de um *scorekeeping* deontico que registra as alterações de compromissos e prerrogativas e autorizações entre os participantes. Por fim, enfatizei que o modelo apresentado por Brandom, pelo qual o conteúdo conceitual deriva de papéis inferenciais e da rede de compromissos e autorizações monitorada pelo *scorekeeping*, pode ser estendido às práticas normativas em geral, inclusive às morais.

Na última seção, argumentei que a abordagem inferencialista oferece uma melhor compreensão dos traços fundacionais e fenomenológicos da normatividade imperiosa; particularmente, a

censura como a tentativa de especificação da autoria genérica de uma falta e a responsabilização como a tentativa de vincular a agência da autoria genérica a algum agente real. Listei brevemente as diferenças inferenciais da censura e da responsabilização. Em seguida, a partir de semelhante caracterização da culpabilização, evidenciei como os aspectos constitutivos dessa prática específica resultam no rebaixamento, na reconfiguração do status deôntico do agente culpabilizado e num fechamento do espaço de razões. Destaquei que é a impossibilidade da derrotabilidade da norma violada que faz com que ela seja percebida como um imperativo categórico, absoluto e incondicional. A explicitação das normas em uso implicadas na prática de culpabilização mostra pragmaticamente como surgem a autoridade, a objetividade, a exterioridade, a independência de mente e a inescapabilidade da normatividade da moral. Por fim, argumentei que as poucas alegações disponíveis ao culpabilizado, todas voltadas para o reconhecimento e aceitação da culpa, já se apresentam como vinculadas às ações que são exigidas do agente. Sugeri que essa forte limitação do campo proposicional do culpabilizado produz o efeito de uma aparente vinculação necessária entre crença e motivação, como assumido pelo internismo forte.

A elucidação dos aspectos imperiosos da normatividade da moral que oferecemos pela leitura inferencialista das práticas da moral talvez possa ser desenvolvida como uma teoria, uma espécie de constitutivismo de práticas. De todo modo, segue sendo uma tarefa desafiadora apresentar inferencialmente a constituição e a estrutura das práticas morais. Por outro lado, a leitura inferencialista das práticas é uma precaução contra pretensões generalizantes e globalizantes, metafisicamente carregadas, para a explicação da normatividade imperiosa da moral.

Referências

- Bastian, L. (2020) *Attitudes First: Rationality Attributions and the Normativity of Rationality*. PhD Thesis. St. Andrews: University of St Andrews.
- Blackford, R. (2016) *The Mystery of Moral Authority*. Hampshire: Palgrave Macmillan.
- Brandon, R. (1994) *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard: Harvard University Press.
- Brandom, R. (2018) *Inferentialism, Normative Pragmatism, and Metalinguistic Expressivism*. Robert Brandom's 2018 John Dewey Lectures at the Dewey Center of Fudan University. Disponível em:
[https://sites.pitt.edu/~rbrandom/Texts/Inferentialism Normative Pragmatism and.pdf](https://sites.pitt.edu/~rbrandom/Texts/Inferentialism%20Normative%20Pragmatism%20and.pdf). Acesso em: 02 de fevereiro de 2025.
- Brown, C. (2011) *Consequentialize This*. *Philosophical Quarterly*, v. 61, n. 243, pp. 601-620. Disponível em:
<https://www.jstor.org/stable/10.1086/660696>. Acesso em: 20 de maio de 2025.
- Chang, R. (2010) *Value in Practical Reasoning*. In: *Oxford Studies in Metaethics*, v. 5. Oxford: Oxford University Press.
- Cosker-Rowland, R. (2023) *Value-First Accounts of Reasons and Fit*. In: Howard, C.; Cosker-Rowland, R. (Orgs.) *Fittingness*. Oxford: Oxford University Press.
- Copp, D.; Morton, J. (2022) *Normativity in Metaethics*. The Stanford Encyclopedia of Philosophy (Fall 2022 Edition). Ed. Edward N. Zalta; Uri Nodelman. Disponível em:
<https://plato.stanford.edu/archives/fall2022/entries/normativity-metaethics/>. Acesso em 02 de fevereiro de 2025.

- Darwall, S. (2006) *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
- Haase, M.; Mayr, E. (2019) *Introduction: Varieties of Constitutivism*. Philosophical Explorations, An International Journal for the Philosophy of Mind and Action, v. 22, n. 2, pp. 95-97. Disponível em: <https://doi.org/10.1080/13869795.2019.1601754>. Acesso em: 20 de maio de 2025.
- Horgan, T.; Timmons, M. (2020) *Gripped by Authority*. In: Bedke, M. S.; Sciaraffa, S. (Eds.) *Normativity and the Problem of Representation*. Routledge: Taylor and Francis.
- Howard, C.; Rowland, R. A. (2022) *Fittingness: A User's Guide*. In: Howard, C.; Rowland, R. A. (Eds.) *Fittingness: Essays in the Philosophy of Normativity*. Oxford: Oxford University Press.
- Howard, N. R.; Laskowski, N. G. (2025) *Robust vs Formal Normativity II, or: No Gods, No Masters, No Authoritative Normativity*. In: Copp, D.; Rosati, C. (Eds.) *Oxford Handbook of Metaethics*. Oxford: Oxford University Press [no prelo].
- Hursthouse, R. (1999) *On Virtue Ethics*. Oxford: Oxford University Press.
- Joyce, R. (2003) *The Myth of Morality*. Cambridge: Cambridge University Press.
- Joyce, R. (2006) *The Evolution of Morality*. Cambridge: MIT Press.
- Katsafanas, P. (2016) *The Nietzschean Self: Moral Psychology, Agency, and the Unconscious*. Oxford: Oxford University Press.
- Korsgaard, C. M. (2009) *Self-Constitution: Agency, Identity, and Integrity*. Oxford: Oxford University Press.
- Korsgaard, C. M. (2019) *Constitutivism and the Virtues*. Philosophical Exploration: An International Journal for the Philosophy of Mind and Action, v. 22, n. 2, pp. 98-116. Disponível em:

- <https://doi.org/10.1080/13869795.2019.1599050>. Acesso em: 20 de abril de 2025.
- Leary, S. (2020) *Choosing Normative Properties: A Reply to Eklund's Choosing Normative Concepts*. *Inquiry: An Interdisciplinary Journal of Philosophy*, v. 63, n. 5, pp. 455-474.
- Mchugh, C.; Way, J. (2016) *Fittingness First*. *Ethics*, v. 126, n. 3, pp. 575-606. Disponível em: <https://doi.org/10.1086/684712>. Acesso em: 20 de abril de 2025.
- Mchugh, C.; Way, J. (2022) *Getting Things Right: Fittingness, Value, and Reasons*. Oxford: Oxford University Press.
- McPherson, T. (2018) *Authoritatively Normative Concepts*. In: Shafer-Landau, R. (Ed.) *Oxford Studies in Metaethics*, v. 13. Oxford: Oxford University Press, pp. 253-277.
- Miller, D. (2015) *Guilty Subjects: The Biopolitical Function of Guilt in Neoliberal States*. University of Western Ontario. Thesis and Dissertation Repository. 2992. Disponível em: <https://ir.lib.uwo.ca/etd/2992>. Acesso em: 22 de abril de 2025.
- Nichols, S. (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford: Oxford University Press.
- Plunkett, D. (2019) *Robust Normativity, Morality, and Legal Positivism*. In: Plunkett, D.; Schapiro, S. J.; Toh, K. (Eds.) *Dimensions of Normativity*. Oxford: Oxford University Press.
- Portmore, D. W. (2022) *A Comprehensive Account of Blame: Self-Blame, Non-Moral Blame, and Blame for the Non-Voluntary*. In: Carlsson, A. B. (Ed.) *Self-Blame and Moral Responsibility*. Cambridge: Cambridge University Press, pp. 48-76.
- Prinz, J. (2015) *An Empirical Case for Motivational Internalism*. In: Bjornson, G. et al. (Eds.) *Motivational Internalism*. Oxford: Oxford University Press, pp. 61-84.

- Rosati, C. (2003) *Agency and the Open Question Argument*. *Ethics*, v. 113, n. 3, pp. 490-527. Disponível em: <https://www.journals.uchicago.edu/doi/10.1086/345625>. Acesso em: 08 de fevereiro de 2025.
- Rosati, C. (2016) *Moral Motivation*. In: Zalta, E. N. (Ed.) *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition. Disponível em: <https://plato.stanford.edu/archives/win2016/entries/moral-motivation/>. Acesso em: 15 de março de 2025.
- Ruse, M. (1998) *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*, v. 2. Ed. Amherst, NY: Prometheus Books.
- Scanlon, T. (2014) *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Scheffler, S. (1994) *The Rejection of Consequentialism*. Oxford: Oxford University Press.
- Schroeder, M. (2007) *Slaves of the Passions*. Oxford: Oxford University Press.
- Schroeder, M. (2021) *Reasons First*. Oxford: Oxford University Press.
- Sladek, M. (2019) *Virtue and Meaning: A Neo-Aristotelian Perspective*. New York: Routledge.
- Smith, A. M. (2005) *Responsibility for Attitudes: Activity and Passivity in Mental Life*. *Ethics*, v. 115, n. 2, pp. 236-271. Disponível em: <https://www.jstor.org/stable/10.1086/426957>. Acesso em: 12 de dezembro de 2024.
- Smith, M. (1994) *The Moral Problem*. Oxford: Blackwell.
- Sterelny, K. (2012) *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge: MIT Press.
- Strawson, P. (2008) *Liberty and Resentment*. In: *Liberty and Resentment and Other Essays*. London: Cambridge University Press.

- Tappolet, C.; Voizard, A. (2011) *The Philosophy of Normativity, or How to Try Clearing Things Up a Little*. Dialogue, v. 50, n. 2, pp. 233-238.
Disponível em: <https://doi.org/10.1017/S0012217311000291>.
Acesso em: 20 de novembro de 2024.
- Thomson, J. J. (2008) *Normativity*. Chicago: Open Court.
- Tomasello, M. (2016) *A Natural History of Human Morality*. Cambridge, MA: Harvard University Press.
- Velleman, J. D. (2000) *The Possibility of Practical Reason*. Oxford: Oxford University Press.
- Wallace, R. J. (1994) *Responsibility and Moral Sentiments*. Cambridge, MA: Harvard University Press.
- Zimmerman, M. J. (2008) *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.

Moralidade como solo normativo comum: primeiros passos de uma teoria

Vitor Sommavilla¹

1. Introdução

Este texto é uma primeira tentativa de se pensar sobre o que é específico à moralidade, em contraste com outros sistemas de regras de conduta, tais como o direito, as convenções e a política. Que ela não se confunde com esses sistemas alternativos pode parecer algo óbvio, à primeira vista. A moralidade não é apenas o objeto privilegiado de estudo de eticistas; ela não se constitui como um discurso puramente teórico, como é o caso de outros conjuntos de conceitos e termos que são artifícios da prática científica, mas que não fazem parte da linguagem natural. Pelo contrário, o vocabulário moral – as noções de bom e mau moral, de moralmente justo ou injusto, de certo ou errado moral, para citar algumas – assim como as práticas morais – a exemplo da repreensão por comportamentos moralmente reprovados e o

¹ Professor do departamento de filosofia da Universidade Federal de Minas Gerais.

elogio a atitudes moralmente admiráveis – integram o dia a dia das sociedades ao longo dos tempos. Mesmo que esse recurso cotidiano ao vocabulário e às práticas morais seja tipicamente inconsistente e impreciso, ele não se deixa reduzir ou equivaler facilmente às operações daqueles sistemas de regras alternativas mencionados acima. À primeira vista, portanto, a prática e o discurso morais ordinários sugerem a existência independente do sistema moral.

Em contrapartida, pode-se pensar que aqueles sistemas alternativos se deixam mais facilmente discernir. O direito geralmente está expresso na forma de leis, explícitas ou implícitas. As convenções são de mais fácil delimitação por serem explicitamente construções humanas resultantes de acordos. E mesmo a política pode ser entendida como dizendo respeito especificamente às instituições políticas, notadamente àquelas ligadas ao Estado.² Em contraste, a ética pode estar ou não consignada na forma de leis, ela pode ser ou não convencional e, a princípio, aplica-se às práticas internas a todas as instituições humanas, inclusive as políticas. Desse modo, sua especificidade parece se perder.

Outro fator dificultador para o entendimento da moralidade em seu caráter distintivo advém da tendência, comum em alguns círculos intelectuais, de associar moralidade com moralismo, compreendido

² Todas essas descrições são no máximo aproximações, que podem ser legitimamente questionadas. Por exemplo, para a tradição jusnaturalista a conexão entre direito e moralidade tende a ser bastante forte, a ponto de alguns deles defenderem que leis (moralmente) injustas não são leis. No caso da política, realistas costumam considerar que a política tem regras e princípios próprios, possivelmente fora da alçada da ética. Por outro lado, diversos autores pensam que a (filosofia) política nada mais é que uma ética aplicada a temas políticos, portanto, que não há regras de conduta especificamente políticas. Essas questões são interessantes, mas escapam dos propósitos deste texto. Explorei a ideia de que a política possui uma normatividade específica em Sommariva (2025).

como uma tendência de emitir julgamentos moralizantes sobre as condutas dos outros que acaba por funcionar como um instrumento opressor ou limitador das liberdades legítimas dos alvos das críticas. De fato, como afirmado acima, a repreensão e o elogio de práticas consideradas significativas são parte central do fenômeno moral. E essas práticas de repreensão e elogio, quando eficazes, operam produzindo maior conformidade social às regras de conduta avalizadas. O que a crítica da moralidade como moralismo faz é considerar como sempre ilegítima a função da moralidade de produzir conformidade a normas socialmente aceitas.

As motivações para essa crítica podem ser pelo menos de dois tipos. De um lado, identifica-se, nas forças de contenção de desvios de condutas, impulsos conservadores em defesa de regras antiquadas ou simplesmente inadequadas à vida humana contemporânea. De outro lado, rejeita-se por completo a ideia de um sistema de condutas para seres humanos, sob a justificativa de que todo e qualquer sistema desse tipo é limitador das potencialidades humanas e que o correto é promover a liberdade em máximo grau, sem as amarras de sistemas de regras. Veremos abaixo uma variante do primeiro tipo de crítica. Brevemente, a reação mais correta parece ser que, de um sistema de regras presente que seja inadequado, não se segue que todos os sistemas de regras serão inadequados. Assim, pode muito bem ser o caso de que a moral vigente deva ser desbancada, sem que isso signifique que devemos abrir mão por completo de um sistema moral.

Quanto ao segundo argumento, em uma leitura ele pode ser entendido como propondo um sistema moral em lugar de outros, um sistema de máxima liberdade, o que teria implicações importantes para as práticas e instituições sociais. Neste caso, não haveria na verdade um abandono da ideia de sistema moral. Em outra leitura,

propõe-se abdicar de quaisquer regras morais de conduta. Essa pode ser a melhor teoria, mas, antes de me comprometer com uma abordagem tão radicalmente revisionista, acredito que vale a pena explorar outras interpretações do fenômeno moral.

Nas páginas seguintes, examino quatro conjuntos de abordagens sobre o fenômeno moral, começando pela ideia de que ele nada tem de específico. Depois de expressar minhas reservas com relação às três primeiras tradições analisadas, ofereço o esboço de uma quarta perspectiva, que incorpora elementos das teorias consideradas anteriormente, mas pretende evitar alguns de seus problemas. Não se almeja oferecer uma apresentação completa e bem-acabada da posição. Na melhor das hipóteses, ao final do texto, teremos à disposição o esboço de uma teoria que merece ser explorada em trabalhos posteriores.

2. Negação da especificidade da moralidade

A primeira alternativa que, biograficamente, explorei no terreno da ética contemporânea foi a de uma filosofia moral que nega que haja qualquer coisa de especial ou distintivo na moralidade. É a essa abordagem que sigo “oficialmente” associado, na medida em que foi essa a posição que expressei na minha tese de doutorado (Sommavilla, 2019). Em linhas gerais, a proposta consistia em argumentar que o foco relevante da reflexão em filosofia prática deveria ser voltado para uma noção mais geral do que a de moralidade, a saber, a ideia de normatividade prática. E que razões, preocupações ou emoções morais (se é que era possível falar delas de forma isolada) nada tinham de especial. No máximo, haveria uma tendência convencional a chamar certas diretrizes normativas de morais e outras, de outros nomes, mas nada de substantivo se seguiria daí. Passei então a falar

em razões, valores, identidades práticas e normas sem a aparentemente desnecessária adição do adjetivo “morais”.

Entender a moralidade como não possuindo nada de especial e preferir discutir mais amplamente sobre razões, valores e identidades práticas acaba por conceder à metaética uma posição de destaque inédito na filosofia. Isso se deve ao fato de que, nessa abordagem, a metaética não trataria mais primordialmente do tema um tanto restrito da análise semântica do discurso moral. Antes, ela se voltaria a uma tarefa muito mais abrangente, a da fundamentação da normatividade. Nas últimas décadas, em conformidade com essa tendência de reflexão mais ampla sobre os fundamentos da normatividade, a metaética passou a fertilizar outras subáreas da filosofia, como a epistemologia e mesmo a estética, dando origem a subsubáreas como a metaepistemologia e a metaestética. Nesse sentido, pode-se dizer que a metaética (uma reflexão de segunda ordem sobre a ética) transformou-se em uma metanormatividade (uma reflexão de segunda ordem sobre a normatividade em geral, ou talvez apenas sobre a normatividade prática), visto que não haveria motivo para restringir as considerações de fundamento a uma área que careceria de especificidade: a moralidade.

Inicialmente, dois motivos me fizeram rever essa posição. Em primeiro lugar, a ideia de que não há nada de especial na normatividade moral e de que o foco da reflexão deve recair sobre potenciais razões ou normas práticas de todo tipo, sejam elas convencionalmente consideradas morais ou não, corre o risco de produzir uma forma de ceticismo diante de vários, quando não todos, os temas clássicos da ética normativa. Se não faz sentido falar em normas morais especificamente, parece seguir-se que não faz sentido também falar em responsabilidade moral, emoções morais, dignidade

e respeito morais, e também de uma série de temas da chamada ética aplicada.

Em segundo lugar, a ética, ou filosofia moral, é uma área bastante consolidada da filosofia, com uma tradição própria que remonta ao menos aos gregos. Por isso, mesmo que em última instância se revele verdadeira a tese de que a moral não tem algo de peculiar e que a ética deve ceder seu espaço a reflexões filosóficas mais amplas, parece justificado conceder a essa tradição ética o benefício da dúvida e examinar algumas de suas teorias. Assim, movido por considerações como essas, passei a experimentar entendimentos alternativos sobre o fenômeno da moralidade. Em especial, passei a me interessar por visões que promovessem a moralidade em sua particularidade e relevância dentro do debate filosófico contemporâneo.

3. Críticas à moral moderna

Meu passo seguinte no terreno da ética contemporânea se deu por meio da obra de autores como Gertrude Elizabeth Margaret Anscombe e Bernard Williams. Embora diferentes em vários aspectos, Anscombe e Williams têm traços de fundo comuns em suas concepções sobre a moralidade. Para Anscombe, assim como para Sidgwick antes dela, o conceito de moralidade é um conceito moderno ausente na ética anterior. Naturalmente, é possível usar o termo “moral” como sinônimo de “ético”, mas a tese desses autores é que a concepção propriamente filosófica do que seria a moralidade é fruto da tradição da lei natural a partir de autores como Grotius, Pufendorf e subsequentes. Como Darwall explica didaticamente, “a lei natural, como teorizada por esses autores, é um conjunto de normas jurídicas e obrigatórias sobre o que é certo e errado ao qual todo agente moral humano está sujeito, independentemente das contingências de

qualquer ordem social ou política dada. Ela é o que os filósofos chamam agora de ‘moralidade’” (Darwall, 2013, p. 7). Ou, como sustenta Anscombe, os conceitos modernos de “obrigação *moral*” e “dever *moral*” e, portanto, “do que é certo ou errado *moralmente*, e do sentido *moral* de ‘dever’” não podem ser encontrados na filosofia antiga (Anscombe, 1998, p. 26). Pelo contrário, eles se conformaram a partir da ideia de lei, por meio de noções “jurídicas” como a de obrigação, dever ou culpa.

Nessa leitura, embora os gregos tivessem reflexões muito ricas em relação à ética em sentido mais amplo, isto é, por exemplo sobre a questão socrática por excelência de como viver a vida, sobre a vida boa, sobre as virtudes e sobre ações “excelentes”, eles não teriam produzido reflexões acerca do sentido especificamente moderno e filosófico de moralidade, ou seja, acerca do que é moralmente obrigatório para todo ser humano, independentemente das circunstâncias sociais ou do que este considere bom ou esteja em seu interesse.

Anscombe não observou esse desenvolvimento com bons olhos. Para ela, os filósofos morais modernos lançaram mão dessa concepção legalista da moral sem mais terem à sua disposição o recurso metafísico a Deus para fundamentar a autoridade de suas pretensas obrigações morais. Sem o recurso teológico, seria melhor à filosofia retornar à reflexão ética mais ampla característica da ética de virtudes antiga. Bernard Williams não diferia muito disso quando sustentava que o destaque grande concedido à moralidade moderna corria o risco de nos alienar de nós mesmos e dos outros. Também, para ele, seria melhor retomarmos um engajamento com a pergunta ética mais ampla de Sócrates sobre como devemos viver e abandonarmos nossa

obsessão com “esse sistema da moralidade” moderna, que tende a nos subjugar e alienar (Williams, 1983, 2011).

Um “sistema moral” pode nos subjugar e alienar fundamentalmente porque é inflexível e desconectado de nossas motivações subjetivas. Se o sistema é inflexível, agentes não estão livres para desrespeitá-lo por razões socioculturais ou de interesse pessoal. E, se é desconectado do “conjunto motivacional do sujeito”, como diria Williams, ele não depende de o sujeito endossar de qualquer forma o sistema moral, para que a obediência a ele seja obrigatória. Desse modo, instaura-se um potencial conflito entre as demandas morais e os projetos particulares do sujeito, de modo que as únicas alternativas passam a ser a alienação subjugada à obrigação moral imposta externamente, ou o imoralismo. Some-se a isso o fato de que, para os autores considerados, os sistemas morais apresentados pelos primeiros filósofos morais são em grande medida tentativas de oferecer uma fundamentação filosófica para um conjunto de obrigações derivadas de um contexto religioso, mas que agora não mais podem ser fundadas com o recurso teológico (como argumentado por Anscombe). Se esse é o caso, a opção pelo imoralismo não parece tão problemática. Será melhor abdicarmos de um sistema normativo que tende a nos subjugar, mas cuja justificação depende de uma visão de mundo moribunda: o Cristianismo. Explica-se, assim, a preferência por rejeitar o “sistema da moralidade” em favor de um retorno à questão socrática.

Além de fornecer um entendimento plausível sobre a natureza específica da moralidade, a posição de Williams também é compatível com (de fato é manifestação clara de) o antirrealismo normativo que eu considero bastante plausível, por razões independentes. Contudo, como tentarei mostrar a seguir, essa visão de Williams nunca me

pareceu completamente satisfatória. Isso por duas principais razões: em primeiro lugar, posições como essa parecem frequentemente não fazer uma separação suficientemente clara e sistemática entre o que podemos chamar, seguindo diversos autores, de o conceito de moralidade e concepções de moralidade. Em narrativas históricas sobre o desenvolvimento da moralidade tais como a de Williams, mesmo que se aceite a ideia de que a moralidade é uma forma de invenção moderna fruto da secularização de teorias do comando divino, tende-se a se confundir as coisas criadas (a moralidade ela mesma, como prática, e a filosofia moral, como reflexão e fundamentação dessa prática) com as teses e visões específicas adotadas por autores e indivíduos em um primeiro momento (isto é, tipicamente posições que são versões mais ou menos secularizadas de uma moral cristã). No entanto, por mais que o conceito e a prática da moralidade possam ter sido inventados em associação ou mesmo em identidade com uma concepção cristã ou criptocristã de moral, o conceito de moralidade ele mesmo não é necessariamente dependente dessa concepção de extração cristã. Se esse é o caso, porém, a eventual rejeição de uma moralidade de inspiração cristã não precisa acarretar a rejeição de todo e qualquer “sistema da moralidade”.

A segunda razão para a minha resistência diante de posições como a de Williams diz respeito ao fato de que, mesmo que queiramos reservar o termo “moralidade” para esse sistema de obrigações e demandas que pretendemos rejeitar, ainda assim haverá um conjunto significativo de práticas humanas de coordenação intersubjetiva carentes de uma análise aprofundada e de fácil incorporação dentro das tradicionais reflexões próprias à filosofia moral. Quer dizer, o

objeto é real. Cabe a nós estudá-lo. Esse último ponto me leva a uma terceira abordagem no terreno da ética contemporânea.

4. A moralidade de segunda pessoa

Segundo a definição de Stephen Darwall, a moralidade diz respeito às “reivindicações que fazemos às condutas e à vontade uns dos outros” (Darwall, 2006, p. 3). Nessa concepção de moralidade focada nas relações de segunda pessoa, as demandas de um agente sobre o outro dependem fundamentalmente das relações normativas existentes entre eles. Isto é, dependem de um ter uma autoridade legítima para cobrar do outro determinados comportamentos. São quatro as noções centrais para entender o cerne da proposta de Darwall, as quais formam um “círculo interdefinível, com cada uma delas implicando as restantes” (Darwall, 2006, p. 12): reivindicação (*claim*), *accountability* (*grosso modo*, a ideia de que podemos ser cobrados pelo que fazemos), razão de segunda pessoa e autoridade de segunda pessoa. Quando eu tenho autoridade de segunda pessoa sobre sua conduta e sua vontade, isso significa que minhas reivindicações fornecem a você razões de segunda pessoa para cumprir minhas demandas, caso contrário eu exigiria de você uma prestação de contas ou o responsabilizaria de alguma forma por descumprir suas obrigações. Claramente, esse modelo funciona sob a pressuposição de que eu e você compartilhamos um solo comum normativo, isto é, que seja ponto pacífico entre nós que, para esta demanda em particular, eu tenho autoridade de segunda pessoa sobre você, como você terá sobre mim em contextos análogos.

Para ilustrar sua visão, Darwall dá o exemplo de uma pessoa que pisa o pé de outra. Quando eu piso seu pé, eu o coloco em uma posição normativa tal que você tem uma reivindicação legítima de que eu

remova meu pé de cima do seu. Você possui, para esses fins, autoridade normativa de segunda pessoa sobre mim. Através dessa reivindicação legítima, que é legítima em razão de sua autoridade de segunda pessoa, você me fornece razões de segunda pessoa para mover meu pé. Se eu falho em atender a essas razões, eu passo a ser um alvo justificado de atitudes reativas tais como a culpabilização ou o ressentimento.

Assim, o modelo apresentado por Darwall, ao qual acabo de fazer apenas uma referência sumária, permite-nos entender a especificidade da moralidade de uma forma que promove sua importância para as mais diversas relações interpessoais. Além disso, na medida em que entende a moralidade a partir das relações de segunda pessoa características das práticas engajadas dos agentes morais, ele abre espaço para um entendimento natural e rico das emoções morais e de outras atitudes reativas tão presentes na vida moral ordinária e na reflexão filosófica sobre a moralidade.

Não obstante, por mais interessante e sofisticada que possa ser a concepção de moralidade oferecida por Darwall, a meu ver, seu projeto não é capaz de fundamentar a moralidade exclusivamente em razões de segunda pessoa, como pretendido. Como o próprio Darwall afirma, existe uma diferença entre “de um lado, a coerção e, de outro, a livre autodeterminação por meio da aceitação interna de uma demanda autoritativa” (Darwall, 2006, p. 23). Em outras palavras, para que uma demanda de segunda pessoa ofereça de fato uma razão de segunda pessoa para mim, sendo normativa sobre mim, é preciso que eu tenha aceitado essa demanda como autoritativa. Sem que eu tenha oferecido anteriormente o meu endosso às demandas normativas do meu interlocutor, a exigência de obediência a essas demandas seria

uma forma de coerção. Ou, como sintetizou muito bem Pauer-Studer, em um preciso artigo de comentário à obra de Darwall:

Darwall pressupõe o internismo da parte do sujeito individual: o agente reconhece a força das obrigações. Portanto, mesmo que a fundação da lei moral repouse em um acordo contratualista de segunda pessoa com os outros, deve haver uma correspondente fonte da normatividade em primeira pessoa. (Pauer-Studer, 2010, p. 303)

Outra insatisfação que tenho com o projeto de Darwall concerne a sua tendência a associar moralidade com objetividade e objetividade com universalismo. Segundo Darwall, a moralidade é necessariamente objetiva. A objetividade faria parte do próprio conceito de moralidade. Em certo sentido, isso me parece bastante plausível. Por exemplo, para Darwall, uma ação é obrigatória para um agente se, ao deixar de executá-la, este seria merecedor de culpabilização. Ora, para que seja possível avaliar se o agente merece ser culpado pelo que fez ou deixou de fazer, é preciso que haja um parâmetro de comparação estabelecendo o que seria a ação obrigatória. Esse parâmetro é um padrão objetivo. Em condições ideais ao menos, ele é acessível ao agente e ao avaliador que considera culpá-lo. No entanto, segundo Darwall, ao decidir o que deve ser feito, agentes em situações interpessoais necessariamente pressupõem uma perspectiva de segunda pessoa mais ampliada, que ultrapassa a relação imediata em que eles estão circunstancialmente envolvidos. Esse ponto de vista de segunda pessoa expandido é o ponto de vista da comunidade moral que, para Darwall, é a comunidade universal de seres racionais. Como ficará mais claro adiante, quando eu apresentar a concepção de

moralidade que favoreço, não me parece justificado pressupor que, necessariamente, relações morais intersubjetivas precisam de uma validação ou referencial externo no ponto de vista da comunidade moral (mesmo que, na prática, isso ocorra frequentemente). Ademais, mesmo nesses contextos em que há um recurso implícito ou explícito a um padrão normativo comunitário, também não me parece justificado pressupor que a comunidade moral de referência é sempre a comunidade universal de seres racionais. Pelo contrário, frequentemente nossas expectativas normativas intersubjetivas são orientadas por expectativas muito mais localizadas culturalmente.

Desse modo, embora Darwall tenha dado passos largos na direção de um aprimoramento em nossa compreensão do que seja o fenômeno da moralidade, ele não foi capaz de oferecer uma fundamentação da força normativa das demandas morais, ao fazer um recurso meramente sub-reptício e, portanto, indevidamente tematizado à necessidade de uma fundação da normatividade moral na aceitação individual das demandas externas da moralidade. Também em seu próprio entendimento da natureza do fenômeno moral, para além de sua fundamentação normativa, Darwall me pareceu se equivocar ao fazer da associação real entre moralidade e objetividade um passaporte para uma defesa da universalidade da moralidade ou de determinado código moral. Repetindo-me: embora haja uma pressuposição de objetividade intrínseca à prática moral, essa objetividade é compatível com uma validade contextualmente circunscrita.

5. Uma proposta alternativa

Tendo explorado algumas visões e as considerado insatisfatórias, gostaria agora de esboçar uma quarta abordagem para a ética

contemporânea, assumindo como mandato a tarefa de conceber uma concepção da moralidade que preserve sua especificidade (como fez Williams) e sua importância (como fez Darwall), mas que não se comprometa conceitualmente com alguma forma de realismo ou universalismo morais. Para tanto, inspirei-me tanto no internismo de Williams quanto na visão intersubjetiva ou de segunda pessoa sobre a moralidade, defendida por Darwall, mas propus alterações motivadas por outros compromissos e leituras.

Aqui está a minha definição alternativa:

Definição: moralidade é o conjunto de tratamentos que nós nos vemos como devendo a – e tendo o direito de receber de – aqueles que consideramos que compartilham conosco a comunidade moral.

A moralidade não é um conjunto de regras ou princípios. Ela é uma série de expectativas ou projeções (de comportamento mútuo, de atribuições de status moral etc.), em suma, uma série de pressuposições de um *common ground*,³ de um solo normativo comum. A moralidade é em essência um fenômeno aspiracional. Ao emitir um juízo moral, o que faço é dizer ao meu interlocutor: “Eis aqui o padrão de comportamento que espero que você endosse e eis aqui minha aprovação ou reprovação por seu cumprimento ou não desse padrão.” Ou, mais precisamente, no juízo moral, eu comunico minha pressuposição ou expectativa de que você adote determinado padrão normativo, e o que é dito é minha aprovação ou reprovação, quando

³ Para a noção de solo comum (*common ground*), ver os artigos presentes na Parte I da seguinte coleção: Stalnaker, 1999.

seu comportamento esteve ou não à altura do padrão normativo projetado sobre você.

Partícipes em um “encontro moral”, quer dizer, em uma situação intersubjetiva carregada de expectativas morais recíprocas, estão constantemente “negociando” suas posições e isso em diferentes níveis. Em primeira instância, eles negociam a obediência ou cumprimento das expectativas morais uns dos outros. Nesse passo, cada um, por assim dizer, “lança a sua rede” e anseia por uma pesca farta, isto é, espera observar nos outros, sobre os quais lançou sua rede, aquele comportamento requerido por suas expectativas ou projeções (ou seja, por sua “rede”). Assim, eles negociam e demandam uns dos outros, de forma mais ou menos explícita, um comportamento obediente. Mas, frequentemente, interlocutores morais negociam diretamente sobre os princípios a serem pressupostos na relação. Tipicamente, quanto mais eles fazem isso (negociar sobre os princípios), mais vai ficando claro que o comportamento obediente das partes não vai se realizar. Em outras palavras, na prática moral nós atuamos simultaneamente na busca pela inclusão de mais pessoas dentro de nossa comunidade moral (quer dizer, obtendo mais comportamentos que cumprem as expectativas normativas que projetamos sobre elas) e na busca pela definição dos termos de nossa comunidade moral (quer dizer, pela definição de quais pressuposições normativas poderemos ter uns sobre os outros). A vida moral é marcada por uma incerteza recorrente sobre quais são as pessoas de quem podemos esperar o endosso a determinados valores e sobre quais valores especificamente podemos esperar que as pessoas endossem. É preciso saber transitar pela vida moral diante dessa precariedade de informações

São essa incerteza e essa precariedade que nos fazem perceber mais uma vez a insuficiência do modelo de Darwall. Para ele, “quando você demanda que alguém remova o pé de cima do seu, você pressupõe [que você possui] um status irredutivelmente de segunda pessoa para endereçar a ele essa razão de segunda pessoa” (Darwall, 2006, p. 13). De fato, ao demandar de você que remova seu pé de cima do meu, eu pressuponho ter autoridade para essa demanda, ou seja, que minha demanda gera uma razão de segunda pessoa para você. Mas, ao contrário do que afirma Darwall, essa razão de segunda pessoa não é normativamente irredutível. Antes, ela pode ser reduzida ao fato de que você compartilha um solo normativo comigo, segundo o qual um não deve pisar o pé do outro. À luz desse solo comum, eu lhe dou uma razão para afastar o pé. Mas ela apenas é efetivamente uma razão para você por causa dos padrões normativos que você endossa. A razão de segunda pessoa tem que ser “traduzida” ou ao menos “traduzível” em termos de razões de primeira pessoa para que ela possa ter força normativa. Se você não endossa aquele código moral, nós não compartilhamos um solo normativo e, portanto, minha reivindicação para que você remova seu pé não terá força normativa, assim como não terão força minhas atitudes reativas moralizantes, como expressar reprovação e experimentar ressentimento. O “cão” das atitudes reativas moralizantes apenas tem dentes quando há compartilhamento de um solo normativo. O cão mostra seus dentes quando você falha em estar à altura de *nossos* padrões normativos comuns. Se a pessoa não considera que fez algo errado, não adianta sentir raiva ou culpabilizá-la: ela não experimentará remorso ou culpa, porque não terá falhado à luz dos padrões que ela mesma endossa.

Entendida dessa forma, e contra Williams, a moralidade não pode ser alienante. Ela nunca é um sistema de diretrizes ao qual eu

devo obediência independentemente dos meus valores e projetos e que frequentemente exigiria de mim que sacrificasse esses mesmos valores e projetos. Pelo contrário, nesse entendimento, a moralidade possui uma dimensão intrinsecamente recíproca: o que eu moralmente espero dos outros é o que eu moralmente espero de mim, se as circunstâncias ou papéis forem os mesmos. Por quê? Porque a moralidade é precisamente “o que eu assumo que nós compartilhamos”. Mas, se esse é o caso, já está pressuposto o meu endosso dos princípios que espero reconhecer em seu sistema de valores e em seu comportamento. De resto, esse ponto ajuda a explicar por que é parte conceitual da moralidade que a hipocrisia é criticável: como parte da própria ideia de moralidade, eu também devo fazer aquilo que demando moralmente de você (mais uma vez, supondo circunstâncias e papéis iguais).

Não devemos, porém, confundir o que acabo de dizer com a ideia de que as considerações morais “derrotam” todas as outras considerações normativas. Essa ideia conta com ilustres defensores, mas não me parece justificada. Segundo esses autores, sempre que uma obrigação ou demanda moral pesa sobre um agente moral, mas conflita com outras demandas normativas (por exemplo, ligadas aos projetos individuais do agente em questão), a demanda normativa moral deve prevalecer. A moralidade sempre teria mais peso na balança normativa. Do meu ponto de vista, considerações normativas têm o peso que o agente em questão a elas atribui. Portanto, não há nada de contraditório na noção de que algumas pessoas, reais ou imaginárias, ponham sempre no topo de sua hierarquia de valores aquelas considerações que têm a ver com demandas e expectativas interpessoais de comportamento (isto é, a moralidade como eu proponho entendê-la). Mas isso não é uma necessidade. Quanto a isto,

vejo-me compelido a concordar com o próprio Williams e também com Susan Wolf (1982), para quem, em vários momentos, outros projetos ou coisas com as quais nos importamos poderão ter mais peso em nossa tomada de decisão sobre o que fazer.

6. Aprofundando a posição

Na seção anterior, apresentei informalmente as linhas gerais da concepção de moralidade que gostaria de avançar. Nesta parte final do texto, gostaria de retomar e aprofundar alguns pontos dessa abordagem, com a intenção de deixá-la mais clara, não sem reconhecer o caráter um tanto incipiente da apresentação. Em um primeiro momento, discutirei mais alguns elementos do enfoque stalnakeriano em que tenho me apoiado. Em seguida, no final do texto, farei breves comentários sobre um caso paradigmático de evolução de uma interação moral marcada por discordâncias, tendo como ponto de partida algumas considerações da obra de Kant.

Quando um falante sinceramente endereça a um ouvinte, por exemplo, o juízo moral “Não deixe de visitar os seus pais semanalmente”, ele comunica tanto o conteúdo explícito de que indivíduos situados como o ouvinte devem visitar seus pais semanalmente quanto a pressuposição de que este último compartilha o mesmo valor em questão. Na abordagem que estou propondo, a força normativa do juízo para o ouvinte depende da confirmação da pressuposição de que esse indivíduo endossa o valor expresso pelo conteúdo explícito do juízo. O valor que o juízo pressupõe estar presente no corpo valorativo do ouvinte pode nunca ser articulado ou mobilizado e pode mesmo ser inconsciente. A

presença da pressuposição se faz sentir por uma disposição do agente para agir ou falar conforme o conteúdo pressuposto.⁴

Em condições normais (ou, talvez mais precisamente, em condições ideais), falantes e ouvintes pressupõem o mesmo corpo de informações. Estou propondo chamar de moralidade em um dado contexto conversacional (ou, mais amplamente, interativo) o conjunto de pressuposições normativas compartilhadas, isto é, o solo normativo comum da interação. Juízos morais são juízos normativos dirigidos por interlocutores uns aos outros, direta ou indiretamente, explícita ou implicitamente, e que lançam mão, em condições normais, desse conjunto compartilhado de pressuposições normativas.

Quando as pressuposições normativas do falante acerca dos compromissos normativos do ouvinte se mostram inverídicas, o contexto da interação é dito defeituoso.⁵ No programa de Stalnaker, contextos defeituosos são atípicos e existe uma pressão sobre os interlocutores para que eles alinhem as suas pressuposições, a fim de comunicarem efetivamente. Tal ajuste pode se dar através do abandono da pressuposição, por parte do falante, ou através da adoção da pressuposição, por parte do ouvinte. Não parece haver uma justificativa geral para preferir uma ou outra dessas duas soluções. No caso das pressuposições normativas que caracterizam a interação moral, é razoável supor que as partes terão, *ceteris paribus*, elevado interesse em fazer com que a outra parte aceite a sua pressuposição. Isso se deve ao fato de que, via de regra, agentes se importam com a

⁴ Na verdade, na abordagem de Stalnaker existe uma diferença entre as predisposições linguísticas que compõem o solo comum e as efetivas crenças ou valores dos agentes (Stalnaker, 1999, p. 52). Essa diferença não é importante para os propósitos deste texto.

⁵ Ver Peet, 2021.

implementação dos seus valores, tanto por eles mesmos quanto, em muitos casos, por outros agentes. Tomemos o exemplo de duas pessoas que discordam sobre a necessidade de uma distribuição equitativa de renda e riqueza na sociedade. O igualitarista pode ajustar sua pressuposição normativa abandonando seu igualitarismo e, assim, alinhando seus valores aos do seu interlocutor. Acontece que, pode-se assumir, é importante para o igualitarista que medidas sejam implementadas na sociedade no sentido de reduzir desigualdades. E tais medidas possivelmente precisarão contar não apenas com a contribuição do próprio igualitarista, mas também do seu interlocutor. Disso se explica que o igualitarista tenha interesse em resolver a discordância inculcando a sua pressuposição no interlocutor em vez de abandonando-a.

Segundo a proposta que venho apresentando, o propósito geral de juízos morais é produzir comportamentos em linha com as pressuposições normativas do falante. Quando o contexto é não defeituoso e as pressuposições normativas destacadas na interação são compartilhadas, o juízo moral é recebido com força normativa pelo ouvinte e, em casos felizes, contribui para a ação considerada moralmente adequada. Nesses contextos não defeituosos, todo um conjunto de emoções morais pode ser experimentado pelos interlocutores, que podem ser agrupadas em termos de culpabilização e elogio, em casos de respeito ou desrespeito às normas compartilhadas, respectivamente. Tais reações emotivas podem originar novos juízos morais, com o intuito de novamente influir sobre os comportamentos moralmente relevantes das partes.

É importante perceber que, segundo a abordagem que venho propondo, juízos morais possuem um lado de expectativa (ou previsão) e um lado normativo. O lado da expectativa diz respeito à

pressuposição de que o interlocutor endossa o conjunto relevante de valores ou normas de conduta. O lado normativo elogia ou repreende o interlocutor por estar à altura ou deixar de estar à altura das exigências normativas que (supõe-se) incidem sobre ele, em virtude do que se acredita que ele endossa. Na medida em que afirmações de teor moral podem ser convertidas em imperativos sobre a conduta alheia, a posição aqui esboçada sugere entender todos esses imperativos como hipotéticos, isto é, como tendo sua validade e autoridade sujeitas à condição de determinados princípios normativos constarem em ou se seguirem dos compromissos já assumidos pelo interlocutor alvo do imperativo.

Uma forma alternativa de compreender o que está sendo proposto aqui faz recurso, de modo um tanto mitigado, a alguns elementos da filosofia kantiana. Em uma nota de rodapé de “Que significa orientar-se pelo pensamento?”,⁶ Kant afirma:

Servir-se de sua *própria* razão não quer dizer outra coisa senão, em tudo aquilo que devemos admitir, perguntar a nós mesmos: achamos possível estabelecer como princípio universal do uso da razão aquele pelo qual admitimos alguma coisa ou também a regra que se segue daquilo que admitimos? (8:146n/Kant, 1985, p. 98n)

Isto é, para Kant somente devemos pensar algo (“admitir alguma coisa”) que possa valer universalmente. Da mesma forma, na *Antropologia*, o filósofo descreve três máximas da razão necessárias na busca pela sabedoria no uso prática da razão: “1. pensar por si mesmo,

⁶ As obras de Kant são citadas conforme a edição da Academia, em que o primeiro número indica o volume e, após os dois pontos, o número indica a paginação. Em seguida, após a barra, estão citadas as traduções utilizadas com as respectivas páginas.

2. colocar-se no lugar do outro (na comunicação com seres humanos),
3. pensar sempre em concordância consigo mesmo” (7:200/Kant, 2006, p. 99). A segunda dessas máximas claramente introduz a necessidade de se acomodar os demais seres humanos (seres racionais) no bom uso da minha razão. Por fim, a *Terceira crítica*, discutindo a razão em seu uso teórico, recorre à noção de *sensus communis* para explicar a comunicabilidade (ou a pretensão de validade universal) do juízo de gosto. Como esclarece Kant, o *sensus communis* é:

uma faculdade de julgamento que em sua reflexão toma em consideração (*a priori*) o modo de representar de todos os demais, para *como que* vincular o seu juízo à razão humana como um todo, escapando assim à ilusão que, a partir de condições subjetivas privadas – que podem facilmente ser tomadas por objetivas –, tivesse uma influência negativa sobre o juízo. (5:293/Kant, 2016, p. 191)

Em todas essas passagens está em jogo a ideia, que para Kant caracteriza o uso adequado da razão, de que devemos adotar princípios para o pensamento e a ação que possam ser universalmente aceitos e que sobrevivam ao diálogo crítico com outras pessoas. A universalidade é um traço forte da filosofia moral kantiana, algo que se depreende da própria primeira formulação do imperativo categórico, na *Fundamentação da metafísica dos costumes*. Contudo, o que as passagens acima permitem perceber é que, mesmo em contextos nos quais a ideia de universalidade não pode ser derivada do próprio conceito em questão (como ela pode ser no caso do conceito de moralidade para Kant), contextos estes como o do ajuizamento de gosto ou da formação de opiniões, Kant pensa que um

compromisso com a validade universal das reivindicações práticas ou teóricas está embutido no bom uso da razão. Colocar-se no lugar dos demais seres racionais e julgar como se fosse um representante de todos eles faz parte do que significa ser racional para Kant.

Não compartilho de toda essa ambição kantiana, mas gostaria de sugerir que algo nesse sentido está em operação quando recorremos a razões morais. Podemos assumir como ponto de partida (ou como ideal regulativo) a noção de que os compromissos morais que adotamos ou veiculamos em nossas ações ou discursos têm a pretensão de validade universal nos moldes kantianos. Na prática, porém, e felizmente, agentes morais não se comportam sempre sob essa pressuposição forte e universalista. Isso se deve ao fato, aludido acima, de que a pressuposição embutida na prática e no discurso moral é de natureza dupla. De um lado, ela define um padrão normativo, por meio do qual a ação de terceiros poderá ser avaliada. Mas, por outro, ela comporta uma dimensão epistêmica que chamei acima de expectativa ou previsão. Agentes minimamente racionais ajustam as suas previsões a respeito dos compromissos de outros agentes à luz das informações que recebem do ambiente e do comportamento prévio desses mesmos agentes. Dessa maneira, embora eu queira concordar com Kant que o recurso a razões morais envolva a pretensão de estender a validade dessas razões a terceiros e, em última instância (idealmente), universalmente, quero também recomendar que a extensão dessa pretensão seja regulada pelo que as evidências contextuais indicam acerca dos compromissos de interlocutores.

Assim, compreende-se em que sentido as regras morais tanto transcendem os compromissos individuais do agente que a elas recorre quanto carecem de força racional para, por si sós, produzirem

uma aproximação valorativa entre indivíduos que habitam mundos valorativos contrastantes. Esse modelo também sugere que a interação moral é caracterizada fundamentalmente por uma lógica da descoberta. Agentes em interação moral implicitamente lançam hipóteses sobre o conjunto valorativo uns dos outros e, na medida em que recorrem ao vocabulário moral, lançam hipóteses minimamente otimistas, no sentido de que pressupõem que suas expectativas morais encontrarão confirmação nos valores dos seus interlocutores. O progresso da interação marca a evolução do conhecimento que as partes têm sobre os valores uns dos outros e, conseqüentemente, sobre quais reivindicações morais possuirão força argumentativa, se é que alguma.

Não cabe aqui explorar os diferentes estágios da interação moral à luz da evolução do conhecimento das partes. De forma muito geral, podemos pensar em um caso idealizado que começa com um agente fazendo uma reivindicação moral sincera a toda a humanidade e sem qualquer conhecimento sobre os compromissos morais das demais pessoas. Em seguida, constata-se que algumas (ou várias) das suas expectativas de compartilhamento de solo normativo se frustram à luz dos fatos. Ele então passa a focar naqueles poucos elementos que lhe parecem ser comuns. Mas, se também estes se revelam elusivos, o passo seguinte pode ser mirar em outros elementos morais ou normativos compartilhados pelas partes, mas não pertinentes à questão original, com a intenção de induzir argumentativamente, a partir da estreita ponte comum entre as partes, a um pouco mais de convergência.⁷ Se também isto se mostra inviável, provavelmente as

⁷ Tenho em mente aqui, por exemplo, um tipo de culpabilização explorado por Bernard Williams, na qual quem culpabiliza não precisa supor que a parte culpada deixou de fazer algo que tinha razões para fazer. Nesse caso, o alvo da culpabilização

alternativas para a concordância sobre princípios de ação não estão mais abertas e, na medida em que esses agentes são levados pelas circunstâncias a interagir, a interação não se dará na forma da apresentação de razões com a pretensão de comum acordo, mas sim na forma da tentativa de imposição, mais ou menos forçada, de comportamentos vantajosos uns sobre os outros.

Essa trajetória imaginária e simplificada descreve o processo de perda de solo comum normativo ou mesmo de perda de mútua inteligibilidade. Também se pode descrever essa evolução como a passagem inevitável do engajamento participativo entre agentes para uma atitude meramente objetificadora entre eles, no sentido classicamente explorado por Strawson (1974). Raramente uma interação moral real se encaixa nos extremos da pressuposição de universalidade, de um lado, e da completa objetificação do interlocutor, de outro. Mas é em todo caso interessante ter em vista esses casos-limite na hora de interpretar as situações reais.

Referências

Anscombe, G. E. M. (1998) *Modern Moral Philosophy*. In: Crisp, R.; Slote, M. (Orgs.) *Virtue Ethics*. Oxford: Oxford University Press, pp. 26-44.

é uma razão de natureza mais geral, esta sim pressuposta como constituinte do conjunto valorativo da parte culpada, a saber, seu desejo de ser vista como uma pessoa merecedora do respeito do agente que a culpabiliza. Com recurso a essa razão mais geral, o culpabilizador visa a produzir uma razão na parte culpabilizada que antes não existia, mas que, em caso de sucesso, resultará do reconhecimento, pela parte culpabilizada, de que ela estava aquém das expectativas do culpabilizador, cuja estima ela valoriza. Ver Williams, 1995.

- Darwall, S. (2006) *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, Massachusetts and London, England: Harvard University Press.
- Darwall, S. (2013) *Morality, Authority, and Law: Essays in Second-Personal Ethics I*. Oxford, UK: Oxford University Press.
- Kant, I. (1985) *Immanuel Kant: Textos Seletos*. Trad. Raimundo Vier e Floriano de S. Fernandes. Petrópolis: Vozes.
- Kant, I. (2006) *Antropologia de um Ponto de Vista Pragmático*. Trad. Clélia A. Martins. São Paulo: Iluminuras.
- Kant, I. (2016) *Crítica da Faculdade de Julgar*. Trad. Fernando C. Mattos. Petrópolis: Vozes.
- Pauer-Studer, H. (2010) The Moral Standpoint: First or Second Personal? *European Journal of Philosophy*, v. 18, pp. 296-310.
- Peet, A. (2021) Defective Contexts. In: Khoo, J.; Sterken, R. (Orgs.) *The Routledge Handbook of Social and Political Philosophy of Language*. New York, London: Routledge, pp. 193-207.
- Sommavilla, V. (2019) *Constructing Ourselves as Moral Agents*. Tese de doutorado. Belo Horizonte: Universidade Federal de Minas Gerais.
- Sommavilla, V. (2025) A Constructivist Account of Political Normativity. *Topoi*, v. 44, pp. 65-75.
- Stalnaker, R. C. (1999) *Context and Context: Essays on Intentionality in Speech and Thought*. New York: Oxford University Press.
- Strawson, P. F. (1974) Freedom and Resentment. In: *Freedom and Resentment and Other Essays*. New York: Routledge, pp. 1-28.
- Williams, B. (1983) *Moral Luck: Philosophical Papers 1973-1980*. Cambridge, UK: Cambridge University Press.

- Williams, B. (1995) Internal Reasons and the Obscurity of Blame. In:
Making Sense of Humanity. Cambridge, MA: Cambridge
University Press, pp. 35-45.
- Williams, B. (2011) *Ethics and the Limits of Philosophy*. Oxford, UK:
Routledge.
- Wolf, S. (1982) Moral Saints. *Journal of Philosophy*, v. 79, n. 8, pp. 419-
439.

Parte 3

Motivação moral e normatividade

Razões no realismo moral naturalista: uma interpretação alternativa

Adelino Ferreira¹

1. Introdução

Este capítulo se insere no debate sobre normatividade a partir da discussão sobre razões no âmbito do realismo moral naturalista. Tal discussão parece, à primeira vista, algo com o qual um realista está mais à vontade do que alguns de seus adversários no domínio da metaética. Isso porque o realismo oferece, *grosso modo*, explicações objetivas sobre aspectos da moralidade que possuem caráter normativo para o agente, ou seja, se impõem a ele independentemente de sua vontade ou aquiescência. Veremos, contudo, que o realismo moral do tipo naturalista nem sempre é visto como um defensor da relação conceitual entre o conteúdo da moralidade e o fornecimento de razões para ação.

O realismo moral naturalista, em especial a vertente não reducionista dos autores de Cornell, advoga uma versão da moralidade em que propriedades morais são propriedades naturais, embora não

¹ Professor do Instituto Federal de Educação, Ciência e Tecnologia Baiano (IF Baiano).

possam ser reduzidas a estas. Além disso, alia uma versão de realismo metafísico a uma visão coerentista de justificação, busca não se comprometer com teorias intuicionistas e afirma que propriedades morais são propriedades naturais sem que haja a necessidade de uma apreensão direta de sua existência. Tal vertente metaética tem alicerces fortemente fixados e é capaz de oferecer uma explicação bastante plausível para o fenômeno moral ordinário.

Uma vez estabelecida como teoria, é possível partir dela para identificar e descrever as propriedades morais existentes no mundo. A abordagem nos apresenta uma compreensão da realidade que envolve propriedades morais entendidas como propriedades ordinárias, estando presentes, assim, em fatos passíveis de descrição (compreendidos e estudados pelas melhores teorias naturais sobre como o mundo funciona) que figuram necessariamente em nossas melhores explicações. Pretende, com isso, não reivindicar para a moralidade um reino à parte, acessado por alguma capacidade misteriosa ou diferente de todas as outras capacidades ordinárias dos agentes morais.

Um crítico poderia, contudo, levantar a seguinte objeção: é certamente possível compreender a metafísica defendida pelos teóricos de Cornell, mas a moralidade não é uma questão apenas de identificação de propriedades. Há, contudo, um elemento adicional a ser considerado se quisermos ser fiéis ao fenômeno moral. Tal crítica apela a certas platitudes referentes à moralidade que envolveriam seu caráter prático e uma certa disposição para a ação. O realista seria capaz de fazer jus a tais supostas platitudes? É sobre esse debate que este capítulo se debruça.

Ao realizar tal discussão, será possível compreender o que está envolvido no debate sobre razões, de que tipo elas são e como estão

envolvidas na compreensão mais ampla da natureza da moralidade. Para isso, partiremos de uma distinção iniciada na metaética por David Brink, qual seja, aquela entre internismo e externismo, ou seja, sobre a natureza da relação entre moralidade e razões e/ou emoções. Tal relação pode se dar, assim, em dois âmbitos: no motivacional e no das razões. Juntos, estes trazem uma versão mais robusta sobre o caráter prático da moralidade. Nosso foco, contudo, será apenas o âmbito das razões, deixando a discussão sobre motivação para um outro momento.

Após lidar com o debate sobre razões de maneira canônica na obra dos realistas de Cornell, em especial David Brink e Richard Boyd, pretendemos esboçar uma interpretação alternativa dos compromissos principais de um realismo naturalista como o deles. Se inicialmente a leitura tradicional os vincula a uma defesa do externismo de razões, almejamos mostrar que é possível nuançar tal compreensão e pensar em uma via alternativa, que não seja nem puramente externista nem totalmente internista.

2. Internismo e externismo: o debate sobre o caráter prático da moralidade

Feitos os debates metafísicos robustos sobre a natureza das propriedades morais, realistas morais naturalistas, como os da vertente de Cornell, têm o ônus de mostrar como a moralidade pode exercer um papel prático, ou se assim quisermos chamar, normativo na experiência humana. Uma vez que defendem que propriedades ou conceitos morais são elementos do mundo com certo grau de objetividade e independência de evidência, há que se mostrar como tais elementos podem, de algum modo, ser normativos para seres como nós.

A forma mais usual de lidar com esse debate é a partir da distinção entre internismo e externismo. Tal diferenciação pode ser pensada principalmente a partir de duas esferas da moralidade: a motivação e o fornecimento de razões. A discussão sobre motivação nos leva a pensar se um indivíduo que compreende o que é a moralidade está sempre motivado, de algum modo, a fazer o que julga correto. O debate sobre razões, por sua vez, questiona se a moralidade e seu conteúdo fornecem, por si sós, razões para ação ou se é necessário algo a mais para isso.

Internistas são aqueles que defendem que há uma relação necessária, conceitual, ou *a priori*, entre moralidade e motivação e/ou moralidade e razões para agir. Externistas são aqueles que negam o internismo e afirmam que tal relação, embora possa existir, é dada de forma contingente e *a posteriori*. Embora o debate sobre internismo e externismo na moralidade tenha se dado inicialmente no campo da motivação, este capítulo se concentrará, como já dito, no campo das razões. Isso porque encontra, a partir do realismo de Cornell, um caminho mais auspicioso para traçar as distinções mais específicas do campo normativo.

2.1 A natureza da moralidade

A pergunta sobre ética e racionalidade revela indícios de estar relacionada a uma pergunta mais geral sobre a natureza da moralidade. Isso porque, ainda que seja possível que um sujeito possa ser capaz de compreender o que está em jogo na moralidade e não se motivar, parece que ainda assim a moralidade está ligada de alguma forma à ideia de razões para agir. É, pois, difícil negar a natureza prática do debate moral e um certo caráter vinculante de sua existência.

Quando alguém defende que a moralidade é, em essência, uma prática que visa prescrever ações, manifestar sentimentos ou estados subjetivos (como defendem os não-cognitivistas),² a explicação de seu caráter prático parece, à primeira vista, mais simples de ser realizada. Tal explicação, contudo, foca mais nas paixões (no caso, os desejos) e esclarece mais o caráter motivacional do que o racional da moralidade. Já aqui podemos antever que uma teoria da racionalidade que vincule razões a desejos precisa estar em curso para que sejamos capazes de passar de motivação para racionalidade. Voltaremos a essa questão no decorrer deste capítulo.

Um cognitivista, por sua vez, precisa contar uma história sobre o caráter prático da moralidade e nos esclarecer os motivos pelos quais a moralidade não é apenas uma questão de descobrir fatos morais, sem qualquer vinculação prática. Esse é um problema que nos levará a pensar, de fato, sobre a natureza da moralidade e acerca de como tratá-la enquanto fenômeno complexo e multifacetado. Há várias soluções possíveis e vale a pena passarmos rapidamente por elas.

Uma primeira resposta cognitivista poderia ser aquela apresentada pelos não naturalistas: fatos morais são, de algum modo, *suis generis*. É próprio do conceito de moralidade que fatos morais são vinculantes e, por sua natureza, fornecem razões para agir. Em suma, a resposta é que, se compreendemos o que significa a moralidade, praticamente não há espaço para uma pergunta tal como a de onde vêm as razões.³ Isso parece fazer sentido à primeira vista. John Mackie, contudo, em sua teoria do erro, acusa os cognitivistas de quererem todos os bônus de uma argumentação metafisicamente inflacionada

² Ver Stevenson, 1944; Blackburn, 1984.

³ Ver Moore, 1993.

sem explicarem bem como tudo cabe no conceito de moralidade e em uma mesma propriedade.

Cognitivistas de inspiração kantiana tentam esclarecer a natureza da moralidade a partir do conceito de racionalidade. Trata-se aí de assumir uma noção muito típica, baseada em uma forte dimensão categórica do conceito de moralidade. Segundo tal projeto, a moralidade seria uma empreitada que se impõe acima dos desejos contingentes e garante uma universalidade adquirida por meio da razão prática.⁴ A ideia de racionalidade kantiana e de sua relação com a moralidade passa por uma série de comprometimentos e encontra dificuldades em especial no modo como lida com o papel das emoções no fenômeno da moralidade.

O que temos até agora é um conjunto de posições que, cada uma à sua maneira, tenta vincular o conceito de moralidade ao fornecimento de razões para ação. Com isso, espera-se alcançar, de saída, a explicação de por que a moralidade é uma empreitada prática e vinculante. Cada uma apresenta dificuldades e todas elas são alternativas internistas ao debate. Isso porque buscam conectar internamente o conceito de moralidade ao conceito de razões para ação.

Brink define assim o internismo de razões em relação a seu paralelo motivacional:

[...] o internismo de razões diz respeito à conexão entre considerações morais e boas razões ou razões justificadoras da ação. O internismo sobre motivação sustenta que é uma verdade conceitual que as considerações morais motivam, enquanto o internismo

⁴ Ver Smith, 2004.

sobre as razões afirma que é uma verdade conceitual que as considerações morais fornecem aos agentes uma razão para a ação. (Brink, 1989, p. 40)

Já aparecem aqui várias questões que precisaremos detalhar para deixar mais clara a discussão. São elas, por exemplo: o que são boas razões ou razões justificadoras? O que são verdades conceituais? Qual é, de fato, a natureza das razões? Antes de respondê-las, vale a pena tentar deixar o debate mais palpável por meio de alguma discussão inicial sobre moralidade e razões para ação. Visando isso, podemos nos perguntar inicialmente: o que está em jogo, em última análise, nesse debate?

Pensemos, pois, sobre o caso de uma obrigação moral qualquer, tal como a de socorrer, dando alguns tapas nas costas, uma pessoa que está engasgada. Imaginemos, para o bem do argumento, que essa seja a atitude moralmente correta naquela ocasião.

Para que o internismo motivacional seja verdadeiro, a pessoa que compreende que dar um tapa nas costas da pessoa é uma obrigação moral deve se sentir motivada a realizar tal ação. Temos, contudo, bons motivos para, ao menos, conceber pessoas que são capazes de entender que há ali uma obrigação moral, mas não se sentem motivadas para tal ato. Agora pensemos no caso das razões. Um internista defende ao menos duas teses: (i) que a moralidade fornece razões para ação de forma necessária; e (ii) que a relação entre obrigação moral e o fornecimento de razões é interna, ou seja, conceitual (*a priori*).

Considerando o que diz o exemplo acima, podemos alegar que a compreensão de que é correto ajudar uma pessoa engasgada fornece uma razão para de fato ajudá-la. Assim, podemos dizer que, quando

nos deparamos com uma obrigação moral, estamos diante de uma razão para ação, e, caso o sujeito não a reconheça, ele estará confuso acerca do que está em jogo na moralidade. Isso é basicamente o que diz (i). O ponto (ii), contudo, afirma que quem faz esse papel de fornecer as razões deve ser o próprio conceito de moralidade (o conceito de “bom”, por exemplo), uma vez que é uma verdade *a priori*, conceitual, que a moralidade fornece razões para ação. Não bastaria ser necessário, mas deveria ser também *a priori*.

Isso parece, de fato, exigir muito do conceito de moralidade para um realista. O fato, ou a propriedade moral, parece ter que dar conta de todos os elementos do fenômeno moral. Seria essa uma exigência justa ao realista? Podemos questionar tal demanda, pois o realista naturalista não quer adicionar “poderes” estranhos aos fatos morais. Fatos morais são fatos naturais, como quaisquer outros no mundo, e isso é importante para o projeto realista naturalista. Um internista parece forçar o realista a se comprometer com algo extravagante para depois criticar a própria extravagância da resposta.

A saída seria, então, abandonar o realismo e abraçar o antirrealismo ou o não-cognitivismo? Essa não é uma alternativa para naturalistas como os autores de Cornell. Uma primeira opção de resposta seria de fato a de abandonar totalmente o caráter normativo do fenômeno moral, recusar o papel prescritivo da moralidade e negar que ele forneça razões para ação. Isso resolveria de um lado o problema das razões, mas pareceria, para muitos, uma medida bastante revisionista da ética. Assim, ainda que o realismo naturalista mantivesse o caráter de superfície da linguagem moral, ele revisaria totalmente seu aspecto de guia para ação.

Brink busca conservar o caráter prático para a ação e diz que, diferentemente do que se pensa, é o realista que consegue garantir isso

e não seus concorrentes. O modo como ele parece querer construir a resposta é a partir da defesa do externismo de razões, em que, embora o fornecimento de razões seja externo ao conceito de moralidade, ele está presente na explicação de seu caráter de guia para ação. Brink apresenta inicialmente sua defesa do externismo em um artigo chamado “Externalist Moral Realism” (1986). Nele estão expostos os principais argumentos que serão reproduzidos em *Moral Realism and the Foundations of Ethics* (1989). Ao final do artigo, ele afirma que:

Porque alguns pensam que o internismo é a forma correta de representar o caráter prático da moralidade, eles afirmam que tal caráter conta contra o realismo. Mas o realismo é perfeitamente compatível com o caráter prático da moralidade. Isso porque é o externismo, mais do que o internismo, a forma apropriada de representar o aspecto prático ou de guia para a ação da moralidade. (Brink, 1989, p. 37)

Para chegar a essa conclusão, Brink irá apresentar uma discussão que sustenta que o internista tenta resolver muito facilmente a questão do caráter prático da moralidade, ao estabelecer sua relação de forma *a priori* e, assim, não levar a sério o ceticismo sobre a moralidade. Brink pretende explicar de forma substantiva o caráter prático da moralidade a partir de uma discussão que envolve três principais elementos: o conceito de racionalidade, o conceito de valor e uma noção de agente moral. Ele afirma que, só a partir daí, é possível compreender todas as questões envolvidas no fenômeno moral.

Brink caracteriza assim o externismo:

O externismo é a negação do internismo; o externismo afirma que a força motivacional e a racionalidade das considerações morais dependem de fatores que são externos às próprias considerações morais. É possível ser um externista negando qualquer uma das três afirmações que o internismo implica. O externista pode alegar, primeiro, que considerações morais apenas motivam ou justificam contingentemente, segundo, que o poder motivacional ou racionalidade da moralidade, seja necessário ou contingente, pode ser conhecido apenas *a posteriori*, ou terceiro, que o poder motivacional ou racionalidade da moralidade, seja necessário ou contingente, *a priori* ou *a posteriori*, depende de outras coisas além do conceito de moralidade, como o conteúdo da moralidade, uma teoria substantiva de razões para ação ou fatos sobre agentes, como seus interesses ou desejos. (Brink, 1986, p. 28)

No parágrafo anterior, Brink afirmara que o internismo tem três componentes, sendo o primeiro deles o de que: “[...] considerações morais necessariamente motivam ou fornecem razões para ação” (1986, p. 28). Em uma nota sobre as teses, Brink diz que irá explicar por que o realista pode resistir aos três componentes (1986, p. 39). Se é assim, ele parece defender o caráter contingente da motivação (isso está bem estabelecido) e também das razões para ação. Nossa discussão se concentrará nos últimos dois componentes, que versam mais especificamente sobre razões.

Ao apresentar o seu argumento, procuraremos mostrar, contudo, que ele pode não ser capaz de sustentar essa tese mais forte. Assim, caso a única reivindicação do internista de razões for uma relação necessária entre o fenômeno da moral e o fornecimento de razões,

Brink talvez não seja capaz de negar tal vínculo. Vejamos então como ele constrói sua argumentação.

2.2. *Razões, valores e agentes*

O argumento de Brink em prol do externismo moral tem como base a afirmação de que, somente após uma discussão substantiva sobre o conteúdo da moralidade (não apenas sobre o seu conceito), uma compreensão sobre a natureza das razões e fatos sobre o agente, é que podemos responder se a moralidade fornece razões ou não para a ação. Esse parece um caminho interessante para lidar com o debate sobre razões para ação. Isso porque, como ele mesmo diz, sustentar que a relação entre moralidade e o fornecimento de razões é algo *a priori* é tentar resolver muito facilmente um debate complicado.

Afirmar a ligação entre o conceito de moralidade e de razões para ação sem maiores detalhes de fato parece passar por cima de algumas questões importantes como por exemplo: com qual noção de racionalidade estamos lidando? Essa talvez seja a questão mais importante para iniciarmos a reconstrução do argumento de Brink. Isso porque, a depender da noção de racionalidade a ser assumida, chegaremos a respostas distintas sobre a compatibilidade do realismo moral com o fornecimento de razões.

Inicialmente é preciso pensar em que tipo de razões estamos pensando quando tratamos desse debate específico. Em uma citação apresentada algumas páginas atrás, vimos que Brink afirma que o externismo de razões “[...] diz respeito à conexão entre considerações morais e boas razões ou razões justificadoras da ação” (1989, p. 40). Ele distingue, assim, razões explanatórias de razões justificadoras, ou boas razões. É sobre razões justificadoras que esse debate irá se travar.

Alguém poderia dizer que a razão para uma pessoa tomar água do mar é o desejo de saciar sua sede. Essa, pois, é uma má razão para se tomar água do mar, já que ela não cumprirá sua função adequadamente, pelo contrário. A essa razão não chamaríamos de razão justificadora, ou uma boa razão. Ela seria, no máximo, uma razão explanatória a depender da teoria das razões adotada. Precisaremos, então, compreender qual a natureza das razões justificadoras e, somente a partir disso, sua relação com a moralidade.

Brink (1989) também estabelece três distinções no debate: o internismo do agente, o internismo do avaliador (*appraiser*) e o internismo híbrido. A diferença aqui está no grau de objetividade das razões. No caso do internismo do avaliador, as razões em voga são aquelas que o sujeito julga possuir, mesmo que, no fim das contas, sejam más razões, isto é, mesmo que não haja tais (boas) razões. Já no caso do internismo do agente, estão em jogo as razões que o indivíduo efetivamente possui, mesmo que, do ponto de vista subjetivo, ele não as reconheça. Por fim, o internismo híbrido diz respeito às razões que o sujeito reconhece como suas. Nesse caso, as razões em questão são aquelas que o indivíduo efetivamente possui e reconhece (ficam de fora as razões que, embora existam, este não as reconhece como tais).

No caso do debate sobre razões, a discussão abarca o internismo do agente e o híbrido, uma vez que, novamente, estamos buscando razões de fato justificadoras e não razões que podem ser avaliadas como tais, sendo, na verdade, más razões. No contexto do debate sobre motivação, isso seria até possível, pois elas seriam capazes de explicar as ações. No debate sobre razões justificatórias, isso não é cabível. Se a moralidade está ligada a razões, essas razões não podem ser ilusórias ou malformadas.

Feitos esses esclarecimentos, Brink propõe a discussão sobre a natureza das razões, pois ele afirma que muitos críticos do realismo partem, sem maiores explicações, de uma teoria da racionalidade baseada nos desejos (*desire-based theory*). Essa talvez seja a teoria sobre a racionalidade mais difundida e sustenta, em rápidas palavras, que um sujeito tem uma razão para fazer algo se esse algo satisfaz algum de seus desejos. É a clássica ideia humiana de que a razão é escrava das paixões. Brink afirma, contudo, que mesmo uma teoria *desire-based*, no fim das contas, não enfraquece a objetividade moral. Ela apenas exigiria levar em consideração que, entre os desejos contingentes do agente, houvesse um desejo de ser moral.

Uma discussão sobre desejos que levasse em conta os desejos refletidos do agente, ou aqueles que ele, em condições ideais, desejaria possuir, poderia ajudar a pensar que, no fundo, o desejo de ser moral estará entre o bojo de crenças de todos os seres como nós. E que, como a moralidade é para seres como nós, caso haja alguém com um conjunto mental tão diferente do nosso, isso não seria, de fato, um problema. Embora a ideia de que a moralidade é para seres como nós nos aparente adequada, essa parece uma tentativa um tanto desesperada de salvar a objetividade da moral, caso uma teoria *desire-based* esteja correta.

Ocorre que há pelo menos outras teorias da racionalidade disponíveis no debate. São elas o egoísmo racional e o altruísmo racional. O egoísmo racional, em uma breve síntese, diz que “[...] um agente tem motivos para fazer x apenas no caso, e na medida, em que x contribuiria para seu próprio interesse, bem-estar ou felicidade” (Brink, 1989, p. 67). Vemos aqui uma diferença importante em relação à teoria anterior. Isso porque a racionalidade não está baseada (ou é escrava) das paixões do sujeito, mas sim de uma determinada

concepção de valor (que pode variar de acordo com a teoria do valor assumida). Podemos entender teorias *desire-based* como versões subjetivas do egoísmo, mas, quando estamos tratando de egoísmo racional, estamos falando de uma teoria que está fundamentada em uma noção objetiva do valor. Assim:

De acordo com o egoísmo racional, se as considerações morais fornecem razões para ação, depende não apenas do conteúdo da moralidade, mas também dos fatos sobre os interesses dos agentes. De acordo com versões objetivas do egoísmo racional, a racionalidade das considerações morais não depende de “inclinações fortuitas e evitáveis”, mas das exigências da moralidade e dos constituintes objetivos do bem de um agente. (Brink, 1986, p. 34)

Vemos, então, que teorias distintas vão afirmar diferentes relações do sujeito com as razões. Assim é possível que, por uma teoria, alguém tenha razões para ação em uma determinada situação e, por outra, não. Isso já mostra que o debate sobre as razões se torna mais substantivo do que apenas formal. Quando falamos em razões, precisamos fazer menção a – e defender – uma teoria específica, e só aí poderemos nos questionar sobre sua relação com a moralidade.

No caso do egoísmo racional, por se tratar de uma teoria objetiva, é possível dizer que um sujeito pode ter razões ainda que não as reconheça e que, talvez, haja razões que são vinculantes a todos os seres humanos. Caso o realismo moral seja verdadeiro e seja possível encontrar um valor objetivo na natureza, este poderá estar ligado ao fornecimento de razões independentemente dos desejos individuais ou da diferença entre os sujeitos. O ponto de Brink aqui, contudo, é apenas estipular essa possibilidade como forma de sustentar o seu

argumento. Mais à frente, quando falarmos das teorias do valor, poderemos pensar em como elas podem ser conjugadas com as teorias de racionalidade.

Por fim, há ainda uma terceira família de teorias da racionalidade, que é o altruísmo racional. Tal abordagem, diferente do egoísmo racional, sustenta que alguém tem motivos para fazer algo apenas no caso e na medida em que algo contribuiria para o “interesse”, “bem-estar” ou “felicidade” de um outro alguém ou que inclua um outro alguém. Trata-se de uma versão que exige uma imparcialidade e uma objetividade ainda maior do que a teoria do egoísmo racional, pois se compromete com uma abordagem imparcial que leva em conta não o indivíduo, mas alguma ideia de coletivo. O que estaria incluído nessa comunidade moral parece ser um fato a ser pensado. Apenas sujeitos racionais? Hoje, ao nos depararmos com a ideia de ampliação do círculo moral, a questão do altruísmo racional se torna ainda mais complexa e interessante.

Aqui novamente Brink tenta mostrar como o fornecimento ou não de razões envolve muito mais do que uma concepção formal de moralidade e de razões para ação. Isso porque “[...] a racionalidade das exigências morais depende do conteúdo da moralidade, de uma defesa substantiva do altruísmo racional, e do fato adicional e contingente de que o solipsismo é falso” (1986, p. 34). Assim, tanto em versões focadas no sujeito, como o egoísmo racional, quanto em versões imparciais, como o caso do altruísmo racional, ele chama a atenção para o fato de que a relação entre a moralidade e as razões não só não precisa ser dependente de desejos subjetivos e variáveis, como também não pode ser dada *a priori*.

Por fim, entre os elementos a serem considerados na relação entre moralidade e racionalidade, resta tratar das teorias do valor, ou

seja, daquelas que dizem respeito ao conteúdo da moralidade. Quando falamos em teoria do valor, estamos buscando teorias sobre o valor intrínseco, ou seja, abordagens que tratam de coisas que são valiosas em si. Assim, para um externista, apenas se o conteúdo da moralidade for convergente com a promoção ou realização de algum valor mais básico não moral é que se poderá dizer se os juízos morais fornecem razões ou não.

As teorias do valor podem ser pensadas tanto do ponto de vista subjetivo quanto do ponto de vista objetivo. Teorias subjetivas do valor afirmam que os componentes de uma vida valiosa consistem ou dependem de certos estados psicológicos (como o prazer ou a satisfação de desejos, por exemplo). Por sua vez, teorias objetivas irão alegar que os elementos constitutivos de uma vida valiosa estão relacionados ao desenvolvimento de certas capacidades, determinados traços de caráter e ao exercício de habilidades pessoais e interpessoais, e esses elementos não consistem ou dependem necessariamente de estados psicológicos específicos. Ou seja, certas características seriam desejáveis em si, independentemente de serem prazerosas ou atenderem a nossas preferências subjetivas (Brink, 1989, pp. 221-222).

Certamente as teorias subjetivistas têm a vantagem de ser mais simples ao fazerem referência direta a propriedades naturais que são relativamente fáceis de mensurar. As teorias objetivas, por sua vez, são mais complexas e apresentam elementos que precisam ser bem detalhados em especial no que tange à sua descrição em termos eminentemente não morais. Elas, contudo, parecem responder melhor aos contraexemplos levantados às teorias subjetivistas do valor, como os casos de ações que, mesmo causando dor, podem ser

consideradas valiosas, ou daquelas que, mesmo sendo prazerosas, não bastam sozinhas para garantir uma vida valiosa.⁵⁶

Ao descrever uma proposta de teoria objetiva do valor, Brink afirma que:

Uma concepção objetiva plausível de bem-estar deve, penso eu, reconhecer vários componentes distintos, mesmo que relacionados, do bem-estar humano. Deve também distinguir entre componentes do bem-estar humano que são intrinsecamente valiosos, bens que são extrinsecamente valiosos e condições necessárias para o bem-estar humano (que também podem ter muito valor extrínseco). Proponho discutir uma teoria do bem-estar que considera a busca reflexiva e a realização de projetos

⁵ Vide, por exemplo, a Máquina de Experiência, de Robert Nozick, um contraexemplo bastante influente no hedonismo. Nozick nos faz pensar a existência de uma suposta máquina que manipulasse nossa mente de modo a termos as sensações mais prazerosas que desejássemos. Bastaria que a programássemos com aquilo que quiséssemos experimentar e nos plugássemos nela. Ligados a ela, esqueceríamos que estamos em uma máquina e sentiríamos todos os prazeres como se fossem reais. Estaríamos, contudo, realmente apenas parados, conectados a uma máquina, e nada daquilo seria verdadeiro. Trocaríamos nossa vida real, com seus prazeres e sofrimentos, por uma vida totalmente prazerosa, mas falsa? Nossas intuições parecem dizer que não, o que levaria à conclusão de que prazer não é a única coisa que importa. Ver Nozick, 1974, p. 42.

⁶ Sobre teorias da satisfação de desejos, podemos dizer que elas resistem ao teste da Máquina de Experiência, uma vez que seus defensores podem replicar dizendo que tal máquina não é capaz de satisfazer certos desejos de um sujeito, como, por exemplo, ter experiências reais de prazer. Um crítico da teoria poderia criticá-la, contudo, alegando que o teórico da satisfação de desejos inverte a ordem de explicação dos fenômenos. As coisas não seriam valiosas porque as pessoas as desejam, mas as pessoas as desejam exatamente por as verem como valiosas. A teoria precisaria, então, de uma explicação de por que essas coisas são valiosas para além dos desejos dos sujeitos, e aí algum tipo de objetivismo sobre o valor seria requerido. Ver Brink, 1989, pp. 224-226.

razoáveis dos agentes e certas relações pessoais e sociais como componentes primários de vidas valiosas. Liberdade, educação e as condições de bem-estar básico estão entre os bens que servem como condições necessárias para esses bens intrínsecos e que possuem valor extrínseco como contribuintes causais para o bem-estar humano, assim interpretado. (Brink, 1989, p. 231)

A teoria objetiva do valor precisa construir uma grande teia de explicações para mostrar como aquilo que ela entende como valioso de fato pode ser entendido dessa maneira. Para isso, ela precisa de uma teoria de como a sociedade funciona, de como se dão as relações entre os indivíduos e de como os sujeitos se constituem. Ela leva em conta não apenas questões imediatas ou estados psicológicos, mas uma série de elementos que, em equilíbrio, parecem ser capazes de sustentar uma vida valiosa em termos objetivos de bem-estar.

2.3. O caráter prático da moralidade: discussões e caminhos possíveis

Tendo apresentado e debatido os elementos envolvidos no processo de deliberação racional, vemos que a argumentação de Brink consiste em apresentar essas várias questões para afirmar que a discussão sobre razões e moralidade tem um caráter externista, uma vez que é substantiva, mais do que formal ou *a priori*. Em relação a isso, ele elenca bons argumentos ao discutir as diferentes teorias e alegar que só podemos chegar a uma conclusão após a análise destas. Vale discutirmos, a partir desses elementos, as diferentes possibilidades de abordagem disponíveis para um realista naturalista. Como dissemos algumas páginas atrás, Brink afirma que um internista defende ao menos duas teses: (i) que a moralidade fornece razões para ação de forma necessária; e (ii) que a relação entre obrigação moral e

o fornecimento de razões é interna, ou seja, conceitual (*a priori*). Ele parece querer negar ambas as teses.

As críticas ao ponto (ii) parecem bem estabelecidas, como vimos. Sobre (i), contudo, ele parece se mostrar ambivalente em alguns momentos de sua argumentação, ou poderíamos talvez ler o naturalismo como podendo se comprometer com uma perspectiva alternativa. Por exemplo, ao falar de versões objetivas do egoísmo racional, Brink diz: “Tais concepções objetivas de egoísmo racional implicarão que os agentes normalmente, *talvez sempre*, tenham pelo menos alguma razão (embora não necessariamente uma razão conclusiva) para cumprir até mesmo os requisitos morais que acarretam compromissos com os outros” (Brink, 1989, p. 74, grifo nosso). Em outra passagem ele afirma:

Existem traços de caráter e atividades que são valiosos e cujo valor independe de serem desejados. E, enquanto houver importantes componentes sociais ou relacionados a outros entre esses constituintes objetivos de uma vida valiosa, os agentes normalmente, *talvez sempre*, terão motivos para cumprir exigências morais – mesmo quando essas exigências forem relativas a outros. (Brink, 1989, p. 73, grifo nosso)

E ainda:

[...] em visões objetivas plausíveis sobre valor que reconhecem componentes sociais importantes no bem de uma pessoa, em teorias de racionalidade relativas a agentes ou neutras a agentes, e em visões comuns sobre o conteúdo dos requisitos morais (isto é, sobre o que a moralidade exige dos agentes), todos normalmente, *talvez*

sempre, terão pelo menos alguma razão para cumprir esses requisitos morais e frequentemente terão razões conclusivas para fazê-lo. (Brink, 1989, p. 76, grifo nosso)

Como vimos, Brink defende uma visão objetiva de valor. Se esse é o caso, parece ser possível dizer que a moralidade fornece razões para ação necessariamente, ainda que essa necessidade só seja descoberta *a posteriori*. Brink parece utilizar expressões como “normalmente” e “talvez sempre” para suavizar a tese. Mas parece razoável imaginar que o cenário final de sua descrição do fenômeno moral inclui ao menos uma concessão a uma tese menos forte de internismo, ou, como chamaremos aqui, uma tese híbrida sobre razões e moralidade.

Se essa leitura estiver correta, é viável dizer que Brink pode aceitar (i) que a moralidade fornece razões para ação de forma necessária, enquanto poderia negar (ii) que a relação entre obrigação moral e o fornecimento de razões é interna, ou seja, conceitual (*a priori*). É possível a ele dizer que a relação entre moralidade e razões para ação é necessária, mas que ela só pode ser estabelecida *a posteriori*, por investigação empírica. Ou seja, a ele é permitido alegar que os preceitos da moralidade, por fim, realmente espelham valores não morais de desenvolvimento e bem-estar humano que são de nosso interesse racional, mas que isso só poderia ser afirmado após uma investigação profunda sobre a nossa natureza e sobre como o mundo funciona.

Sua argumentação parece nos permitir uma leitura nesses termos, em especial por levarmos em conta que estamos falando de razões justificatórias e não apenas de razões explanatórias, e pensando no egoísmo racional como teoria da razão. Assim, seria

possível falar em razões para sujeitos mesmo que eles não as reconheçam ou até mesmo as neguem. Isso tornaria inteligível a discussão do amoralista,⁷ uma vez que a motivação parece decorrer de um reconhecimento das razões que é plenamente possível (ou ao menos concebível, como precisa o argumento) um sujeito não possuir.

O amoralista, contudo, no fim das contas, acabaria por se revelar irracional, se a tese da relação necessária entre razões e moralidade for verdadeira. E Brink considera isso ao dizer que: “Não há nada acerca de externismo sobre motivos que nos impeça de mostrar que um amoralista é irracional por não se importar com considerações morais. Se o amoralista é irracional, depende da racionalidade, não da força motivacional, das considerações morais” (Brink, 1986, p. 31). Assim, podemos dizer que, decerto, Brink sustenta fortemente o externismo motivacional, mas, quando pensamos no externismo de razões, talvez encontremos uma versão fraca, ou híbrida da teoria.⁸

⁷ O amoralista é uma figura capaz de compreender os enunciados morais, mas que não se move em direção aos ditames éticos. Brink (1989) constrói tal figura como forma de mostrar a plausibilidade do externismo motivacional. Parece ser amplamente razoável, de fato, ao menos conceber um sujeito que compreenda o que está envolvido na empreitada moral e que, ainda assim, não se sinta minimamente motivado a agir de acordo com ela. Este talvez seja um indivíduo com o qual não gostaríamos de conviver e que certamente teria problemas no âmbito social, mas não parece incoerente ou ilógico afirmar a possibilidade de sua existência. Brink sustenta que, ainda que ele não exista na realidade, a mera plausibilidade de sua existência já indica que somos capazes de pensar a moralidade dissociada da motivação. Motivação e moralidade guardariam, então, segundo Brink, uma relação íntima, mas não conceitual ou necessária.

⁸ A depender da compreensão do que é internismo e externismo, alguém poderá dizer que essa seria uma tese internista. Isso porque, de fato, o que conta realmente é a relação necessária entre moralidade e razões para ação. Contudo, se pensarmos que o externismo é normalmente descrito como exigindo uma relação conceitual e *a priori* entre moralidade e motivação, a proposta não seria totalmente internista. Daí a ideia de pensar em algo híbrido ou em uma versão fraca de externismo (ou internismo).

Uma crítica que, logo de início, pode ser levantada é a de que, para garantir a relação entre razões e moralidade (algo que parece ser importante se se quer manter uma visão não revisionista do fenômeno moral), é necessário percorrer um longo caminho que passa por comprometimentos substantivos em relação à natureza dos valores, da racionalidade e do agente. Há muito o que se provar para que a estrutura pare em pé. De fato, algumas escolhas e comprometimentos teóricos que levem ao fornecimento de razões para ação de forma substantiva e não *a priori* serão bastante complexos e muitas investigações deverão ser feitas a fim de se provar o que se espera provar. Isso é algo com o que uma teoria realista naturalista deve realmente se comprometer, mas está longe de ser uma derrota do argumento.

Brink, quando apresenta sua teoria do valor objetivo, assume que os elementos de uma vida valiosa dizem respeito ao exercício de certas capacidades humanas e que: “a avaliação das atividades e das capacidades exercidas será guiada, entre outras coisas, por concepções normativas da natureza humana e ideais de pessoa” (Brink, 1989, p. 232). Podemos pensar, então, que uma leitura neoaristotélica estaria disponível a um realista moral naturalista. Haveria, pois, uma relação mais forte e necessária entre moralidade e razões para ação, uma vez que, se a moralidade de algum modo realiza a nossa natureza, as razões morais podem ser coincidentes com nossas razões para agir.⁹

⁹ Richard Boyd, outro realista de Cornell, segue mais claramente por essa vertente, ver Boyd, 2003a e 2003b. Ele propõe compreender a moralidade como envolvendo um agregado homeostático de propriedades de florescimento. Entendida por esse viés, a teoria de Boyd sustenta que a moralidade, uma vez que diz respeito às nossas capacidades humanas, consistiria em uma forma de realização da nossa natureza. Assim, quando pensamos no que seria uma teoria acerca da natureza humana, suas

Talvez Brink evitasse explicitamente esse caminho, a despeito da citação acima, mas ela parece estar disponível, em especial, em uma descrição da moralidade tal qual a apresentada por Richard Boyd e sua versão de florescimento. O que importa, contudo, é que uma saída mais internista poderia ser pensada e desenvolvida em uma perspectiva naturalista como a dos realistas de Cornell. Ela teria a vantagem de mostrar como os deveres morais podem ser deveres genuínos; estaria empiricamente justificada e afastaria os fantasmas do relativismo. Precisaria, contudo, lidar com questões de caráter teleológico e sua relação com a moralidade. Isso, contudo, não faria com que essa opção estivesse descartada, de saída, para um realista como Boyd.

3. Considerações finais

A questão da normatividade é, sem dúvida, um dos maiores desafios para um naturalista na moral. Isso porque o caráter prático parece ser, de fato, algo muito próprio do fenômeno moral. O realista não precisa, contudo, assumir o ônus de abrir mão totalmente da explicação do âmbito prático da ética. Essa já nos pareceu uma opção bastante convincente e atrativa. Ela é, contudo, bastante extrema. Uma visão que exclua a dimensão prática do discurso moral parece revisionista em um nível que não é necessário para um realista. Ele não precisa olhar para a normatividade com tamanho peso como fazem alguns de seus adversários, mas pode ser capaz de dar uma boa resposta para o caráter prático do discurso moral.

capacidades e sobre como ela pode melhor se organizar e se desenvolver em suas dimensões pessoais e sociais, temos um caminho (ainda que bastante complexo) que possibilita abarcar alguma normatividade que leve um realista a aceitar uma relação mais fundamental entre razões e moralidade.

Para fazer isso, um naturalista precisa observar com cuidado cada elemento que constitui o lado prático da moralidade e perceber qual é constitutivo do fenômeno e qual é próximo e familiar, embora não constitutivo. Ao distinguir motivação e razões e tratar especificamente da segunda categoria, nossa intenção foi pensar separadamente problemas de naturezas distintas que, muitas vezes, são tratados e pensados conjuntamente. É plenamente possível dar respostas distintas à questão da motivação e do fornecimento de razões. Em relação à motivação, de fato uma leitura canônica dos realistas de Cornell como defensores estritos do externismo parece ser bastante estabelecida e incontroversa. Contudo, no âmbito das razões certamente há, ao menos, alguma margem para pensar alternativas possíveis a um externismo estrito.

Como vimos, os realistas naturalistas parecem ser vistos como defensores estritos também do externismo de razões. Sendo essa visão bastante radical e um pouco revisionista dos compromissos gerais da moralidade, um realista naturalista parece, de fato, não precisar assumi-la necessariamente. Assim, neste capítulo, buscamos problematizar a questão e mostrar que, embora Brink (o principal expoente do externismo moral) seja visto corretamente como um defensor do externismo motivacional, ele é ambivalente em relação ao externismo de razões. Após longas digressões, ele sempre parece estar aberto à possibilidade de haver uma relação necessária (ainda que *a posteriori*) entre a moralidade e o fornecimento de razões e poderia ser lido como defendendo uma versão fraca de externismo ou mesmo uma versão híbrida entre internismo e externismo.

Buscamos mostrar, então, com a análise da resposta de Brink – e também com a sinalização da proposta de Boyd de um realismo baseado em propriedades de florescimento –que é possível considerar

algum nível de internismo no naturalismo. Isso nos levaria a postular que a moralidade pode sim fornecer razões necessariamente, dependendo de como compreendemos o que são razões para ação e a que tipo de agente estamos nos referindo. Talvez, de fato, o externista motivacional seja possível, mas considerado irracional, uma vez que não reconhece razões que deveria acolher.

A moralidade, como um realista naturalista busca mostrar, é a moralidade para seres como nós e se relaciona com nossas capacidades humanas, com nossas possibilidades enquanto seres biológicos, racionais e sociais, e é utilizada para marcar certos fenômenos que fazem sentido dada nossa interação uns com os outros e com o mundo físico. Somente levando todos esses elementos em consideração é possível construir um panorama em que a moralidade possa ser compreendida como parte do tecido do mundo e tomada, de alguma forma, como normativa para o tipo de pessoas que nós somos.

Referências

- Blackburn, S. (1984) *Spreading the Word*. Oxford: Clarendon Press.
- Boyd, R. (1988) How to Be a Moral Realist. In: Sayre-McCord, G. (Org.) *Essays on Moral Realism*. Ithaca: Cornell University Press, pp. 187-228.
- Boyd, R. (1989) What Realism Implies and What It Does Not. *Dialectica*, v. 43, n. 1/2, pp. 5-29.
- Boyd, R. (2003a) Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part I. *Philosophy and Phenomenological Research*, 66, pp. 505-553.
- Boyd, R. (2003b) Finite Beings, Finite Goods: The Semantics, Metaphysics and Ethics of Naturalist Consequentialism, Part II. *Philosophy and Phenomenological Research*, 67, pp. 24-47.

- Brink, D. O. (1986) Externalist Moral Realism. *The Southern Journal of Philosophy*, 26, Supplement.
- Brink, D. O. (1989) *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Dancy, J. (1993) *Moral Reasons*. Oxford: Blackwell.
- Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Clarendon Press.
- Mackie, J. (1977) *Ethics: Inventing Right and Wrong*. London: Penguin Books.
- Moore, G. E. (1993) *Principia Ethica. A Revised Edition*. Cambridge: Cambridge University Press.
- Nozick, R. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- Shafer-Landau, R. (2003) *Moral Realism: A Defence*. Oxford: Clarendon Press.
- Smith, M. (2004) *The Moral Problem*. Oxford: Blackwell Publishing.
- Stevenson, C. (1944) *Ethics and Language*. New Haven: Yale University Press and London: Humphrey Milford, Oxford University Press.
- Stocker, M. (1979) Desiring the Bad: An Essay in Moral Psychology. *Journal of Philosophy*, v. 76, n. 12, pp. 738-753.

A imaginação e o agir moral¹

Claudia Passos-Ferreira²

1. Introdução

A ideia de que a imaginação desempenha um papel crucial na psicologia moral tem uma longa história. Ela percorre uma extensa tradição filosófica sentimentalista que remonta a David Hume ([1739] 1978) e Adam Smith ([1759] 2015). Mais recentemente, Alvin Goldman (2006), Stephen Darwall (1998) e Antti Kauppinen (2010, 2014) enfatizaram a função de imaginar os sentimentos do outro no julgamento moral. Muitos estudos empíricos também têm sido usados para sustentar a visão de que as capacidades imaginativas e a adoção da perspectiva afetiva do outro (*affective perspective-taking*) executam papéis centrais na moralidade e em comportamentos pró-sociais e altruístas (Eisenberg; Strayer, 1987; Batson *et al.*, 1981; Batson; Shaw, 1991; Zahn-Waxler; Radke-Yarrow, 1990; Hoffman, 2000; Vaish;

¹ Este capítulo é uma tradução do seguinte texto original em francês: Passos-Ferreira, C. (2020) L'imagination et l'agir morale. In: Beziau, J-Y.; Schulthess, D. (Orgs.) *L'Imagination. Actes du 37e Congrès de l'ASPLF* (Rio de Janeiro, 26-31 março 2018). London: College Publications, Academia Brasileira de Filosofia, v. 1, pp. 583-600.

² Professora do Centro de Bioética da New York University.

Carpenter; Tomasello, 2009; Vaish; Warneken, 2011; Decety, 2004; Tomasello; Vaish, 2013).

Na filosofia moral, tem renascido o interesse por posições sentimentalistas da moralidade, para as quais processos imaginativos exercem uma função central na deliberação moral. Para sentimentalistas, nossos julgamentos morais são constituídos por sentimentos de aprovação e desaprovação, e, apenas a partir deles, podemos formular nossos julgamentos morais. Segundo Hume e Smith, esses sentimentos se baseiam em nossa capacidade de sentir empatia (*simpatia*)³. A empatia é uma tendência da imaginação de compartilhar os estados afetivos de outrem. Nossa apreciação de ações requer que nos coloquemos imaginativamente no lugar dos outros e sintamos o que os afetados pela ação possam sentir ou pensar.

Mesmo alguns racionalistas passaram a considerar processos imaginativos como cruciais para o raciocínio moral em situações de indeterminação moral e dilemas morais. A imaginação pode desempenhar muitos papéis diferentes na moralidade. Ela pode ser usada para pensar cenários futuros, para se colocar no lugar do outro, para idealizar ideais de vida, para usar metáforas morais ao raciocinar e para projetar uma autoimagem moral ou uma imagem moral dos outros (Coeckelbergh, 2007). Neste capítulo, estou sobretudo interessada no papel de imaginar os sentimentos dos outros na moralidade.

Ao mesmo tempo, é um assunto controverso se a imaginação de fato exerce uma função crucial na psicologia moral. Uma razão para

³ “*Simpatia*” é o termo original em Hume ([1739] 1978) e Smith ([1759] 2015). Conforme a literatura contemporânea, “*empatia*” é a palavra utilizada para designar a capacidade de compartilhar estados afetivos de outrem e “*simpatia*” para o sentimento de preocupação moral por outrem.

rejeitar esse papel deriva de uma perspectiva racionalista, segundo a qual a moralidade é compreendida como um sistema governado por princípios morais complexos que definem o que é certo e o que é errado (Johnson, 2016). Em um sistema racionalista, princípios morais racionais que determinam proibições e prescrições complexas derivam do entendimento e do raciocínio puro. Não há muito espaço para estados subjetivos da imaginação.

Recentemente, diversos filósofos (Nichols, 2004; McGeer, 2008; Goldie, 2011; Maibom, 2010) têm argumentado que a imaginação e a habilidade de assumir a posição afetiva do outro não desempenham uma tarefa essencial no entendimento moral ou na motivação moral. Eles recorrem a casos de pessoas com autismo e de crianças pequenas, utilizando-os como contraexemplos à ideia de que a imaginação e a adoção da perspectiva do outro executam uma função importante na moralidade. Os dois grupos têm problemas com a imaginação e a adoção da perspectiva do outro, mas, ainda assim, eles parecem entender as normas morais e a desaprovação moral. Os cenários das pessoas com autismo e das crianças pequenas mostrariam que explicações do julgamento moral com base na ideia de adotar a perspectiva do outro, tais como a explicação sentimentalista, seriam empiricamente inadequadas para descrever o julgamento moral.

Para mim, em contraste, a imaginação constitui um aspecto distintivo do julgamento moral e a teoria sentimentalista é capaz de explicar esse tipo de julgamento, inclusive no caso de pessoas com autismo e crianças pequenas. Ela ocupa uma posição central não apenas em comportamentos altruístas e pró-sociais e em ações de ajuda e cooperação, mas também em normas e restrições morais explícitas – tais como normas sobre causar dano e matar – e na internalização dessas regras explícitas.

A abordagem neossentimentalista de Antti Kauppinen (2010) permite explicar os julgamentos de crianças pequenas e pessoas com autismo, cujas atitudes de desaprovação moral estão ligadas à aplicação rígida de regras. Segundo Kauppinen, os julgamentos morais desses indivíduos não resultam de uma competência moral genuína. Eles são dependentes do julgamento moral comum. Por sua vez, o julgamento moral comum decorre do processo sentimentalista de sentir empatia pelas emoções do outro. O exercício imaginativo e a simulação da perspectiva afetiva do outro se estabelecem como um elemento central nesse processo.

Argumento que crianças pequenas, embora ainda não manifestem uma competência moral plena, manifestam uma forma básica de competência moral que lhes permite expressar julgamentos morais de modo a satisfazer determinadas condições sentimentalistas. Essa forma básica de competência moral é a capacidade de adotar afetivamente a perspectiva do outro e de se preocupar empaticamente, algo que elas manifestam em situações em que uma vítima é prejudicada. Sustento que essa capacidade básica de adotar a perspectiva do outro é o que permite a crianças pequenas internalizarem regras morais durante o seu processo de educação moral.

2. Posições antiempáticas

A empatia tem sido conceituada de diferentes maneiras (Eisenberg; Strayer, 1987; Batson, 2011; Batson *et al.*, 1997; Hoffman, 2000; de Vignemont; Singer, 2006; Decety; Jackson, 2006). Em linhas gerais, empatia significa “sentir como o outro sente”. Ela é uma reação afetiva vicária que ocorre como resposta a sinalizações explícitas e perceptíveis, que indicam o estado afetivo do outro, ou como uma

consequência de se inferir o estado afetivo do outro a partir de sinalizações indiretas (Eisenberg; Strayer, 1987). A literatura psicológica contemporânea distingue dois processos: a *empatia afetiva*, que consiste em compartilhar estados afetivos com outra pessoa, e a *empatia cognitiva*, que se resume em adotar cognitivamente a perspectiva dos pensamentos e crenças de outra pessoa. Eu proponho uma distinção adicional entre *empatia perceptual* e *empatia imaginativa*, focada não nas reações empáticas (cognitivas ou afetivas), mas nos mecanismos psicológicos subjacentes necessários para acessar os estados afetivos dos outros (Passos-Ferreira, 2015). A empatia perceptual é uma resposta a uma percepção direta das emoções de outrem. A empatia imaginativa é uma resposta à simulação imaginativa ou projetiva da perspectiva afetiva dos outros.

O tema da empatia suscita várias questões éticas. Fomenta problemas normativos a respeito de sua parcialidade, seus perigos e seus limites. Estimula indagações metaéticas relativas ao seu papel constitutivo no julgamento moral. Ante o entusiasmo recente a respeito do papel das emoções e da empatia na psicologia moral, tem-se desenvolvido um movimento contrário, mesmo em meio àqueles que endossam uma explicação sentimentalista da moralidade. Alguns defendem um “otimismo prudente” acerca da empatia (Holton; Langton, 1998) e outros expressam ceticismo e adotam uma posição antiempática (Prinz, 2011a, 2011b; Bloom, 2016). A posição antiempática argumenta de diferentes maneiras contra o papel necessário da empatia na moralidade. Alguns argumentam que a empatia não é requerida para o julgamento moral ou a motivação moral (Kennett, 2002; McGeer, 2008; Prinz, 2011a, 2011b; Maibom, 2009, 2010). Outros alegam que a empatia tem uma função negativa

na moralidade e não deveria ser usada como um guia moral (Maibom, 2009; Prinz, 2011a, 2011b; Bloom, 2016; Goldie, 2011).

Muita da crítica recente à empatia tem se dirigido contra o papel da empatia afetiva. Em trabalho anterior, discuti os problemas da posição antiempática de Prinz, a qual rejeita a empatia afetiva como uma condição e um guia para a moralidade (Passos-Ferreira, 2015). Argumentei em favor de uma visão mais otimista sobre a empatia, na qual ela executa uma função primordial no desenvolvimento moral e é nosso melhor guia para a moralidade, em determinados contextos.

Meu foco neste capítulo será a empatia imaginativa, cujo papel na moralidade é rejeitado por várias posições antiempáticas recentes. Tais concepções sustentam que a empatia imaginativa – nossa capacidade de adotar a perspectiva do outro e de imaginar os sentimentos e pensamentos do outro – não é necessária para a moralidade. A tese central aqui é que nossas habilidades imaginativas de pensar sobre outrem não são necessárias para o julgamento moral. Examinarei especialmente duas abordagens: a crítica de Heidi Maibom ao papel da imaginação da perspectiva do outro na moralidade⁴ e a visão cética de Shaun Nichols a respeito de explicações da moralidade a partir da ideia da adoção da perspectiva do outro.

⁴ Maibom (2010) argumenta que “imaginar-se na pele de outra pessoa” não é eficaz para a motivação moral e pode causar sofrimento nos indivíduos. Simpatizar por pessoas com sofrimento levaria ao sofrimento pessoal, isto é, uma forma empática de adoção da perspectiva do outro poderia levar ao mesmo sofrimento pessoal. Ela sustenta que imaginar os pensamentos e sentimentos dos outros é algo enviesado e impreciso e que se deveria evitar como guia moral. Embora estes sejam argumentos interessantes, eu discuti tais temas normativos relacionados aos perigos da empatia em trabalho anterior e os deixarei de lado aqui, ver Passos-Ferreira, 2015.

Maibom (2010) não adota uma posição particular acerca de como nós formamos os nossos julgamentos morais. Em sua visão instrumentalista da moral, o que conta como ações morais são aquelas que resultam na adesão a normas morais ou atos supererrogatórios que não são motivados pelo interesse próprio, sob a compreensão de que prescrições morais são universais e não estão sujeitas à mudança por uma autoridade específica e de que sua transgressão é passível de punição. Maibom distingue dois tipos de normas morais: as *normas supererrogatórias* e as *normas de proibição*. Normas supererrogatórias são normas que são mais elogiáveis do que obrigatórias, tais como as ações de ajuda e de tentar ser uma boa pessoa, em que a imaginação pode desempenhar um papel positivo. Normas de proibição são “normas morais brutas” que incluem restrições explícitas, tais como proibir prejudicar ou matar os outros, em que a imaginação não tem qualquer função. Maibom sustenta que nós não precisamos de habilidades imaginativas para pensar sobre o que é certo ou errado. Imaginar o que os outros estão sentindo ou pensando não seria central para aspectos básicos da competência moral, tais como compreender o que é certo ou errado.

Maibom discute evidências psicológicas substanciais de que imaginar estar no lugar de outra pessoa estimula a motivação pró-social e promove ações voltadas para ajudar os outros. Estudos conduzidos por Batson e colegas (2003) mostram que imaginar estar na posição de outrem pode promover ações pró-sociais que beneficiam os outros. Assim, conclui Maibom, se a imaginação dos sentimentos alheios desempenha algum papel na moralidade, trata-se mais de um papel em promover ações de ajuda e estabelecer relações sociais do que em compreender normas de proibição.

Nichols (2004) propõe uma complexa teoria sentimentalista da moralidade. Ele defende que nós possuímos um mecanismo psicológico responsável por “julgamentos morais básicos”, que nos permite julgar uma classe específica de ações como não permissíveis ou transgressoras em um sentido específico. Nós distinguimos entre normas convencionais e normas morais. Normas convencionais são normas em resposta a determinada autoridade. Normas morais são regras sentimentais. Isto é, normas morais envolvem regras prescritivas que são acompanhadas por respostas afetivas. Nichols argumenta que a capacidade para julgamentos morais básicos surge antes da capacidade de adotar a perspectiva dos outros e ser capaz de fazer simulações mentais dos estados mentais alheios. Se este for o caso, a empatia imaginativa não desempenha um papel nos julgamentos morais básicos.

Nichols recorre a pessoas com autismo e crianças pequenas para sustentar sua posição. Ambos os grupos populacionais têm problemas para imaginar e para adotar a perspectiva do outro. Embora pessoas com autismo e crianças pequenas sejam capazes de se engajar em processos afetivos empáticos, elas parecem falhar ao tentarem imaginar os estados mentais e afetivos de outrem. Elas têm dificuldade para pensar sobre o que os outros sentem ou pensam. Contudo, parecem compreender normas morais e a desaprovação moral; são capazes de entender o que é certo e o que é errado e de seguir regras morais. Tais casos sugerem que nós nem sempre chegamos aos julgamentos morais da forma padrão que as explicações sentimentalistas sugerem – através de exercícios imaginativos e processos de simulação. Às vezes, nossos julgamentos morais resultam de reações afetivas imediatas ou de um processo não emocional de

respeito por regras explícitas – tais como regras sobre prejudicar e matar.

A partir desse corpo de evidências, a perspectiva antiempática conclui que o julgamento moral não requer imaginação e empatia. Gostaria de argumentar contra essa posição. Na minha visão, a imaginação é crucial não apenas em ações supererrogatórias de ajuda, como os atos altruístas, mas também na compreensão e internalização de normas e regras morais explícitas. Isto permite que uma teoria sentimentalista explique os julgamentos morais de crianças pequenas e de pessoas com autismo.

3. Posições pró-empáticas: o sentimentalismo moral

Para o sentimentalismo em filosofia moral, emoções e sentimentos estão no cerne de nossas avaliações morais. Hume e Smith, em particular, enfatizam o papel da empatia e da imaginação no julgamento moral. Sentimentalistas morais sugerem que nós precisamos da imaginação para desaprovar moralmente pessoas e ações. O que significa “desaprovar moralmente pessoas e ações”? Podemos desaprovar uma ação de várias maneiras: esteticamente, por meio do gosto, por meio de preferências. O que caracteriza a desaprovação moral? A fim de desaprovar moralmente uma ação, eu me abstraio dos meus gostos e preferências particulares e adoto uma perspectiva imparcial. Para ser capaz de me alhear das minhas preferências e pensar sobre quais seriam as reações de um ser humano normal qualquer, eu preciso fazer uso da minha imaginação e me colocar na perspectiva de um observador imparcial.

Meu foco aqui é a teoria histórica de atitudes morais de Antti Kauppinen, a qual ele chama de *sentimentalismo histórico*. Nos termos de Kauppinen (2010), sentimentalistas sobre julgamento moral

explicam a desaprovação moral como uma atitude que resulta de se colocar imparcialmente na pele daqueles afetados pela ação e compartilhar suas atitudes reativas negativas. Assim, ela envolve a empatia afetiva,⁵ isto é, o compartilhamento de sentimentos e atitudes reativas daqueles afetados pela ação, mas envolve também e, sobretudo, a empatia imaginativa, ou seja, a simulação da reação de um espectador imparcial afetado pela ação.

A atitude da desaprovação moral envolve três tipos de processos: busca de informação, empatia (*simpatia*) e correção reflexiva de distorções de perspectiva típicas. Deste modo, *o processo histórico* através do qual chegamos a um julgamento moral é determinante para que se considere ou não o julgamento como sendo moralizante. Uma atitude apenas é moral caso ela resulte de um processo histórico que transcenda uma perspectiva egocêntrica pura. Para sentimentalistas como Kauppinen, moralizar é o processo de formação de um julgamento que transcende a própria perspectiva egocêntrica através da simulação de reações emocionais – como a raiva, a gratidão, a culpa e o desgosto – que qualquer participante informado e não enviesado teria nas circunstâncias daqueles afetados pela ação ou de um espectador, caso ninguém seja afetado diretamente (Kauppinen, 2010). Esse processo requer um exercício imaginativo.

Tal exercício imaginativo é uma ferramenta epistêmica poderosa. Ele permite compreender a perspectiva dos outros e evitar distorções de pontos de vista. Ademais, ele pode ser uma fonte crucial

⁵ Para uma visão alternativa sobre a empatia afetiva, ver Kauppinen, 2014. O que interessa para o estudo dos julgamentos morais não é o compartilhamento imediato e não controlado dos estados afetivos *efetivos* do outro, mas a *empatia regulada* (*regulated empathy*), isto é, a empatia tomada como uma forma de regulação emocional.

de informação em *situações opacas*,⁶ que são aquelas em que a informação sobre os estados mentais alheios não é transparente para o espectador. Nesses casos, o espectador precisa compreender a situação sem perceber diretamente sinalizações afetivas. A imaginação e a simulação podem desempenhar um papel central nesses casos.

No processo de simulação, quando atribuímos estados mentais aos outros, nossos mecanismos psicológicos operam de forma desconectada, na medida em que usam aspectos da situação da pessoa-alvo em vez de aspectos da nossa própria situação como fontes de informação. O que é simulado não é uma decisão real, mas uma reação hipotética a determinada situação. A simulação envolvida na geração de um sentimento moral nos força a transcender nossas próprias reações imediatas: nós não apenas imaginamos como nos sentiríamos na situação de outra pessoa, mas também como nos sentiríamos caso fôssemos livres das distorções potenciais que derivam da nossa posição particular.

Qual tipo de habilidade imaginativa está envolvido no julgamento moral? Imaginar a situação dos outros objetivamente sem engajamento empático não suscita sofrimento e, portanto, evita os problemas da empatia, mas também não motiva o sujeito a ajudar ou a agir em benefício de outrem. A adoção da perspectiva do outro que parece relevante para a moralidade é colocar-se na posição de outra pessoa e imaginar ser outra pessoa, o que mobiliza a empatia. Portanto, a imaginação empática que mobiliza a capacidade de agir moralmente consiste em se identificar com o outro, imaginando-se em

⁶ Para uma discussão detalhada da função da imaginação empática em situações de opacidade, nas quais as reações emocionais do outro não são diretamente perceptíveis ou acessíveis, ver Passos-Ferreira, 2015.

sua situação. Isto parece compatível com a posição de Smith, segundo a qual a empatia não precisa envolver de fato os sentimentos de outrem. Como ele exprime em uma famosa passagem:

Embora nosso irmão esteja sendo torturado, enquanto nós mesmos estamos tranquilos, nossos sentidos jamais nos informarão sobre o que ele sofre. Pois não podem e jamais poderão levar-nos para além de nossa própria pessoa, e apenas pela imaginação nos é possível conceber em parte quais as suas sensações. Tampouco essa faculdade nos pode ajudar senão representando para nós as próprias sensações se nos encontrássemos em seu lugar. Nossa imaginação apenas reproduz as impressões de nossos sentidos, e não as alheias. Por intermédio da imaginação, podemos nos colocar no lugar do outro, concebemo-nos sofrendo os mesmos tormentos, *é como se entrássemos no corpo dele e, de certa forma, nos tornássemos a mesma pessoa*, formando, assim, alguma ideia das suas sensações, e até sentindo algo que, embora em menor grau, não é inteiramente diferente delas. Assim incorporadas em nós mesmos, adotadas e tornadas nossas, suas agonias começam finalmente a nos afetar, e então trememos, e sentimos calafrios, apenas à imagem do que ele está sentindo. (Smith, [1759] 2015, I, i, 1, § 2, pp. 73-74, grifos nossos)

Portanto, na imaginação empática nós nos colocamos na situação do outro para aprendermos seu sentimento, sua experiência, seu sofrimento.

4. Tipos de julgamento moral: três casos salientes

No entanto, a imaginação não está envolvida ativamente em todos os casos de julgamento moral. Nossa experiência moral e nosso discurso moral cotidianos sugerem haver julgamentos morais que não resultam da simulação de sentimentos, pensamentos ou da perspectiva de outros. Tanto o senso comum quanto a pesquisa psicológica mostram que somos capazes de desaprovar moralmente ações e pessoas sem passarmos pelo processo da imaginação e da adoção afetiva da perspectiva alheia. A desaprovação moral pode resultar de reações afetivas imediatas (sentimentos viscerais) ou regras explícitas (normas sobre prejudicar e matar). Tais casos constituem um desafio para a teoria sentimentalista histórica.

Além disso, há evidências empíricas de que crianças pequenas e pessoas com autismo conseguem compreender o que é certo e o que é errado e seguir regras morais, mesmo que tenham problemas para simular a perspectiva de outros. Como apontado por Maibom (2010) e Nichols (2004), embora pessoas com autismo frequentemente se comportem de forma a não considerar os outros adequadamente, elas raramente agem de forma drasticamente imoral e parecem ser capazes de reconhecer as noções de certo e errado. Elas aparentam entender que prejudicar e matar os outros é errado. Este fato nos dá razões para duvidar que a habilidade para simular os sentimentos alheios seja necessária para a moralização.

Esses casos sugerem que nem sempre chegamos aos julgamentos morais da forma padrão proposta por teorias sentimentalistas – através de exercícios imaginativos e processos de simulação. Às vezes, nossos julgamentos morais resultam de reações afetivas imediatas ou de um processo de seguir regras explícitas, como as normas sobre prejudicar, matar e mentir.

Acredito que uma teoria sentimentalista do julgamento moral pode acomodar esses cenários. É possível que as desaprovações morais desencadeadas por reações afetivas e as formadas a partir da aplicação de regras pertençam à mesma classe de julgamentos morais sentimentalistas. Todos esses tipos de julgamentos morais dependem de processos que se conformam a condições sentimentalistas. O que conta como moralizar depende do *modo* como chegamos ao julgamento moral.

Processos sentimentalistas caracterizam a desaprovação moral, da qual toda atitude moral resulta (Kauppinen, 2010). O fato de que nem todos os julgamentos morais sejam produto da empatia imparcial sentimentalista não solapa uma teoria sentimentalista histórica das atitudes morais. Atitudes morais desencadeadas por sinalizações afetivas são dependentes de processos sentimentalistas ou *parasitam* esses processos. Reações viscerais ou sentimentos não corrigidos *dependem assimetricamente* de atitudes morais que se seguem da empatia e da correção reflexiva. Para que conte como possuindo atitudes morais, uma pessoa deve ao menos às vezes chegar à aprovação ou à desaprovação como resultado da adoção de uma perspectiva de observador imparcial. Indivíduos que formam atitudes exclusivamente com base em reações afetivas ou na aplicação de regras não estão moralizando. Em contraste, sujeitos que formam atitudes exclusivamente com base na imaginação de uma perspectiva imparcial ainda estão moralizando. Assim, atitudes morais dependem ao menos indiretamente da imaginação.

A seguir, discuto os três casos mais difíceis para uma posição centrada na imaginação: reações viscerais e regras explícitas tanto para autistas quanto para crianças pequenas. Mostro que os cenários do julgamento moral resultante de reações viscerais e de pessoas com

autismo seguindo regras explícitas podem ser explicados pela ideia de dependência assimétrica proposta por Kauppinen (2010). Nessas conjunturas, julgamentos morais dependem de atitudes morais como resultado de um processo sentimentalista. Ofereço uma explicação diferente para o caso de crianças pequenas que seguem regras explícitas. Existem evidências recentes de que crianças pequenas são capazes de adotar a perspectiva do outro e de manifestar preocupação em situações em que uma vítima é prejudicada. Argumento que esse comportamento pode valer como precursor de um processo sentimentalista de desaprovação moral através da imaginação afetiva da perspectiva dos outros.

4.1 Reações viscerais

Reações afetivas podem desencadear atitudes de desaprovação moral. Podemos desaprovar moralmente outras pessoas sem termos que passar pelo processo de imaginação e adoção afetiva da perspectiva alheia. Os estudos psicológicos de Haidt, Koller e Dias (1993) mostram que julgamentos morais podem resultar de reações afetivas automáticas como o nojo ou a raiva. Eles constataam que pessoas de status socioeconômico baixo tendem a considerar moralmente erradas ações que lhes causam nojo, sem serem capazes de explicar as razões para tal. Por exemplo, é mais provável que elas considerem que indivíduos que fazem sexo com uma galinha morta ou que limpam vasos sanitários com a bandeira nacional deveriam ser impedidos de fazê-lo ou punidos. Este resultado indica que podemos formar julgamentos morais à base de sentimentos viscerais, usando sentimentos como o nojo e a raiva como guias para a ação moral.

Julgamentos morais baseados exclusivamente em sentimentos viscerais poderiam pôr em xeque o papel necessário da imaginação no

juízo moral. Contudo, não é óbvio que esse tipo de comportamento de fato conta como moralização genuína. Em uma teoria sentimentalista como a de Kauppinen (2010), o que figura como moralizante envolve simular as reações emocionais que qualquer outra pessoa teria nas mesmas circunstâncias. Sentimentos viscerais somente contam como moralizantes quando expressam atitudes de desaprovação provenientes de um agente capaz de imaginar e adotar a perspectiva afetiva de outros. Os traços característicos do juízo moral dependem do modo como chegamos às atitudes morais. Um agente moral pode, por vezes, alcançar a desaprovação moral como consequência de um processo de simulação da perspectiva afetiva alheia. Nesse contexto, nossos sentimentos viscerais se apoiam nas atitudes que emergem do processo sentimentalista. Eles parasitam a capacidade do juízo moral normal. Contudo, um agente que nunca reflete sobre ou corrige seus sentimentos viscerais não profere julgamentos morais genuínos. Assim, julgamentos morais baseados em sentimentos viscerais dependem, em última instância, da imaginação.

4.2 Regras explícitas e autismo

Alguns filósofos (Nichols, 2004; Maibom, 2010) argumentam que a imaginação desempenha um papel limitado naquilo que eles chamam de “moralidade bruta”, “julgamentos a frio” ou “seguir regras explícitas”, que são categorias que incluem, por exemplo, princípios que nos impedem de prejudicar ou matar outros.

Maibom argumenta que a capacidade de imaginar os sentimentos ou estados mentais do outro tem uma função crucial apenas em normas que essencialmente se referem a estados mentais representacionais do sujeito (tais como ajudar alguém caso a pessoa

esteja passando por necessidade), sem que ela exerça qualquer papel na maioria das normas sobre o que é moralmente desaprovado.

Como Kauppinen aponta, um sentimentalista pode aceitar que alguns dos nossos julgamentos sejam baseados em regras e princípios, desde que tal regra seja uma generalização de vereditos sentimentais em casos particulares. Isto é, em última instância, um julgamento moral frio e baseado em regras deve se fundamentar em julgamentos sentimentais.

O caso mais desafiador para a posição sentimentalista envolve pessoas com autismo, que são incapazes de simular a reação dos outros e possuem problemas em suas habilidades de leitura da mente alheia, mas que ainda assim parecem ser capazes de desaprovação moral. Elas aparentam ser aptas para distinguir entre violações de normas morais e violações de normas convencionais. Pessoas com autismo, que têm problemas para adotar a perspectiva do outro e para imaginar os estados afetivos alheios, não parecem ter dificuldades para entender e aderir às normas morais centrais. A partir dessa evidência, filósofos como McGeer (2008), Nichols (2004), Maibom (2009) e Kennett (2002) concluem que a empatia imaginativa não é necessária para a competência moral.

Uma maneira de a posição sentimentalista explicar as evidências relacionadas às “regras frias” que pessoas com autismo conseguem seguir é mostrar que essas regras não sentimentais, de fato, se apoiam em regras sentimentais normais. Nas palavras de Kauppinen, “a capacidade dos autistas de fazer julgamentos morais, quero sugerir, é parasitária em relação à capacidade dos julgadores morais ‘normais’ [...] Eles precisam recorrer aos outros para decidir, por exemplo, quando a culpa é apropriada. (2010, p. 251). Assim, pessoas com autismo podem depender psicologicamente da aprovação e da

desaprovação morais de pessoas que não possuem alguma deficiência. Outra hipótese proposta por Kauppinen (2010) é a de que pessoas com autismo manifestam uma capacidade preservada para respostas emocionais e que elas podem estar utilizando seus próprios sentimentos e atitudes reativas para gerar algumas regras morais.

Uma terceira hipótese é sugerida por de Vignemont e Frith (2008).⁷ Elas mostram que não há indícios fortes de que a capacidade aparente de pessoas com autismo de proferirem julgamentos morais seja o resultado de elas aplicarem regras morais ou experimentarem preocupação moral. Ademais, existem evidências mostrando que pessoas com autismo apresentam limitações na habilidade de experimentar e manifestar sentimentos centrados em outros sujeitos, como a simpatia e a preocupação. Além disso, elas raramente expressam sentimentos por ou em relação a outros indivíduos (Hobson *et al.*, 2009). Pessoas com autismo conseguem detectar a transgressão de regras normativas e o sofrimento em outrem. No entanto, elas não parecem ser capazes de detectar violações morais. A hipótese das autoras do estudo é a de que sujeitos com autismo simplesmente percebem as regras que seguem como regras convencionais e que sua aptidão aparente para o julgamento moral resulta da aplicação dessas regras convencionais.

4.3 Crianças pequenas e a empatia imaginativa

Psicólogos do desenvolvimento frequentemente sustentam que crianças até três ou quatro anos de idade não possuem uma teoria da mente desenvolvida e não são capazes de simular a perspectiva ou as reações de outra pessoa (Rochat, Passos-Ferreira, 2009; Rochat *et al.*

⁷ Ver Passos-Ferreira, 2015, para uma discussão desta hipótese.

2009). Como consequência, elas teriam dificuldade de se engajar na adoção do ponto de vista do outro ou no exercício imaginativo da perspectiva do outro. Não obstante, estudos indicam que essas crianças conseguem expressar preocupação, exibir comportamentos pró-sociais, distinguir entre regras convencionais e não convencionais e desaprovar certos comportamentos.

Tais resultados são, por vezes, captados como evidências de que crianças pequenas são capazes de moralizar. Contudo, elas não parecem chegar ao julgamento moral por meio do processo tipicamente sentimentalista de imaginar e simular a perspectiva afetiva dos outros. Suas atitudes parecem expressar respostas afetivas automáticas ou a aplicação de regras explícitas. Segundo uma interpretação desses resultados, crianças pequenas primeiro desenvolvem uma moralidade baseada em normas, seguindo regras explícitas, e somente mais tarde desenvolvem a aptidão mais complexa de imaginar a perspectiva alheia (Maibom, 2010). Outra possibilidade é a de que elas sigam regras sentimentais básicas (*core moral judgments*) (Nichols, 2004). Se essa interpretação estiver correta, será preciso concordar com a concepção de Maibom de que a capacidade de imaginar os sentimentos dos outros é relevante para a interação social e contribui para melhorar nossas relações sociais, mas não é necessária para a formulação e compreensão de julgamentos morais.

Outra interpretação é a *dependência assimétrica* proposta por Kauppinen (2010). Kauppinen rejeita a ideia de que as atitudes das crianças pequenas contam como moralizantes. Visto que elas apenas exibem respostas afetivas, mas não exibem atitudes reativas (tais como a culpa, o ressentimento e a gratidão) formadas através do processo sentimentalista, seus julgamentos ainda não são plenamente

morais. Seus julgamentos não convencionais dependem dos julgamentos morais que resultam da empatia sentimentalista imparcial. Isto é, os julgamentos formulados pelas crianças pequenas seriam derivados da competência moral dos adultos, cujos julgamentos decorrem de um processo sentimental histórico.

Eu discordo tanto de Maibom quanto de Kauppinen. Embora eu concorde com Kauppinen que crianças pequenas não são capazes de raciocínio moral pleno, eu sustento que elas exibem uma forma básica de julgamento moral que expressa uma aptidão mínima para a empatia e para a adoção afetiva da perspectiva do outro. Esta habilidade é crucial para a sua educação moral.

Tradicionalmente, estudos experimentais sobre o desenvolvimento da simpatia, da empatia e de atitudes pró-sociais testam as habilidades de crianças de três a quatro anos de demonstrar simpatia na presença de sinalizações perceptuais visíveis (Eisenberg *et al.* 2006). Eles mostram que, quando se apresenta a crianças dessa faixa etária uma pessoa visivelmente em sofrimento, elas são capazes de sentir empatia pela vítima e exibir atitudes pró-sociais em relação a ela. A interpretação usual desse experimento é a de que crianças pequenas sentem empatia através de contágio emocional ou identificando sinais emocionais. Contudo, essa via para a empatia poderia estar baseada inteiramente na leitura das sinalizações emocionais visíveis por parte da vítima, sem que a imaginação desempenhe qualquer papel.

Recentemente, um novo paradigma experimental foi desenvolvido por Hobson, Harris, García-Pérez e Hobson (2009) e adaptado por Vaish, Carpenter e Tomasello (2009) para testar a simpatia (sentimento de preocupação com o outro) sem leitura emocional perceptual.

Eles verificaram se crianças pequenas e pessoas com autismo são capazes de simpatizar com um sujeito que passou por algo negativo, mas que não expressa qualquer emoção. Nesse experimento, cada um dos participantes e dois pesquisadores desenharam uma imagem. Em seguida, um pesquisador (o agressor) ou bem rasgou inesperadamente o desenho do outro pesquisador (a vítima), no caso da condição experimental, ou rasgou uma folha de papel em branco, na condição de controle. Em ambos os casos, a vítima observa a ação do agressor de forma neutra. Foram analisadas as feições e a preocupação das crianças. Este paradigma experimental permite que se avaliem a resposta emocional das crianças (sua preocupação com a vítima) a uma situação negativa e o modo como elas chegam ao julgamento moral (resposta afetiva), sem que haja qualquer aporte afetivo visível (sem sentimentos viscerais ou contágio emocional).

Os dados mostram que crianças muito novas (de 18 a 25 meses de idade) são capazes de simpatizar com uma vítima na ausência de um sinal emocional explícito (Vaish *et al.* 2009). Esse resultado sugere que elas sentem simpatia através de um tipo de adoção afetiva da perspectiva alheia, ao imaginarem o que a outra pessoa está sentindo com base em várias sinalizações não emocionais e ao se colocarem no lugar do outro.

Através de quais mecanismos as crianças simpatizam? Nesse contexto, a simpatia não parece resultar diretamente da exposição às sinalizações emocionais da vítima (tais como a imitação de sinalizações emocionais ou o contágio emocional). Ela parece decorrer da imaginação e da adoção da perspectiva afetiva do outro. Esse processo envolve considerar a fonte do sentimento de outrem, o que implica um foco sobre o outro. Em uma teoria sentimentalista, isso inclui inferir o estado afetivo do outro ao se colocar no lugar dele

e basear a própria resposta nessa inferência. Na ausência de sinalizações emocionais, uma maneira de construir essa inferência é a simulação, que consiste em imaginar-se na situação de outrem. Esses estudos indicam que crianças com menos de três anos de idade, isto é, que ainda não desenvolveram habilidades cognitivas e linguísticas sofisticadas, podem possuir a capacidade de adotar a perspectiva afetiva do outro. Embora não haja consenso se crianças de dois anos de idade dispõem de uma teoria da mente, é consensual que as aptidões imaginativas das crianças aparecem aos 24 meses de idade (Passos-Ferreira, 2018).

Portanto, as habilidades imaginativas e de adoção da perspectiva do outro emergem bem cedo no desenvolvimento ontogenético, e permitem explicar como crianças pequenas conseguem expressar desaprovações morais.

Em resumo, reações viscerais e o respeito por regras explícitas sem atuação da imaginação podem ser explicados pelo fato de o julgamento moral depender assimetricamente da imaginação, tal como proposto por Kauppinen. Contudo, para crianças pequenas a moralização pode envolver um papel mais direto para a imaginação. Elas parecem dispor da capacidade mínima para a adoção da perspectiva do outro requerida para moralizar, segundo a teoria sentimentalista de Kauppinen. A agência moral surge gradativamente, à medida que outras habilidades cognitivas são adquiridas, expandindo suas capacidades de moralização.

Uma questão final é saber se as ações supererrogatórias de ajuda e cooperação, por um lado, e a moralidade baseadas em normas, por outro, estão associadas a duas rotas distintas para a moralidade, como a abordagem de Maibom sugere. Tomasello e Vaish (2013) propõem que a emergência ontogenética da moralidade humana pode se dar

paralelamente às origens filogenéticas da moralidade. Começamos com a cooperação mútua e com interações motivadas pró-socialmente e direcionadas a indivíduos específicos, depois desenvolvemos uma moral impessoal fundada sobre convenções sociais e normas morais. Nessa abordagem, crianças menores de três anos de idade não são na verdade capazes de compreender normas morais brutas. O que seu comportamento revela não é o respeito por normas morais, mas uma resposta a imperativos impostos por adultos. Se isto estiver correto, os primeiros estágios do julgamento moral podem se iniciar com a empatia imaginativa e a adoção afetiva da perspectiva do outro.

5. Conclusão

Neste capítulo, defendi uma posição sobre a moralidade centrada na imaginação, segundo a qual a empatia imaginativa e a adoção da perspectiva do outro são centrais para o raciocínio moral. Com recurso à teoria sentimentalista histórica de Kauppinen, argumentei que a moralização por via de sentimentos viscerais e pelo uso de regras explícitas em pessoas com autismo e em crianças pequenas não solapa a posição centrada na imaginação. Existem explicações alternativas para tais casos que são consistentes com a ideia de que o julgamento moral depende de processos sentimentalistas em que a empatia imaginativa e a adoção da perspectiva do outro desempenham um papel central.

Também considerei uma nova maneira de explicar por que mesmo crianças pequenas podem ter a capacidade para julgamento moral fundada na imaginação. Se essa visão estiver correta, encorajar as crianças a imaginarem os sentimentos e pensamentos dos outros pode se mostrar como a melhor forma de contribuir não somente para

sua socialização e interação cooperativa, como também para sua educação moral e internalização de regras morais.

Referências

- Batson, C. D. (2011) *Altruism in Humans*. New York: Oxford University Press.
- Batson, C. D. et al. (1981) Is Empathic Emotion a Source of Altruistic Motivation? *Journal of Personality and Social Psychology*, v. 40, n. 2, pp. 290-302.
- Batson, C. D. et al. (1997) Is Empathy-Induced Helping Due to Self-Other Merging? *Journal of Personality and Social Psychology*, v. 73, n. 3, pp. 495-509.
- Batson, C. D. et al. (2003) As You Would Have Them Do Unto You: Does Imagining Yourself in the Others Place Stimulate Moral Actions? *Personality and Social Psychology Bulletin*, v. 29, n. 9, pp. 1190-1201.
- Batson, C. D.; Shaw, L. L. (1991) Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry*, v. 2, pp. 107-122.
- Bloom, P. (2016) *Against Empathy: The Case for Rational Compassion*. New York: Ecco.
- Coeckelbergh, M. (2007) *Imagination and Principles: An Essay on the Role of Imagination in Moral Reasoning*. London: Palgrave-Macmillan.
- Darwall, S. (1998) *Empathy, Sympathy, Care*. *Philosophical Studies*, v. 89, n. 2-3, pp. 281-282.
- Decety, J. (2004) L'empathie est-elle une simulation mentale de la subjectivité d'autrui. In: Berthoz, A.; Jorland, G. (Orgs.) *L'empathie*. Paris: Odile Jacob, pp. 53-88.

- Decety, J.; Jackson, P. L. (2006) A Social-Neuroscience Perspective on Empathy. *Current Directions in Psychological Science*, v. 15, n. 2, pp. 54-58.
- De Vignemont, F.; Frith, U. (2008) Autism, Morality and Empathy. In: Sinnott-Armstrong, W. (Org.) *Moral Psychology. Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development*. Cambridge, MA: MIT Press, pp. 273-280.
- De Vignemont, F.; Singer, T. (2006) The Empathic Brain: How, When and Why? *Trends in Cognitive Sciences*, v. 10, n. 10, pp. 435-441.
- Eisenberg, N. (2000) Emotion, Regulation, and Moral Development. *Annual Review of Psychology*, v. 51, n. 1, pp. 665-697.
- Eisenberg, N.; Spinrad, T. L.; Sadovsky, A. (2006) Empathy-Related Responding in Children. In: Killen, M.; Smetana, J. G. (Eds.) *Handbook of Moral Development*. Mahwah, NJ: Erlbaum, pp. 517-549.
- Eisenberg, N.; Strayer, J. (1987) Critical Issues in the Study of Empathy. In: Eisenberg, N.; Strayer, J. (Orgs.) *Empathy and Its Development*. Cambridge: Cambridge University Press, pp. 3-13.
- Goldie, P. (2011) Anti-Empathy: Against Empathy as Perspective Shifting. In: Coplan, A.; Goldie, P. (Orgs.) *Empathy: Philosophical and Psychological Perspectives*. New York: Oxford University Press, pp. 302-317.
- Goldman, A. I. (2006) *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. New York: Oxford University Press.
- Haidt, J. (2001) The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, v. 108, pp. 814-834.

- Haidt, J.; Koller, S. H.; Dias, M. G. (1993) Affect, Culture, and Morality or Is It Wrong to Eat Your Dog? *Journal of Personality and Social Psychology*, v. 65, n. 4, pp. 613-628.
- Hobson, J. A. et al. (2009) Anticipatory Concern: A Study in Autism. *Developmental Science*, v. 12, n. 2, pp. 249-263.
- Hoffman, M. L. (2000) *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge: Cambridge University Press.
- Holton, R.; Langton, R. (1998) Empathy and Animal Ethics. In: Jamieson, D. (Org.) *Singer and His Critics*. Oxford: Blackwell, pp. 209-232.
- Hume, D. ([1739] 1978) *A Treatise of Human Nature*. Nidditch, P. H. (Org.). Oxford: Oxford University Press.
- Johnson, M. (2016) Moral Imagination. In: Kind, A. (Org.) *The Routledge Handbook of Philosophy of Imagination*. Abingdon, New York: Routledge, pp. 255-367.
- Kauppinen, A. (2010) What Makes a Sentiment Moral? In: Shafer-Landau, R. (Org.) *Oxford Studies in Metaethics, Vol. 5*. Oxford, GB: Oxford University Press, pp. 225-256.
- Kauppinen, A. (2014) Empathy, Emotion Regulation, and Moral Judgment. In: Maibom, H. L. (Org.) *Empathy and Morality*. New York: Oxford University Press, pp. 97-121.
- Kennett, J. (2002) Autism, Empathy and Moral Agency. *The Philosophical Quarterly*, v. 52, n. 208, pp. 340-357.
- Maibom, H. L. (2009) Feeling for Others: Empathy, Sympathy, and Morality. *Inquiry*, v. 52, n. 5, pp. 483-499.
- Maibom, H. L. (2010) Imagining Others. *Atelier de l'éthique*, v. 5, n. 1, pp. 34-49.
- McGeer, V. (2008) Varieties of Moral Agency: Lessons from Autism and Psychopathy. In: Sinnott-Armstrong, W. (Org.) *Moral*

- Psychology. Volume 3. The Neuroscience of Morality: Emotion, Brain Disorders, and Development.* Cambridge, MA: MIT Press, pp. 227-257.
- Nichols, S. (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment.* New York: Oxford University Press.
- Pacherie, É. (2004) L'empathie et ses degrés. In: Berthoz, A.; Jorland, G. (Orgs.) *L'empathie.* Paris: Odile Jacob, pp. 149-181.
- Passos-Ferreira, C. (2015) *In Defense of Empathy: A Response to Prinz.* Abstracta, v. 8, n. 2, pp. 31-51. Reprinted in Portuguese: Passos-Ferreira, C. (2018) Em defesa da empatia: uma resposta a Prinz. In: Copp, David; Williges, Flavio (Eds.) *O lugar das emoções na ética e na metaética.* Pelotas: Ed. UFPel, pp. 137-176.
- Passos-Ferreira, C. (2018) Imagination in Infants. *The Junkyard: A Scholarly Blog Devoted to the Study of Imagination.* Disponível em: <https://junkyardofthemind.com/blog?author=5a39c170419202030ee8459a>. Acesso em: 21 de março de 2018.
- Prinz, J. (2011a) Is Empathy Necessary for Morality? In: Coplan, A.; Goldie, P. (Orgs.) *Empathy: Philosophical and Psychological Perspectives.* New York: Oxford University Press, pp. 211-239.
- Prinz, J. (2011b) Against Empathy. *Southern Journal of Philosophy*, v. 49, pp. 214-233.
- Rochat, P.; Passos-Ferreira, C. (2009) From Imitation to Reciprocation and Mutual Recognition. In: Pineda, Jaime A. (Org.) *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition.* New York: Humana Press, Springer Science, pp. 191-212.
- Rochat, P.; Passos-Ferreira, C.; Salem, P. (2009) Three Levels of Intersubjectivity in Early Development. In: Carassa, A.; Morganti, F.; e Riva, G. (Eds.) *Enacting Intersubjectivity: Paving the Way for*

- a Dialogue between Cognitive Science, Social Cognition and Neuroscience*. Lugano: Università della Svizzera Italiana, pp. 173-190.
- Smith, A. ([1759] 2015) *Teoria dos sentimentos morais*. Trad. Lya Luft. São Paulo: Editora WMF Martins Fontes.
- Tomasello, M.; Vaish, A. (2013) Origins of Human Cooperation and Morality. *Annual Review of Psychology*, v. 64, pp. 231-255.
- Vaish, A.; Carpenter, M.; Tomasello, M. (2009) Sympathy through Affective Perspective Taking and Its Relation to Prosocial Behavior in Toddlers. *Developmental Psychology*, v. 45, n. 2, pp. 534-543.
- Vaish, A.; Warneken, F. (2011) Social-Cognitive Contributors to Young Children's Empathic and Prosocial Behavior. In: Decety, J. (Org.) *Empathy: From Bench to Bedside*. Cambridge, MA: MIT Press, pp. 131-146.
- Zahn-Waxler, C.; Radke-Yarrow, M. (1990) The Origins of Empathic Concern. *Motivation and Emotion*, v. 14, n. 2, pp. 107-130.

A moralidade da ansiedade: problemas e perspectivas¹

Flávio Williges²

1. Introdução

Nos últimos anos, o tema da ansiedade emergiu como um foco central na tradição analítica da filosofia (Prinz, 2024; Kurth, 2018a, 2018b, 2024; Chopra, 2024) e na fenomenologia das emoções (Aho, 2019; Glas, 2020; Ratcliffe, 2017). Esse interesse disseminado pelo fenômeno foi influenciado por diversos fatores. Talvez o principal

¹ Este capítulo é resultado das pesquisas realizadas no âmbito do projeto “Memória e Tecnologias”, do Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Maria (UFSM), durante meu período como Professor Visitante Sênior na University of York, na Inglaterra. Nessa temporada, contei com o apoio da bolsa CAPES-Print, processo nº 88887.935197/2024-00, concedida pela CAPES/UFSM. Agradeço ao Prof. Dr. César de David pelas valiosas orientações que possibilitaram minha participação no Programa de Internacionalização da UFSM. Manifesto, também, minha gratidão a Leonardo de Mello Ribeiro, Vitor Somavilla e Rafael Graebin Vogelmann pelas críticas e sugestões oferecidas às versões anteriores deste trabalho.

² Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal de Santa Maria. Bolsista de produtividade em pesquisa do CNPq.

deles seja a frequente discussão da ansiedade no discurso social contemporâneo. Numerosos livros e artigos acadêmicos caracterizaram nossa era atual como a “era da ansiedade” (Horwitz, 2013; Haidt, 2024; Chopra, 2024). Não é surpreendente, portanto, que filósofos também venham dando atenção ao tema.

O interesse filosófico pela ansiedade tomou, contudo, diferentes direções. Filósofos analíticos voltaram suas investigações para a ansiedade entendida como uma emoção episódica, de curta duração, e abordaram principalmente seus aspectos epistemológicos e morais, como o papel que ela tem na formação de crenças (Vazard, 2024) e seu impacto na deliberação e juízo moral (Kurth, 2018a, 2018b). Na fenomenologia, por outro lado, o interesse foi direcionado para formas abrangentes e disruptivas de ansiedade, ou seja, para aquelas que alteram a experiência de si e do mundo. A ansiedade não é entendida nesse caso como uma emoção, mas como um humor ou estado de ânimo pervasivo (Aho, 2019). De fato, fenomenólogos como Kevin Aho (2019) e Gerrit Glas (2020) desenvolveram uma descrição detalhada das estruturas da experiência vivida da ansiedade, investigando as rupturas que ela promove na vivência da temporalidade, corporeidade, relacionalidade, e com o próprio self.

Embora os aportes teóricos dessas duas tradições sejam valorizados neste capítulo, o objetivo fundamental será explorar o tipo de ansiedade que tem sido o interesse dominante na filosofia da mente e psicologia moral analítica. É principalmente a ansiedade episódica que será abordada aqui. O foco recairá, em particular, sobre as questões morais que surgem quando sentimos ansiedade diante de conflitos normativos em contextos de escolha moral, bem como sobre a influência dessa emoção em nossas deliberações e no próprio caráter moral.

Para começar a explorar o tema, é importante notar que a preocupação com o significado moral da ansiedade é parte de um movimento filosófico mais amplo de reabilitação das emoções negativas na ética. Estudos sobre inveja (Protasi, 2021), desprezo (Mason, 2003; Bell, 2013), raiva (Cherry, 2021; Srinivasan, 2018) e nojo (Kumar, 2017; Kurth, 2021) têm desafiado abordagens clássicas e modernas que julgavam essas emoções indignas de consideração moral ou simples ameaças à moralidade. Contra a assunção de que tais emoções são más ou influências perniciosas que deveríamos simplesmente suprimir de nossas almas, esses estudos defendem sua contribuição moral e, em alguns casos, seu papel positivo no desenvolvimento da excelência do caráter moral.

Nesse contexto, devemos a Charlie Kurth a apresentação de uma defesa ampla do valor moral da ansiedade. Ele propôs, em oposição às tradições morais que remontam a Platão, Sêneca e Kant, que essa emoção possui tanto valor instrumental, quanto aretaico em nossas vidas morais (Kurth, 2018a, 2018b, 2024). O valor instrumental da ansiedade está vinculado com contribuições metamorais ligadas à cautela e precaução na formulação de juízos. O valor aretaico é relativo à conexão interna entre a virtude da preocupação moral e a ansiedade. Essas duas contribuições morais serão objeto de atenção detalhada no decorrer do capítulo.

Embora influente e inovadora no domínio da ética normativa, a defesa do valor moral da ansiedade de Kurth foi alvo de um número significativo de objeções (Shepherd, 2019; Munch-Juriscic, 2021; Maibom, 2022). Neste capítulo, além de reconstruir os passos fundamentais da argumentação de Kurth em favor do valor moral da ansiedade, eu discuto também as principais objeções que foram lançadas ao seu projeto. Na parte final do capítulo, eu extraio algumas

conclusões sobre as limitações que o estudo das emoções na ética normativa pode enfrentar quando seguimos a tendência de isolar uma emoção qualquer para investigar sua função e seu significado moral. Kurth adota essa tendência em sua análise da ansiedade. Eu mostro, em conexão com a abordagem das emoções na ética do cuidado, que tal orientação obscurece uma série de problemas colocados pela ansiedade em nossas vidas morais. Antes de tudo, no entanto, é importante precisar o que é a ansiedade e quais seus componentes fundamentais.

2. O que é a ansiedade?

Uma abordagem recorrente na filosofia analítica contemporânea das emoções consiste em isolar um componente da psicologia moral e investigar sua função e relevância moral. Kurth faz isso em sua análise da emoção da ansiedade. Ele não examina a ansiedade como parte de uma engrenagem complexa de apegos e conflitos internos da mente, muitos deles inconscientes, que desembocam em questões mais amplas ligadas à moralidade, ameaçando até mesmo a estabilidade do self. A emoção da ansiedade figura, em sua abordagem, como parte de uma psicologia filosófica funcionalista, que segue um modelo biocognitivo das emoções (Kurth, 2018a, pp. 7-8). Nesse modelo, emoções existem como parte de uma equipagem biológica e são ferramentas com funções bem precisas em nossa interação com o mundo e os outros. Seguindo a teoria biocognitiva das emoções, episódios emocionais, como experiências de desconforto e ansiedade diante de uma situação de risco para nossos projetos pessoais, são entendidos como o resultado de dois sistemas interativos: um sistema “central” biologicamente inato e um sistema “de controle” flexível, influenciado culturalmente. O sistema central consiste em programas

afetivos específicos de cada emoção que despertam automaticamente comportamentos estereotipados em resposta a desafios e oportunidades básicas. Essas respostas são biologicamente enraizadas e ocorrem em reação a estímulos específicos. O sistema de controle, por outro lado, molda os estímulos que ativam esses programas afetivos e influencia os comportamentos subsequentes, com base em processos cognitivos e experiências culturais. No caso da ansiedade, o sistema central torna o indivíduo sensível a ameaças incertas no ambiente, desencadeando comportamentos como minimização de riscos e busca de informações. O sistema de controle refina essa resposta incorporando fatores culturais e experiências que moldam a maneira como a incerteza é percebida e enfrentada (Kurth, 2018a, p. 8).

O sistema de controle é um sistema secundário que responde pela dimensão cognitiva da ansiedade. De fato, nesse modelo as emoções são vistas como formas de “avaliação”, na medida em que apreendem seus objetos segundo propriedades “axiológicas” (Deonna; Teroni, 2012; Tappolet, 2016). Kurth segue essa tradição ao conceber a ansiedade como um tipo de “consciência avaliativa” (2018b). O tipo de consciência avaliativa na ansiedade difere daquele presente em emoções como a raiva e o medo. Enquanto o medo é uma resposta a uma ameaça concreta e imediata, a ansiedade tem como foco intencional ameaças potenciais ou incertas. Ela funciona como um sistema de alarme na detecção de riscos potenciais do nosso ambiente físico e social, ou seja, riscos de dano físico, como a ansiedade provocada por dirigir numa estrada movimentada, bem como danos sociais, como a ansiedade desencadeada por uma apresentação ou encontro importante, no qual nossos propósitos pessoais poderão ser de algum modo prejudicados se não causarmos uma “boa impressão”.

Kurth articula esse ponto afirmando que a ansiedade é “uma resposta à incerteza sobre uma ameaça ou desafio potencial” (2018a, p. 96). O objeto ameaçador é percebido como tendo “incerteza problemática” (Kurth, 2018b, p. 96; 2015, p. 5). Sentimos ansiedade *antes* de um exame médico, por exemplo, pois tememos receber um diagnóstico de doença grave. Esse é o sentido de dizer que a ansiedade apresenta cenários de ameaça envolvidos numa avaliação de “incerteza problemática”. Como Kurth coloca: “quando alguém se sente ansioso, percebe a situação em que se encontra como envolvendo uma ameaça ou desafio cujo potencial é imprevisível, incontrolável ou questionável” (2018b, p. 96).

Além da dimensão intencional e avaliativa, a ansiedade vem acompanhada de sentimentos desprazerosos de preocupação e aflição, bem como de tendências comportamentais de cautela e precaução:

A ansiedade desse tipo tipicamente traz pensamentos sobre a própria situação, como “Estou preocupado”, “O que devo fazer?”, sentimentos de valência negativa ligados à inquietação e preocupação, e uma tendência motivacional de cautela diante da ameaça ou desafio incerto que se enfrenta. (Kurth, 2018b, p. 96)

Em síntese, a ansiedade episódica discutida por Kurth tem como marca distintiva um senso elevado de antecipação e preocupação com possíveis danos ao nosso bem-estar, assim como tendências comportamentais que, em contextos que sinalizam risco e incerteza, motivam cautela, prudência, investigação ou observação – tendências de ação que, como veremos, contribuem na deliberação moral e formação de juízos morais.

Tendo como base essa concepção da ansiedade como uma consciência avaliativa ligada a ameaças potenciais desencadeada em contextos pontuais, Kurth sustenta que a má fama dessa emoção no mundo moral é infundada. Vejamos melhor esse ponto.

3. O valor moral da ansiedade: seu papel epistêmico e aretaico

Como vimos, a ansiedade rastreia cenários de ameaça potencial. Em contextos morais, essa característica é epistemicamente relevante e faz da ansiedade uma sensibilidade para saliências morais. Um exemplo histórico relevante de como essa emoção pode ser benéfica nesse sentido foi indicado por Kurth a partir da autobiografia de John Woolman, um auxiliar de escritório *quaker* que, após ser solicitado por seu patrão a redigir um anúncio da venda de uma mulher escravizada, experimentou uma grande aflição e ansiedade, o que acabou desencadeando um conflito interno entre suas crenças e princípios morais que o levou a tornar-se abolicionista e engajar-se num esforço contínuo de oposição à escravidão. Woolman sentiu-se dividido entre atender as demandas de seu empregador e participar de práticas reprováveis e geradoras de sofrimento. A aflição ansiosa experimentada atuou como um alarme moral, sinalizando que sua ação era eticamente duvidosa e impulsionando uma reavaliação de normas pessoais. A ansiedade produziu esse resultado, pois ela vem acompanhada da *tendência motivacional* de cautela e prudência diante da ameaça ou desafios incertos que enfrentamos em ambientes morais (Kurth, 2018b). Nesse caso, ela representou: “uma forma importante de metacognição: uma fonte que nos ajuda a identificar e resolver conflitos em nossas crenças, atitudes e valores” (Kurth, 2018a, p. 6). Sua contribuição instrumental é especialmente de ordem epistêmica. A ansiedade mostra-se um tipo de sensibilidade que

sintoniza agentes morais com situações de risco e incerteza moral. Woolman supostamente experimentou essa emoção por observar que atender o pedido do patrão resultaria num tipo grave de ofensa moral: tratar humanos como artigos de comércio, desrespeitando sua dignidade fundamental.

Mas a defesa do valor moral da ansiedade de Kurth não se restringe ao domínio epistêmico de investigação e resolução de conflitos entre crenças e atitudes morais. Kurth observa que as tendências de ponderação e cautela reveladas por indivíduos como Woolman são próprias de pessoas que possuem a virtude da preocupação moral (*moral concern*). A virtude da preocupação moral envolve, segundo Kurth, duas atitudes: (1) defender o que é certo e bom, especialmente quando o certo está sob risco; e (2) manifestar uma sensibilidade à incerteza, assumindo atitudes de reflexão e ponderação em nossas decisões morais, especialmente quando não está claro como devemos prosseguir (Kurth, 2018a). Em outras palavras, a preocupação moral (*moral concern*) é uma virtude que expressa a atitude moral de se importar com questões morais, mantendo-se vigilante no rastreo de situações morais potenciais e em condutas reflexivas em relação às decisões morais, reconhecendo que posturas de fechamento mental podem ser fontes de erro moral. (Kurth, 2018b, p. 100).

A ansiedade possui valor aretaico justamente porque constitui parcialmente a preocupação moral. Ela responde pela dimensão afetiva da preocupação moral. As pessoas dotadas desse tipo de preocupação tendem a experimentar ansiedade diante situações de incerteza moral. E, como a preocupação moral é uma virtude e os componentes de uma virtude compartilham o valor dessa virtude, Kurth defende que a ansiedade possui valor aretaico.

Para reforçar a relevância da ansiedade para a virtude da preocupação moral, Kurth discute uma crítica que ele chama de “objeção do Xanax”. A objeção do Xanax postula que a preocupação moral poderia existir sem a ansiedade se fosse possível remover os sentimentos negativos a ela associados (através da ingestão de um medicamento ansiolítico como o Xanax). Kurth argumenta que é empiricamente improvável que se possa obter os benefícios epistêmicos na deliberação moral e na tomada de decisões sem também experimentar os custos associados a ela: aflição e desconforto. Ou seja, a ansiedade está profundamente embutida nos mecanismos que permitem a sensibilidade e a responsividade moral que caracterizam a virtude da preocupação moral e, como não teríamos esses ganhos sem ela, então a ansiedade tem valor aretaico por ser parte da virtude da preocupação moral.

Em síntese, a ansiedade é uma emoção valiosa moralmente porque ocasiona uma sensibilidade e uma capacidade de resposta únicas, características que a tornam importante não apenas por seus efeitos epistêmico-instrumentais, mas também por sua relevância na constituição de uma psicologia virtuosa. Com base nessa argumentação, Kurth sustenta que devemos cultivar a ansiedade, assim como recomendamos o cultivo de outras emoções (como compaixão, empatia e, em certos contextos, a raiva) que ajudam a sermos agentes virtuosos e a rastrear questões normativas. No entanto, ele enfatiza que isso não significa experimentar mais ansiedade ou senti-la de maneira mais intensa. A questão central é aprender a sentir ansiedade nos momentos certos e de forma apropriada. Esse ponto desafia teorias tradicionais da virtude e da agência, que tendem a subestimar o papel das emoções negativas, mas também abre espaço para objeções significativas.

4. Críticas ao valor moral da ansiedade

Como vimos, Kurth argumenta que a ansiedade constitui parcialmente a virtude da preocupação moral. Dado que a preocupação moral é uma virtude, a própria ansiedade deve carregar o mesmo valor moral da virtude que ela ajuda a constituir. Ele também vai mais longe e alega que uma ansiedade adequada deve ser cultivada por contribuir para uma psicologia moral virtuosa. Assim como outros filósofos defenderam que a raiva na medida certa tem uma função moral de proteção e defesa contra ofensas (Cherry, 2021; Srinivasan, 2018), a ansiedade na medida certa pode contribuir para nossa responsabilidade e para a própria constituição do caráter moral, uma vez que a virtude da preocupação moral está ligada a ela. Joshua Shepherd (2019) contestou essas duas teses. Em primeiro lugar, ele critica o argumento de que a virtude da preocupação moral depende da ansiedade. Shepherd aponta que a preocupação moral pode ser concretizada de maneiras diferentes entre indivíduos, o que significa que certos componentes, como a ansiedade, podem não ser essenciais para sua manifestação. Ele sugere que é possível imaginar agentes que exibam preocupação moral sem necessariamente experimentarem ansiedade, o que enfraqueceria a tese de que a ansiedade é um componente indispensável e, portanto, portadora de valor aretaico.

A segunda crítica de Shepherd aborda a proposta de Kurth de que a ansiedade merece ser cultivada por seu valor moral. Shepherd destaca que essa emoção é amplamente suscetível a falhas de regulação, como evidenciado pela prevalência de transtornos de ansiedade em diferentes populações. Ele argumenta que casos de ansiedade mal regulada tornam difícil aceitar a ideia de que a ansiedade, por si só, possui valor aretaico. Nesses cenários, o mérito

parece estar menos na ansiedade em si e mais nas capacidades do agente de equilibrar suas emoções de forma eficaz. Shepherd enfatiza que o valor de pacotes funcionais complexos, como a preocupação moral, não se transfere automaticamente para seus componentes individuais, especialmente quando estes últimos são suscetíveis a falhas. Se a ansiedade é positiva apenas quando bem regulada, isso parece ser um detrator da alegação de “a ansiedade *qua* ansiedade tem valor não instrumental (Shepherd, 2019, p. 1.340). Como alternativa, ele sugere que, em vez de cultivar a ansiedade, seria mais produtivo concentrar-se em estratégias para regulá-la adequadamente, minimizando seus aspectos prejudiciais.

Um terceiro aspecto discutido diz respeito à defesa do lugar das emoções na ética desenvolvida por Kurth. Como defensor das emoções na ética, Kurth recusa modelos superintelectualizados e baseados na reflexividade e deliberação racional em torno de normas. Ele se apoia num “modelo de habilidades para a agência virtuosa” (Shepherd, 2019). Este é estabelecido a partir daquele proposto por Julia Annas, porém com algumas diferenças. Annas considera que, quando uma habilidade é adquirida, a atividade do agente habilidoso ocorre automaticamente, sem a necessidade de deliberação consciente. Kurth critica o modelo “subintelectualizado” de Annas, ao mesmo tempo que deseja preservar a tese que as respostas rápidas de ansiedade desempenham um papel essencial ao permitir que o agente virtuoso reavalie e revise seus planos em situações novas ou desafiadoras. Ele sustenta, assim, que a ausência de preocupação ou desconforto nessas circunstâncias seria um sinal de deficiência na agência virtuosa. Kurth propõe, portanto, um modelo que busca maior integração entre mecanismos cognitivos e habilidades.

Shepherd aceita esse aspecto integrativo do projeto de Kurth, mas nega que a ansiedade seja o tipo desejável de resposta às incertezas morais num modelo de habilidades para agência virtuosa. Ele imagina um filósofo, Manuel, que inicialmente experimentava grande ansiedade ao escrever artigos, mas que, com a prática, adquiriu confiança e habilidade, eliminando a sensação de desconforto sem que isso compromettesse sua capacidade de identificar falhas ou de se engajar em um trabalho rigoroso. Isso parece sugerir que respostas não ansiosas à incerteza podem ter resultados positivos. Além disso, Shepherd aponta que há evidências empíricas, especialmente na psicologia do esporte, de que a ansiedade pode prejudicar o desempenho, enquanto a autoconfiança tende a correlacionar-se positivamente com resultados de performance. Ele também observa que o modelo de Kurth não prevê adequadamente esses efeitos adversos da ansiedade. Em outras palavras, Shepherd questiona se a ansiedade é de fato o principal meio de reconhecimento e resposta à incerteza moral, sugerindo que outros mecanismos poderiam executar esse papel. Ele sugere que outras experiências – como o tédio, a frustração, o esforço, e até mesmo aquelas tidas como positivas, como a alegria ou o entusiasmo – também desempenham papéis importantes no enfrentamento de incertezas, como o tipo de contexto de incerteza moral discutido por Kurth.

Em última análise, Shepherd conclui que, embora o trabalho de Kurth sobre o papel da ansiedade seja valioso, ele não abrange toda a complexidade da forma como a mente humana lida com incertezas de tipo moral, assim como não garante o valor aretaico da ansiedade, dado que esse valor parece estar inextricavelmente ligado a mecanismos de regulação emocional. Num nível mais geral, essas dificuldades parecem comprometer até mesmo a ideia de que a

ansiedade pode ter um papel importante no progresso moral. Afinal, a aflição experimentada pode levar tanto à revisão de crenças e valores quanto a resultados negativos, como falhas morais ou paralisia decisória, dependendo do contexto e da capacidade de regulação emocional do agente.

Esse último ponto também foi discutido de forma ampliada por Munch-Jurisc (2021). Sua crítica à abordagem de Charlie Kurth sobre a moralidade da ansiedade se concentra em dois aspectos principais: a dependência de “ferramentas hermenêuticas” para que a ansiedade realize seu potencial benéfico e a necessidade de uma compreensão mais abrangente dos fatores ambientais e epistemológicos que moldam as experiências individuais dessa emoção.

Munch-Jurisc reconhece que Kurth apresenta um caso convincente para o papel positivo da ansiedade, particularmente em relação à “ansiedade prática”, que funciona como um alarme, sinalizando incertezas ou possíveis inconsistências morais e incentivando o agente a reavaliar suas normas ou escolhas. O exemplo do abolicionista John Woolman, citado por Kurth, é emblemático para a compreensão desse papel. Woolman experimentou uma sensação de desconforto ao redigir um documento relacionado à escravidão, o que, segundo Kurth, levou-o a uma reflexão mais profunda e, eventualmente, a uma vida dedicada à causa abolicionista. No entanto, Munch-Jurisc questiona se a ansiedade, por si só e nos termos da psicologia moral de Kurth, é suficiente para produzir esses benefícios. Ele observa que muitos contemporâneos de Woolman provavelmente sentiram um desconforto semelhante, mas não mudaram suas práticas ou crenças; pelo contrário, muitos se reconciliaram com a instituição da escravidão para preservar seus interesses econômicos (Munch-Jurisc, 2021, p. 13.588). Essa

disparidade aponta para a importância das ferramentas hermenêuticas disponíveis ao agente, ou seja, os recursos epistemológicos e interpretativos que permitem a uma pessoa dar sentido à sua ansiedade e utilizá-la para uma mudança positiva. Esses recursos englobam um conjunto de conceitos, normas culturais, práticas sociais e conhecimentos que ajudam os agentes a identificar, interpretar e responder de maneira construtiva aos seus estados emocionais e aos desafios que enfrentam. A ansiedade, segundo Munch-Jurisic, só pode produzir benefícios epistemológicos ou morais quando o indivíduo possui esses instrumentos para compreender sua experiência emocional. Tais ferramentas são moldadas por diversos fatores. O contexto cultural e social desempenha um papel fundamental, uma vez que ele influencia o significado atribuído às emoções. Por exemplo, em uma sociedade que estigmatiza a ansiedade e aflição diante do sofrimento de escravizados, o sujeito pode interpretá-la como um sinal de fraqueza, uma reação incômoda que deve ser rapidamente desprezada, enquanto, em culturas que valorizam a compaixão e respeito à dignidade da vida, a ansiedade tem potencial de ser vista como uma oportunidade de reflexão e crescimento moral. Além disso, as ferramentas hermenêuticas também ajudam as pessoas a compreender e regular sua ansiedade, permitindo-lhes, assim, processar suas emoções e experiências de forma produtiva. É possível que a ausência desses instrumentos, no entanto, leve a consequências negativas. Indivíduos que não conseguem interpretar e regular adequadamente sua ansiedade podem experimentar conformismo ou até mesmo resolver seu desconforto de maneira moralmente questionável.

5. Ansiedade relacional, preocupação e cuidado apropriado

Nas seções anteriores apresentei uma reconstrução da abordagem do valor moral da ansiedade de Charlie Kurth e descrevi um conjunto de críticas que sua interpretação recebeu, destacando falhas fundamentais em seu programa de defesa dessa emoção. Embora Kurth apresente a ansiedade como uma emoção com valor instrumental e aretaico, sustentando que ela contribui para a deliberação em contextos de conflitos morais e como parte da virtude da preocupação moral, as críticas analisadas exploram um elemento comum: uma certa limitação da psicologia moral filosófica adotada por Kurth. Primeiramente, sua abordagem trata a ansiedade como um estado psicológico isolado, desconsiderando seu entrelaçamento com outras emoções e, em particular, a facilidade com que ela pode assumir formas intensas e desreguladas. Além disso, a abordagem “atomística” ignora as complexas dinâmicas emocionais que frequentemente acompanham a ansiedade e sua relação com fatores interpretativos mais amplos, como as “ferramentas hermenêuticas” destacadas por Munch-Jurisc (2021). Essa redução analítica, embora metodologicamente útil, pode fornecer uma perspectiva favorável à ansiedade bem regulada, mas parece implicar, em última instância, uma defesa um tanto quanto “artificial” do significado moral da ansiedade, uma vez que o tratamento adotado isola a emoção da vulnerabilidade a falhas epistêmicas e motivacionais.

Além disso, a abordagem cognitiva de Kurth enfrenta dificuldades práticas: em muitos casos, é quase impossível separar a ansiedade de outros estados como precaução, cautela, pensamentos repetitivos e ruminações. Esses elementos frequentemente introduzem hesitação e dificuldades nas decisões práticas, minando a ideia de que a ansiedade, por si só, gera benefícios morais. Como

Shepherd (2019) aponta, os resultados positivos atribuídos à ansiedade parecem depender não da emoção em si, mas de habilidades de regulação emocional e recursos cognitivos externos à esfera puramente afetiva. Isso sugere que a virtude moral e as contribuições epistêmicas de sentimentos como a ansiedade dependem mais de um “bom temperamento” e da capacidade de controlar e integrar emoções do que da ansiedade propriamente dita. Embora Kurth defenda que a ansiedade promove prontidão para a ação e coordene respostas complexas a desafios significativos, essa visão fica enfraquecida quando reconhecemos as complicações colocadas pela facilidade com que ela se sujeita a descontrolo.

Essas considerações críticas sugerem uma conclusão ligeiramente distinta daquela professada por Kurth: a ansiedade tem um *estatuto moral ambíguo*. Para ilustrar essa tese, eu discutirei a seguir a conexão da ansiedade com a percepção moral (sensibilidade ética e emocional), particularmente em relações pessoais próximas, que são paradigmas da moralidade segundo a tradição da ética do cuidado.

Kurth estrutura sua defesa do papel moral da ansiedade em formas controladas e moderadas da emoção, especialmente no contexto de agentes que buscam alcançar decisões justas ou imparciais, tal como ilustra o caso Woolman. Podemos dizer que sua abordagem se inscreve em uma “concepção jurisdicional da moralidade”, centrada no problema do respeito ou violação de normas morais e direitos individuais, típica do modelo liberal de justiça. Contudo, à luz da ética do cuidado, como proposta por Carol Gilligan (1982) e retomada por Annette Baier (1987), a ansiedade pode ser pensada de maneira distinta: como expressão relacional de uma preocupação moral genuína pelo bem-estar dos outros, enraizada em

laços de interdependência, e não apenas em exigências de imparcialidade. Ficamos ansiosos quando pessoas que amamos podem estar passando por alguma dificuldade psicológica ou doença física. Tal forma de ansiedade, direcionada ao outro, é parte de um modelo ético que dá primazia a relações afetivas e responsabilidades não escolhidas e expressa um ideal moral não exclusivamente direcionado para autonomia e igualdade formal entre agentes. Uma moralidade centrada exclusivamente em direitos pode ser cega a valores interpessoais essenciais – como o cuidado, a confiança e a responsabilidade mútua –, e tais valores não são meramente opcionais, mas indispensáveis à manutenção de vínculos humanos significativos. Nesse modelo, a ansiedade voltada ao cuidado evidencia-se como uma emoção moralmente relevante, precisamente porque expressa uma sensibilidade ao risco de abandono e desconexão, preocupações centrais para uma moralidade comprometida com a sustentação de vínculos de cuidado e atenção, e não apenas com o respeito a direitos. Assim, diferentemente da ansiedade apresentada no modelo “jurisdicional” de Kurth, no qual ela aparece como uma emoção voltada para o monitoramento de riscos morais e como sensibilidade a falhas na preservação de direitos imparciais, a ansiedade relacional está direcionada para nossa condição como seres vulneráveis e dependentes. Como parte de uma ética do cuidado, a ansiedade tem um claro *papel epistêmico* ligado à responsividade moral. Ela atua do mesmo modo que fenômenos como o amor e a preocupação (*care about*), tal como discutidos por Margareth Little (1995) e, antes dela, por Iris Murdoch (1970), que são emoções essenciais para perceber nuances ou saliências relevantes para as pessoas de nossos círculos morais mais próximos. O pressuposto aqui é que ser capaz de perceber o que está em jogo num

contexto moral e fazer a coisa certa depende, em grande medida, de um tipo especial de conexão emocional. Little formula esse ponto nos seguintes termos: “às vezes a verdade é melhor revelada, a paisagem é vista de modo mais claro, a partir de uma perspectiva que tem sido chamada de ‘percepção amorosa’ ou da perspectiva de ‘pensamento empático’” (Little, 1995, p. 118). Mas esse é um ponto que não precisa ficar restrito a emoções como a empatia, amor ou cuidado amoroso. Podemos pensar na ansiedade também como uma sensibilidade moral, na medida que a ansiedade e preocupação com os outros manifestam que nos importamos com eles. Cuidar (*care about*), lembra Baier, “é tornar algo importante; investimos, colocamo-nos naquilo com que nos importamos” (Baier, 1982, p. 272). Parece plausível, nesse sentido, pensar que a ansiedade é parte de um conjunto de atitudes emocionais da família do cuidado, justamente porque, como vimos, a ansiedade reage a potenciais ameaças àquilo que é significativo e de importância para nós mesmos.

À luz dessas considerações, fica evidente que a ansiedade desempenha um papel contributivo para a percepção e ação moral em situações de cuidado interpessoal, quando expressa um comprometimento afetivo com o bem-estar do outro. Nesses casos, sua presença pode, de fato, indicar uma sensibilidade moral genuína. No entanto, a propensão da emoção à desregulação torna difícil sustentar a tese de que a ansiedade possui valor moral intrínseco, como quer Kurth. Vale lembrar que as considerações positivas acerca do valor moral da ansiedade na percepção e na expressão de que nos importamos (ao, por exemplo, nos mostrarmos ansiosos e preocupados com algo importante para alguém que amamos) não podem ser pensadas independentemente de aspectos externos à ansiedade, como a capacidade de regulação emocional e de

ferramentas hermenêuticas para interpretação de seu sentido. No contexto das relações interpessoais com outros que são significativos para nós, é evidente que os padrões comportamentais associados à preocupação constante e intensa podem conduzir a formas desreguladas de preocupação e consideração, as quais possivelmente limitam a autonomia e a expressão pessoal daqueles sob nossos cuidados. Um pai ou mãe excessivamente ansiosos podem, através de sua aflição e medos constantes, gerar insegurança ou indivíduos excessivamente ansiosos e incapazes de reconhecer aspectos da vida (presentes ou passados) que merecem ser celebrados. Além disso, a apreensão constante e falta de calma impedem um tipo de conexão humana profunda que é essencial para uma vida boa. Como Kambartel (2017) argumenta, a calma e a confiança são essenciais para evitar o controle excessivo e permitir formas de interação que promovam abertura e confiança mútua. Em contraste, a ansiedade exacerbada pode desencadear barreiras emocionais, limitando não apenas a expressão de emoções positivas, como orgulho e encorajamento, mas também a capacidade de oferecer um cuidado genuíno e respeitoso. A facilidade de a ansiedade ganhar vida própria sugere que, mesmo que a ansiedade episódica possa, em alguns casos, contribuir para uma atenção positiva e cuidadosa, ela é ainda instável demais para se configurar como emoção moral. Mesmo pais preocupados e atenciosos estão sujeitos a formas de ansiedade que possivelmente assumem um padrão de descontrole. Tudo isso sugere que a ansiedade é muito instável para receber a credencial de uma emoção que merece ser cultivada para nosso próprio benefício moral.

6. Considerações finais

Este capítulo passou em revista os aspectos fundamentais da defesa do valor moral da ansiedade de Charlie Kurth, bem como as principais críticas dirigidas à sua abordagem. A revisão crítica revelou que a ansiedade, longe de ser apenas um estado emocional disfuncional, pode, sob certas condições, desempenhar um papel crucial na moralidade. Contudo, seu valor depende da regulação adequada e articulação com outras disposições morais e fatores pessoais. Por outro lado, quando analisada de forma atomística, temos um retrato irrealista ou artificial da ansiedade como uma emoção funcional e fácil de controlar e compreender. Nesse sentido, a consideração de fatores externos à ansiedade episódica, como componentes hermenêuticos que interferem na forma como a ansiedade é apreendida e elementos ligados à própria dinâmica psicológica do controle emocional, sugerem uma conclusão mais fraca do que aquela defendida por Kurth: *a ansiedade tem um valor moral ambíguo*. Seu significado moral depende de aspectos contextuais, da capacidade do agente para regular suas emoções e da disponibilidade de quadros interpretativos, que muitas vezes estão associados à fragilidade existencial. Nessas condições, discutidas aqui no quadro da ética do cuidado, em vez de ampliar a perspectiva moral do agente acerca da situação de seus parceiros morais, a ansiedade frequentemente distorce a percepção dos riscos e obscurece o que realmente está em jogo. Em vez de favorecer ações cuidadosas, a ansiedade possivelmente oblitera ou distorce aspectos de nossa realidade: o foco real da emoção pode ser o próprio medo de falhar ou de ser julgado e não o bem-estar do outro em si. Assim, o outro aparece menos como um fim e mais como um reflexo do próprio desconforto.

Nessas circunstâncias, torna-se difícil sustentar que a ansiedade possua valor moral instrumental ou intrínseco.

Referências

- Aho, K. (2019) *Contexts of Suffering: A Heideggerian Approach to Psychopathology*. London: Rowman & Littlefield.
- Baier, A. (1982) Caring about Caring. *Synthese*, n. 53, pp. 273-290.
- Baier, A. (1987) The Need for More than Justice. *Canadian Journal of Philosophy*, n. 17, Supplementary, v. 13, pp. 41-56.
- Bell, M. (2013) *Hard Feelings: The Moral Psychology of Contempt*. Oxford: Oxford University Press.
- Brady, M. S. (2024) Anxiety's Allure. In: Rondel, D; Chopra, S. (Orgs.) *The Moral Psychology of Anxiety*. New York: Lexington Books, pp. 197-212.
- Cherry, M. (2021) *The Case for Rage*. New York: Oxford University Press.
- Chopra, S. (2024) *Anxiety: A Philosophical Guide*. Princeton: Princeton University Press.
- Deonna, J.; Teroni, F. (2012) *The Emotions: A Philosophical Introduction*. London: Routledge.
- Gilligan, C. (1982) *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Glas, G. (2020) An Enactive Approach to Anxiety and Anxiety Disorders. *Philosophy, Psychiatry, & Psychology*, v. 27, n. 1, pp. 35-50.
- Haidt, J. (2024) *The Anxious Generation: How the Great Rewiring of Childhood Is Causing an Epidemic of Mental Illness*. New York: Penguin.

- Horwitz, A. (2013) *Anxiety: A Short History*. Baltimore: The Johns Hopkins University Press.
- Kahn, J. (2013) *Angst: Origins of Anxiety and Depression*. Oxford: Oxford University Press.
- Kambartel, F. (2017) On Calmness. *Philosophia*, n. 45, pp. 1.613-1.619.
- Kumar, V. (2017) Foul Behavior. *Philosopher's Imprint*, v. 17, n. 15, pp. 1-17.
- Kurth, C. (2015) Moral Anxiety and Moral Agency. In: Timmons, M. (Org.) *Oxford Studies in Normative Ethics*, Vol. 5. Oxford: Oxford University Press, pp. 171-195.
- Kurth, C. (2018a) *The Anxious Mind: An Investigation into the Varieties and Virtues of Anxiety*. Cambridge, MA: MIT Press.
- Kurth, C. (2018b) Anxiety: A Case of Study on the Value of Negative Emotion. In: Tappolet, C.; Teroni, F.; Ziv, A. K. (Orgs.) *Shadows of the Soul: Philosophical Perspectives on Negative Emotions*. London: Routledge, pp. 95-104.
- Kurth, C. (2021) Cultivating Disgust: Prospects and Moral Implications. *Emotion Review*, v. 13, n. 2, pp. 101-112.
- Kurth, C. (2024) Moral Anxiety: A Kantian Perspective. In: Rondel, D; Chopra, S. (Orgs.) *The Moral Psychology of Anxiety*. New York: Lexington Books, pp. 149-172.
- Little, M. O. (1995) Seeing and Caring: The Role of Affect in Feminist Moral Epistemology. *Hypatia*, v. 10, n. 3, pp. 117-137.
- Maibom, H. L. (2022) Don't Worry, Be Happy? *Synthese*, n. 200, 67. Disponível em: <https://doi.org/10.1007/s11229-022-03597-y>.
- Mason, M. (2003) Contempt as a Moral Attitude. *Ethics*, v. 113, n. 2, pp. 234-272.
- Mitchell, J. (2021) Affective Shifts: Mood, Emotion and Well-Being. *Synthese*, n. 199, pp. 11.793-11.820.

- Munch-Jurisc, D. M. (2021) Lost for Words: Anxiety, Well-Being, and the Costs of Conceptual Deprivation. *Synthese*, 199, pp. 13:583-13:600.
- Murdoch, I. ([1970] 2021) *The Sovereignty of Good*. London: Routledge.
- Prinz, J. (2024) Generalized Anxiety Disorder: Natural, Normative or Neither? In: Rondel, D; Chopra, S. (Orgs.) *The Moral Psychology of Anxiety*. New York: Lexington Books, pp. 99-124.
- Protasi, S. (2021) *The Philosophy of Envy*. Cambridge: Cambridge University Press.
- Ratcliffe, M. (2017) *Real Hallucinations: Psychiatric Illness, Intentionality and the Interpersonal World*. Cambridge, MA: MIT Press.
- Rondel, D.; Chopra, S. (2024) *The Moral Psychology of Anxiety*. New York: Lexington Books.
- Rossi, M.; Tappolet, C. (2024) The Anguish of Ivan Karamazov. In: Rondel, D; Chopra, S. (Orgs.) *The Moral Psychology of Anxiety*. New York: Lexington Books, pp. 63-83.
- Shepherd, J. (2019) Review of *The Anxious Mind: An Investigation into the Varieties and Virtues of Anxiety*, by Charlie Kurth. *Mind*, v. 128, n. 512, pp. 1:337-1:344.
- Srinivasan, A. (2018) The Aptness of Anger. *The Journal of Political Philosophy*, v. 26, n. 2, pp. 123-144.
- Tappolet, C. (2016) *Emotions, Values, and Agency*. Oxford: Oxford University Press.
- Vazard, J. (2024) The Epistemic Virtue of Anxiety. In: Rondel, D; Chopra, S. (Orgs.) *The Moral Psychology of Anxiety*. New York: Lexington Books, pp. 173-196.

Parte 4

Justificação moral e normatividade

Quem tem competência ética?

Leonardo de Mello Ribeiro¹

1. Como identificar competência ética?

Consideremos a informação de um físico ou astrônomo de que as leis dos movimentos planetários de Kepler estabelecem, entre outras coisas, que a Terra e os demais planetas movem-se em elipses em torno do Sol e em variadas velocidades; ou a informação de um virologista de que a forma mais eficaz e rápida de erradicação de um determinado vírus se dará através da aplicação em massa de uma vacina; ou a informação de um médico que, através de exames, detecta um problema cardíaco em um paciente; ou a informação da previsão do tempo local de que irá chover torrencialmente na sua cidade nas próximas horas; ou a informação de um engenheiro de que um avião pesando 640 toneladas é capaz de voar; ou ainda a informação de um matemático de que, após 358 anos, o último teorema de Fermat foi finalmente provado por Andrew Wiles.

Com exceção de algumas visões implausíveis sobre a natureza da justificação de nossas crenças, todas as informações que acabamos de descrever estão aptas a se tornar objeto de uma crença justificada

¹ Professor do departamento de filosofia e do programa de pós-graduação em filosofia da Universidade Federal de Minas Gerais.

mesmo para alguém que não seja capaz de compreender como aquelas informações e conclusões foram obtidas. Se um especialista em física e astronomia tentasse demonstrar matematicamente a mim as três leis de Kepler, mas eu fosse incapaz (por limitação de raciocínio matemático e conhecimento de física) de acompanhar todos os passos da demonstração, eu ainda assim estaria justificado a acreditar na opinião desse especialista. Tenho razão para tal crença justamente porque ela foi transmitida por um especialista, que detém competência e autoridade no assunto. Há duas explicações por trás dessa razão.

Em primeiro lugar, a fonte da informação precisa ser *confiável*. Ser um especialista, obviamente, não garante competência, nem carrega por si só credenciais de confiabilidade. O que conta como testemunho confiável pode variar em cada um dos casos listados acima, mas inevitavelmente envolve uma dimensão social. Nas sociedades contemporâneas, existe uma instituição social do conhecimento científico, organizada na forma de comunidades de especialistas, que é responsável pela produção desse conhecimento e por torná-lo publicamente acessível através de educação formal, divulgação, implementação de políticas públicas estatais, desenvolvimento tecnológico etc. A dimensão social da confiabilidade, aquilo que Annette Baier (1986) identificou como pré-condição de qualquer atividade cooperativa que está na fundação das sociedades humanas, é uma das bases da razão para se acreditar em especialistas. Isso é simplesmente parte do modo como as sociedades humanas contemporâneas estão estruturadas, que pressupõe uma divisão social do trabalho na produção e distribuição do conhecimento. Dessa perspectiva, a identificação de critérios de competência e autoridade em qualquer uma das áreas anteriormente

mencionadas envolve, também, juízos de confiabilidade social. Mas confiabilidade social, nesse sentido específico, não será o principal foco de nossa discussão.

Uma segunda explicação para a razão de acreditarmos nas informações de especialistas está relacionada à própria natureza da investigação daquelas áreas científicas. Todos aqueles casos podem ser tomados como fornecendo dados de especialistas cuja competência pode ser verificada através de *critérios objetivos* oferecidos pelas comunidades científicas de cada uma das áreas referidas. Isso é verificável de modo objetivo não apenas no sentido de que as predições de físicos, astrônomos, virologistas, médicos, engenheiros etc. podem ser confirmadas pela realidade empírica de modo independente (ou, no caso da matemática, que cálculos produzidos sejam confirmados por modelos do que conta como prova e sejam consistentes com axiomas), mas também no sentido de que há *suficiente coesão teórica e convergência de opiniões* naqueles âmbitos científicos, de modo que suas pretensões de conhecimento podem ser ditas objetivas.² Dessa maneira, essa razão é mais básica, pois sua questão central é saber *como identificar* critérios de competência, perícia e autoridade naqueles domínios científicos. O tema de nossa discussão será primariamente sobre essa segunda explicação.

² A matemática, como disciplina *a priori*, pode ser questionável nessa lista. Mas nada dito aqui depende de uma defesa desse ponto. A inserção da matemática se dá por ela exibir maior coesão teórica e convergência de opiniões entre seus especialistas do que, como veremos, a ética. Nesse aspecto específico, sendo uma forma de ciência, ela está mais próxima das ciências empíricas do que da ética. Ademais, a produção e distribuição do conhecimento matemático (sobretudo em nível avançado e altamente especializado) depende também de uma dimensão social de confiabilidade, tal como a das ciências empíricas. Sobre este último ponto, ver Hardwig, 1991.

Antes de prosseguirmos, cabe tornar claro o conceito de objetividade aqui em uso. Como a nossa discussão será essencialmente *epistêmica*, o conceito de objetividade empregado será igualmente epistêmico, e não metafísico. Por esse conceito, deve-se entender que, em um determinado universo de discurso, há suficiente coesão teórica, convergência de opiniões e concordância entre seus usuários de tal modo que é possível estabelecer um núcleo comum de crenças compartilhadas a partir das quais os enunciados do discurso possam ser confirmados ou verificados. Assim, esse conceito de objetividade não implica qualquer compromisso metafísico realista (ou antirrealista) do universo de discurso em questão, embora não o exclua e possa ser compatível com ele. Nesse sentido epistêmico de objetividade, tanto a física quanto a matemática são disciplinas objetivas, ainda que os métodos de verificação de seus enunciados possam ser radicalmente distintos. Independentemente das implicações metafísicas de seus discursos, físicos e matemáticos são capazes de resolver suas disputas (se não todas, um número substancial delas) recorrendo a crenças compartilhadas.³

³ Essa afirmação não pretende minimizar a existência de discordâncias entre cientistas, muitas delas possivelmente insuperáveis e radicalmente conflitantes. Há um sentido trivial da tese da subdeterminação de teorias científicas pela observação/experiência que confirma tais conflitos. Mas deve-se também admitir que, apesar de discordâncias no nível teórico, cientistas dispõem de um aparato conceitual, metodológico e instrumental compartilhado (envolvendo o domínio de certos conceitos centrais, padrões de investigação e testagem de hipóteses, mecanismos de controle experimental e observacional etc.) que torna possível estabelecer uma linguagem comum, produzir substanciais convergências de crenças e justificar suas pretensões de conhecimento. Assim, discordâncias existentes são explicadas e inteligíveis (muitas, superadas) com base em um pano de fundo de ampla convergência. Mas, como veremos, nada similar a tal unidade básica conceitual, metodológica e instrumental parece estar disponível na ética. Nesta, encontramos discordâncias aparentemente insuperáveis não apenas no nível teórico, mas também

Mas nada do que dissemos até agora parece se aplicar ao domínio da ética. Não parecemos possuir critérios objetivos e independentes de nossas próprias opiniões éticas pessoais e/ou comunitárias, que tendem a divergir em aspectos cruciais, para identificar competência ética. Na ética, não parecemos possuir nem a unidade básica conceitual e metodológica nem o tipo de predição e confirmação empírica que encontramos nas ciências. Assim, na ética, ao contrário daquelas áreas científicas, parecemos carecer de recursos para conferir a alguém o status de especialista ou perito ético e, por extensão, autoridade no assunto. Diferentemente do caso retratado sobre as três leis de Kepler, se alguém me fornecesse uma descrição fidedigna da *Fundamentação da metafísica dos costumes*, de Kant, ou do *Utilitarismo*, de Mill, e eu não fosse capaz de acompanhar os passos dos argumentos desses autores que os levam, respectivamente, a defender o imperativo categórico e o princípio da utilidade, eu não estaria justificado a acreditar em um ou outro. Por que não? Uma explicação possível tem relação com o tipo de restrição que estamos dispostos a estabelecer ao conceito de conhecimento quando aplicado ao domínio da ética. Ao contrário de conhecimento físico, astronômico, médico, matemático etc., reagimos com suspeição ao se falar de conhecimento ético. Se não há critérios objetivos e independentes para se especificar competência, perícia e autoridade éticas, não parecemos estar justificados em postular pretensão de conhecimento ético.

no nível “intuitivo”, de modo que sequer conseguimos isolar intuições comuns, com suficiente refinamento e substância, que sirvam de dado para a confirmação de teorias éticas. Ver Lillehammer, 2007, cap. 6, para uma discussão sobre várias aproximações entre ética e ciência, mas que reconhece uma assimetria entre as áreas como resultado do que podemos justificadamente esperar da resolução de certas discordâncias.

Antes de tentarmos responder por que não parece haver critérios objetivos e independentes para se especificar competência, perícia e autoridade éticas, vale considerar uma resposta negativa mais direta à ideia de *deferência* e *testemunho* éticos, que poderia explicar aquela suposta assimetria entre as ciências e a ética. Uma hipótese comum em meios filosóficos é a de que a justificação de crenças éticas exige a satisfação de um ideal de autonomia. Esse ideal envolve uma dupla dimensão: capacidade de autodeterminação e de compreensão.⁴ Ecoando a visão kantiana, alguns autores argumentam que, em assuntos éticos, devemos ser capazes de endossar, através de nossas próprias capacidades de reflexão e deliberação, os princípios e valores constitutivos de nossas crenças éticas. Não devemos, portanto, ter nossas crenças éticas determinadas por uma autoridade externa a nós. Além dessa capacidade de autodeterminação, devemos também ser capazes de *compreensão* de nossas crenças éticas e daquilo que as justifica (isto é, as suas razões), sob pena de não realizarmos a virtude suprema do agente ético responsável, a saber, a capacidade de agir conscientemente pelas razões certas.⁵

⁴ Nguyen (2010) fornece uma interpretação nesses termos. Mas nem todos compreendem autodeterminação e compreensão como parte de um mesmo ideal de autonomia. Enoch, 2014 e Sliwa, 2012, por exemplo, distinguem estritamente a dimensão epistêmica da dimensão ética do testemunho ético.

⁵ O debate recente sobre deferência e testemunho éticos pode ser dividido, em termos gerais e um tanto imprecisos, entre *pessimistas*, *céticos moderados* e *otimistas*. Entre os pessimistas, que rejeitam deferência e testemunho éticos, por considerá-los deficientes epistêmica e eticamente, podemos listar Nickel, 2001; Hopkins, 2007; Hills, 2009 e Crisp, 2014. Entre os céticos moderados, que admitem a possibilidade de testemunho ético em certos contextos específicos, encontramos Jones, 1999; McGrath, 2008, 2009 e 2011; Driver, 2006. Entre os otimistas, em diferentes sentidos, podemos incluir Singer, 1972; Enoch, 2014; Sliwa, 2012 e Lillehammer, 2014. Embora seja um dos nossos pressupostos aqui que crenças formadas por testemunho podem estar justificadas em certos contextos (e, assim, restaria saber se a ética é um desses

Há certamente algo correto nessa explicação. Afinal, por ser a ética um fenômeno essencialmente *prático*, e nossas ações serem inescapavelmente individuais (embora algumas possam também ser coletivas), precisamos, com efeito, de algum nível de compreensão do que está em jogo em nossos processos de reflexão, formação de crenças e deliberação éticas sobre como agir. Além disso, dada a importância da ética na vida da maior parte de nós, não se importar com autodeterminação e compreensão de nossas crenças e ações seria correr risco de incoerência e, na pior das hipóteses, de sujeição à dominação. Por fim, precisamos também, como uma necessidade social, de algum critério de responsabilização por ações, que parece exigir algum nível de autodeterminação e de compreensão de por que alguém age como age. A despeito disso, o ideal de autonomia pode ser questionado, ao menos, por duas razões.

Primeiro, as explicações que acabamos de fornecer – para algum nível de autodeterminação e compreensão na formação de crenças éticas – podem ser interpretadas em termos puramente *pragmáticos*, e não segundo a pressuposição de um *dever ético*. Mas, segundo o ideal de autonomia, é um dever ético (de ordem superior) exercermos nossas capacidades de autodeterminação e de compreensão de nossas razões para agir.⁶

Segundo, o ideal de autonomia é a expressão de uma certa interpretação do fenômeno ético que está longe de ser obviamente verdadeira e consensual. O ideal kantiano de interpretar a experiência

contextos), vários detalhes daquela literatura não serão importantes para os propósitos deste capítulo. Supondo que crenças possam estar justificadas com base em testemunhos, há ainda uma pergunta mais básica sobre os critérios para identificar testemunhos competentes. É essa pergunta mais básica nosso foco aqui.

⁶ Ver Nickel, 2001; Hopkins, 2007 e Hills, 2009 para defesas dessa ideia.

ética como irredutivelmente individual e intelectual, cujos processos de formação de crenças, deliberação e ação devem estar sob controle absoluto do indivíduo (expresso pelo seu eu racional), é apenas uma visão possível sobre como compreender o fenômeno ético. Em contraposição, para alguns (por exemplo, Lillehammer, 2014), deferência e testemunho são inevitavelmente parte do modo como formamos nossas crenças éticas, pois há uma dimensão de dependência social no fenômeno ético. Para outros (por exemplo, Enoch, 2014; Sliwa, 2012; Brink, 1989), se faz sentido falar de conhecimento ético (algo que o ideal de autonomia não rejeita), então deve também fazer sentido falar de especialistas em ética e, por extensão, validar a ideia de que deferência e testemunho podem ser fontes justificadoras de crenças éticas. Essas teses favoráveis à ideia de deferência e testemunho éticos não são obviamente falsas.

Seja qual for o resultado desse debate, parecemos dispor de uma explicação alternativa da assimetria entre as ciências e a ética. Os critérios disponíveis às ciências não parecem estar disponíveis à ética porque o universo de discurso desta última é profundamente marcado por *discordâncias*.⁷ Ao contrário das ciências, na ética não parece haver suficiente coesão teórica, convergência de opiniões e concordância de tal modo a tornar possível estabelecer um núcleo comum de crenças compartilhadas a partir das quais os enunciados do discurso ético possam ser confirmados ou verificados.⁸ Portanto, se há

⁷ Como veremos, essa é uma hipótese explicativa mais *inclusiva* do que aquela baseada no ideal de autonomia, pois ela não exclui *a priori* a possibilidade de crenças éticas poderem ser justificadas por deferência e testemunho. Ela lança dúvida sobre a resposta a algo mais básico: “quais são os critérios para se identificar competência e autoridade éticas?”.

⁸ A ideia de que discordâncias são um elemento constitutivo e mesmo insuperável do fenômeno ético é defendida, de diferentes maneiras, por Berlin, 1997; Hampshire, 1983

algum sentido possível de pretensão de conhecimento ético, ele parece ser distinto do modo como atribuímos conhecimento àquelas outras áreas científicas.

É por essa razão que Bernard Williams (1995b) diz ser inevitável reconhecer como paradoxal a instituição social, comum no mundo contemporâneo, dos comitês de ética –embora Williams também reconheça a necessidade social da existência desses comitês. Haja vista ser necessário que instituições públicas tomem decisões sobre questões éticas controversas e respondam por elas, torna-se inevitável apelar à ideia de especialidade e perícia para orientar tais decisões, sob a pressuposição de algum senso de neutralidade e tecnicidade na governança pública. Mas isso, de forma alguma, remove o ar paradoxal da sugestão comum de que tais comitês devam ser ocupados, por exemplo, por pessoas que possuam um título de especialidade (tal como o acadêmico) em ética.⁹

Essa é a nossa imagem inicial, por assim dizer. A questão aqui é sobre como podemos responder à pergunta “quem tem competência ética?”. No que se segue, desenvolveremos a ideia de que não possuímos critérios inteiramente objetivos e independentes para julgar competência, perícia e autoridade éticas. Na melhor das hipóteses, possuímos pretensão de conhecimento ético em um

e 2000; Kekes, 1993; Rorty, 1989; Williams, 1985, 1995a e 1995b; Walzer, 1987; Wong, 2006. Ver Hall, 2020, para uma discussão detalhada da ideia em Berlin, Hampshire e Williams.

⁹ Desconforto similar ao de Williams é expresso no relato concreto de Mary Warnock sobre sua participação no Committee of Inquiry into Human Fertilisation and Embryology. Warnock diz: “[...] esse não foi, e não poderia ser um comitê de peritos [*experts*]. Nossa tarefa [...] foi primariamente aconselhar os ministros, com base em juízos morais; pois não existe essa coisa de perito ético (*moral expert*)” (Warnock, 1985, p. 95).

sentido *deflacionado*, cuja atribuição só faz sentido a partir de uma perspectiva ética comum *local* e *contingente*. Mas, como passos para o desenvolvimento dessa ideia, abordaremos preliminarmente três respostas problemáticas à questão mencionada.

2. Três respostas problemáticas

2.1. *Progresso ético*

Talvez a impressão inicial descrita acima não passe de uma especulação empírica infundada. Juntamente com Parfit (1984), podemos acreditar que não temos *ainda* critérios seguros para identificar competência, perícia e autoridade éticas, e que nosso estudo secular da ética é incipiente e, como consequência, nossa pretensão de conhecimento ético ainda carecem de desenvolvimentos mais precisos. Em algum momento futuro, talvez possamos falar de uma ciência da ética e ser capazes de identificar objetivamente peritos e autoridades éticas. Como sabemos, essa não é uma ideia nova, a despeito do que Parfit diz. Ela está presente na *República*, de Platão, aquela que talvez seja a versão mais ambiciosa da ideia, com filósofos ocupando a posição de autoridade ética para governar as sociedades humanas.

Como deve estar claro, a ideia de Parfit é também baseada, parcialmente, em uma especulação empírica. (“Parcialmente” porque, sendo justo com Parfit, o argumento empírico não é o único que ele fornece a favor de seu otimismo.) A questão, então, consiste em saber se a especulação empírica de Parfit está mais bem fundada do que a que descrevemos antes. Se levarmos em consideração dados empíricos amplamente conhecidos sobre o fenômeno da discordância ética, parece que temos razões para duvidar seriamente do otimismo de Parfit em acreditar em uma maior convergência futura de opiniões

e maior precisão nos critérios sobre o que conta como competência ética. Questões éticas são, provavelmente, objeto de reflexão e práticas humanas desde, ao menos, o momento em que seres humanos desenvolveram capacidade de organização social mínima e articulação linguística. Mas, do ponto de vista empírico, o que podemos detectar, com alguma segurança, são amplas e persistentes discordâncias que atravessam tempos históricos, sociedades e culturas.

Parfit (1984, p. 453) diz que a ética secular não começou a ser estudada sistematicamente antes de cerca de 1960. Esse diagnóstico é um tanto incompreensível historicamente. Parfit, obviamente, não ignora parcelas significativas das éticas seculares da Antiguidade, mas ele afirma (de modo puramente especulativo, sem fornecer qualquer dado empírico) que uma visão secular da ética é uma tendência minoritária na história da humanidade até os anos sessenta. Uma sugestão mais plausível, mas ao mesmo tempo concedendo algo a Parfit, seria propor, como recorte histórico, a modernidade e o iluminismo no Ocidente como o período do surgimento dos primeiros movimentos de éticas seculares na história recente da humanidade. Mesmo com esse recorte, não parecemos ter “progredido” em torno de uma convergência significativa sobre o que conta como eticamente adequado e sobre o que poderia ser considerado como critérios objetivos de competência ética. Discordâncias amplas e persistentes (interculturais e intraculturais) sobre questões como exigências de justiça, promoção do bem comum, punição, monogamia, comportamento sexual, aborto, eutanásia, guerras, tratamento de animais etc. parecem estar longe de ser superadas. Essas discordâncias, devemos notar, não se restringem ao nível, digamos, ordinário da prática e do discurso éticos, mas se estendem a disputas

entre acadêmicos. Fazendo uso de um exemplo óbvio, embora um tanto retórico e *ad hominem* no contexto em questão, mesmo entre seus pares filosóficos, a visão ética do próprio Parfit tem, provavelmente, mais críticos do que adeptos.

Talvez alguns contestem esse diagnóstico, argumentando que atualmente convergimos mais em nossas visões éticas de mundo, em escala global, do que no passado. Por exemplo, alguns poderiam considerar um progresso ético a tendência majoritária global dos tempos atuais de indignação e aversão a práticas excludentes sociais (por exemplo, com base em critérios sexistas e racistas) e suas associadas desigualdades. Suponhamos que esse seja, de fato, um diagnóstico correto da atualidade. O problema que persiste é o de que, como argumento empírico, essa suposta maior convergência na época atual pode ser explicada tanto como resultado de um processo de maior aproximação (ou rastreamento) daquilo que se pleiteia ser conhecimento ético (genuíno e objetivo) quanto como resultado da expansão global política, social e cultural da visão de mundo ocidental e liberal, inaugurada com a modernidade, que carrega consigo aqueles compromissos igualitários (ainda que, por certo, o conceito de igualdade tenha sofrido alterações significativas e sido objeto de disputas desde então). Talvez boa parte do que Parfit entenda por “éticas seculares” seja justamente a expressão dos valores preponderantes da visão de mundo ocidental e liberal.

Algo similar pode ser dito da onda mais recente de modelos realistas naturalistas em ética, defendida por autores como Boyd (1988) e Brink (1989). Esses autores postulam que propriedades éticas são simplesmente propriedades naturais e empíricas, e que, portanto, a investigação ética deve ser compreendida como uma forma de investigação científica. Boyd e Brink rejeitam a ideia de que sejamos

capazes de, apenas por meio de análise conceitual, descobrir quais propriedades naturais e empíricas instanciam propriedades éticas. Ainda que propriedades éticas sejam propriedades naturais e empíricas, a relação aqui seria dada na forma ou de identidades *a posteriori* (reduativas, não analíticas) ou de múltipla realização (não reduativas, constitutivas, supervenientes).¹⁰ Essa é essencialmente uma tese metafísica e semântica sobre a natureza da ética. Nesse sentido, é de se esperar que ela encontre espaço para a ideia de conhecimento e progresso éticos e, por extensão, de competência, perícia e autoridade éticas como algo genuinamente objetivo. Se a investigação ética é, em seus aspectos essenciais, uma forma de investigação científica, devemos poder falar de especialistas e autoridades éticas da mesma forma que o fazemos com as demais ciências. Na esteira disso, poderíamos ter razão para concordar com Parfit que o desenvolvimento futuro mais amplo e preciso dessa “ciência da ética” nos colocará em situação melhor do que a atual para a identificação de especialistas e autoridades éticas.

Na medida em que essa tese é essencialmente metafísica e semântica, ela não faz parte do foco central de nossa discussão que, como vimos, pressupõe uma noção *epistêmica* de objetividade. Mas alguns comentários sobre ela serão úteis para o que veremos adiante. Ainda que a tese de Boyd e Brink possa, de uma perspectiva *teórica e abstrata*, postular a existência de especialistas e autoridades éticas, ela enfrenta a mesma dificuldade epistêmica de ser capaz de especificar os critérios segundo os quais poderíamos identificar tais especialistas

¹⁰ Ver Copp, 2012, para um mapeamento detalhado das diferentes versões de realismo naturalista ético.

e autoridades.¹¹ O modelo epistêmico de justificação de crenças adotado por Boyd e Brink é um modelo coerentista (uma versão de equilíbrio reflexivo), e veremos que, dessa perspectiva epistêmica, a tese desses autores não está livre de dificuldades. Nesse sentido, do ponto de vista epistêmico, Boyd e Brink estão sujeitos à mesma dificuldade que direcionamos a Parfit: a de que o desenvolvimento de uma suposta “ciência da ética” possa não passar da expressão dos valores preponderantes da visão de mundo ocidental e liberal como um fenômeno histórico e cultural.

2.2. *Intuicionismo*

Outro caminho possível de resposta ao nosso problema inicial ancora-se não em um argumento empírico, mas em um argumento *a priori*. Moore (1903) defendeu a existência da propriedade *sui generis* da bondade, cujo conhecimento de sua instanciiação seria acessível por algo como uma intuição racional ou intelectual. Embora Moore seja tomado aqui como exemplo, suporemos que, feitas as devidas adaptações, tudo o que for dito aqui vale também para intuicionistas de outras estirpes. Como sabemos, intuicionistas éticos divergem sobre o que é o objeto de uma intuição: estados de valor intrínseco, deveres, direitos, razões, princípios etc.

Sejam quais forem os detalhes da melhor teoria intuicionista, essa corrente de pensamento filosófico poderia alegar que conhecimento ético é possível, mas que, dado o caráter *sui generis* das propriedades éticas e sua não menos *sui generis* associada epistemologia, os critérios para se julgar competência, perícia e

¹¹ A distinção conceitual entre a *existência* e o *conhecimento* de especialistas e autoridades éticas é também defendida por Driver (2006) e McGrath (2008, 2009, 2011) no debate sobre deferência e testemunho éticos.

autoridade éticas seriam, na melhor das hipóteses, *indiretos*. Afinal, como diz a epígrafe do *Principia ethica*, de Moore, tomando emprestada uma frase de Butler, “tudo é o que é, e não outra coisa”. A sentença de Butler, interpretada no contexto do *Principia*, quer dizer que o conhecimento ético só pode ser obtido de modo direto, pois a propriedade da bondade é não analisável e irreduzível. Assim, nesse modelo, os critérios para se julgar competência, perícia e autoridade éticas não estariam disponíveis senão através de juízos irreduzivelmente éticos, senão através da minha perfeita concordância com a opinião ética expressa por um interlocutor. Se nenhum conjunto de descrições não éticas da realidade (seja ela natural, supranatural ou metafísica) é suficiente para figurar como premissas em um argumento que gere, como resultado, uma conclusão ética, e, se não há implicações semânticas entre termos ou conceitos não éticos e termos ou conceitos éticos, não parece ser possível dispor de critérios objetivos e independentes do que alguém já julga eticamente para se identificar competência ética. Na melhor das hipóteses, podemos *indiretamente* almejar identificar competência ética através de características como sensibilidade reflexiva, agudeza de pensamento, capacidade de raciocínio lógico acurado, experiência de vida etc. Mas, como deve estar claro, nenhuma dessas características ou uma conjunção exaustiva delas seriam capazes de fornecer critérios inteiramente objetivos e independentes de competência ética.

Devemos notar que essa é uma conclusão que não precisa demover Moore de sua posição. Não há nenhuma incoerência em sustentar que é possível haver conhecimento ético, mas não haver critérios inteiramente objetivos e independentes para se julgar competência ética. Mas há, claro, um grande ônus nisso. Não apenas a

ideia de propriedades valorativas *sui generis* e algo como uma faculdade ou capacidade de intuição racional igualmente *sui generis* soam suspeitas, mas a própria ideia de conhecimento ético como algo esotérico torna-se também questionável por sua vagueza. Se, como professam alguns, há algo de útil e social no conceito de conhecimento – que cumpre a função de atender a propósitos sociais humanos contextualizados e pressupõe uma divisão social da sua produção e distribuição –, o tipo de conhecimento que a tese de Moore parece implicar, ao contrário, torna-o publicamente inacessível e, assim, impossível de satisfazer essa dimensão pragmática e social do conhecimento. A rigor, a tese de Moore não parece capaz de fornecer os recursos para se fundamentar qualquer sentido possível de testemunho ético confiável como uma empreitada social. Eu não pareço ter razão para acreditar nas intuições de outrem porque esse outrem não possui meios de me convencer delas.

A despeito disso, parece haver algo correto no pensamento de Moore quando ele sugere que não há critérios objetivos e inteiramente independentes para identificação de competência, perícia e autoridade éticas. Mas o autor explica isso, como vimos, através da postulação do suposto caráter não analisável e irreduzível dos termos ou conceitos éticos associado a uma epistemologia intuicionista. Assim, para Moore, a investigação ética é passível de conhecimento genuíno e objetivo, ainda que sejamos epistemicamente limitados pelos usos das nossas capacidades intuitivas, que podem ser falhas e não são, a rigor, compartilháveis. É isso o que explica discordâncias éticas e introduz uma assimetria entre as minhas intuições éticas e as de outrem, de tal modo que, *prima facie*, tenho razão para confiar nas

minhas intuições e não nas de outrem.¹² Obviamente, posso ter evidência indireta de que minhas capacidades intuitivas podem estar a falhar. Por exemplo, posso reconhecer que não estou suficientemente bem-informado, que estou excessivamente tomado por emoções, que estou sendo parcial ou simplesmente que estou deveras cansado para refletir, julgar e deliberar sobre uma situação com implicações éticas etc. Mas, ainda assim, isso não tornaria justificável a deferência às intuições éticas de outrem. Na melhor das hipóteses, supondo aquelas condições subótimas para reflexão, julgamento e deliberação, as intuições éticas de outrem, conflitantes com as minhas, levar-me-iam a suspender meu juízo.

Veremos abaixo que há um outro caminho possível para preservar a ideia de que não possuímos critérios inteiramente objetivos e independentes para julgar competência, perícia e autoridade éticas, mas sem associar isso a uma explicação intuicionista de nossas capacidades de reflexão, julgamento e deliberação éticas.

2.3. *Equilíbrio reflexivo*

Uma terceira resposta possível pode ser encontrada na tese de que nossas crenças em equilíbrio reflexivo amplo constituem ou são evidência de que possuímos conhecimento ético. Com base nessa tese, seríamos capazes de identificar competência, perícia e autoridade éticas através daqueles que exibem equilíbrio reflexivo amplo entre suas crenças.¹³

¹² Para uma defesa desse status epistêmico de intuições, ver Wedgwood, 2010.

¹³ Boa parte dos detalhes das discussões sobre o método do equilíbrio reflexivo – como a transição de uma versão estreita (*narrow*) para uma versão ampla (*wide*) e os níveis de ajuste entre juízos ponderados (*considered judgments*) em seus variados graus de

Em uma versão que podemos chamar de “constitutiva” da tese do equilíbrio reflexivo, justificação de crenças e conhecimento/verdade se equivalem. Supondo que a tese do equilíbrio reflexivo seja uma tese coerentista da justificação de crenças, isso equivaleria a dizer que ela é também uma versão da tese da verdade e do conhecimento éticos como coerência. Algumas interpretações de Rawls (1971, 1974), que enfatizam sua rejeição do intuicionismo e a defesa da ideia de construtivismo, atribuem-lhe essa posição.¹⁴

Por outro lado, versões da tese do equilíbrio reflexivo que podemos chamar de “evidencialistas” distinguem justificação e verdade/conhecimento éticos. Segundo essa concepção, é sempre possível que o conjunto de nossas crenças esteja justificado em relações de coerência, ainda que ele seja (substancialmente) falso. A justificação de crenças em equilíbrio reflexivo amplo seria apenas evidência, mas não garantia, da verdade de tais crenças e, por extensão, de conhecimento ético. Essa versão da tese pode ser encontrada em autores (tão distintos) como Daniels (1979), Brink (1989) e Scanlon (2014).¹⁵

generalidade, princípios éticos e teorias amplas (científicas e filosóficas) – não será importante para nossos propósitos aqui.

¹⁴ Ver Scanlon, 2003 e Daniels, 1979, para essa possibilidade de interpretação de Rawls. Daniels (1979, pp. 277ss.) descreve tal versão do equilíbrio reflexivo como “acordo intersubjetivo” (no que diz respeito à objetividade da ética) e “eliminativa” (no que diz respeito à noção realista de verdade ética). Scanlon (2003, pp. 141ss.) chama essa forma de interpretação daquilo que o método produz de “descritiva” (e não normativa) do que uma pessoa ou grupo acredita eticamente (em particular, descritiva de suas *concepções* de justiça).

¹⁵ Enquanto Scanlon (2014) é simpático a uma interpretação fundacionalista e intuicionista dessa versão do método, Daniels e Brink são antifundacionalistas e anti-intuicionistas. Como veremos adiante, há sérias dúvidas de que o método do equilíbrio reflexivo possa justificar convergência de crenças éticas em indivíduos, sociedades e culturas distintas sem apelar a alguma forma de fundacionalismo e

Dado que o método do equilíbrio reflexivo envolve um processo de idealização na justificação de nossas crenças, algumas condições gerais devem ser satisfeitas por todo conjunto coerente de crenças, de tal modo a traçar uma distinção entre as crenças atualmente em coerência, que podem estar sujeitas a vários tipos de vieses e distorções, e crenças idealmente em coerência, com a supressão de tais vieses e distorções. Assim, um indivíduo que tenha suas crenças em equilíbrio reflexivo amplo deve estar ciente de todos os fatos não éticos relevantes, não possuir informações não éticas falsas relevantes, não ter cometido nenhum erro de raciocínio ou incorrer em falácias, não estar tomado emocionalmente pela questão de modo a comprometer o seu julgamento, não ser influenciado por causas estranhas ao que está em discussão, e, ainda, possivelmente, ser dotado de uma capacidade imaginativa ampla, alguma capacidade de empatia, ser capaz de deliberar a partir de um senso de imparcialidade etc. É difícil estabelecer uma lista exaustiva dessas condições. Mas, seja como for, a ideia básica de estabelecer uma distinção entre um conjunto de crenças atualmente em coerência e um conjunto contrafactual de crenças, capaz de corrigir distorções e vieses das crenças atuais, é certamente uma ideia plausível sobre qualquer processo de justificação de crenças.

O problema, seja qual for a melhor versão da tese do equilíbrio reflexivo, é que não parecemos possuir, mais uma vez, critérios

intuicionismo. Na versão de Brink (1989), que postula uma metafísica e semântica realista naturalista associada, a dificuldade é saber como garantir (de uma forma que vá além de uma mera projeção otimista) que as crenças justificadas em equilíbrio reflexivo *rastreiem* fatos éticos. Na versão de Daniels (1979), que não postula explicitamente uma metafísica e semântica associada, a tese de que crenças em equilíbrio reflexivo são evidência de verdade (independente) não parece passar de uma espécie de “carta de intenções”.

inteiramente objetivos e independentes do que já julgamos eticamente para estabelecer o escopo e os limites daquela lista de condições. Podemos concordar que estar ciente de todos os fatos não éticos relevantes, não possuir informações não éticas falsas relevantes, não ter cometido nenhum erro de raciocínio lógico são pré-requisitos que, de fato, devem estar presentes em todos os conjuntos de crenças em equilíbrio reflexivo amplo. Mas, no que diz respeito às demais condições listadas acima, discordâncias começariam a surgir. Alguns poderiam argumentar que, embora estar tomado emocionalmente por uma questão leve tipicamente a distorções, isso nem sempre é o caso. Poderia ser algo essencial para promover mudanças em um rígido *status quo*. Mesmo uma capacidade imaginativa ampla e uma de empatia, embora tipicamente condições favoráveis à formação de crenças éticas, podem ser questionáveis: não é claro quanta imaginação e quanto exercício empático são necessários para se chegar a um julgamento ético adequado em certos contextos. Algo similar poderia ser dito sobre imparcialidade. Ainda que esta seja uma exigência padrão de um julgamento ético, não são poucas as situações que desafiam esse padrão e propiciam julgamentos de parcialidade.

Alguns de inclinação cética ainda mais radical poderiam argumentar que até mesmo saber quais são os fatos não éticos relevantes já envolve julgamento ético, pois há o risco de o conceito de relevância empregado nesse contexto já possuir significado ético. Afinal, para sermos capazes de dizer o que conta como relevante eticamente, precisamos, ao mesmo tempo, ser capazes de identificar quais similaridades *importam eticamente*. Em outros termos, esses céticos diriam que, em ética, sequer temos como garantir a ausência de discordâncias sobre o que conta como evidência.

É justamente por essas razões que parte dos defensores do método do equilíbrio reflexivo admite que o método, interpretado literalmente, não garante convergência entre indivíduos com crenças ou pontos de partida distintos, mesmo supondo que essas crenças sejam “refletidas” ou ponderadas (*considered judgments*) e revisáveis.¹⁶ Alguns dos defensores do método (Scanlon, 2014; DePaul, 1993) tentam superar essa dificuldade insistindo que é possível estabelecer uma distinção entre crenças que ocupam um papel fundacional ou hierarquicamente superior em relação a outras crenças e, com base nisso, fixar algum critério de comparação e avaliação entre conjuntos coerentes de crenças que sejam conflitantes. Apenas aqueles conjuntos que confeririam espaço a tais crenças fundacionais seriam candidatos genuínos a verdade e conhecimento éticos. Assim, além daquelas condições listadas acima, uma série de *platitudes* do discurso ético podem ser postuladas para ocupar tal papel fundacional.¹⁷ Por exemplo, parecemos estar justificados a encontrar espaço, em qualquer modelo de ética que julgamos adequado, para crenças substanciais que apontam para algum senso de reciprocidade, igualdade, imparcialidade, promoção do bem-estar, utilidade social, benevolência, autonomia, respeito mútuo, florescimento pessoal etc.

Mas, mesmo que estejamos dispostos a admitir tais platitudes como constitutivas do discurso ético (sejam elas constitutivas em sentido semântico ou metafísico), todas estas parecem poder ocupar

¹⁶ Sobre essa conclusão, ver de Maagt, 2017; Scanlon, 2003 e 2014; Kelly e McGrath, 2010 e Daniels, 1979.

¹⁷ A ideia de que o método do equilíbrio reflexivo consiste em um procedimento de justificação de nossas crenças éticas que dependem da identificação de um conjunto de platitudes do conceito de ética ou moralidade é defendida por Smith (1994, pp. 39ss.). Smith identifica platitudes de acordo com as seguintes dimensões: praticidade, superveniência, objetividade, substância e procedimento.

aquele papel justificador apenas em sentido muito geral e abstrato, que é insuficiente para garantir *conteúdo normativo genuíno* ou servir de guia para nossas ações. Além disso, precisaríamos de critérios de hierarquização e atribuição de pesos distintos aos valores éticos expressos por tais platitudes, supondo que a realização de todos eles em igual proporção ou livre de conflitos seja impossível. Mas aqui esbarramos em visões de mundo distintas, que fornecerão interpretações específicas diferentes de cada uma daquelas platitudes, além de postular esquemas distintos de hierarquização e pesos a elas.¹⁸

Se esse for um diagnóstico plausível sobre o método do equilíbrio reflexivo, a versão constitutivista da tese precisaria simplesmente aceitar a ideia de que pode haver tantas verdades e objetos de conhecimento éticos quanto há conjuntos coerentes de crenças em equilíbrio reflexivo amplo. Por outro lado, a versão evidencialista poderia almejar alguma solução buscando uma conciliação entre um modelo de justificação coerentista de crenças e a relevância de algumas intuições fundamentais. De fato, alguns autores preferem interpretar o método do equilíbrio reflexivo como uma versão de fundacionalismo moderado, segundo a qual, a despeito de o modelo ser substancialmente coerentista na sua *estrutura* de justificação, há espaço nele para intuições que ocupariam um *papel epistêmico positivo* de justificação fundacional. Nenhuma dessas intuições está livre de revisões, mas, a menos que haja razões robustas para sua revisão, elas devem ser preservadas.¹⁹ Mas, como está óbvio, o modelo

¹⁸ A ideia de que platitudes universais abstratas precisam de interpretação e ordenação particulares, em contextos concretos históricos, sociais e culturais, para constituírem-se como conteúdo de uma ética é defendida por Wong (2006).

¹⁹ Ver Scanlon, 2014 e DePaul, 1993. DePaul desenvolve a versão que chama de “radical”, segundo a qual o método do equilíbrio reflexivo associado ao status epistêmico positivo fundacional de juízos ponderados pode promover alterações radicais nos

evidencialista que adere ao fundacionalismo moderado, se levado às últimas consequências – isto é, propondo superar o tipo de pluralidade que sugerimos acima sobre como interpretar as platitudes do discurso moral e hierarquizá-las –, não parece ser capaz de evitar incorrer em uma versão de intuicionismo ético.²⁰

Portanto, o método do equilíbrio reflexivo, em sua versão mais ambiciosa na postulação de conhecimento ético – a versão evidencialista –, enfrenta as mesmas dificuldades do intuicionismo ético para a identificação de critérios objetivos de competência, perícia e autoridade éticas.

3. Competência ética e cultura ética

Exploraremos agora o caminho de resposta proposto por Bernard Williams (1995b) à nossa questão inicial. Estritamente falando, Williams é cético sobre a possibilidade de formularmos critérios objetivos e independentes para julgar competência, perícia e autoridade éticas. Mas, por outro lado, ele admite um sentido específico, “deflacionado”, de conhecimento ético. No que se segue, forneceremos uma resposta no espírito de Williams, ainda que ele mesmo, talvez, não a endossasse por completo.

pontos de partida de julgamento ético de um indivíduo. DePaul defende ainda que juízos ponderados que cumprem um papel epistêmico fundacional são um tipo de *intuição*, ainda que não deva ser interpretada como não revisável, incorrigível e infalível.

²⁰ Scanlon, por exemplo, propõe a seguinte formulação do que é definidor de um juízo ponderado (*considered judgment*) com status epistêmico positivo: “deve ser algo que me parece (*seems to me*) ser claramente verdadeiro quando estou pensando sobre o assunto sob boas condições para formar julgamentos do tipo em questão” (Scanlon 2014, p. 82).

Como um passo instrutivo para compreender o que está em jogo, retomemos a nossa conclusão anterior sobre o método do equilíbrio reflexivo. Se uma conclusão plausível for a da versão constitutiva, de que há tantas verdades éticas quanto há conjuntos coerentes de crenças em equilíbrio reflexivo amplo, então, de uma perspectiva *interna* a uma visão ética comum, seria talvez possível estabelecer critérios (se não inteiramente precisos, ao menos aproximados) do que contaria como competência, perícia e autoridade éticas. Obviamente, esses critérios não seriam válidos para toda e qualquer comunidade ética. As considerações e condições que figurariam como critérios de competência para um determinado conjunto coerente de crenças poderiam não figurar para um outro conjunto coerente de crenças.

Em grande medida, o sentido segundo o qual Williams admite ser possível falar de conhecimento ético se aproxima disso, pois depende da pressuposição de um repertório comum de conceitos, valores e práticas que constituem conjuntos de crenças éticas relativamente unificados. Mas nada do entendimento de Williams sobre o fenômeno ético nos autoriza a postular meios racionais ou intuitivos que garantam convergência intercultural (e mesmo, em algum nível, intracultural) de visões éticas de mundo. Assim, se há um sentido possível de critérios de competência, perícia e autoridade éticas, ele deve ser *interno* a um ponto de vista ético comum local.²¹

²¹ Essa interpretação de Williams aproxima-o de uma versão de relativismo ético que pode soar mais forte do que algumas passagens de seus textos podem sugerir. Não poderemos desenvolver esse ponto aqui, mas cabe lembrar que Williams (1981, 1985, 1995a) é enfático em sua rejeição de qualquer forma de racionalismo e intuicionismo éticos, assim como do status de necessidade metafísica ao conteúdo da ética e de qualquer método *a priori* para decidir questões éticas. Esses elementos são evidência de que Williams compreende a ética em bases contingentes e como um fenômeno

Em termos mais específicos, Williams (1985) desenvolve essa ideia através da apresentação do uso e domínio dos conceitos éticos que ele denomina de “*thick*”. Na ausência de melhor tradução, chamemos esses conceitos de “particulares”. Conceitos éticos particulares, em oposição a conceitos éticos gerais (*thin*) como “bom”, “mau”, “certo”, “errado”, envolvem um componente fixo como parte de seu significado descritivo e um componente valorativo associado. São candidatos a tais conceitos, por exemplo: generosidade, crueldade, igualdade, sobriedade, privacidade, coragem, modéstia, talvez mesmo justiça etc. Segundo autores como Williams, o trabalho propriamente normativo em nossas vidas éticas é realizado por conceitos éticos particulares e não por conceitos éticos gerais. Nossa compreensão destes seria, por assim dizer, derivada da compreensão daqueles – embora Williams admita que conceitos éticos gerais possam, em certos contextos, tornar possíveis diálogos entre perspectivas éticas distintas.

Há várias formas de se caracterizar a relação entre os componentes descritivo e valorativo desses conceitos. Para nossos propósitos aqui, não será relevante abordar as várias, e conflitantes, teses sobre a natureza de tais conceitos. Basta reconhecermos como um ponto comum a todas elas que esses conceitos envolvem aqueles dois componentes e que há algum tipo de interdependência entre eles. Mas, além disso, é também fundamental para a tese de Williams que tais conceitos recebam interpretações *locais*, variáveis entre culturas

contextualizado social, histórica e culturalmente. Assim, quando ele critica uma versão de relativismo, sua crítica se direciona à tentativa de transformar o relativismo em uma *teoria ética* e, assim, introduzir um método *a priori* para decidir questões éticas e conferir status de necessidade a enunciados éticos. Ver Rovane, 2009, para uma interpretação de Williams como um relativista ético em consonância com a aqui apresentada.

humanas. Assim, o que conta como igualdade e crueldade em uma determinada comunidade ética pode não coincidir inteiramente ou mesmo não contar como igualdade e crueldade em outra comunidade ética. Como diz Rovane (2013, p. 224), ecoando Williams: a vida ética “está sempre situada em condições sociais historicamente específicas nas quais apenas os valores [éticos] particulares [*thick*] podem servir como um guia útil”.

Chamemos – ainda que de modo um tanto impreciso – isso que se configura como o conjunto de conceitos, crenças, valores e práticas comuns, internos a uma determinada comunidade, de uma “cultura ética”. Se podemos circunscrever uma cultura ética, então podemos almejar falar de “conhecimento” em um sentido específico, *de uma perspectiva interna a tal cultura*, daquilo que conta para ela como propriedades, fatos e verdade éticas.

Nesse sentido, algumas habilidades poderiam colocar certos indivíduos em condições mais aptas do que outros para julgamento ético dentro de uma cultura ética. Por exemplo, indivíduos que se dedicassem a estudar a história dessa cultura, assim como suas implicações conceituais e desdobramentos práticos, que possuíssem ampla experiência concreta nesses assuntos e, como resultado, desenvolvessem um tipo de sensibilidade refinada local poderiam ser considerados mais capacitados do que outros a fazer julgamentos sobre o domínio da ética nesse contexto cultural. Isso é algo que podemos conjecturar que tenderia a ser desenvolvido nos contextos profissionais de acadêmicos como filósofos/as, historiadores/as, cientistas sociais, mas também na dimensão concreta da vida ordinária através das atividades de políticos locais, de organizações civis, de jornalistas, de juristas, de críticos literários etc. Em certo sentido, esses indivíduos e grupos poderiam desfrutar de alguma

autoridade local no assunto; poderiam ser considerados algo como *intérpretes qualificados* dessa cultura.

Segundo essa visão, pertencer a uma cultura ética significa possuir um núcleo comum suficientemente compartilhado de *sensibilidades éticas e convergência no uso e entendimento dos contornos de termos e conceitos éticos*, cujo desenvolvimento e construção são parte de um processo histórico, social e cultural. Esse processo pode tornar algumas convicções dessa cultura como suficientemente rígidas e estáveis para resistirem à revisão; outras como exigindo permanente reinterpretação; outras como vagas, incertas e instáveis. Assim, em grande medida, o fazer ético envolveria um processo de busca por uma espécie de *autoconhecimento e interpretação* de uma cultura ética a que se pertence, de busca pelos compromissos fundamentais dessa cultura e suas implicações conceituais e práticas.²² Nesse sentido, instruir-se sobre os conteúdos dos conceitos éticos e sobre as sensibilidades éticas constitutivas de uma cultura ética pode ser interpretado como uma forma de conhecimento na medida em que se configura como uma *descrição* dos compromissos conceituais e práticos desta.

Mas, então, isso significaria falar *genuinamente* de conhecimento e de critérios, ainda que locais, para se identificar competência ética? A questão não é simples. Alguém que possua fortes intuições linguísticas ou semânticas sobre o conceito de conhecimento, poderia de antemão rejeitar qualquer caracterização “não monádica” e não universal dele, relegando tal uso “relativo” do conceito a um sentido meramente pragmático da linguagem. Não teremos espaço aqui para

²² A ideia de que boa parte dos nossos exercícios de justificação de crenças éticas em diálogo com nossos pares sociais é, em essência, um processo de interpretação de uma cultura é defendida por Walzer (1987).

entrar no mérito dessa discussão e em uma disputa de intuições, mesmo porque parece haver outra razão que nos leva a admitir que, se há pretensão de conhecimento ético, mesmo que local, esse conceito de conhecimento é distinto daquele que associamos àquelas disciplinas com as quais iniciamos nossa discussão: física, astronomia, medicina, matemática etc. Isso deve ser suficiente para os nossos propósitos aqui. Vejamos.

Ainda que seja possível estabelecer algum critério de competência, perícia e autoridade éticas de uma perspectiva local, interna a uma determinada cultura ética, sempre parece haver espaço para algum tipo de *contestação* desses critérios. Alguns autores (sobretudo aqueles de orientação prescritivista e expressivista em metaética) endossam a ideia de que conceitos éticos são *essencialmente contestáveis* por revelarem a permanente possibilidade de crítica no seu uso.²³ Prescritivistas e expressivistas parecem pouco dispostos a estabelecer restrições conceituais descritivas ao conteúdo normativo da ética de tal modo que, de fato, “contestabilidade essencial” envolve a possibilidade irrestrita de contestação de *qualquer* pretensão de vinculação conceitual entre quaisquer pares de descrição e prescrição/valoração. Talvez isso seja excessivo, pois, como sugerimos, parecemos precisar circunscrever *substancialmente*,

²³ A tese de que os conceitos centrais de certos universos de discurso (por exemplo, da ética, da política e da estética) possuem a característica comum de serem essencialmente contestáveis em seu uso foi proposta originalmente por Gallie (1956). Ideia similar pode ser encontrada em versões de “não-cognitvismo” como o prescritivismo universal de Hare (1952) e o expressivismo de Blackburn (1993). Para Hare (1952), o único aspecto incontestável do *conteúdo* dos conceitos éticos é seu escopo universal (prescritividade sendo sua *forma lógica*). Em obras posteriores, Hare acrescentou algo como igualdade de consideração de interesses individuais. Não é claro que Blackburn admitisse sequer esses aspectos como incontestáveis.

em termos de conteúdo, uma determinada cultura ética a fim de podermos estabelecer um compartilhamento de visões de mundo e conceitos éticos e, assim, almejarmos obter suficiente convergência e concordância éticas. Mas, a despeito de prescritivistas e expressivistas serem excessivamente liberais no que “essencialmente contestável” quer dizer, parece haver algo correto na visão desses autores. Como explicar isso no contexto do que viemos chamando de uma “cultura ética”?

Como já dito, uma cultura ética envolve não apenas convicções suficientemente rígidas e estáveis que são mais resistentes à revisão, mas também convicções que exigem permanente reinterpretação, e ainda outras que são vagas, incertas e instáveis. É de se esperar, portanto, que crítica, discordância e questionamento clamando por reconfiguração conceitual sejam práticas constitutivas da vivência de uma cultura ética. Como parte desse processo permanentemente aberto de possibilidade de crítica, discordância e reinterpretação, criam-se condições para visões plurais de mundo, mesmo dentro de uma mesma cultura, e para discordâncias que podem ser insuperáveis. Nenhuma cultura ética, a despeito das visões inocentes antropológicas do passado, é constituída de um todo inteiramente coeso e harmônico. Há algo de inescapavelmente conflituoso na experiência ética, no uso dos conceitos e no exercício de nossas sensibilidades éticas.

Assim, além de um processo de busca por autoconhecimento e interpretação dos seus compromissos centrais, uma cultura ética pode, em certos momentos, como consequência de seus conflitos internos, passar por *crises* e *reformas* profundas. Isso é passível de ser explicado como resultado de uma alteração coletiva nas sensibilidades dos membros dessa cultura, nas condições materiais e sociais, do surgimento de oportunidades históricas, de influência e

necessidade pragmática de adaptação à interação com culturas distintas etc. Ao passar por uma crise, uma cultura se vê pressionada a rever alguns ou muitos dos seus compromissos fundamentais e, assim, seus conceitos podem sofrer uma mudança radical. Se essa cultura passará por uma ruptura completa e se será transformada em outra é uma questão em aberto.²⁴

Existe, então, uma espécie de interdependência dinâmica entre o desenvolvimento de certas sensibilidades sociais coletivas e o conteúdo descritivo de conceitos éticos como parte de um processo histórico, social e cultural. É essa compreensão do fenômeno ético como uma espécie de *construção social* ancorada em uma *prática* o que parece não nos permitir dizer que há pretensão justificada de genuíno conhecimento ético e que, mesmo internamente a uma cultura, seja possível estabelecer critérios *definitivos* de competência, perícia e autoridade éticas.

Assim, por um lado, é possível admitir que, dado suficiente compartilhamento de crenças de fundo, práticas sociais e valores éticos, podemos estabelecer critérios internos a essa cultura ética que justifiquem nossa pretensão de conhecimento ético local e, assim, sermos também capazes de identificar pares éticos e epistêmicos. Isso

²⁴ Nesse sentido, a extensão daquilo que estamos chamando de “local” é uma questão em aberto. Nada do que foi dito aqui é incompatível, por exemplo, com a ideia de que todo o mundo pós-iluminista tende a uma forma de homogeneização em torno de uma única cultura ética. Assim, retomando um exemplo anterior, é uma questão difícil saber se a tendência majoritária global dos tempos atuais de indignação e aversão a práticas excludentes sociais e suas associadas desigualdades, como discursos sexistas e racistas, é parte da mesma cultura inaugurada na modernidade e no iluminismo – e, assim, poderia ser considerada uma forma de “progresso ético”, na medida em que envolveria um maior esclarecimento dos compromissos valorativos dessa cultura – ou se tal tendência marca alguma ruptura de paradigma, inaugurando outra cultura em substituição.

parece ser suficiente para, em certos contextos sociais, justificar um sentido deflacionado de competência, perícia e autoridade éticas, tornando deferência e testemunho possíveis fontes de justificação de crenças éticas. Podemos admitir, em vários contextos, satisfeitas aquelas condições sociais locais, que alguns de nossos pares éticos e epistêmicos podem estar mais bem informados do que nós, ter tido mais tempo para refletir, ter se dedicado mais exaustivamente a analisar uma determinada questão, um determinado padrão de argumentação ou os contornos de determinado conceito, estar em condições psicológicas mais aptas, conhecer melhor a história e realidade concreta de um determinado cenário etc., de modo a tornar justificável formar crenças com base em suas opiniões éticas. Mas, por outro lado, como vimos, há sempre espaço para contestação na dimensão ética de nossas vidas, o que torna contingentes quaisquer critérios que venhamos a adotar para a identificação de competência, perícia e autoridade éticas. Assim, se essa descrição se configura como pretensão de conhecimento ético, o conhecimento em questão deve ser significativamente distinto do que chamamos de conhecimento naquelas outras disciplinas científicas.²⁵

²⁵ Agradeço à Universidade Federal de Minas Gerais (UFMG), à Universidade Federal de Santa Catarina (UFSC) e a seus respectivos departamentos de filosofia por tornarem possível o desenvolvimento desta pesquisa no âmbito de um pós-doutorado em 2024. Em especial, agradeço a Darlei Dall'Agnol e Milene Tonetto pelos comentários e apoio. Agradeço também à audiência do 3º Colóquio Internacional de Filosofia Prática do PPGFIL, realizado na Universidade de Caxias do Sul (UCS), no qual uma versão anterior deste texto foi apresentada. Por fim, agradeço aos participantes dos meus cursos “Desacordo Moral” e “Discordância Ética”, ministrados, respectivamente, na UFSC e na UFMG.

Referências

- Baier, A. (1986) Trust and Antitrust. *Ethics*, v. 96, n. 2, pp. 231-260.
- Berlin, I. (1997) *The Proper Study of Mankind*. Hardy, H.; Hausheer, R. (Orgs.). London: Chatto & Windus.
- Blackburn, S. (1993) *Essays in Quasi-realism*. Oxford: Oxford University Press.
- Boyd, R. (1988) How to Be a Moral Realist. In: Sayre-McCord, G. (Org.) *Essays on Moral Realism*. Ithaca, N.Y.: Cornell University Press, pp. 181-228.
- Brink, D. (1989) *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Copp, D. (2012) Normativity and Reasons: Five Arguments from Parfit against Normative Naturalism. In: Nuccetelli, S.; Seay, G. (Orgs.) *Ethical Naturalism: Current Debates*. Cambridge: Cambridge University Press, pp. 24-57.
- Crisp, R. (2014) Moral Testimony Pessimism: A Defence. *Aristotelian Society*, Supplementary Volume 88, pp. 129-143.
- Daniels, N. (1979) Wide Reflective Equilibrium and Theory Acceptance in Ethics. *Journal of Philosophy*, v. 76, n. 5, pp. 256-282.
- De Maagt, S. (2017) Reflective Equilibrium and Moral Objectivity. *Inquiry*, v. 60, n. 5, pp. 443-465.
- DePaul, M. R. (1993) *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*. London: Routledge.
- Driver, J. (2006) Autonomy and the Asymmetry Problem for Moral Expertise. *Philosophical Studies*, v. 128, pp. 619-644.
- Enoch, D. (2014) A Defense of Moral Deference. *The Journal of Philosophy*, v. 111, n. 5, pp. 229-258.
- Gallie, W. B. (1956) Essentially Contested Concepts. *Proceedings of the Aristotelian Society*, v. 56, pp. 167-198.

- Hall, E. (2020) *Value, Conflict, and Order*. Chicago: The University of Chicago Press.
- Hampshire, S. (1983) *Morality and Conflict*. Oxford: Basil Blackwell.
- Hampshire, S. (2000) *Justice Is Conflict*. Princeton: Princeton University Press.
- Hardwig, J. (1991) The Role of Trust in Knowledge. *The Journal of Philosophy*, v. 88, n. 12, pp. 693-708.
- Hare, R. M. (1952) *The Language of Morals*. Oxford: Oxford University Press.
- Hills, A. (2009) Moral Testimony and Moral Epistemology. *Ethics*, v. 120, n. 1, pp. 94-127.
- Hopkins, R. (2007) What Is Wrong with Moral Testimony? *Philosophy and Phenomenological Research*, v. 74, n. 3, pp. 611-634.
- Jones, K. (1999) Second-hand Moral Knowledge. *The Journal of Philosophy*, v. 96, n. 2, pp. 55-78.
- Kekes, J. (1993) *The Morality of Pluralism*. Princeton: Princeton University Press.
- Kelly, T.; McGrath, S. (2010) Is Reflective Equilibrium Enough? *Philosophical Perspectives*, v. 24, pp. 325-359.
- Lillehammer, H. (2007) *Companions in Guilt: Arguments for Ethical Objectivity*. New York: Palgrave MacMillan.
- Lillehammer, H. (2014) Moral Testimony, Moral Virtue, and the Value of Autonomy. *Proceedings of the Aristotelian Society Supplementary Volume* 88, pp. 111-127.
- McGrath, S. (2008) Moral Disagreement and Moral Expertise. In: Shafer-Landau, R. (Org.) *Oxford Studies in Metaethics* 3. Oxford: Oxford University Press, pp. 87-108.
- McGrath, S. (2009) The Puzzle of Pure Moral Deference. *Philosophical Perspectives*, v. 23, n. 1, pp. 321-344.

- McGrath, S. (2011) Skepticism about Moral Expertise as a Puzzle for Moral Realism. *The Journal of Philosophy*, v. 108, n. 3, pp. 111-137.
- Moore, G. E. (1903) *Principia Ethica*. Cambridge: Cambridge University Press.
- Nguyen, C. T. (2010) Autonomy, Understanding, and Moral Disagreement. *Philosophical Topics*, v. 3, n. 2, pp. 111-129.
- Nickel, P. (2001) Moral Testimony and Its Authority. *Ethical Theory and Moral Practice*, v. 4, pp. 253-266.
- Parfit, D. (1984) *Reasons and Persons*. Oxford: Oxford University Press.
- Rawls, J. (1971) *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.
- Rawls, J. (1974) The Independence of Moral Theory. *Proceedings and Addresses of the American Philosophical Association*, v. 48, pp. 5-22.
- Rorty, R. (1989) *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press.
- Rovane, C. (2009) Did Williams Find the Truth in Relativism? In: Callcut, D. (Org.) *Reading Bernard Williams*. London: Routledge, pp. 43-69.
- Rovane, C. (2013) *The Metaphysics and Ethics of Relativism*. Cambridge MA: Harvard University Press.
- Scanlon, T. M. (2003) Rawls on Justification. In: Freeman, S. (Org.) *The Cambridge Companion to Rawls*. Cambridge: Cambridge University Press, pp. 139-167.
- Scanlon, T. M. (2014) *Being Realistic about Reasons*. Oxford: Oxford University Press.
- Singer, P. (1972) Moral Experts. *Analysis*, v. 32, n. 6, pp. 115-117.
- Sliwa, P. (2012) In Defense of Moral Testimony. *Philosophical Studies*, v. 158, pp. 175-195.

- Smith, M. (1994) *The Moral Problem*. Oxford: Blackwell.
- Walzer, M. (1987) *Interpretation and Social Criticism*. Cambridge, MA: Harvard University Press.
- Warnock, M. (1985) *A Question of Life: The Warnock Report on Human Fertilisation and Embryology*. Oxford: Blackwell.
- Wedgwood, R. (2010) The Moral Evil Demons. In: Feldman, R.; Warfield, T. A. (Orgs.) *Disagreement*. Oxford: Oxford University Press, pp. 216-246.
- Williams, B. (1981) The Truth in Relativism. In: Williams, B. *Moral Luck*. Cambridge: Cambridge University Press.
- Williams, B. (1985) *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.
- Williams, B. (1995a) Saint-Just's Illusion. In: Williams, B. *Making Sense of Humanity*. Cambridge: Cambridge University Press, pp. 135-150.
- Williams, B. (1995b) Who Needs Ethical Knowledge? In: Williams, B. *Making Sense of Humanity*. Cambridge: Cambridge University Press, pp. 203-212.
- Wong, D. B. (2006) *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.

O problema da explicação e da justificação nas teorias evolucionárias da moralidade

Matheus de Mesquita Silveira¹

A ética evolucionária constitui uma abordagem naturalista da ética, que estabelece uma conexão entre as ciências naturais e a teoria moral, fundamentando-se na biologia evolucionária e na psicologia cognitiva. Farber (1994, p. 6) caracteriza a evolução histórica dessa abordagem como uma série de “tentativas de entender a natureza humana a partir de uma perspectiva evolutiva [e de] usar a evolução como fundamento da ética”. Embora essa investigação histórica evidencie uma gama de críticas à ética evolucionária, destaca-se um problema recorrente em todas as formulações desse viés naturalista: em vez de derivar a normatividade da natureza, projeta-se nela os valores sociais e culturais de determinado período histórico. Esse problema se agrava com a adoção de narrativas conjecturais

¹ Professor do programa de pós-graduação em filosofia da Universidade do Vale do Rio dos Sinos (UNISINOS) e pesquisador do International Wolf Center (EUA).

associadas à evolução, que acabam sendo empregadas para justificar posições éticas politicamente controversas.² Farber (1994, p. 2) manifesta ceticismo em relação às discussões surgidas no final do século XX, argumentando que “a maioria dos argumentos contemporâneos são simplesmente versões recicladas de argumentos anteriores”. A despeito desse histórico de insucessos, a ética evolucionária permanece no centro dos debates da filosofia moral contemporânea.

Embora muitos aspectos das abordagens contemporâneas em ética evolucionária ainda reflitam concepções historicamente datadas, determinados argumentos presentes nessas perspectivas permanecem relevantes no debate contemporâneo e, portanto, merecem uma consideração mais aprofundada. Diferentemente da maioria das teses anteriores, que visavam alcançar conclusões normativas e prescritivas, o debate atual em ética evolucionária concentra-se predominantemente em questionar fundamentos metaéticos. Essas discussões, que abordam as implicações da evolução para a teoria metaética, permanecem insuficientemente exploradas.

Neste capítulo, focar-se-á nas perspectivas neodarwinianas que sustentam a ideia de que os seres humanos desenvolveram um senso moral inato, as quais serão denominadas como *nativismo moral darwiniano*. Será examinada a natureza do senso moral e suas características fundamentais, conforme descritas por essas teorias,

² Stewart (1863, p. 18) argumenta que, “na falta dessa evidência direta, somos forçados a substituir os fatos por conjecturas; e [...] devemos considerar como eles provavelmente teriam agido, com base nos princípios de sua natureza e nas circunstâncias externas que enfrentavam”. O termo história conjectural introduzido pelo autor é utilizado para designar uma forma de história que se desenvolveu como substituta da história tradicional, quando o método convencional de investigação se mostrava ausente ou insuficiente.

com especial atenção às diversas implicações metaéticas associadas a essas posições. Em seguida, será desenvolvido o argumento de que o *nativismo moral darwiniano* falha em fundamentar de maneira convincente essas alegações metaéticas, pois se apoia em um método genealógico de justificação problemático para estabelecer os alicerces da moralidade. Embora certos aspectos dessas abordagens possam ter utilidade filosófica, concluir-se-á que seu fundamento geral apresenta sérios problemas. Por fim, será desenvolvida uma defesa de que as limitações da referida posição espelham as restrições mais amplas da ética evolucionária tradicional em sua interpretação geral.

A linha metodológica adotada consiste em, primeiramente, analisar os métodos filosóficos empregados na ética evolucionária, bem como as influências que moldaram as teorias contemporâneas nesse campo. Essa discussão permite esclarecer quais são essas teorias, como se assemelham a projetos anteriores de ética evolucionária e, ao mesmo tempo, como diferem de forma significativa, o que justifica uma nova avaliação crítica. A investigação desenvolve-se a partir de uma consideração mais ampla sobre a função da explicação e da justificação, examinando o papel que ambas desempenham nas teorias da ética evolucionária e abordando a questão central que orienta esse debate: afinal, qual é a origem da moralidade?

1. O papel da explicação e da justificação

Uma explicação consiste no fato de um conjunto de proposições esclarecer determinado fenômeno ou objeto. Em outras palavras, oferece uma justificativa sobre o *porquê* e o *como*, servindo de base para elucidar uma questão específica. Sendo assim, o propósito de uma explicação é simplesmente proporcionar uma compreensão mais clara de determinado fenômeno, por exemplo:

- (i) Alice aprecia limões porque eles possuem um sabor cítrico.
- (ii) Limões são cítricos devido a presença de ácidos cítricos na fruta.

Em ambos os exemplos, a proposição contém uma constatação (“Alice aprecia limões” e “limões são cítricos”), a qual possui uma explicação (“porque eles possuem um sabor cítrico” e “devido à presença de ácidos cítricos na fruta”). A partir dessas explicações, podemos compreender por que Alice aprecia limões e a razão pela qual eles são cítricos.

Salmon (2006) alega que a adequação de uma explicação depende de uma série de fatores, muitos dos quais ultrapassam o escopo deste capítulo.³ Todavia, o autor afirma que o êxito de qualquer explicação geralmente está vinculado à investigação inicial. Em outras palavras, a explicação proporcionada satisfaz a questão que se procurava compreender sobre o fenômeno? No exemplo (ii), se o problema inicial fosse *por que os limões são cítricos?*, a explicação provavelmente seria considerada satisfatória. Entretanto, se a questão fosse *o que torna um limão cítrico em nível molecular?*, então a resposta seria insuficiente. Esse exemplo evidencia três características essenciais de uma explicação: (a) existem diferentes níveis de explicação para um mesmo fenômeno; (b) o seu sucesso está relacionado ao nível de investigação que se pretende alcançar; e (c) o nível de complexidade da resposta depende da natureza da investigação inicial e de seus objetivos. Dessa forma, embora o

³ Neste capítulo, não há um compromisso com uma teoria específica ou com os diferentes tipos de explicação que podem estar em disputa. Ao contrário, o foco está no papel desempenhado pelas explicações em qualquer teoria ou narrativa, a fim de explorar a distinção fundamental entre explicação e justificação. Para discussões sobre os critérios para explicações satisfatórias, ver Pitt, 1988; Salmon, 2006 e Ruben, 1993.

objetivo fundamental de uma explicação seja sempre esclarecer o fenômeno, os diferentes níveis de complexidade são determinados pelos objetivos específicos de cada investigação.

As condições informais de adequação fazem parte da clarificação do *explicandum*; a explicação é a definição formal do conceito aprimorado que deve substituir o conceito original vago de explicação. O que talvez seja surpreendente é a complexidade da explicação formal. (Salmon, 2006, p. 23)

Embora os pontos supracitados possam parecer evidentes, a complexidade aumenta ao considerarmos a relação entre explicação e justificação, pois ambas possuem propósitos distintos. Ao contrário das explicações, que se concentram no *porquê* e *como* de um fenômeno, as justificações fornecem uma narrativa sobre o *dever* que o fundamenta. Além disso, ambas possuem objetivos distintos. A explicação tem como finalidade o esclarecimento e a compreensão de um fenômeno, enquanto a justificação procura demonstrar por que algo é correto ou justificável. Consequentemente, uma justificação também atua como uma defesa de uma posição e envolve um argumento, ao passo que uma explicação não necessariamente o faz.⁴ Considere o seguinte exemplo de justificação:

(iii) Alice deve beber limonada porque os ácidos cítricos são saudáveis.

⁴ Desconsiderando os casos nos quais são desenvolvidos argumentos sobre que tipo de explicação é a correta, de modo amplo é possível estabelecer que uma explicação não consiste num argumento *per se*.

A distinção entre os exemplos (i), (ii) e (iii) reside no fato de que as duas primeiras declarações apenas esclarecem o fenômeno (por que Alice aprecia limões e por que a referida fruta é cítrica), enquanto a terceira declaração apresenta um argumento que justifica por que seria correto, ou por que haveria boas razões, para Alice consumir uma limonada.⁵ Embora explicações e justificações sejam conceitualmente distintas devido aos seus objetivos divergentes, existe uma possibilidade de sobreposição entre elas, visto que uma explicação pode ser integrada em uma justificativa em determinados contextos. Para ilustrar esse ponto, recapitular-se-á as declarações anteriores:

- (i) Alice gosta de limões porque eles têm um gosto cítrico.
- (ii) Limões são cítricos porque contêm ácidos cítricos.
- (iii) Alice deve beber limonada porque ácidos cítricos são saudáveis.

As declarações explicativas em (i) e (ii) agora parecem apoiar a justificativa apresentada em (iii). No entanto, essa relação depende do objetivo da investigação original. Caso este seja meramente explicativo, então as declarações (i) e (ii) permanecerão como simples explicações. Por outro lado, caso o objetivo resida na esfera justificativa, então as declarações explicativas passam a integrar uma justificação.

No contexto apresentado, as declarações (i) e (ii) podem atuar em conjunto como parte de uma justificação, mas, isoladamente,

⁵ Bonjour e Sosa (2003) abordam a questão da justificação epistêmica, em particular no debate entre fundacionalismo, coerentismo e confiabilismo. Contudo, a justificação será discutida, neste capítulo, num sentido amplo e não avançará na avaliação das condições epistemológicas que conferem o estatuto satisfatório a uma justificativa.

continuam sendo apenas declarações explicativas. Para que uma explicação se transforme em uma justificação, é necessário um trabalho adicional, a saber, modificar as próprias declarações, o contexto em que se inserem e quaisquer outras características relevantes para o fenômeno investigado. A forma exata dessa transição depende do que está sendo justificado e se as declarações explicativas fornecem ou não boas razões para determinada justificação – cabe destacar que variações no contexto acarretarão diferentes perspectivas sobre o que constitui uma *boa razão*. O ponto é que não há uma transição automática entre uma declaração explicativa e uma justificatória, e um esforço adicional se faz necessário para transformar uma explicação (razão explicativa) em uma justificação (razão justificatória).

Para que uma explicação se transforme em justificação, ela precisa ser incorporada em um argumento, mediante a apresentação de informações adicionais, no intuito de demonstrar o porquê de as declarações explicativas fornecerem suporte à conclusão justificatória. Sem esse esforço adicional, a confusão entre explicação e justificação resulta em um raciocínio circular de natureza cartesiana.⁶ Por exemplo, o argumento de Descartes (1996) pode ser formulado da seguinte maneira:

- (i) Se Deus é benevolente (não enganador), então, quando *eu* percebo algo de forma clara e distinta, isso deve ser verdadeiro.
- (ii) *Eu* percebo clara e distintamente que um Deus benevolente existe.
- (iii) Portanto, um Deus benevolente existe.

⁶ Esse é um exemplo ilustrativo, e foge ao escopo deste capítulo adentrar na discussão acerca das possíveis interpretações da estrutura epistemológica cartesiana. Para um debate sobre esse ponto, ver Doney, 1955; Gewirth, 1970 e Cottingham, 1991.

(iv) Se um Deus benevolente existe, então, quando *eu* percebo algo de forma clara e distinta, isso deve ser verdadeiro.

(v) Portanto, se *eu* percebo algo clara e distintamente, isso deve ser verdadeiro.

O raciocínio de Descartes é circular, pois a explicação de como o meditador sabe que não está sendo enganado, ou seja, o fato de ele perceber algo clara e distintamente, também é usada como justificativa, tanto para sua percepção quanto para a existência de Deus. Em outras palavras, a explicação e a justificação se confundem e o argumento permanece viciosamente circular caso novas premissas não sejam adicionadas.

As explicações na ética evolucionária procuram elucidar *por que* e *como* os indivíduos passaram a ter agência moral. Em linhas gerais, elas visam à compreensão da moralidade em diversos níveis; como o porquê de grupos cooperarem entre si, como o altruísmo se desenvolveu, se os indivíduos possuem ou não um senso moral inato, entre outros. Como mencionado anteriormente, o sucesso de qualquer explicação está vinculado à investigação inicial. A questão primordial que orienta a abordagem evolucionária pode ser caracterizada na seguinte questão: de onde vem a moralidade? Esta é uma pergunta orientada ao campo da explicação. No entanto, há abordagens que utilizam explicações evolucionárias como justificativa para sustentar teses morais, tanto no nível normativo quanto metaético. Tais teorias não apenas investigam *por que* e *como* a moralidade evoluiu, mas também avançam na justificação acerca de quais princípios morais devem ser adotados e, inclusive, de quais comportamentos sociais seriam moralmente justificados.

A transição da explicação para a justificação e, consequentemente, para a prescrição normativa demanda uma análise mais detalhada. O argumento apresentado neste capítulo é o de que essa passagem é complexa e sua natureza bastante problemática. Em outras palavras, o entusiasmo filosófico com a perspectiva evolucionária fez com que explicações e justificativas fossem frequentemente confundidas, sem que se realizasse um trabalho adicional para distinguir tais categorias. A investigação subsequente da questão inicial que orienta a ética evolucionária, a saber, a origem da moralidade, permitirá esclarecer, ao menos em parte, essa confusão.

2. Genealogia evolucionária da moral

No contexto da ética evolucionária contemporânea, Ruse (2009) assinala que o conceito de evolução utilizado se refere à teoria da síntese estendida, que combina o mecanismo darwiniano de seleção natural com a genética moderna. Contudo, a ética evolucionária não possui uma teoria unificada. Richards (1986) destaca que existem várias questões abordadas no campo em questão e diferentes métodos utilizados para lidar com tais problemas. O autor defende que o viés evolucionário na ética possui uma ampla miríade de posições. Em outras palavras, os fundamentos evolucionários moldam de formas distintas as reivindicações metaéticas e normativas de cada uma dessas posições.⁷ Todavia, James (2010) sustenta que essas diversas

⁷ É importante observar que nem todas as posições dentro da ética evolucionária fazem afirmações normativas, seja em termos metaéticos ou prescritivos, enquanto outras alegam ser exclusivamente descritivas. Por exemplo, Hamilton, 1964; Trivers, 1971 e Axelrod, 1984 buscam uma explicação evolucionária do altruísmo adotando uma abordagem puramente descritiva. No entanto, a filosofia frequentemente utiliza

perspectivas têm em comum a alegação de que a teoria evolucionária ajuda a explicar as origens do fenômeno moral, a qual é utilizada como argumento para estabelecer a pedra de toque da moralidade. Sendo assim, em alguma medida a ética evolucionária busca responder à seguinte pergunta: afinal, de onde vem a moralidade?

Julian Huxley queria justificar a ética normativa com base nos fatos e processos da evolução, e, para isso, ele precisava argumentar que o progresso é um fato objetivo da natureza. Ele estava, portanto, disposto a enfrentar de forma direta a falácia naturalista. Simpson queria explicar a ética normativa em termos dos fatos e processos da evolução e, ao fazê-lo, embora fosse progressista, ele queria evitar afirmar que isso é um fato objetivo da natureza. Ele levava a falácia naturalista a sério, como algo a ser evitado, e falava muito sobre a importância da responsabilidade individual como base para a moralidade. (Ruse, 2009, p. 149)

A questão sobre a origem da moralidade é ambígua e pode ser interpretada de duas formas distintas: (i) como uma investigação sobre as origens da moralidade; ou (ii) como uma busca pelos fundamentos da moralidade. Embora essas duas abordagens possam parecer semelhantes, cada uma apresenta objetivos e métodos próprios. Quando se trata de uma investigação sobre as origens da moralidade, conjuntos de explicações são oferecidos para responder à questão de onde a moralidade se origina. Por outro lado, o que a busca

esse tipo de trabalho descritivo como suporte para sustentar posições normativas, tanto em termos metaéticos quanto prescritivos.

pelos fundamentos da moralidade exige são conjuntos de justificações.

Para esclarecer essas distinções, é útil considerar posições filosóficas que adotam tais abordagens. Note que o objetivo não é apresentar uma análise exaustiva de diferentes visões, tampouco defender uma posição específica no campo filosófico. Em vez disso, buscar-se-á oferecer uma categorização geral de algumas posições como forma de ilustrar as diferentes abordagens na resposta à investigação inicial sobre as origens da moralidade. Assim, o foco deste capítulo não está em defender interpretações específicas dessas teorias, mas em mostrar como mapear essas abordagens ajuda a elucidar suas respostas à questão que fundamenta a ética evolucionária.

2.1 Bases históricas do método genealógico

Num sentido amplo, Owen (1997) sustenta que uma investigação sobre as origens da moralidade utiliza o método genealógico para analisar os valores contemporâneos com base em seu desenvolvimento histórico. Nessa perspectiva, Elbe (2001) argumenta que a história do surgimento ou desenvolvimento de uma crença moral oferece uma compreensão das razões pelas quais possuímos os valores que temos. A forma como os adquirimos desempenha um papel crucial tanto em sua sustentação quanto em sua fragilidade. Em outras palavras, nessa abordagem, a história de uma crença ou valor moral é considerada fundamental para a avaliação de sua legitimidade.

A primeira característica de uma abordagem genealógica, então, é que se trata de um estudo histórico que se volta

para o passado, não por si só, mas com o objetivo de explicar algo que continua a ser problemático no presente. Em resumo, é uma história de como nos tornamos o que somos. (Elbe, 2001, pp. 260-261)

A abordagem genealógica se assemelha ao que Sidgwick (1886, p. 211) denominou como *método histórico*, o qual é criticado pelo autor e compreendido de maneira abrangente como “qualquer estudo do passado com o intuito de explicar o presente”. Contudo, diversos teóricos do cânone filosófico adotaram esse método de investigação. Por exemplo, Hobbes (2009) e Rousseau (1923) analisaram os valores morais de sua época à luz de um processo histórico com o objetivo de esclarecer tais valores. Korsgaard (2010) sustenta que ambos os autores não exploram historicamente o estado original da natureza humana, mas o fazem de forma hipotética e imaginada, com o intuito de explicar a subsequente ascensão ou decadência humana, a fim de compreender melhor os seus compromissos morais e práticos. De fato, a posição hobbesiana utiliza essa genealogia do ser humano para fundamentar determinadas estruturas políticas, enquanto Rousseau busca, ao revelar sua origem, questionar os valores que as fundamentam.

[...] um dos modos tradicionais de entendimento filosófico, especialmente no que diz respeito à moralidade, é a história de origem: pense, por exemplo, nos relatos de moralidade que encontramos em Hobbes, Nietzsche ou Rousseau. Todos esses pensadores tentam esclarecer o que significa ser humano contando histórias sobre como os motivos morais, as emoções ou até mesmo as obrigações podem ter surgido a partir de eventos ou processos concebidos como históricos. Portanto, é natural pensar

que uma explicação evolutiva da moralidade possa, de alguma forma, lançar luz sobre o próprio fenômeno. (Korsgaard, 2010, p. 2)

Os objetivos do método genealógico, portanto, são variados e podem ser empregados tanto para justificar quanto para questionar valores contemporâneos. No entanto, o seu uso para justificar diretamente um valor, instituição ou prática frequentemente incorre na falácia genética, que consiste no erro lógico de aceitar ou rejeitar uma alegação com base exclusivamente em sua origem. Em outras palavras, essa forma de aplicação do método genealógico é falaciosa, pois confunde explicações genealógicas com justificativas, seja para apoiar ou para refutar valores morais contemporâneos.

Geuss (1994) destaca que o uso mais frequente da genealogia não é o da justificação, mas o da reflexão crítica, focalizada em questionar e deslegitimar valores, práticas ou instituições contemporâneas. Embora o uso crítico não esteja completamente imune à falácia genética, existem formulações que, à primeira vista, parecem evitar esse erro lógico. Dessa forma, as metodologias genealógicas podem ser divididas em dois tipos: (i) o método dedutivo, que tende a incorrer na falácia genética; e (ii) o método crítico, que aparenta escapar desse erro. Aqui, cabe-se dedicar uma atenção especial, sobretudo porque o método crítico é amplamente empregado em teorias éticas evolucionárias contemporâneas, especialmente ao âmbito das discussões metaéticas. Para ilustrar o argumento, examinar-se-á a genealogia crítica exemplificada pela investigação nietzschiana da moralidade.

Prinz (2009a) sustenta que um dos principais objetivos dos escritos morais nietzschianos era traçar um relato histórico das

origens da moralidade, com o propósito de criticar os valores morais dominantes de sua época. De fato, Nietzsche (1994, 1996) dá especial atenção à influência do cristianismo na construção desses valores. Elbe (2001) define o método genealógico crítico de Nietzsche como “episódico”, uma abordagem que se concentra em eventos específicos em vez de abarcar a história em sua totalidade. O autor usa como exemplo a discussão acerca das crenças religiosas, na qual Nietzsche (1996) busca demonstrar como as origens de uma crença podem impactar a confiança em determinado valor. Dessa forma, a genealogia episódica torna-se uma ferramenta central no pensamento do filósofo alemão para questionar e criticar os valores morais da sociedade na qual vivia.

Além disso, a ideia de um Deus perturba e humilha enquanto é acreditada, mas, dado o estado atual da etnologia comparativa, sua origem já não pode ser questionada; e, com a compreensão dessa origem, a crença desaparece. O cristão que compara sua natureza à de Deus é como Dom Quixote, que subestima sua própria bravura porque está preocupado com os feitos milagrosos de heróis dos romances de cavalaria; em ambos os casos, o padrão de medida utilizado pertence ao reino da fábula. (Nietzsche, 1996, p. 93)

Nietzsche (1994) busca revelar as origens dos valores morais de sua época como uma forma de questionar a legitimidade das crenças que sustentam esses mesmos valores. Em outras palavras, se os valores possuem *origens ignóbeis*, isso sugere que eles devem ser reavaliados. Contrariando a interpretação comum do alegado niilismo nietzschiano, Prinz (2009b) argumenta que não há uma negação de

que os indivíduos possuam um conjunto de valores, reconhecendo nesse sentido a sua existência. De fato, o filósofo alemão explica que seu objetivo é questionar a moralidade com base em suas origens e, portanto, não está debatendo a existência de tais valores, mas sim a correção das crenças sobre eles. Portanto, sua crítica é normativa, ou seja, ele sugere que, ao se compreender as razões pelas quais acredita-se nesses valores, será possível encontrar justificativas para rejeitá-los.

Precisamos de uma crítica dos valores morais, o valor desses valores deve, pela primeira vez, ser examinado – e por isso precisamos saber sobre as condições e circunstâncias sob as quais esses valores surgiram, se desenvolveram e mudaram [...] já que até agora não tivemos esse conhecimento, nem sequer o desejamos. (Nietzsche, 1994, p. 8)

Em sua análise sobre a utilidade da genealogia da moral, Prinz (2009b, p. 237) identifica a posição nietzschiana como operando a partir da tese de que “se a crença em um valor moral surgiu como consequência de eventos históricos vis, então essa crença deve ser abandonada”. De acordo com essa interpretação do método genealógico crítico, a história e as circunstâncias sob as quais determinados valores se formaram, ou seja, as origens de uma crença moral, podem enfraquecer a confiança de alguém neles.⁸ É fundamental destacar que essa não é uma tese de natureza

⁸ Prinz (2009a) destaca que o método genealógico crítico pressupõe a existência de um ponto de vista moral objetivo a partir do qual é possível julgar as narrativas históricas. Segundo o autor, isso permitiria a identificação de certos valores como provenientes de origens ignóbeis, o que justificaria sua rejeição. Para um debate mais aprofundado sobre essa questão, ver Leiter, 2000.

epistemológica ou metafísica; em outras palavras, não questiona diretamente a validade ou justificação das crenças morais nem a verdade ou existência dos valores em si. Com referência a Nietzsche (2001), Geuss (1994, pp. 286-287) argumenta que “o propósito e o efeito de uma genealogia não podem ser o de criticar valores ou valorações diretamente, mas uma genealogia pode enfraquecer várias crenças sobre as origens de diferentes formas de valoração”. O ponto é que, se o objetivo fosse atacar determinados valores morais diretamente, então a falácia genética seria inevitável.

Prinz (2009b, p. 239) propõe uma versão mais moderada da tese nietzschiana, defendendo que “se uma crença em um valor moral surgiu como consequência de eventos históricos ignóbeis, deveríamos considerar abandonar essa crença”. Embora essa variante do método genealógico crítico ainda funcione como uma ferramenta de análise, o autor alega que ela é mais segura e evita a falácia genética. O argumento é o de que a explicação genealógica não se confunde com uma justificativa para aceitar ou rejeitar valores morais contemporâneos. Em outras palavras, em vez de rejeitar uma crença com base exclusivamente em suas origens, a proposta do filósofo estadunidense apenas alerta sobre tais crenças, exigindo uma investigação adicional para determinar se elas devem ser abandonadas.⁹

Elbe (2001) resume o método genealógico crítico, tal como empregado nas investigações nietzschianas, a partir de três características centrais. Segundo o autor, tal metodologia é histórica,

⁹ É possível destacar que Nietzsche (1996, 2001) oferece outros argumentos além da genealogia da moral para questionar determinados valores, juntamente com uma proposta de visão moral positiva. Todavia, tal análise foge ao escopo deste capítulo, e uma discussão aprofundada poderá ser encontrada em Leiter (2002).

crítica e episódica, pois examina a trajetória dos valores morais, reflete sobre sua presença na sociedade contemporânea e concentra-se em eventos históricos específicos, em vez de abarcar a história como um todo. Além disso, na forma aplicada por Nietzsche (1994, 1996), o método opera em um nível normativo, sem se envolver diretamente com questões epistemológicas ou metafísicas (como garantia, justificação, verdade ou existência).¹⁰

Retornando à pergunta inicial, a saber, *de onde vem a moralidade?*, evidenciou-se que o método genealógico busca responder a ela por meio de uma investigação histórica, na qual a crítica aos valores é realizada com base em suas origens. Chamar-se-á essa primeira interpretação da origem da moralidade de *abordagem genealógica*. Como mencionado anteriormente, existem dois tipos de abordagens genealógicas: a dedutiva e a crítica. Em suma, as tentativas dedutivas buscam apoiar ou rejeitar diretamente uma visão moral específica com base em sua origem. Dado que a aproximação dedutiva é claramente falaciosa, e a abordagem crítica é a mais comumente utilizada nas teorias éticas evolucionárias contemporâneas com foco em questões metaéticas, concentrar-se-á a análise na metodologia crítica.

2.2 Bases metafísicas da abordagem teórica

A fundamentação metafísica é uma alternativa ao método genealógico enquanto uma forma de compreender a questão acerca

¹⁰ Leiter (2002) comenta que tais reivindicações normativas têm implicações metaéticas, mas isso seria apenas de maneira indireta e, portanto, não se constituiriam como foco imediato das críticas apresentadas. Em outras palavras, são posições desenvolvidas em projetos adicionais e utilizam uma metodologia distinta da discutida neste capítulo.

das origens da moralidade. Essa abordagem busca identificar a fonte das obrigações morais a partir da determinação de seu fundamento, o qual pode estar na mente, na natureza, na sociedade, entre outros. Em linhas gerais, esses fundamentos esclarecem a natureza das obrigações morais e podem ser utilizados para legitimar ou contestar os valores existentes, e não apenas as crenças sobre eles. Diferentemente de uma análise histórica, essa investigação não trata das origens cronológicas, mas sim do fundamento a partir do qual derivamos as obrigações morais. Nesse sentido, o relato das origens da moralidade, entendido no sentido de uma fundamentação, justifica diretamente obrigações morais ou valores específicos. Em linhas gerais, o foco não é apenas explicar de onde vêm as obrigações ou por que se acredita em determinados valores morais (como em uma genealogia), mas examinar como estes são justificados. Portanto, a busca pelos fundamentos da moralidade geralmente tem como objetivo justificar um conjunto de valores, em vez de questionar ou minar os presentes na sociedade, sendo, em regra, um método não crítico.

A filosofia moral de Kant (1993) exemplifica a abordagem metafísica, na qual ele busca estabelecer o imperativo categórico como um princípio fundamental para a moralidade, a qual é derivada da razão pura e consiste na base do desenvolvimento de sua teoria normativa. Para o autor, a descoberta de tal fundamento não pode ser um processo histórico, pois está vinculada à lei moral, caracterizada como a dimensão racional da metafísica dos costumes. Dessa forma, os valores não pertencem à empiria, a qual faz referência a uma antropologia prática, sendo histórica no mesmo sentido do que o método genealógico apresentado na seção anterior. Em outras palavras, na posição kantiana a moralidade não deve se basear em considerações empíricas e contingentes, mas em princípios

necessários de natureza lógica ou conceitual. De fato, Kant (1993, p. 2) coloca que “[...] o fundamento da obrigação aqui deve, portanto, ser buscado não na natureza do homem nem nas circunstâncias do mundo em que o homem se encontra, mas deve ser buscado *a priori* exclusivamente nos conceitos da razão pura”. Assim, o filósofo prussiano procura justificar os valores a partir dos fundamentos ou origens daquilo que denomina como lei moral.

Ao buscar uma base metafísica para a moralidade, a posição kantiana considera questões epistemológicas e metafísicas. O foco central dessa abordagem está nas condições de fundamento e justificação de crenças morais, na verdade dos valores existentes e em seu estatuto ontológico. Consequentemente, Kant (1993) lida diretamente com questões metaéticas e, por ser uma abordagem a-histórica e não empírica, denominar-se-á essa segunda interpretação da origem da moralidade como *abordagem teórica*. Na próxima seção, mostrar-se-á que o método genealógico crítico e a abordagem teórica não apenas possuem objetivos distintos, pois o primeiro visa criticar os valores morais vigentes, enquanto a segunda está focada na sua justificação (ainda que não necessariamente dos presentes em determinada sociedade), mas também podem ser vistos como concepções opostas sobre como conduzir investigações filosóficas sobre tais temas.

2.3 Método genealógico e abordagem teórica

Nietzsche (1994) entendia a metodologia genealógica como uma rejeição das abordagens teóricas kantianas que predominavam em seu tempo. A posição nietzschiana é a antítese da apresentada por Kant (1993), no sentido de que a sua visão constitui uma rejeição da metafísica. Hoy (1986, pp. 252-253) descreve tal projeto moral como

marcadamente não metafísico, afirmando que “[...] Nietzsche usará a genealogia para destruir a metafísica por completo. A genealogia se torna, assim, uma maneira de fazer filosofia não metafísica.” Segundo o autor, na posição nietzschiana não existem características transcendentais à humanidade.¹¹

Pode-se distinguir dois tipos de projetos éticos que ajudam a clarificar as diferenças entre a o método genealógico crítico e a abordagem teórica. Segundo Sober (1994), as perguntas (i) *por que as pessoas têm os pensamentos e sentimentos morais que possuem?*, e (ii) *qual o problema de determinar qual é o estatuto desses pensamentos e sentimentos?* são questões distintas e não estão necessariamente conectadas. Note que a primeira é de natureza explicativa, enquanto a segunda converge para o problema de justificação. De fato, o autor aponta que não há conexão necessária entre (i) e (ii), ou seja, entre explicação e justificação. Portanto, investigar as origens da moralidade a partir do método genealógico crítico se alinha com o projeto explicativo descrito em (i), enquanto tratar a mesma questão sob a abordagem teórica corresponde à busca pela justificação mencionada em (ii).

A abordagem teórica busca justificar o estatuto dos valores morais, estabelecendo seus fundamentos metafísicos. Por outro lado, o método genealógico crítico usa a história como base para explicar como as pessoas passaram a sustentar as crenças morais que possuem. Prinz (2009b, p. 242) afirma que “quando vemos a moralidade historicamente, deixamos de lado as questões de justificação e nos

¹¹ Foge ao escopo deste capítulo apresentar as sutilezas dessa discussão. Entretanto, caso o leitor deseje aprofundar a questão, encontrará em Bailey (2013) uma interpretação alternativa, na qual a vontade de poder seria uma característica metafísica.

envolvemos em uma autoantropologia. Não perguntamos ‘Por que devo acreditar em P?’, mas sim ‘Por que acredito em P?’.” Portanto, a metodologia genealógica pode levar ao questionamento das crenças e valores morais, porém não justifica nem compromete automaticamente seu estatuto.

Como discutido anteriormente, explicações e justificativas diferem por atenderem a propósitos distintos, ou seja, não há uma transição automática entre uma declaração explicativa e uma justificatória. No entanto, na ética evolucionária, o método genealógico crítico e a abordagem teórica tendem a se fundir, pois ela não apenas fornece uma genealogia da moralidade, mas também é usada para estabelecer os seus fundamentos metafísicos. Em outras palavras, as explicações evolucionárias são frequentemente utilizadas na filosofia como justificativas para determinadas posições normativas. Note que tal movimento não diz respeito exclusivamente a uma distinção entre fato e valor no contexto da ética evolucionária. A questão central é o colapso da diferença entre explicação e justificação, e a necessidade de explorar em que ponto uma explicação se transforma em justificação. Embora seja possível considerar essas abordagens equivalentes, tal confusão parece consistir num erro (mesmo que uma possa se transformar na outra em determinados contextos) e, portanto, vai além da simples questão da falácia naturalista de Moore (2004). Apesar de essa discussão ser capaz de levantar preocupações mais amplas sobre a viabilidade de uma ética naturalizada, esse não é o foco do argumento apresentado no capítulo, que se restringe a questionar o caso específico do problema entre explicação e justificação.

A teoria evolucionária pode ser vista como a extensão final de uma metodologia histórica na filosofia e nas ciências naturais.

Sidgwick (1886, p. 12) conclui que, ao aplicar esse método ao campo da ética, o objetivo não é mais apenas explicar o presente, mas também estabelecer “um método para determinar [...] até que ponto elas [as explicações e justificações morais] são verdadeiras ou falsas”. Em linhas gerais, as teorias evolucionárias contemporâneas combinam abordagens históricas e empíricas para justificar a moralidade, sustentando o realismo moral, ou para questioná-la, levando ao antirrealismo acerca do fenômeno moral. No último caso, Sober (1994) observa que os argumentos na ética evolucionária muitas vezes buscam provar o subjetivismo ético, baseando-se no fato de que as crenças morais derivam de um processo evolucionário – uma abordagem explicativa clara, mas que encontra problemas para avançar na justificação de valores.

3. Conclusão

Afinal, uma genealogia da moral baseada na teoria evolucionária pode realmente determinar os fundamentos da moralidade? Em outras palavras, podem as explicações evolucionárias servir como justificativas para uma visão metaética específica? Ainda não há uma resposta filosófica definitiva para tais perguntas, e apresentar o amálgama entre o método genealógico crítico e a abordagem teórica é um desafio a ser superado pelas abordagens evolucionárias. Todavia, cabe-se também ponderar se as bases filosóficas estabelecidas como condições necessárias à justificação são adequadas para dar conta dos fenômenos presentes no mundo – dentre eles, o da moralidade. Em conclusão, enfrentar o desafio imposto pela separação entre explicação e justificação é tão necessário quanto questionar os cânones sobre os quais se debruça para estabelecer tais distinções.

Referências

- Axelrod, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.
- Bailey, T. (2013) Nietzsche the Kantian? In: Richardson, J.; Gemes, K. (Orgs.) *The Oxford Handbook on Nietzsche*. Oxford: Oxford University Press, pp. 134-159.
- Bonjour, L.; Sosa, E. (2003) *Epistemic Justification: Internalism vs Externalism, Foundations vs Virtues*. Malden: Blackwell.
- Cottingham, J. (1991) *Descartes*. Malden: Blackwell.
- Descartes, R. (1996) *Meditations on First Philosophy*. In: Cottingham, J. (Org.) *Descartes: Meditations on First Philosophy: with Selections from the Objections and Replies*. Cambridge: Cambridge University Press, pp. 1-62.
- Doney, W. (1955) The Cartesian Circle. *Journal of the History of Ideas*, v. 16, n. 3, pp. 324-338.
- Elbe, S. (2001) "We Good Europeans...: Genealogical Reflections on the Idea of Europe". *Millennium: Journal of International Studies*, v. 30, n. 2, pp. 259-283.
- Farber, P. L. (1994) *The Temptations of Evolutionary Ethics*. Berkeley: University of California Press.
- Geuss, R. (1994) Nietzsche and Genealogy. *European Journal of Philosophy*, v. 2, n. 3, pp. 274-292.
- Gewirth, A. (1970) The Cartesian Circle Reconsidered. *The Journal of Philosophy*, v. 67, n. 19, pp. 668-685.
- Hamilton, W. D. (1964) The Genetical Evolution of Social Behaviour. *Journal of Theoretical Biology*, v. 7, n. 1, pp. 1-16.
- Hobbes, T. (2009) *Leviathan*. Oxford: Oxford University Press.
- Hoy, D. C. (1986) Nietzsche, Hume and the Genealogical Method. In: Yovel, Y. (Org.) *Nietzsche as Affirmative Thinker: Papers Presented*

- at the Fifth Jerusalem Philosophical Encounter, April 1983.
Dordrecht: Kluwer Academic Publishers, pp. 20-38.
- James, S. M. (2010) *An Introduction to Evolutionary Ethics*. Malden: Blackwell.
- Kant, I. (1993) *Grounding for the Metaphysics of Morals*. Indianapolis: Hackett Publishing Company.
- Korsgaard, C. M. (2010) Reflections on the Evolution of Morality. *The Amherst Lecture in Philosophy*, n. 5, pp. 1-29.
- Leiter, B. (2000) Nietzsche's Metaethics: Against the Privilege Readings. *European Journal of Philosophy*, v. 8, n. 3, pp. 277-297.
- Leiter, B. (2002) *Nietzsche on Morality*. London: Routledge.
- Moore, G. E. (2004) *Principia Ethica*. New York: Dover Publications.
- Nietzsche, F. (1994) *On the Genealogy of Morals*. Cambridge: Cambridge University Press.
- Nietzsche, F. (1996) *Human, All Too Human: A Book for Free Spirits*. Lincoln: University of Nebraska Press.
- Nietzsche, F. (2001) *The Gay Science*. Cambridge: Cambridge University Press.
- Owen, D. (1997) *Maturity and Modernity: Nietzsche, Weber, Foucault, and the Ambivalence of Reason*. New York: Routledge.
- Pitt, J. (1988) *Theories of Explanation*. New York: Oxford University Press.
- Prinz, J. J. (2009a) Against Moral Nativism. In: Murphy, D.; Bishop, M. (Orgs.) *Stich and His Critics*. Malden: Wiley-Blackwell, pp. 167-189.
- Prinz, J. J. (2009b) *The Emotional Construction of Morals*. New York: Oxford University Press.
- Richards, R. (1986) A Defense of Evolutionary Ethics. *Biology and Philosophy*, v. 1, pp. 263-293.

- Rousseau, J. (1923) *The Social Contract and Discourses*. Toronto: J. M. Dent and Sons.
- Ruben, D. (1993) *Explanation*. Oxford: Oxford University Press.
- Ruse, M. (2009) *Defining Darwin*. Amherst: Prometheus Books.
- Salmon, W. (2006) *Four Decades of Scientific Explanation*. Pittsburgh: University of Pittsburgh Press.
- Sidgwick, H. (1886) The Historical Method. *Mind*, v. 11, n. 42, pp. 203-219.
- Sober, E. (1994) *From a Biological Point of View*. Cambridge: Cambridge University Press.
- Stewart, D. (1863) Account of the Life and Writings of Adam Smith. In: Hamilton, S. W. (Org.) *The Collected Works of Dugald Stewart*. Edinburgh: Thomas Constable and Co., pp. 1-203.
- Trivers, R. L. (1971) The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, v. 46, n. 1, pp. 35-57.

