

An abstract painting of a face, rendered in dark, expressive brushstrokes. The face is composed of various shades of grey, black, and white, with some areas of red and orange. The eyes are closed or heavily shadowed, and the mouth is open, showing a dark interior. The background is a mix of light and dark tones, with some splatters and textures. The overall style is reminiscent of expressionism or abstract art.

Agência, Deliberação e Motivação

Volume 2

João Hobuss

Evandro Barbosa (Orgs.)

φ

Em “Agência, Deliberação e Motivação”, Volume I, reunimos as principais conferências apresentadas ao longo do V Congresso Internacional de Filosofia Moral e Política (V CIFMP): Agência, Deliberação e Motivação. O evento, realizado de 06 a 09 de novembro de 2017, foi organizado pelo Programa de Pós-Graduação em Filosofia da Universidade Federal de Pelotas (UFPel). Dando continuidade às discussões, a presente obra Agência, Deliberação e Motivação – Volume II foi elaborada com o intuito de fornecer a especialistas, iniciantes e demais interessados no assunto a oportunidade de ler em nossa língua materna traduções de textos de renomados autores na área, acompanhados pela contribuição dos organizadores desta coletânea sobre o tema. Temos, assim, um movimento de “mão dupla”. Por um lado, o V CIFMP se consolidou como um espaço profícuo para a discussão e a divulgação de resultados parciais de pesquisas de excelência realizadas por pesquisadores do Brasil em contato direto com outros centros de pesquisa em filosofia na América do Norte e Europa. Por outro lado, a tradução para o português de textos filosóficos na área proporcionará à comunidade filosófica brasileira – especialmente aos alunos de graduação e pós-graduação, que ainda sentem o desafio da leitura em outra língua – a possibilidade de acesso a temas relevantes, antes inacessíveis. Esperamos, dessa forma, ampliar o alcance dessas discussões, estabelecendo um fórum de debate no qual todos se sintam incluídos. Este volume contém oito capítulos que representem discussões recentes sobre o tema da agência, deliberação ou motivação provindo de textos selecionados de revistas de filosofia especializadas e traduzidos do inglês para o português, tomando como critério a relevância dos textos para as discussões na temática. Além deles, dois capítulos deste livro são resultado das pesquisas dos organizadores do presente volume.



Agência, Deliberação e Motivação

Agência, Deliberação e Motivação

Volume 2

Organizadores:

João Hobuss

Evandro Barbosa

***φ** editora fi*

Diagramação: Marcelo A. S. Alves

Capa: Carole Kümmecke - <https://www.behance.net/CaroleKummecke>

Arte de Capa: Eric Lacombe

Revisão textual: Stephany Barbosa

O padrão ortográfico e o sistema de citações e referências bibliográficas são prerrogativas de cada autor. Da mesma forma, o conteúdo de cada capítulo é de inteira e exclusiva responsabilidade de seu respectivo autor.



Todos os livros publicados pela Editora Fi estão sob os direitos da Creative Commons 4.0
https://creativecommons.org/licenses/by/4.0/deed.pt_BR



Associação Brasileira de Editores Científicos

<http://www.abecbrasil.org.br>

Dados Internacionais de Catalogação na Publicação (CIP)

HOBUSS, João; BARBOSA, Evandro (Orgs.)

Agência, Deliberação e Motivação — Volume II [recurso eletrônico] / João Hobuss; Evandro Barbosa (Orgs.)
-- Porto Alegre, RS: Editora Fi, 2018.

315 p.

ISBN - 978-85-5696-388-8

Disponível em: <http://www.editorafi.org>

1. Agência. 2. Deliberação. 3. Motivação. 4. Filosofia. 5. Ética. I. Título

CDD: 170

Índices para catálogo sistemático:

1. Ética 170

Sumário

Apresentação	9
Sobre os autores	11
Capítulo 1	15
Cognitivismo, motivação e abordagens de processo-duplo ao juízo normativo Brendan Cline	
Capítulo 2.....	61
Obrigação moral e motivação moral* David Copp	
Capítulo 3.....	99
Dando razões práticas David Enoch	
Capítulo 4.....	147
O ponto de vista da segunda-pessoa, parte II: obrigação moral e o problema da responsabilidade Evandro Barbosa	
Capítulo 5.....	173
Pessoas virtuosas podem fazer coisas más? Aristóteles entre a idealização e o realismo João Hobuss	
Capítulo 6.....	185
A questão do realismo Kit Fine	
Capítulo 7.....	243
Correção dos sentimentos em Hume: Intersubjetividade sem objetividade Max Barkhausen	
Capítulo 8	277
Valor e implicatura Stephen Finlay	

Apresentação

Em “Agência, Deliberação e Motivação”, Volume I, reunimos as principais conferências apresentadas ao longo do *V Congresso Internacional de Filosofia Moral e Política (V CIFMP)*: Agência, Deliberação e Motivação. O evento, realizado de 06 a 09 de novembro de 2017, foi organizado pelo Programa de Pós-Graduação em Filosofia da Universidade Federal de Pelotas (UFPel). Dando continuidade às discussões, a presente obra *Agência, Deliberação e Motivação – Volume II* foi elaborada com o intuito de fornecer a especialistas, iniciantes e demais interessados no assunto a oportunidade de ler em nossa língua materna traduções de textos de renomados autores na área, acompanhados pela contribuição dos organizadores desta coletânea sobre o tema.

Temos, assim, um movimento de “mão dupla”. Por um lado, o *V CIFMP* se consolidou como um espaço profícuo para a discussão e a divulgação de resultados parciais de pesquisas de excelência realizadas por pesquisadores do Brasil em contato direto com outros centros de pesquisa em filosofia na América do Norte e Europa. Por outro lado, a tradução para o português de textos filosóficos na área proporcionará à comunidade filosófica brasileira – especialmente aos alunos de graduação e pós-graduação, que ainda sentem o desafio da leitura em outra língua – a possibilidade de acesso a temas relevantes, antes inacessíveis. Esperamos, dessa forma, ampliar o alcance dessas discussões, estabelecendo um fórum de debate no qual todos se sintam incluídos.

Este volume contém oito capítulos que representam discussões recentes sobre o tema da agência, deliberação ou motivação provindo de textos selecionados de revistas de filosofia

especializadas e traduzidos do inglês para o português, tomando como critério a relevância dos textos para as discussões na temática. Além deles, dois capítulos deste livro são resultado das pesquisas dos organizadores do presente volume.

Somos gratos às revistas *Philosophers' Imprint*, *Canadian Journal of Philosophy* e *Ergo Journal of Philosophy*, bem como aos autores dos textos – Brendan Cline (*Grand Valley State University*), David Copp (*University of California, Davis*), David Enoch (*The Hebrew University*), Kit Fine (*New York University*), Max Barkhausen (*New York University*) e Stephen Finlay (*University of Southern California*) pela permissão e incentivo à tradução e publicação de seus textos. Parcela importante deste trabalho foi desenvolvido pelos membros do *GOAT - Grupo de Análise e Tradução de Textos Filosóficos*, coordenado pelo professor Evandro Barbosa na UFPel. Gostaríamos de agradecer pelo minucioso e incansável trabalho desenvolvido nas traduções e discussões sobre a terminologia dos textos do inglês para o português, especialmente a Celina Alcântara Brod, Guilherme Meneghello, Gustavo Generaldo, Jade Arbo, Luca Igansi, Mariana Burkle e Thaís Alves Costa. Somos gratos também a Stephany Barbosa pela criteriosa revisão do português e a Carina Dalsoto, da Editora Fi, pelo incentivo e paciência no desenvolvimento deste trabalho."

Os organizadores

Sobre os autores

BRENDAN CLINE é *Florence Levy Kay Fellow* em Filosofia e Neurociência na *Brandeis University*. Suas áreas de pesquisa incluem discussões sobre Metaética e Psicologia Moral e sua pesquisa atual concentra-se em questões na interseção da metaética e da psicologia moral, incluindo a evolução da moralidade, a relação entre o julgamento normativo e a motivação e a semântica do pensamento normativo e do discurso. Suas publicações recentes incluem: “The Tale of a Moderate Normative Skeptic” (*Philosophical Studies*, 2018); “Smith’s Practicality Requirement Meets Dual-Process Models of Moral Judgment” (*Philosophical Psychology*, 2017); “Against Deliberative Indispensability as an Independent Guide to What There Is” (*Philosophical Studies*, 2016); “Nativism and the Evolutionary Debunking of Morality” (*Review of Philosophy and Psychology*, 2015).

DAVID COPP é *Distinguished Professor of Philosophy* na Universidade da Califórnia - *Campus Davis*. Ele lecionou anteriormente na Universidade Simon Fraser, na Universidade de Illinois em Chicago, na Universidade Estadual de Bowling Green (BGSU) e na Universidade da Flórida. David Copp foi presidente da Divisão do Pacífico da Associação Filosófica Americana (2014-2015) e foi professor visitante na Universidade de Stanford e na Universidade de Leiden, na Holanda. Já atuou como editor do *Canadian Journal of Philosophy* (1982-1992), como editor associado da renomada revista *Ethics* (1991-2008) e como editor da área de meta-ética da *Stanford Encyclopedia of Philosophy* (2003-2013). Atualmente, é editor da *Oxford Ethical Theory* da

Oxford University Press. É autor de *Morality, Normativity, and Society* (Oxford University Press, 1995) e *Morality in a Natural World* (Cambridge University Press, 2007) e editou várias antologias sobre ética e metaética, incluindo *The Oxford Handbook of Ethical Theory* (OUP, 2006).

DAVID ENOCH é *Rodney Blackman Chair* no Departamento de Filosofia na Universidade Hebraica de Jerusalém e co-diretor do *Center for Moral and Political Philosophy* na mesma instituição. Suas áreas de interesse incluem moral, política e filosofia do direito (tradição analítica). Ele é autor de *Taking Morality Seriously: A Defense of Robust Realism* (Oxford University Press, 2011); “Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Agency?” (*Philosophical Review*, 2006); “Reason-Giving and the Law” (*Oxford Studies in the Philosophy of Law* 1, 2011); “Moral Luck and the Law” (*Philosophy Compass* 5, 2010).

EVANDRO BARBOSA é professor adjunto no Departamento de Filosofia na *Universidade Federal de Pelotas*, atuando nos Programas de Graduação e Pós-Graduação em Filosofia. Desenvolveu estágio pós-doutoral na Universidade Federal de Pelotas (UFPel) com bolsa CAPES/Fapergs na área de metaética e foi visiting scholar (short term) na University of California, sob orientação de David Copp. Suas áreas de interesse incluem teorias da justiça (Rawls e Scanlon, metaética e motivação moral (Darwall), normatividade e naturalismo moral (Copp e Korsgaard). Discute também a relação entre ética normativa e metaética a partir da imbricação entre o método construtivista e a base contratualista no pensamento contemporâneo. Suas publicações incluem *Temas de Filosofia Política Contemporânea* (Organizou com K. Bresolin, 2017); *Direito e Moral em Kant* (2015); *Temas de Filosofia do Direito: Debates Contemporâneos* (Org. com C. A. Ferraz, 2014); *Ensaios de Filosofia Política: reflexões contemporâneas* ((Org. – com C. Andrade, 2013).

JOÃO HOBUSS é professor associado no Departamento de Filosofia na *Universidade Federal de Pelotas*. Foi Diretor de Pós-Graduação da Universidade Federal de Pelotas, coordenou o Mestrado em Filosofia da Universidade Federal de Pelotas, área de concentração em Ética e Filosofia Política, de 2009 a 2013. Atualmente é Diretor do Instituto de Filosofia, Sociologia e Política da UFPel e membro do GT Aristóteles/ANPOF. Tem experiência na área de Filosofia, com ênfase em História da Filosofia Antiga, atuando principalmente nos seguintes autores e temas: Aristóteles, estoicos, ética, política, a discussão contemporânea acerca da liberdade da vontade e responsabilidade moral. Foi o fundador e Editor da *Revista Dissertatio de Filosofia* - Qualis A2 (1995-2010) e é Editor-Chefe da Série *Dissertatio* de Filosofia. Publicou, dentre outros, os livros *Eudaimonia e auto-suficiência em Aristóteles* (2009), *Virtude e mediedade em Aristóteles*; (2009), *Ética das Virtudes* (Org., 2011), *Introdução à História da Filosofia Antiga* (2014) e *Sobre Responsabilidade* (Org., 2014).

KIT FINE (B.A., Oxford; Ph.D., Warwick) é *University Professor* e *Silver Professor* de Filosofia e Matemática na *New York University*. Ele é especialista em Metafísica, Lógica e Filosofia da Linguagem, membro da Academia Americana de Artes e Ciências, membro correspondente da Academia Britânica e ex-editor do *Journal of Symbolic Logic*. Além de suas principais áreas de pesquisa, ele escreveu artigos sobre filosofia antiga, linguística, ciência da computação e teoria econômica. Ele é autor de *Semantic Relationism* (Blackwell, 2007); *Modality and Tense: Philosophical Papers* (Oxford University Press, 2005); *The Limits of Abstraction* (Oxford University Press, 2002); *Reasoning With Arbitrary Objects* (Blackwell, 1986); *Worlds, Times, and Selves* - com A. N. Prior (University of Massachusetts Press, 1977).

MAX BARKHAUSEN terminou recentemente seu doutorado em Filosofia pela *New York University* com área de especialização em ética e metaética. Seus trabalhos incluem “Reductionist Moral Realism and the Contingency of Moral Evolution” (*Ethics*, 2016) e “Hume’s Correction of the Sentiments. Intersubjectivity without Objectivity.” (*Ergo*, 2016).

STEPHEN FINLAY é professor de Filosofia na *University of Southern California*. Sua pesquisa é focada em metaética e ética, especialmente a natureza da normatividade, relativismo de valor, raciocínio prático, psicologia moral e os fundamentos semânticos e metafísicos da ética. Sua produção acadêmica inclui *Confusion of Tongues: A Theory of Normative Language* (Oxford University Press, 2014); “Disagreement Lost and Found” (*Oxford Studies in Metaethics* 12, 2017); “Deontic Modality” – com M. Schroeder (Eds.), (*Pacific Philosophical Quarterly*, 2014).

Capítulo 1

Cognitivismo, motivação e abordagens de processo-duplo ao juízo normativo

Brendan Cline¹

Uma fonte central de embasamento para concepções expressivistas do discurso normativo é o relacionamento íntimo entre juízo normativo e motivação. Expressivistas argumentam que juízos normativos devem ser estados não cognitivos, análogos ao desejo para que possam ser tão intimamente conectados com a motivação. Declarações normativas são então colocadas como expressões destes estados não-cognitivos. Neste texto, me utilizarei de modelos de processo-duplo da psicologia cognitiva para responder a este argumento. De acordo com minha proposta, juízos normativos são crenças ordinárias que são tipicamente produzidas por dois tipos de processos: processos intuitivos-afetivos e raciocínio de domínio geral. Quando produzidos pelo primeiro tipo de processo, motivação e juízos tendem a se alinhar. Quando produzidos pelo segundo tipo, motivação e juízos podem não se alinhar. Já que o primeiro tipo de processo é o caminho mais comum para a formação de crenças normativas, juízos normativos são tipicamente acompanhados por motivação alinhada. Essa proposta permite que cognitivistas expliquem a ligação íntima entre juízo normativo e motivação, portanto

¹ Grand Valley State University. Tradução: Luca Igansi - Revisão: Jade Arbo.

removendo o maior obstáculo para interpretar declarações normativas em condição de verdade.

1. O Argumento da Motivação para o Expressivismo

Parece que a gramática superficial de declarações normativas apoia uma interpretação de condição de verdade (Glassen 1959). Locuções como “Isto está errado” e “Ele é mau” se parecem muito com asserções ordinárias, factuais feitas no modo indicativo tais como “Isto é azul” e “Ele está vivo.” E, tal como a observação célebre de Geach (1960; 1965), termos normativos parecem funcionar muito como predicados descritivos ordinários quando inseridos em sentenças complexas. Uma interpretação natural dessas (dentre outras) características de nosso discurso normativo é inequivocamente condicional de verdade: declarações normativas são asserções ordinárias sobre o papel normativo de atos, indivíduos ou estados-de-casos. Uma afirmação companheira natural é cognitivista: juízos normativos são crenças ordinárias sobre o papel normativo de atos, indivíduos ou estados-de-casos².

O internalismo sobre juízos normativos (doravante “internalismo”) provou-se um obstáculo persistente a este par natural de afirmações. Quando pessoas julgam que devem ϕ , são tipicamente motivadas a ϕ , e mudanças no juízo são normalmente acompanhadas por mudanças apropriadas na motivação. Também parece que há algo de equivocado sobre uma pessoa que mantém que deve ϕ mas não é nem um pouco motivada a ϕ . Estas observações são normalmente levantadas para apoiar alguma

² Utilizarei “declarações normativas” como um termo neutro para se referir aos atos de fala – independentemente de sua natureza – que envolvem elocuições sinceras de sentenças declarativas que contém termos normativos. “Juízos normativos” será um termo neutro para se referir aos estados mentais – independentemente de sua natureza – que são expressos por elocuições normativas. Eu contrasto perspectivas *apofânticas* com perspectivas *expressivistas* sobre o significado de declarações normativas, e contrasto perspectivas *cognitivistas* com perspectivas *não cognitivistas* sobre a natureza de juízos normativos. Esta terminologia útil é adotada de Wedgwood (2007.)

versão de internalismo – a visão que há uma conexão necessária entre juízos normativos e motivação. Uma versão irrestrita de internalismo afirma que, necessariamente, se uma agente julga que deve fazer ϕ em um contexto C, então ela está ao menos minimamente motivada a ϕ em C. Porém, é mais comum atualmente que internalistas apoiem versões mais fracas da perspectiva que afirma que esta conexão necessária é apenas obtida em agentes plenamente racionais, ou em condições normais, ou algum outro conjunto restrito de circunstâncias³.

O problema é que, de acordo com a perspectiva popular humeana da motivação, crenças ordinárias não podem nos motivar a agir por elas próprias. Na verdade, estados independentes, não cognitivos e análogos ao desejo são a fonte de nossa motivação, e crenças meio-fim meramente nos guiam a ações particulares identificando quais podem ser instrumentos efetivos para satisfazer nossos desejos. Mas se crenças são incapazes de nos mover para agir por elas próprias, então crenças normativas não poderiam nos mover a agir por elas próprias. Assim parece que cognitivistas estão comprometidos com um externalismo do juízo normativo (doravante “externalismo”) – a perspectiva que qualquer relacionamento entre juízo normativo e motivação é uma característica contingente de nossa psicologia. Pensa-se frequentemente que externalistas são incapazes de oferecer uma explicação das observações acima acerca da ligação íntima entre motivação e juízo normativo. Parece, portanto, que juízos normativos devem na verdade serem estados análogos ao desejo, não-cognitivos, para que possam manifestar o tipo de ligação íntima com a motivação observado acima.

De acordo com não cognitivistas, não é frutífero interpretar juízos normativos em termos de condições de verdade robustas, já que ao contrário das crenças ordinárias, os estados não cognitivos

³ Ver Björnsson, Strandberg, Francén Olinder, Eriksson e Borklund (2015b) para uma revisão das várias formas que o internalismo recentemente tomou.

que constituem juízos normativos não são robustamente aptos à verdade, e são melhores entendidos em termos de seus efeitos motivacionais. A partir disso, uma visão natural da semântica do discurso normativo é que os significados de declarações normativas devem ser explicados em termos dos estados mentais não cognitivos que expressam, da mesma forma que os significados de declarações descritivas podem ser explicados em termos das crenças que expressam. E já que declarações normativas expressam estados análogos ao desejo, não cognitivos, são eles também não frutiferamente interpretados em termos de condições de verdade robustas. Este é o *argumento da motivação para o expressivismo*⁴.

Neste texto, me utilizarei de recentes trabalhos na área da psicologia cognitiva para desenvolver uma resposta cognitiva para este argumento. Em particular, mostrarei como uma abordagem de processo-duplo ao juízo normativo permite que externalistas cognitivistas expliquem o relacionamento íntimo entre motivação e juízo enquanto aceitam que crenças por si são motivacionalmente inertes. De acordo com minha proposta, juízos normativos são crenças (robustas, aptas à verdade) que são tipicamente produzidas por dois tipos distintos de processos cognitivos. Por um lado, crenças normativas podem ser o produto de processos rápidos, involuntários, motivacionalmente eficazes e intuitivo-afetivos. Por outro, crenças normativas podem ser o produto de

⁴ O foco central de minha discussão será em teorias expressivistas “puras” (e.g. Blackburn 1984; 1993; 1998; Gibbard 1990; 2003), embora eu brevemente discuta teorias híbridas na seção 6 abaixo. O relacionamento entre as afirmações centrais envolvidas neste argumento é explorado em detalhe em Smith (1994). Note que nem todas as formas de internalismo apoiam o não-cognitvismo igualmente bem. Versões irrestritas do internalismo proveem o apoio mais forte, embora Gibbard e Blackburn tenham apelado a versões restritas desta perspectiva. Por exemplo, Gibbard (1993: 318-319) mantém que juízos normativos genuínos tipicamente nos motivam em circunstâncias normais. Apesar desta afirmação mais fraca, a frequência relativa do alinhamento entre juízo normativo e motivação ainda engendra um enigma para cognitivistas que aceitam a teoria humeana da motivação, e é facilmente explicada por não-cognitivistas. Se o internalismo torna-se muito enfraquecido, no entanto, então não é mais muito útil para depender o não-cognitvismo (e.g. Smith 1994; Tresan 2006.)

raciocínio lento, deliberado, explícito e consciente. Já que o primeiro tipo de processo é o caminho padrão para a formação de crenças normativas, há uma correlação muito forte entre acreditar que deve-se ϕ e que se está motivado à ϕ . A explicação para esta correlação, no entanto, tem a ver com os mecanismos tipicamente envolvidos na produção de juízos normativos, e não com a natureza de tais juízos. Esta estrutura permite aos externalistas cognitivistas explicar todos os fenômenos relevantes ao juízo e à motivação, bloqueando então o argumento da motivação para o expressivismo.

Início a seção 2 caracterizando modelos de processo-duplo na psicologia cognitiva e sugerindo uma abordagem para aplicá-los à psicologia do juízo normativo. Na seção 3, ilustrarei esta abordagem mostrando como a estrutura do processo-duplo pode permitir que cognitivistas desenvolvam uma versão convincente de externalismo motivacional sobre juízos de perigo. Na seção 4, estendo esta estrutura para juízos normativos e apresento algumas evidências para suporte. Na seção 5 argumento que minha proposta é, de forma explicativa, superior àquelas oferecidas por não-cognitivistas, e na seção 6 argumento que este resultado deixa o expressivismo em maus lençóis. Na seção 7 defendo a metodologia *a posteriori* aplicada em meu argumento. O resultado é uma nova abordagem ao externalismo cognitivista que solapa a sabedoria comum acerca dos limites explanatórios da perspectiva.

2. Modelos de Processo-Duplo e Juízo Normativo

Modelos de processo-duplo na psicologia postulam dois tipos distintos de processos cognitivos para explicar vários fenômenos psicológicos (e.g. Evans & Frankish 2009; Kahneman 2011; Sherman, Gawronski, & Trope 2014)⁵. Enquanto detalhes variam

⁵ Para nossos propósitos, não é significativo que a palavra “cognitivo” é utilizada para caracterizar estes diferentes processos. Neste contexto filosófico, “cognitivo” significa, a grosso modo, “mental”. Isto é ortogonal à distinção filosófica entre estados mentais cognitivos (i.e. análogos à crença) e não-

entre diferentes perspectivas, é comum distinguir entre processos que são (1) rápidos, inconscientes, involuntários, sem esforço e independente da memória de trabalho, e aqueles que são (2) lentos, conscientes, deliberativos, difíceis e dependentes da memória de trabalho. A nomenclatura de “Tipo 1” e “Tipo 2” são normalmente utilizadas para marcar a distinção entre o primeiro e segundo tipos de processo⁶.

Uma característica chave dos processos de Tipo 1 é que suas operações tomam forma abaixo do limiar da consciência. É o resultado de processos de Tipo 1, na verdade dos passos dados dentro deles, que são disponíveis à consciência. Pensa-se que estes processos são responsáveis por um amplo panorama de pensamentos que podemos ter, como reconhecer que uma sentença é não-gramatical, ter uma boa sensação sobre ações na bolsa de valores, reconhecer um rosto familiar ou perceber que uma parceira conversacional contou uma piada, para mencionar apenas alguns. Em vários casos, pensamentos sobre estes temas apenas “vêm à mente.” E enquanto frequentemente confiamos nestes pensamentos, não podemos explicar facilmente como neles chegamos – nós “apenas vemos” a resposta. Tipicamente, não experienciamos nós mesmos como voluntariamente exercendo esforço ou exercendo agência quando chegamos a conclusões desta forma. Na verdade, os processos que regurgitam estas conclusões frequentemente operam involuntariamente e automaticamente, de

cognitivos (i.e. análogos ao desejo). Isto é simplesmente uma consequência infeliz da terminologia na qual este debate foi situado. Obrigado à Tongli Wu por me pressionar a clarificar este tema.

⁶ Ocasionalmente, prosseguem-se afirmações sobre sistemas integrados compostos de processos do Tipo 1 e do Tipo 2. Aqui, podemos deixar de lado os problemas envolvidos com estas afirmações. Estou de acordo com Evans e Stanovitch (2013) que a característica central de processos do Tipo 1 é sua automaticidade, e a característica central dos processos do Tipo 2 é sua dependência na memória de trabalho. Os outros traços associados com diferentes tipos de processamento são provavelmente melhor vistos como correlatos comuns, e não como critérios necessários. Embora minha caracterização de processos de Tipo 1 e Tipo 2 destoe de algumas formas dos trabalhos de outros (e.g. Mercier & Sperber 2009; 2011), acredito que em maior parte, seria relativamente fácil traduzir minha proposta em formas que são mais amenas para interpretações alternativas dos dois tipos de processos.

alguma forma semelhante à forma que nossas pupilas involuntariamente e automaticamente retraem-se quando uma lanterna é apontada a elas e dilatam quando é afastada. Isto não é dizer que não temos controle sobre estes processos, claro. Por exemplo, podemos dirigir a atenção em várias formas para “liberá-los” sobre problemas particulares, de uma maneira semelhante a como podemos voluntariamente mover uma lanterna para controlar os padrões de constrição em um olho saudável.

Em contraste, as operações de processos do Tipo 2 são geralmente disponíveis à atenção consciente, e requerem esforço para que aconteça. A disparidade é fácil de ilustrar com problemas aritméticos (Kahneman 2011: 20-21). Considere o problema: $2 + 1 = \underline{\quad}$. É bastante provável que você possa imediatamente “apenas ver” a solução. Agora considere um diferente problema: $53 \times 34 = \underline{\quad}$. Se você é capaz de encontrar a solução para isto, muito provavelmente não envolverá “apenas ver” de imediato a solução. Na verdade, você precisará exercer esforço em arranjar algumas linhas de problemas mais simples, mantendo o resultado de cada subproblema na memória de trabalho, e então integrando-as para uma solução. De fato, talvez você sentiu aversão para investir o esforço, e nem se preocupou em achar a solução. Mas se você o fez, você estaria ciente de cada passo no caminho, e poderia possivelmente recontar cada um dos passos se pedissem a você. Você se sentiria como se estivesse exercendo sua volição em planejar como aproximar-se do problema e enquanto executava o plano.

Processos de Tipo 2 não são restritos à aritmética. De fato, todos os exemplos do Tipo 1 mencionados acima podem possuir contrapartes do Tipo 2. Por exemplo, embora normalmente reconheçamos nossa própria face imediatamente ao olhar nosso reflexo, um paciente com dano ao seu giro fusiforme não tem esse luxo e pode na verdade precisar explicitamente pensar para si mesmo: “há uma face naquele espelho, e ninguém está nesta sala comigo, então esta deve ser minha face no espelho”. É tipicamente

pensado que processos de Tipo 2 são mais flexíveis, e em grande parte envolvem mecanismos de domínio geral que mantêm e manipula diferentes tipos de informação na memória de trabalho⁷. No entanto, pessoas geralmente parecem ser contrárias a investir o esforço associado com estes processos e preferem evitá-lo. É, portanto, pensado que processos de Tipo 2 são primariamente recrutados quando se espera recompensa pelo esforço, por exemplos em casos onde processos de Tipo 1 geram resultados conflitantes, ou quando há muito em jogo (Braver & Botvinick 2015; Shenhav, Botvinick, & Cohen 2013.) Como um resultado, modelos de processo-duplo frequentemente tratam a atividade rápida, contínua e especializada dos processos de Tipo 1 como o caminho “padrão”, e processos de Tipo 2 como exerceedores de supervisão sobre processos de Tipo 1 e ocasionalmente intervindo quando necessário (e.g. Evans & Stanovitch 2013; Kahneman 2011).

A velocidade e falta de esforço de processos de Tipo 1 se dá em parte à sua especificidade de domínio, e processos de Tipo 1 especializados através de diferentes domínios possuem distintas características que dependem da sua implementação e conteúdo proprietário. Por exemplo, computações realizadas pelos mecanismos envolvidos em reconhecimento facial são claramente distintas daquelas realizadas pelos mecanismos envolvidos na compreensão da linguagem. Isto significa que modelos de processo-duplo de diferentes domínios irão preencher alguns dos detalhes relevantes de maneiras diferentes. Por exemplo, recentes aplicações da estrutura de processo-duplo à psicologia do juízo moral sugerem que os mecanismos envolvidos no caminho de Tipo 1 são implementados por circuitos que são envolvidos na motivação e estímulo afetivo (e.g. Cushman, Young, & Greene 2010; Greene 2014a; Haidt 2001). Estudos de imageamento têm

⁷ Vide Chein, Moore e Conway (2011), Fedorenko, Duncan, e Kanwisher (2013) e Barbey, Koenigs, and Grafman (2013). Cf. Mercier e Sperber (2009; 2011).

associado estruturas tais como o córtex pré-frontal ventromedial (CPFvm) e a amígdala com certos padrões de juízo moral característicos de processos de Tipo 1 (Greene, Nystrom, & Engell 2004; Greene, Sommerville, Nystom, & Cohen 2001; Shenhav & Greene 2014), e foi encontrado que danos à estas estruturas perturbam este padrão (Ciaramelli, Muccioli, Iàdavas, & di Pellegrino 2007; Koenigs et al. 2007). Os mecanismos envolvidos no pensamento moral de Tipo 2, por outro lado, não são similarmente conectados ao afeto e motivação, mas na verdade parecem ser constituídos por mecanismo de controle de domínio geral. E de fato, estudos de imageamento associaram estruturas envolvidas na memória de trabalho e controle cognitivo – tal como o córtex pré-frontal dorsolateral (CPFdl) e córtex cingular anterior (CCA) – com padrões distintos de juízo moral, e tarefas de carga cognitiva que perturbam a memória de trabalho perturbam este padrão (Greene, Morelli, Lowenberg, Nystrom, & Cohen 2008; ver também Paxton, Ungar, & Greene 2012).

Filósofos recentemente dedicaram muita atenção para a aplicação da estrutura de processo-duplo de psicólogos ao juízo moral. No entanto, parece menos frequentemente notado que esta aproximação foi altamente inspirada pelo sucesso da estrutura em explicar o pensamento avaliativo ordinário e processos de decisão de forma mais geral (ver, e.g., Bargh & Chartrand 1999; Haidt 2001: 815; Zajonc 1980). E, de fato, parece agora que muitos dos mecanismos centrais neurais implicados no pensamento moral são em sua maioria os mesmos implicados no pensamento avaliativo geral e em processos de tomada de decisão (Fehr & Krajbich 2014; Greene 2014b; Shenhav & Greene 2010). Então enquanto cientistas que estudam o pensamento avaliativo e de tomada de decisão de uma perspectiva de processo-duplo nem sempre explicitamente estudam coisas como juízos de dever, parece plausível que a estrutura ainda assim é aplicada apropriadamente ao pensamento normativo em geral. No que se seguirá, partirei deste pressuposto em minha discussão da conexão entre motivação e juízos

normativos, embora reconheça que será importante para a pesquisa futura engajar mais explicitamente com, por exemplo, pensamentos acerca do que se deve fazer, para que a relevância desta perspectiva nas questões filosóficas tratadas aqui esteja estabelecida.⁸

Antes de seguir, gostaria de brevemente qualificar minha discussão abaixo. Primeiramente, é importante notar que minha proposta é completamente compatível com a noção que processos do Tipo 1 sejam capazes de aprender baseados na experiência (Horgan & Timmons 2007; Wielenberg 2014: cap. 3) e interação continuamente com processos de Tipo 2 (Van Bavel, Xiao, & Cunningham 2012). De forma mais importante, pretendo manter um nível relativamente alto de generalidade na proposta básica que desenvolverei. Pretendo que minha proposta seja algo como um esquema de explicação que é compatível com um número de formas distintas nas quais pode-se preencher os detalhes relevantes sobre a natureza do desejo, o relacionamento entre emoção e motivação, e assim por diante. Na seção 4 sugiro uma hipótese acerca de como aspectos de minha proposta podem ser implementadas em cérebros humanos, e apresento algumas evidências a favor disso. Mas é importante manter em mente que minha proposta principal não é contingente à esta hipótese em particular.

Com estas qualificações colocadas, refiro-me à próxima seção para ilustrar minha proposta geral ao examinar a conexão entre juízo e motivação no domínio dos juízos de perigo.

⁸ Não penso que esta presunção seja injustamente hostil às perspectivas que critico. Desde que os não-cognitivistas insistem na conexão íntima entre motivação e juízo normativo, parece que estarão inclinados a concordar que juízos normativos estão no cerne daquilo que é estudado pela ciência da avaliação e tomada de decisão. Esta inexistência das questões explícitas feitas por pesquisadores não tornariam seu trabalho irrelevante à psicologia do juízo normativo.

3. Juízos de Perigo e Motivação

Parto de que o perigo de uma entidade ou atividade é um fato ordinário, “prosaico”, que pode ser mais ou menos corretamente ou incorretamente representado por crenças ordinárias.⁹ A semântica de elocuções como “Isto é perigoso” se encaixa no modelo padrão de condição de verdade, e atribuem a propriedade de perigo ao referente da declaração¹⁰. Como formamos crenças sobre quais entidades ou atividades são perigosas? O paradigma do processo-duplo oferece uma perspectiva plausível (Slovic, Peters, Finucane, & Macgregor 2005).

Em alguns casos, processos de Tipo 2 são claramente responsáveis. Organizações como o Centros de Controle e Prevenção de Doenças, por exemplo, regularmente se engaja em sofisticados recolhimentos de dados e análise estatística para determinar que entidades e atividades são perigosas ao público, e o quão perigosas são. Similarmente, é possível para um indivíduo formar juízos de perigo baseado em um cuidadoso pesar de evidências contidas em relatórios publicados por tais organizações. Em muitas situações, no entanto, crenças de perigo são produzidas por processos de Tipo 1 intuitivo-afetivos. Uma jovem criança, por exemplo, pode vir a acreditar que fogareiros quentes são perigosos

⁹ Uma posição plausível sobre o perigo pode ser algo em termos de se ter uma disposição a causar dano. McDowell (1988: 175–177) parece pensar da propriedade do perigo em termos normativos, como a propriedade que exige respostas de medo. No entanto, se a periculosidade de algo exige uma resposta de medo ou não é um outro fato “acima e além” da periculosidade da coisa. Medo não é sempre uma resposta apropriada ao perigo, e a periculosidade de coisas afiadas não depende de nossas razões para temê-las. Portanto, penso que o perigo é melhor pensado como um fato ordinário, “prosaico”. Porém, mesmo que se prefira pensar do perigo em termos normativos, juízos e declarações de perigo ao menos na superfície parecem ser crenças e asserções ordinárias, e a discussão abaixo ilustra como a motivação que acompanha estes juízos pode ser tomada dentro de uma perspectiva cognitivista humeana.

¹⁰ Acredito que asserções contendo a palavra “perigoso” podem ser utilizadas para se expressar medo, ou instruir-se um locutor a comportar-se de certas formas. Mas estes casos podem ser interpretados de forma útil em termos pragmáticos, da mesma forma que asserir que um objeto é caro pode servir para expressar-se preocupação sobre o objeto em questão, e dirigir-se à uma audiência para tratar o objeto com cuidado.

imediatamente após queimar sua mão em um. De forma semelhante, a saliência afetiva de alguma coisa pode influenciar as crenças de pessoas sobre o quão perigosa ela é. Em casos como estes, um estado motivacionalmente análogo ao desejo eficaz (e.g. dor, medo) precede e diretamente contribui à formação da crença de perigo.¹¹

De acordo com esta abordagem, então, o juízo que algo é perigoso é uma crença que pode ser produzida por dois tipos distintos de processo. Uma vez que tais crenças são formadas, são tornadas disponíveis para a memória de trabalho e podem ser envolvidas no raciocínio da mesma forma que outras crenças ordinárias podem. A semântica de declarações de perigo parece de fato relevantes às escolhas que pessoas fazem. As pessoas são tipicamente ao menos de alguma forma motivadas a afetar coisas que elas têm como perigosas, e seria uma personagem esquisita quem sistematicamente negasse todos os seus juízos de perigo em sua tomada de decisão. Este seria um problema sério com a afirmação que juízos de perigo são crenças? Será que talvez fosse melhor considerar perseguir uma semântica expressivista para declarações de perigo? Penso que não. A perspectiva de processo-duplo nos permite entender a motivação associada com juízos de perigo de uma perspectiva cognitivista, mesmo com os limites do paradigma humeano da motivação.

Se crenças sobre o que é e o que não é perigoso são tipicamente geradas ou ativadas por processos de Tipo 1 intuitivo-afetivos motivacionalmente eficazes, então podemos dar uma explicação cognitivista das razões pelas quais motivações

¹¹ Não defendo que é impossível que outros processos de Tipo 1 resultem no *tokening* de uma crença de perigo. Por exemplo, o recordar da memória de longo prazo de uma crença de perigo pode envolver processos de Tipo 1, e o juízo recordado pode ter sido inicialmente via processos de Tipo 2. Para ajudar a clarificar, utilizarei a frase “intuitivo-afetivo” para eliciar os subtipos de processos de Tipo 1 onde estados motivacionalmente eficazes estão envolvidos no surgimento de pensamentos conscientes que são *outputs* do processo. Irei ocasionalmente utilizar este qualificador para trazer atenção a este caráter preciso. Mas exceto onde notado, quando eu discutir o caminho de Tipo 1 da formação de juízos normativos abaixo, me referirei à processos de Tipo 1 intuitivo-afetivos.

específicas tendem a acompanhar crenças sobre algo ser ou não perigoso. Se minha crença que *x* é perigoso foi produzida ou ativada por processos de Tipo 1 intuitivo-afetivos, então serei motivado a evitar *x* ou proceder com cuidado quando lidar com *x*. Isso é por que estados motivacionalmente eficazes estavam envolvidos na produção desta crença. E já que os caminhos de Tipo 1 intuitivo-afetivos são os processos-padrão mais comuns envolvidos na geração de crenças sobre perigo, as pessoas tipicamente são motivadas a agir de maneiras que se alinhem à suas crenças sobre o quão perigosas as coisas são. (Esta motivação será modulada pelo nível de aversão de risco do indivíduo, é claro.)¹²

Ao mesmo tempo, seria implausivelmente forte afirmar que pessoas são *sempre* (mesmo que apenas um pouco) motivadas a agir de acordo com seus juízos sobre o que é e o que não é perigoso. Por exemplo, algumas pessoas têm muito medo de viajar em aviões e recusam a fazê-lo, apesar de sinceramente acreditar que viajar em aviões não é muito perigoso. Na outra direção, algumas pessoas são desinibidas sobre engajar-se em atividades – como fumar ou dirigir um carro – que sabem ser perigosas. Qualquer perspectiva plausível da psicologia de juízos de perigo precisa dar conta de tais fenômenos.¹³ Afortunadamente, o paradigma de processo-duplo nos oferece recursos para fazê-lo.

Quando a formação de um juízo de perigo não engaja com processos de Tipo 1 intuitivo-afetivos, pode-se criar a crença que *x*

¹² Existem fenômenos próximos que são relevantes também. Se acredito que é provável que algo cause a minha morte, e desejo continuar vivendo, então serei motivado a evitar aquela coisa. Plausivelmente, tais desejos precisamente são o que normalmente dão origem a nossos juízos de perigo. No entanto, estes tipos de casos não apresentam enigmas sobre motivação para o humano. Os casos-problema são aqueles que certas classes de juízo parecem ser ambos análogos à crença e suficientes para a motivação. É por isto que juízos de perigo apresentam um enigma tão interessante para humanos motivacionais. Agradeço à James Beebe por me pressionar neste ponto.

¹³ Muitos destes tipos de casos simplesmente envolverão motivações conflitantes, onde alguma motivação *pro tanto* a voar ou evitar fumar simplesmente é sobreposta por desejos conflitantes mais fortes. No entanto, é também plausível que ao menos ocasionalmente, um agente faça juízos de perigo genuínos apesar de não ter absolutamente nenhuma motivação acompanhante.

é (não é) perigoso sem ter nenhuma motivação correspondente apropriada à (falta de) periculosidade de x. De fato, o que é comum aos exemplos de incompatibilidade entre juízo e motivação que acabei de descrever é que são todos normalmente casos onde as pessoas dão à si mesmas (através de processos de Tipo 2) razões para conclusões acerca da periculosidade de algo, mas falham em engajar à processos relevantes de Tipo 1. Alguém que lê o texto de avisos sobre os perigos de fumar, por exemplo, pode ser menos inibido sobre isto do que alguém que é repetidamente exposto a imagens gráficas dos efeitos negativos de se fumar (Borland et al. 2009). De forma semelhante, pensar sobre a comparação de segurança entre andar de avião e andar de carro normalmente não ajudará muito uma pessoa que possui uma forte fobia de voar, mas não de dirigir. No entanto, se você colocar tal pessoa em um avião vazio no chão e balançá-lo de forma a associar a turbulência com um ambiente seguro, você pode ter algum sucesso em gradualmente mudar sua disposição de voar (Hajek 2016).

Esta perspectiva não implica que seja impossível que processos de Tipo 2 influenciem a motivação, é claro. Enquanto concordo com a tese humeana que a crença por si só é motivacionalmente inerte, é possível voluntariamente reenquadrar uma situação ou direcionar a atenção de tal forma que operações de processos de Tipo 1 mudem a motivação de alguém. Por exemplo, você pode intencionalmente se sujeitar a imagens de pessoas que estão visivelmente sofrendo dos efeitos de doenças causadas pelo fumo. Semelhantemente, se você experiencia terror à três quilômetros de altura você pode não ser capaz de interrompê-lo voluntariamente. Mas você pode fechar seus olhos e se imaginar em algum outro lugar agradável, ou tentar imaginar que você está de fato no chão e as pessoas estão apenas balançando o avião – oferecendo um novo *input* para um processo involuntário Tipo 1. Estes tipos de intervenções podem agir de uma forma análoga a mover uma lanterna com suas mãos de forma a controlar a dilatação e constrição involuntárias de sua pupila.

A crença de alguém sobre o perigo de algo, então, é insuficiente para determinar sua motivação. O que determina sua motivação em última instância é a fonte daquela crença. Se é produzida por processos de Tipo 1 intuitivo-afetivos, é provável que a motivação e juízo sejam compatíveis. Se é produzida em maior parte por processos de raciocínio de Tipo 2 sem o envolvimento de processos de Tipo 1 intuitivo-afetivos, é possível que motivação e juízo não se alinharem. O ponto importante a se manter em mente é que é nos processos que são responsáveis pela formação do juízo, ao invés de no próprio juízo, que se encontra a história sobre motivação.

4. Estendendo a Abordagem ao Juízo Normativo

Mantenho que uma história similar à contada acima acerca dos juízos de perigo pode ser contada sobre juízos normativos. A primeira coisa a se notar é que, se correto, o rascunho de juízos de perigo da seção anterior nos mostra que há um mecanismo que transforma respostas intuitivo-afetivas em crenças sobre coisas no mundo, e produtos deste mecanismo são disponibilizados à memória de trabalho para o raciocínio de domínio geral. Se aceitarmos esta história, então já estamos comprometidos com tal coisa ocorrendo em nossas mentes ao menos em alguns casos. Minha proposta agora é que um mecanismo similar está envolvido em nossos juízos normativos da mesma forma.

Como pode tal mecanismo ser implementado no cérebro? Pesquisas recentes gerara um número de resultados sugestivos. Parece agora que sinais transmitidos de (a) circuitos envolvidos em aprendizado de recompensas e seleção de ação (e.g. núcleos da base, núcleos dopaminérgicos do mesencéfalo) e (b) circuitos envolvidos no estímulo emocional e motivação (e.g. a amígdala, hipotálamo, lobo da ínsula anterior) ao (c) CPFvm são chave para entender o caminho de Tipo 1 intuitivo-afetivo da formação do juízo normativo (para detalhes, ver Camerer, & Montague 2008;

Craig 2009; Haber & Knutson 2010; Hare, Camerer, Knoepfle, O'Doherty, & Rangel 2010; Rangel & Clithero 2014; Rangel, Railton 2014; Sanfey, rilling, Aronson, Nystrom, & Cohen 2003; Smith et al. 2010). Parece que uma vez que esta informação alcance (c) o CPFvm, é integrada em pensamentos explícitos normativos que podem ser disponibilizados à atenção consciente e memória de trabalho, e podem ser verbalmente expressas por declarações normativas (vide Basten, Biele, Heekeren, & Fiebackh 2010; Roy, Shohamy, & Wager 2012; Shenhav & Greene 2014). Há motivos para se pensar que nossos estados análogos ao desejo motivacionalmente eficazes são realizados pelos circuitos (a) e (b) (Arpaly & Schroeder 2014: capítulo 6; Schroeder 2004; Schroeder, Roskies, & Nichols 2010). Consistente com esta perspectiva, pesquisa em imageamento sugere que estas estruturas continuamente e automaticamente avaliam ações potenciais e de fato e estados-de-casos (Lebreton, Jorge, Michel, Thirion, & Pessiglione 2009), e que comportamento futuro pode ser previsto na base da atividade automatizada nestas estruturas (Abe & Greene 2014; Haruno & Frith 2010; Tusche, Bode, & Haynes 2010). Estas estruturas (a) e (b) parecem então simultaneamente (i) regular o afeto e guiar a seleção de ações, e (ii) gerar pensamentos normativos conscientemente acessíveis. Minha sugestão é que é o fato de que estados mentais que ocorrem nos circuitos (a) e (b) são causas comuns de (i) e (ii) que explica a conexão íntima entre crença normativa e motivação.

Ao mesmo tempo, pesquisas apontam que avaliações e escolhas são ocasionalmente influenciadas por (d) estruturas envolvidas na memória de trabalho e controle cognitivo (como o CPFdl e o CCA). Parece que estes circuitos constituem distintos sistemas de controle que podem competir ou influenciar a atividade dos circuitos (a)–(c). Por exemplo, propositalmente se focar no quão saudáveis são suas opções de comida pode influenciar as escolhas de alimentação, e este efeito parece ser mediado pela atividade no CPFdl ao modular atividade no CPFvm (Hare, Malmaud, & Rangel 2011;

vide também Baumgartner, Knoch, Hot, Eisenegger, & Fehr 2011; Botvinick & Braver 2015; Hare et al. 2009; Hutcherson, Plassmann, Gross, & Rangel 2012; Ochsner & Gross 2005; Plassmann, O'Doherty, Shiv, & Rangel 2008; Rudorf & Hare 2014; Shenhav et al. 2013; Sokol-Hessner, Hutcherson, Hare, & Rangel 2012). Estas mudanças podem ser transmitidas de volta aos circuitos (a) e (b), assim influenciando-se a motivação (Pujara & Koenigs 2014). Este tipo de efeito é análogo aos meus exemplos iniciais de intencionalmente se apontar uma lanterna ao olho para controlar a constrição da pupila, ou intencionalmente imaginar um lugar agradável e seguro para mitigar-se o medo de voar. Minha interpretação preferida destas “intervenções de lanterna” é que a atividade esforçada de processos de Tipo 2 oferece novos estímulos para avaliação por processos de Tipo 1 intuitivo-afetivos.

No entanto, embora seja possível para a ativação esforçada de circuitos (d) para influenciar a motivação e o comportamento, parece que estes efeitos são indiretos, e que os mecanismos envolvidos em raciocínio consciente e memória de trabalho são indissociáveis daqueles mais diretamente envolvidos em memória de trabalho de domínio-geral – tal como o CPFdl – não parecem possuir nenhum relacionamento especial com a motivação, já que estão envolvidos em manipular informação na memória de trabalho através de um número de domínios (Barbey et al. 2013; Botvinick & Braver 2015; Chein et al. 2011; Fedorenko et al. 2013). Além do mais, onde estudos de pensamento avaliativo normalmente implicam circuitos (a)–(c), encontrou-se que circuitos (d) são apenas ocasionalmente ativados. Isto sugere que processamentos do Tipo 2 não são necessários para a formação de juízos avaliativos e decisões, mas ao invés uma camada extra que pode ser recrutada quando necessário (Haber & Knutson 2010).

Muitos agora aceitam que, ocasionalmente, pessoas falham em ser motivadas a agir de acordo com seus juízos normativos. E sugiro que processos de Tipo 2 implementados no CPFdl também podem ser capazes de dar conta destes casos. Através do raciocínio explícito,

um agente pode chegar à conclusão sobre o que ela deve fazer, e ainda assim isto pode ser motivacionalmente epifenomenal, uma vez que atividade no CPFdl pode não interagir com ou influenciar circuitos envolvidos na motivação e no desejo afetivo.

Evidência para a existência desta atividade motivacionalmente epifenomenal de Tipo 2 vem de estudos de pacientes com danos no CPFvm. Um estudo icônico observou um grupo de pacientes de CPFvm que manifestavam uma inabilidade de integrar seus pensamentos explícitos com estratégias de decisão otimizada em uma tarefa de aposta com sua performance naquela tarefa (Bechara, Damasio, Tranel, & Damasio 1997). Isto foi em contraste com sujeitos saudáveis que eram capazes de aprender e implementar estratégias otimizadas antes de se tornarem capazes de explicitamente articular a estratégia ou explicar as razões de ser otimizada. De fato, estratégias otimizadas estavam sendo aplicadas mesmo por sujeitos saudáveis que nunca se tornaram capazes de explicitamente declarar os méritos da estratégia. Parece que sujeitos saudáveis poderiam apenas “imediatamente ver” a melhor estratégia via processos de Tipo 1 intuitivo-afetivos, e isto explica as razões de suas motivações terem se alinhado aos seus juízos sobre qual estratégia era melhor. Por outro lado, parece que os pacientes de CPFvm podem ter explicitamente raciocinado sobre a tarefa, formado uma crença sobre qual estratégia era melhor, e ainda assim terem continuado a perseguir a estratégia inferior por seu raciocínio explícito (de Tipo 2) não engajou com o circuito motivacional subjacente do caminho de Tipo 1.

Estudos subsequentes de pacientes de CPFvm continuaram a documentar notáveis anomalias no comportamento em jogos experimentais (e.g., Krajcich, Adolphs, Tranel, & Denburg 2009; Koenigs & Tranel 2007) e em respostas à cenários afetivamente carregados (Koenigs et al. 2007). No entanto, investigações de atitude explícitas de pacientes de CPFvm sobre cenários de escolha hipotéticos e abstratos encontrou sobreposição considerável com controles saudáveis (Leland & Grafman 2005). Adicionalmente,

pacientes de CPFvm são frequentemente indistinguíveis de sujeitos saudáveis em uma gama de medidas de inteligência e habilidade cognitiva em geral, sugerindo que memória declarativa e memória de trabalho são preservadas (Damasio 1994). Não parece que o problema com pacientes de CPFvm seja que são incapazes de atuar emocionalmente e realizar ações motivadas, nem que são incapazes de explicitamente raciocinar sobre assuntos práticos, mas que na verdade que estes dois aspectos de sua psicologia não consigam interagir fluentemente. O dano ao CPFvm parece irromper a conexão entre o raciocínio consciente explícito e motivação. Considerações como estas apoiam a interpretação do CPFvm como uma estação intermediária para a transmissão entre os processos de Tipo 1 envolvidos na motivação e afeição e os processos de Tipo 2 envolvidos no raciocínio normativo (vide Roy et al. 2012).¹⁴

Assim como com os juízos de perigo, no entanto, seria implausível manter que juízos normativos são *sempre* acompanhados de (mesmo alguma) motivação respectiva. E novamente, a perspectiva atual oferece uma explicação natural para casos de incongruência motivação-juízo. Em particular, juízos normativos que são produzidos por raciocínios explícitos e conscientes de Tipo 2 podem não se alinhar com a motivação. A explicação para isto é que crenças por si só são motivacionalmente inertes, e processos de Tipo 2 não são implementadas por circuitos diretamente envolvidos na geração de afeto e motivação. Quando o raciocínio e a deliberação de fato influenciam a motivação, não é por que crenças sobre o que se deve fazer possuem uma força motivacional especial por parte de sua natureza essencial. Na verdade, ocorre por causa das formas nas quais os vários conteúdos de nossos pensamentos conscientes (tais como o quão saudáveis são nossas opções de comida) interagem com nossos

¹⁴ Note que não estou afirmando que pacientes de CPFvm são contraexemplos ao internalismo (cf. Roskies 2003). Ao invés, uso estes casos para apoiar a afirmação que os mecanismos ordinários que subjazem o pensamento normativo de Tipo 2 são dissociáveis daqueles que subjazem o pensamento normativo de Tipo 1, e estes mecanismos possuem diferentes perfis motivacionais.

desejos e disposições afetivas quando os conteúdos destes pensamentos são sujeitos à avaliação de Tipo 1 intuitivo-afetiva.¹⁵

O ponto importante para nossos propósitos é o nível abstrato de explicação dado nos últimos dois parágrafos. A afirmação principal é que estados análogos ao desejo motivacionalmente eficazes são a fonte típica de crenças normativas, e esta é a explicação básica para a conexão estreita entre motivação e juízos. Se algo relevantemente semelhante a este esquema-explicação funciona, então cognitivistas podem explicar a conexão íntima entre motivação e juízo normativo dentro de uma perspectiva humeana. E, presentemente, parece que há relativamente boas evidências para esta ideia geral.¹⁶

5. Não-Cognitivistas Fazem Melhor?

Como vimos na seção 1, o argumento para a motivação no expressivismo jaz na afirmação que, em luz da perspectiva humeana da motivação, apenas não-cognitivistas podem adequadamente explicar a relação íntima entre juízos normativos e

¹⁵ O foco principal desta discussão é em juízos de primeira-pessoa da forma “eu devo ϕ ”. O motivo para isto é que estes são juízos que tipicamente aparecem sob o escopo de teses internalistas. No entanto, pensa-se que modelos de processo-duplo são aplicáveis à todos os pensamentos avaliativos que se pode ter, como as rápidas impressões que formamos sobre pessoas que recém conhecemos (Zajonc 1980). Normalmente, avaliações e motivações associadas – e.g. confiar ou não em alguém – ocorrem rapidamente e sem a necessidade para o raciocínio explícito, embora seja possível explicitamente raciocinar sobre tais coisas como confiar ou não em um recém conhecido.

¹⁶ Recentemente, Sinhababu (2017: capítulo 4) independentemente desenvolveu uma aproximação similar para defender o cognitivismo externalista, e tomo que meu trabalho complementar fortemente o seu através da integração de modelos de processo-duplo na psicologia cognitiva, e oferecendo uma hipótese detalhada sobre a implementação neural dos processos relevantes. Além de seu foco principal nos juízos morais (ao invés de juízos normativos em geral), a maior diferença entre nossas perspectivas está nos detalhes do que descrevo como o caminho “intuitivo-afetivo”. Essencialmente, Sinhababu trata as *respostas emocionais/afetivas elas mesmas* como a fonte de crenças morais intuitivo-afetivas, enquanto minha proposta trata a atividade em estruturas envolvidas na emoção – circuitos (b) – como *causas comuns* de mudanças fisiológicas e crenças normativas. Também proponho que atividade nas estruturas envolvidas na recompensa e seleção de ação – circuitos (a) – possuem um papel importante em *diretamente* causar a formação de crenças normativas, e que emoções sentidas ou estados afetivos não são necessários para que isto ocorra. Obrigado a Neil Sinhababu por me ajudar a esclarecer estes detalhes.

motivação. Na seção anterior, desenvolvi uma explicação cognitivista desta conexão íntima dentro dos limites do paradigma humeano. Parto do pressuposto que esta proposta é suficiente para amenizar um pouco da força do argumento da motivação. No entanto, se o não-cognitivismo ainda oferece uma explicação melhor dos fenômenos relevantes, isto ainda seria uma desvantagem do cognitivismo. Nesta seção, argumento que minha proposta cognitivista de fato oferece uma explicação melhor dos fenômenos relevantes do que fazem as não-cognitivistas.

Para iniciar, note que há dois padrões básicos de motivação que uma descrição do juízo normativo necessita explicar. A primeira é que o juízo e a motivação tendem a alinhar-se na maioria das pessoas na maior parte do tempo. Chamemos isto de padrão de alinhamento. O segundo é que este alinhamento nem sempre se dá todas as vezes em todos os agentes. Chamemos isto de padrão de incongruência. Parece que a maioria das partes do debate aceitam ao menos isto.¹⁷

De acordo com não-cognitivistas, julgar que se deve ϕ é estar em um estado mental análogo ao desejo em se está motivada à ϕ . Então, se uma pessoa julga que deve ϕ , está motivada à ϕ . Além do mais, mudanças no juízo implicam em mudanças em estados mentais análogos ao desejo motivacionalmente eficazes, e isto explica a razão de sua covariação. Esta é uma explicação forte do padrão de alinhamento.

De acordo com minha proposta cognitivista, julgar que se deve ϕ é estar em um estado de crença motivacionalmente inerte com o conteúdo que se deve ϕ . No entanto, a causa padrão de crenças normativas são processos de Tipo 1 intuitivo-afetivos que simultaneamente motivam à ϕ e causa a formação da crença que

¹⁷ Em particular, parece que estes padrões são aceitos ambos por Blackburn (1998: 6off) e Gibbard (1993: 318–319; 2003: 153–154; vide também Strandberg 2012), ainda que alguns autores (e.g., Bromwich 2013) tenham recentemente defendido versões irrestritas de internalismo. O alvo de minha argumentação nesta seção será os não-cognitivistas que aceitam versões restritas de internalismo.

se deve ϕ . Além do mais, quando a crença que se deve ϕ é produzida por processos de Tipo 2, frequentemente envolve *tokening* representações afetivamente salientes que servem como *inputs* à processos intuitivo-afetivos de Tipo 1 que influenciam a decisão do agente. O resultado é o padrão de alinhamento.

Portanto parece que ambas perspectivas podem adequadamente capturar o padrão de alinhamento. Mas e o padrão de incongruência?

Minha proposta cognitivista oferece uma explicação direta deste padrão também. Pois permite que crenças normativas sejam ocasionalmente produzidas por processos de Tipo 2 que resultam na formação de uma crença normativa sem engajar com mecanismos diretamente envolvidos em afeto e motivação. E uma vez que crenças por si só são motivacionalmente inertes, meramente formar uma crença normativa é insuficiente para fazer emergir uma motivação correspondente. Temos então o padrão de incongruência.

Não-cognitivistas, em contraste, não explicam tão facilmente casos onde o juízo normativo e motivação não alinham-se, pois seu viés trata juízos normativos como um tipo de estado motivacionalmente eficaz. Por exemplo, Blackburn (1998: 59-68; veja também Bedke 2009; Eriksson 2014) sugerem que embora indivíduos talvez ocasionalmente possam falhar em serem motivados por seus juízos normativos, estes indivíduos podem apenas realizar juízos normativos genuínos participando em uma comunidade em que juízos normativos tipicamente alinham-se à motivação, ou possuindo uma psicologia em que juízos normativos tipicamente não alinham-se com a motivação. Não está claro se estas explicações estão concomitantes com os comprometimentos dos não-cognitivismo, no entanto. Pois elas parecem permitir que uma pessoa pode fazer um juízo normativo apesar de falhar em estar em um estado análogo ao desejo motivacionalmente eficaz relevante. Mas se juízos normativos *apenas são* estados análogos ao desejo motivacionalmente eficazes, parece então que estes

indivíduos devem, ao invés, estar em algum outro tipo de estado. Neste caso, não estariam de fato manifestando o padrão de incongruência. Então, isto de fato não parece ser uma explicação do padrão de incongruência de forma alguma.

Tendo em vista estas preocupações, Gibbard (1993: 318–319; 2003: 152–158) propôs que o padrão de incongruência surge quando estados motivacionalmente eficazes que constituem juízos normativos se tornam de alguma forma distorcidos, de tal forma que seus efeitos motivacionais são bloqueados (ver também Björnsson & McPherson 2014; Eriksson 2014; Toppinen 2015; cf. Ridge 2015; Strandberg 2012). Então, enquanto o juízo do agente é o tipo de estado que é disposto a ter efeitos motivacionais em condições normais, estas disposições meramente são suprimidas em alguns casos anormais. Assim, temos o padrão de incongruência.

Mesmo que isso possa potencialmente explicar o padrão de incongruência, não é uma explicação particularmente forte. Primeiramente, perceba que minha proposta cognitivista coloca que o padrão de incongruência surge apenas em casos pouco comuns onde estados mentais análogos à crença que são tipicamente acompanhados por uma motivação respectiva falham em ser acompanhados de tal forma. Esta hipótese é apoiada pelo fato que o padrão consiste em casos onde há juízo, mas não motivação – a falta de motivação é evidência de falta de desejo. Agora considere: qual a evidência para supor que o estado mental em que um indivíduo está quando ele manifesta o padrão de incongruência é um estado análogo ao desejo não-cognitivo enfraquecido? O perfil motivacional do agente certamente não apoia esta interpretação. Ao invés, parece que não-cognitivistas podem apenas apelar à sua habilidade de explicar o padrão de alinhamento, e então afirmar em bases indutivas que todos os juízos normativos devem ser estados análogos ao desejo não-cognitivos. No entanto, o padrão de alinhamento não necessariamente refere em favor do não-cognitivismo sobre minha

proposta cognitivista, já que o padrão é explicável de forma simples dentro de ambas perspectivas. E não parece haver nenhuma evidência independente para postular estados análogos ao desejo enfraquecidos em indivíduos manifestando o padrão de incongruência. Como resultado, esta explicação não-cognitivista do padrão de incongruência parece servir meramente para acomodar o padrão de incongruência, ao invés de ser apoiado por ele. Em contraste, minha proposta cognitivista é apoiada pelo padrão de incongruência, e pode claramente explicar o padrão em termos da natureza do juízo normativo e a operação da atividade motivacionalmente epifenomenal de processos de Tipo 2.

Concluo que cognitivistas estão em uma posição melhor que não-cognitivistas para explicar os fenômenos relevantes. Armados com o modelo de processo-duplo do juízo normativo, cognitivistas podem oferecer um escopo preciso e informativo de ambos, do padrão de alinhamento e do padrão de incongruência. Por outro lado, embora não-cognitivistas possam dar explicações simples do padrão de alinhamento, devem se esforçar para acomodar o padrão de incongruência. Como um resultado, observações sobre o relacionamento entre juízo e motivação não constituem um obstáculo para interpretações condicionais de verdade do discurso normativo, por que cognitivistas estão na verdade melhores posicionados para explicar estas observações. Como argumento na próxima seção, este resultado deixa o expressivismo em maus lençóis.

6. Consequências da Abordagem

Na seção anterior, argumentei que minha proposta cognitivista oferece uma explicação melhor do padrão de incongruência do que o não-cognitismo. Para se notar a relevância completa de minha proposta, no entanto, vamos brevemente supor que meu caso contra a explicação dos não-cognitivistas acerca do padrão de incongruência não tenha sido

convincente. Suponhamos ao invés que ambas abordagens são igualmente pares em termos de habilidade explanatória.

Mesmo neste caso, o cenário não é muito favorável para o pacote não-cognitivist-expressivista. O problema básico é que a teoria semântica expressivista possui alguns problemas sérios. Como observei na seção 1, a linguagem normativa aparenta, na superfície, muito como a linguagem descritiva ordinária. Superficialmente, a gramática de declarações normativas parece muito como afirmações acerca do status normativo de vários atos, indivíduos ou estados-de-casos. Além do mais, declarações normativas parecem reter um significado unívoco quando incorporados em sentenças complexas (Geach 1960; 1965). E apesar de algumas respostas criativas para o problema Frege-Geach, expressivistas falharam em elaborar um tratamento adequado do significado de sentenças normativas complexas, e falharam em elaborar a aparente continuidade entre declarações normativas e descritivas. De fato, há alguma razão para se pensar que nenhuma abordagem expressivista possa funcionar como um modelo de linguagens naturais como o inglês (Schroeder 2008). No mínimo, perseguir uma semântica expressivista envolve enfrentar cargas teóricas que não são compartilhadas por cognitivistas que defendem uma interpretação de condição de verdade do discurso normativo.

Em face destas dificuldades, o internalismo fora instrumental em preservar o apelo do expressivismo, desde que aparentemente cognitivistas não conseguem dar conta do papel difuso da motivação no pensamento normativo (e.g., Gibbard 1990: 33; 2003: Capítulo 1). No entanto, se a estrutura proposta neste texto está no caminho certo, então esta acusação contra cognitivistas é errônea. Mesmo dentro das restrições da teoria humeana da motivação, cognitivistas são capazes de prover uma explicação simples e direta sobre o papel da motivação no discurso e pensamento normativo. Além do mais, ao tratar juízos morais como crenças, o cognitivismo nos permite especificar os

significados de declarações normativas sem incorrer nos custos teóricos de uma semântica expressivista.

Este resultado deixa o expressivismo em maus lençóis. No melhor caso, cognitivistas e não-cognitivistas estão parelhos em termos de explicações que podem oferecer para a conexão entre juízo e motivação. No pior, os cognitivistas podem oferecer explicações melhores. Assim, uma das motivações-chave para perseguir uma semântica expressivista é eliminada. Mas dado que o expressivismo é uma abordagem problemática do significado de declarações normativas é agora muito menos claro o porquê acharíamos tais aproximações convidativas. Pois quaisquer outros supostos benefícios do expressivismo – tais como legitimar nosso discurso normativo em um mundo sem fatos normativos irreduzíveis – podem ser oferecidos por uma miríade de visões alternativas que aceitam concepções de condição de verdade do discurso normativo (e.g., Joyce 2001: Chapter 8; Olson 2014: capítulo 9; Railton 1986; Schroeder 2007). De fato, mesmo o coordenar das funções do discurso normativo podem ser prontamente explicados sem incorrer nos custos de uma semântica expressivista (Bar-On & Chrisman 2009; Strandberg 2011). Concluo que a abordagem do juízo normativo proposta acima representa uma ameaça significativa para a viabilidade do programa expressivista.

Pode parecer que perspectivas semânticas híbridas que tentam combinar elementos expressivistas e de condição de verdade terão maior sucesso, já que não necessariamente se comprometem com a semântica problemática do expressivismo “puro” (e.g., Joyce 2006: capítulo 2; Ridge 2015). No entanto, creio que o apelo destas perspectivas é também comprometido pela proposta desenvolvida aqui. Tipicamente, o propósito de se construir em um elemento de condição de verdade é evitar os problemas semânticos encarados pelos expressivistas, e o propósito de construir no elemento expressivista é reter o poder explanatório do não-cognitivismo no tópico da motivação. Mas se cognitivistas podem explicar todos os efeitos motivacionais que necessitam explicação, então não mais

parece ter muito propósito em reter o maquinário desnecessário associado com o elemento expressivista.

Considerações similares se aplicam à abordagens híbridas do juízo normativo que procuram misturar elementos de cognitivismo e não-cognitivismo (Kumar 2016; ver também Kauppinen 2015; Kriegel 2012). Kumar (2016), por exemplo, defende a perspectiva que juízos morais são um tipo natural composto de ambos estados análogos à crença e análogos ao desejo, argumentando que esta abordagem promete capturar os benefícios de ambos cognitivismo e não-cognitivismo. No entanto, quando comparados com a proposta que desenvolvi aqui, abordagens híbridas do juízo tais como a de Kumar parecem estar no mesmo barco que as abordagens de semântica híbridas em que parecem construir sobre um maquinário desnecessário. Se aceitarmos ao menos que o componente cognitivista e condicional de verdade destas perspectivas, então não é claro que benefícios adicionais são providos acrescentando o elemento não-cognitivista, dado que minha proposta pode lidar com os efeitos motivacionais relevantes igualmente bem. Assim, a proposta desenvolvida aqui é ameaçadora à viabilidade de perspectivas híbridas do juízo normativo da mesma forma.

Antes de seguir gostaria de brevemente clarificar o papel que a perspectiva de processo-duplo deve ter no meu argumento. Não afirmo que o cognitivismo se segue diretamente de abordagens de duplo-processo ao juízo normativo. Nem penso que modelos de duplo-processo são necessariamente incompatíveis com abordagens não-cognitivistas ou híbridas do juízo normativo. Ao invés, primariamente apelei à abordagem de processo-duplo de uma maneira defensiva, propondo uma interpretação dos modelos empíricos relevantes que permitem externalistas cognitivistas a explicar a conexão íntima entre juízo normativo e motivação. O objetivo deste argumento principal é simplesmente responder ao desafio presente mais sério ao cognitivismo: explicar a conexão íntima entre juízo e motivação dentro da teoria humeana da motivação. No entanto, o que argumentei nesta seção é que

quando este desafio é adequadamente respondido, o balanço de evidências favorece o cognitivismo, dados os problemas independentes com o expressivismo. E uma vez que abordagens híbridas não possuem nenhuma vantagem explanatória sobre minha abordagem puramente cognitivista, não parecem haver razões para ir-se além do cognitivismo puro e construir em um componente expressivista ou não-cognitivista. Assim, quando eu argumento defensivo central é considerado em um contexto dialético mais amplo, o cognitivismo emerge como a perspectiva com maior plausibilidade.¹⁸

7. Internalismo como uma Afirmação Conceitual

Como vimos na seção 1, internalismo é uma premissa chave no argumento da motivação para o expressivismo. Argumento que a abordagem de processo-duplo ao juízo normativo desenvolvida aqui permite que externalistas cognitivistas deem conta de todos os fenômenos relevantes concernentes à conexão íntima entre juízo normativo e motivação, bloqueando então a conclusão do argumento da motivação.

No entanto, pode ser argumentado que meu escopo não pesa no debate entre internalistas e externalistas. Isto o é, pois, meu objetivo primário tem sido iluminar os mecanismos cognitivos que subjazem o pensamento normativo em humanos comuns. Mas o internalismo é normalmente interpretado como uma afirmação conceitual sobre a natureza do juízo normativo, e como tal é imune à refutação empírica. Internalistas não afirmam que por acaso juízos normativos são normalmente conectados à motivação, mas que para que possa contar como um juízo normativo, um estado mental deve ter alguma conexão necessária com a motivação. E uma vez que ofereci nenhum argumento direto à falsidade desta

¹⁸ Sou grato a um revisor anônimo à me pressionar a clarificar este ponto.

afirmação, o argumento da motivação para o expressivismo permanece intocado. (vide Kennet & Fine 2008; Smith 2008).¹⁹

Embora seja verdade que o internalismo é tipicamente entendido como uma afirmação conceitual, não penso que esta linha de objeção seja muito produtiva. Meu objetivo aqui tem sido rascunhar uma abordagem empiricamente motivada do que acontece em nossas cabeças quando formamos pensamentos na forma de “devo ϕ ”. Argumentei que esta perspectiva oferece explicações plausíveis de não apenas a difusão do padrão de alinhamento mas também explica o menos comum padrão de incongruência. Em resposta, o internalista critica que nosso conceito de juízo normativo exclui minha abordagem *a priori*, pois minha abordagem coloca a conexão entre juízo e motivação como contingente ao invés de necessária. Nenhuma abordagem como a minha pode possivelmente ser verdade, pois despreza esta restrição conceitual internalista. Mas que evidência há para se pensar que uma abordagem como a minha poderia possivelmente não estar correta?

Uma linha de evidência poderia ser o padrão de alinhamento descrito acima. Pode-se pensar que apenas se o internalismo é verdadeiro poderia haver tal conexão difusa entre motivação e juízo. Mas esta evidência não é persuasiva, uma vez que já mostrei que externalistas podem explicar este padrão. De forma mais importante, não é claro que estas observações seriam suficientes para mostrar que juízos normativos teriam alguma relação necessária à motivação. Como Smith (2008: 210) coloca, mesmo se

¹⁹ Agradeço a Alex King por me pressionar sobre este tópico. Enquanto alguns internalistas, como Smith (1994: capítulo 3) e Wedgwood (2007: capítulo 1), mantêm que a versão mais plausível do internalismo apela à irracionalidade para dar conta de casos onde a motivação e juízo falham em alinhar, esta manobra não parece ser disponível para expressivistas. Isto é por que eles tipicamente mantêm que juízo normativo e internalismo motivacional pode ser entendido sem o uso de termos normativos como “racional.” Assim, versões do internalismo baseadas em racionalidade estão fora do escopo deste artigo, uma vez que meu objetivo principal aqui é defender o cognitivismo. São as interpretações psicológicas – não normativas – do internalismo que são mais relevantes aqui. Vide line (no prelo) para mais discussões do “requerimento de praticidade” de Smith no contexto dos modelos de processo-duplo do juízo moral.

todas as instâncias reais de juízo fossem combinadas com motivação, isto não mostraria que juízo e motivação estão conectados por uma questão de necessidade conceitual.

Uma linha de evidência mais promissora parece vir do fato que internalistas não sentem-se confortáveis a atribuir um juízo normativo genuíno à um agente que diz ou pensa “eu devo ϕ ” mas que continua completamente não-motivado à ϕ . Isto pode ser tomado como apoio à ideia que nosso conceito de juízo normativo não se aplica legitimamente à estados mentais que não possuem alguma conexão necessária com a motivação.

Esta evidência, no entanto, é bastante fraca, uma vez que nem todas as partes aceitam que este conceito de juízo normativo não pode ser legitimamente aplicado em tais casos. É verdade que ocasionalmente alguns indivíduos podem não entender alguns conceitos, e poderiam então cometer erros em sua aplicação. Por exemplo, uma pessoa bastante confusa pode defender que gatos são um tipo de instrumento musical. Em casos assim, algo claramente deu errado. No entanto, internalismo é uma tese notoriamente controversa, e está longe de ser claro que externalistas estão confusos da mesma maneira que uma pessoa que afirma que gatos são instrumentos musicais o seria (cf. Francén 2012).²⁰ De fato, em uma grande enquête de filósofos profissionais, cerca de 30% dos que responderam afirmaram apoiar ou inclinarem-se ao externalismo, cerca de 35% apoiam ou inclinam-se ao internalismo e 35% responderam “outro” (Bourget & Chalmers 2014: 496). Esta falta de consenso sugere que se deve ser especialmente cauteloso ao pressionar acusações de confusão conceitual a qualquer lado.

²⁰ De fato, a maior parte dos internalistas agora concordam que pode ser ocasionalmente legítimo atribuir um juízo normativo à um agente que não está motivado a agir de acordo com tal, desde que aquele agente seja irracional, ou normalmente motivado à agir de acordo com seus juízos, ou membro de uma comunidade de indivíduos que são normalmente motivados a agir de acordo com seus juízos. No entanto, a profusão de diferentes internalismos parece trazer o questionamento da presunção que há um fato definitivo do tema sobre os limites de um conceito único e claro compartilhado por todas as partes.

Talvez conceitos de filósofos externalistas sobre o juízo normativo são carregados de teoria e tornaram-se enviesado por seus comprometimentos teóricos. No entanto, estas hipóteses não são bem apoiadas por trabalhos na filosofia experimental sobre intuições populares sobre o juízo moral (e.g., Björnsson, eriksson, Strandberg, Francén Olinder, & Björklund 2015a; Strandberg & Björklund 2013). Parece que no máximo algumas pessoas ocasionalmente possuem alguma hesitação em atribuir juízos morais genuínos à indivíduos que não possuem motivação correspondente. Isto não ajuda muito os internalistas, uma vez que certamente não há refutações uniformes para realizar tais atribuições sob as condições que foram estudadas, diferentemente do que esperaríamos de erros óbvios como na tese dos gatos-como-instrumentos-musicais. Além disso, o que há de hesitação pode ser facilmente explicado dentro de um paradigma externalista.

Por exemplo, se minha proposta está no caminho correto, então o padrão de incongruência será relativamente incomum. Como um resultado, há várias formas nas quais mecanismos gerais pragmáticos poderiam explicar por que se expressar crenças morais normalmente transmite informação sobre suas motivações (Strandberg 2011; vide também Bar-On & Chrisman 2009). Isto explicaria nossa expectativa que o juízo de um agente e motivação se alinham mesmo quando não há conexão necessária entre motivação e juízo normativo.

Similarmente, sabemos há algum tempo que algumas pessoas não tratam todos os membros de uma categoria da mesma forma. Ao invés, as pessoas possuem maior facilidade ao identificar instâncias mais prototípicas como membros. Por exemplo, sujeitos são mais velozes em identificar um pintarroxo como um tipo de pássaro do que identificar o pinguim como um tipo de pássaro (Rosch 1978). Dado que a maioria das pessoas são normalmente motivadas a agir de acordo com seus juízos normativos, a motivação pode estar para o juízo normativo como voo está para os pássaros. Assim nossa maior hesitação em atribuir.

Além do mais, há indubitavelmente casos em que pessoas afirmam de forma não sincera acreditar em alguma afirmação normativa. Estamos sempre atentos à pessoas que possam tentar nos manipular (Tooby, Cosmides, & Price 2006), e uma forma que alguém pode fazer isso é afirmar que compartilha de nossos valores. Como um resultado, faz sentido a nós suspeitar de pessoas que não são movidas pelo que dão a entender que pensam ser bom e correto.

À luz de considerações como estas, não é difícil entender as razões de alguns filósofos e leigos terem intuições aparentemente internalistas, mesmo se o externalismo fosse verdade. Penso, portanto, que há poucas razões para manter que minha proposta é irrelevante baseado em que violaria nosso conceito de um juízo normativo. Erros conceituais claros podem ser uma estrutura mais firme para basear uma argumentação. Mas em casos mais controversos, seu oponente não pode ser tão facilmente dispensado. Ao invés, parece mais produtivo se utilizar de abordagens empiricamente informadas sobre os mecanismos subjacentes do pensamento normativo para ajudar a iluminar a questão se juízos normativos são constituídos de estados não-cognitivos (vide Bromwich 2013; Kumar 2015; 2016; Prinz 2015).

De fato, acredito que possuir uma interpretação suficientemente detalhada dentro destas linhas é um pré-requisito para a resolução deste debate. Isto o é dada a afirmação comum que a tese internalista requer apenas que um agente possua um mínimo de motivação para que o internalismo esteja satisfeito. O que é um mínimo de motivação? Como podemos determinar se um mínimo de motivação é presente ou não se nunca influencia uma ação? Não consigo imaginar como interpretações de poltrona de casos imaginários possam oferecer uma história final neste tópico. Ao invés, parece que a história deva ultimamente implicar as operações de algum mecanismo cognitivo ou outro que tenha algum papel na ação humana. Focando nos mecanismos por trás do pensamento normativo e identificando seus perfis funcionais, podemos progredir além dos achismos se um agente imaginário

possuía um mínimo de motivação à ϕ apesar de não realizar ϕ (cf. Botvinick & Braver 2015: 104–105).

Pode-se pensar que uma vantagem de se aceitar o internalismo com uma afirmação conceitual seja a habilidade de apelar ao conteúdo semântico de juízos normativos para explicar as razões da motivação ser relevante à juízos normativos.²¹ Isto pode ser de fato algo que externalistas não podem oferecer (embora vide Sturgeon 1999: 95). No entanto, ao interpretar crenças normativas como tipicamente causadas por estados mentais análogos ao desejo motivacionalmente eficazes, o paradigma do processo-duplo desenvolvido aqui promete explicar por que experienciamos certas ações em potencial como pedindo respostas de nós. Já que, em minha proposta, desejos e respostas afetivas nos causam a ser motivados de várias maneiras, e também nos causar a pensar que temos razões de agir destas maneiras, há uma explicação prontamente disponível dos motivos de haver uma conexão estreita entre juízos normativos e motivação. Mesmo que não se veja esta explicação como suficientemente satisfatória, esta vantagem de aceitar o internalismo como uma afirmação conceitual não parece ser uma consideração decisiva a seu favor. Perspectivas nestes tópicos necessitam ser asseridas por sua plausibilidade geral, e esta vantagem seria apenas uma parte da equação.

Finalmente, mesmo se o internalismo é uma afirmação conceitual verdadeira, é possível que seja uma afirmação conceitual em necessidade de revisão. Não é incomum para avanços científicos nos guiarem ao refinamento de conceitos em resposta de novas evidências. Por exemplo, embora à outrora popular opinião contrária, foi descoberto que baleias na verdade não eram peixes. Se estamos interessados em entender de fato os juízos normativos que pessoas realizam, parece no mínimo concebível que um entendimento mais aprofundado dos mecanismos cognitivos subjacentes poderia nos mostrar que nossos conceitos

²¹ Agradeço a um revisor anônimo por eliciar esta vantagem potencial.

estavam errados (vide Bedke 2009: 202–203; Kumar 2016). Ao invés de prender-se à uma conexão real necessária, talvez estávamos meramente supergeneralizando na base da constante conjunção de juízo e motivação.

A estrutura desenvolvida acima goza de um nível razoável de apoio empírico, e oferece explicações informativas de vários fenômenos-chave. De fato, indiquei agora como pode mesmo permitir que externalistas expliquem como refutar evidências que podem ser tomadas como apoiadoras do internalismo como uma afirmação conceitual. Portanto, não perce que a aproximação tomada aqui possa ser rapidamente dispensada meramente baseado no fato que oferece uma abordagem externalista do juízo normativo.²²

8. Conclusões

Neste texto, parti de modelos de processo-duplo na psicologia cognitiva para oferecer uma abordagem externalista cognitivista da conexão íntima entre juízo normativo e motivação. A inovação chave em minha proposta é que trata estados análogos ao desejo motivacionalmente eficazes como causas típicas de crenças normativas. Na seção 5, argumentei que esta proposta permite que cognitivistas expliquem o padrão de alinhamento tão bem como não-cognitivistas podem, enquanto fazem um serviço melhor que não-cognitivistas ao explicar o padrão de incongruência. Também argumentei que mesmo que ambas perspectivas estejam mais ou menos parelhas em termos de habilidade explanatória em geral, ainda assim este será um forte baque ao não-cognitivismo, uma vez que uma parte considerável do apelo da perspectiva era sua suposta habilidade única de

²² Meu objetivo subjacente neste artigo é defender o cognitivismo mostrando que versões externalistas da perspectiva são mais plausíveis do que normalmente pensado. No entanto, se minha resposta à objeção discutida nesta seção parece insatisfatória, talvez seja possível desenvolver uma versão cognitivista de um internalismo fraco que seja compatível com a estrutura de processo-duplo rascunhada acima (cf. Tresan 2006; vide também nota de rodapé 14).

capturar a conexão entre juízo normativo e motivação dentro de uma estrutura humeana. E sem o apoio do não-cognitivismo, o expressivismo perde muito de seu apelo da mesma forma. Estas são boas notícias para cognitivistas que abraçam uma abordagem de condição de verdade ao discurso normativo.

É importante lembrar que estas boas notícias são apenas condicionais, no entanto. A viabilidade de minha proposta dependerá ultimamente em trabalhos futuros nas ciências do pensamento normativo e tomada de decisão. Ao mesmo tempo, a proposta de fato goza um grau de apoio empírico, e oferece recursos para guiar e interpretar futuros trabalhos.

No mínimo espero ter mostrado que a sabedoria popular sobre a falta de adequação explanatória do cognitivismo não é uma conclusão do passado. A natureza subjacente do pensamento e discurso normativo pode muito bem ser apenas o que é sugerido por suas características superficiais.

Agradecimentos

Por seus comentários prestativos em rascunhos deste texto, sou grato à James Beebe, Alex King, Ken Shockley, Lewis Powell, e dois revisores anônimos por esta revista. Também sou grato pela discussão proveitosa com Ariane Nomikos, Neil Sinhababu, Tongli Wu, e os públicos na Universidade de Shandong e da Universidade em Buffalo.

Referências

- ABE, N.; GREENE, J. D. Response to Anticipated reward in the Nucleus Accumbens Predicts Behavior in an independent Test of Honesty. **The Journal of Neuroscience**, v. 34, n. 32, p. 10564–10572, 2014.
- ARPALY, N.; SCHROEDER, T. In Praise of Desire. In: SHAFERLANDAU, RUSS (Ed.). **Oxford Studies in Metaethics**. Oxford: Oxford University, 2014. v. 4p. 132–165.

BARBEY, A. K.; KOENIGS, M.; GRAFMAN, J. Dorsolateral prefrontal contributions to human working memory. **Cortex**, v. 49, n. 5, p. 1195–1205, maio 2013.

BARGH, J. A. AN. T. L. C. The Unbearable Automaticity of Being. **American Psychologist**, v. 54, n. 7, p. 462–479, 1999.

BASTEN, U. et al. **How the Brain integrates Costs and Benefits During Decision Making**. Proceedings of the National Academy of Sciences. **Anais...**2010

BAUMGARTNER, T. et al. Dorsolateral and Ventromedial Prefrontal Cortex Orchestrate Normative Choice. **Nature Neuroscience**, v. 14, n. 11, p. 1468–1474, 2011.

BAVEL, V. et al. Evaluation is a Dynamic Process: Moving Beyond Dual System Models. **Social and Personality Psychology Compass**, v. 6, n. 6, p. 438–454, 2012.

BECHARA, A. et al. Deciding Advantageously Before Knowing the Advantageous Strategy. **Science**, v. 275, n. 5304, p. 1293–1295, 1997.

BEDKE, M. S. Moral Judgment Purposivism: Saving internalism from Amoralism. **Philosophical Studies**, v. 144, n. 2, p. 189–209, 2009.

BJÖRNSSON, G. et al. Motivational internalism: Contemporary Debates. In: BJÖRNSSON, G. et al. (Eds.). **Motivational Internalism**. Oxford: Oxford University Press, 2015a. p. 1–20.

BJÖRNSSON, G. et al. Motivational internalism and Folk intuitions. **Philosophical Psychology**, v. 28, n. 5, p. 715–734, 2015b.

BJÖRNSSON, G.; MCPHERSON, T. Moral Attitudes for Non-Cognitivists: Solving the Specification Problem. **Mind**, v. 123, n. 489, p. 1–38, 2014.

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

BLACKBURN, S. **Essays in Quasi-Realism**. Oxford: Oxford University Press, 1993.

- BLACKBURN, S. **Ruling Passions: A Theory of Practical Reasoning**. Oxford: Oxford University Press, 1998.
- BORLAND, R. et al. Impact of Graphic and Text Warnings on Cigarette Packs: Findings from Four Countries Over Five Years. **Tobacco Control**, v. 18, n. 5, p. 358–364, 2009.
- BOTVINICK, M.; BRAVER, T. Motivation and Cognitive Control: From Behavior to Neural Mechanism. **Annual Review of Psychology**, v. 66, p. 83–113, 2015.
- BOURGET, D.; CHALMERS, D. J. What do Philosophers Believe? **Philosophical Studies**, v. 170, n. 3, p. 465–500, 2014.
- BROMWICH, D. Motivational internalism and the Challenge of Amoralism. **European Journal of Philosophy**, v. 24, n. 2, p. 452–471, 2013.
- CHEIN, J. M.; MOORE, A. B.; CONWAY, A. R. A. (EDS.). Domain-General Mechanisms of Complex Working Memory Span. **NeuroImage**, v. 54, n. 1, p. 550–559, 2011.
- CIARAMELLI, E. et al. Selective Deficit in Personal Moral Judgment Following Damage to Ventromedial Prefrontal Cortex. **Social Cognitive and Affective Neuroscience**, v. 2, n. 2, p. 84–92, 2007.
- CLINE, B. Smith’s practicality requirement meets dual-process models of moral judgment. **Philosophical Psychology**, v. 30, n. 8, p. 1043–1063, 17 nov. 2017.
- CRAIG, A. D. How Do You Feel—Now? The Anterior insula and Human Awareness. **Nature Reviews Neuroscience**, v. 10, n. 1, p. 59–70, 2009.
- CUSHMAN, F.; YOUNG, L.; GREENE, J. D. Multi-System Moral Psychology. In: DORIS, J. M.; THE MORAL PSYCHOLOGY RESEARCH GROUP (Eds.). **The Moral Psychology Handbook**. Oxford: Oxford University Press, 2010. p. 47–71.
- DAMASIO, A. **Descartes’ Error: Emotion, Reason, and the Human Brain**. New York: Putnam, 1994.
- DORIS, J. M.; THE MORAL PSYCHOLOGY RESEARCH GROUP. **The Moral Psychology Handbook**. [s.l.] Oxford University Press, 2010.

ERIKSSON, J. Elaborating expressivism: Moral Judgments, Desires, and Motivation. **Ethical Theory and Moral Practice**, v. 17, n. 2, p. 253–267, 2014.

EVANS, J. S. B. T. AN. K. F. **In Two Minds: Dual Processes and Beyond**. Oxford: Oxford University Press, 2009. v. 10

EVANS, J. S. B. T.; STANOVICH, K. E. Dual-Process Theories of Higher Cognition: Advancing the Debate. **Perspectives on Psychological Science**, v. 8, n. 3, p. 223–241, maio 2013.

FEDORENKO, EVELINA; DUNCAN, J.; KANWISHER, AND N. **Broad Domain Generality in Focal regions of Frontal and Parietal Cortex**. Proceedings of the National Academy of Sciences. **Anais...**2013

FEHR, E.; KRAJBICH, I. Social Preferences and the Brain. In: FEHER, E.; GLIMCHER, P. W. (Eds.). **Neuroeconomics: Decision Making and the Brain**. 2. ed. Cambridge: Academic Press, 2014. p. 196–218.

FRANCÉN, RAGNAR. Svavarsdóttir's Burden. **Philosophia**, v. 40, n. 3, p. 577–589, 2012.

GEACH, P. T. Ascriptivism. **The Philosophical Review**, v. 69, n. 2, p. 221–225, 1960.

GEACH, P. T. Assertion. **The Philosophical Review**, v. 74, n. 4, p. 449–465, 1965.

GIBBARD, A. **Wise Choices, Apt Feelings: A Theory of Normative Judgment**. Oxford: HarvardUniversity Press, 1990.

GIBBARD, A. Reply to Sinnott-Armstrong. **Philosophical Studies**, v. 69, n. 2–3, p. 315–327, 1993.

GIBBARD, A. **Thinking How to Live**. Harvard: Harvard University Press, 2003.

GLASSEN, P. The Cognitivity of Moral Judgments. **Mind**, v. 68, n. 269, p. 57–72, 1959.

GREENE, J. D. An fMRI Investigation of Emotional Engagement in Moral Judgment. **Science**, v. 293, n. 5537, p. 2105–2108, 14 set. 2001.

- GREENE, J. D. et al. The neural bases of cognitive conflict and control in moral judgment. **Neuron**, v. 44, n. 2, p. 389–400, 14 out. 2004.
- GREENE, J. D. et al. Cognitive load selectively interferes with utilitarian moral judgment. **Cognition**, v. 107, n. 3, p. 1144–1154, jun. 2008.
- GREENE, J. D. The Cognitive Neuroscience of Moral Judgment and Decision Making. In: GAZZANIGA, M. (Ed.). **The Cognitive Neurosciences**. 5th ed., 1013–1024 ed. Cambridge: MiT Press, 2014a.
- GREENE, J. D. Beyond Point and Shoot Morality. **Ethics**, v. 124, n. 4, p. 695–726, 2014b.
- HABER, S. N. AN. B. K. The reward Circuit: linking Primate Anatomy and Human imaging. **Neuropsychopharmacology**, v. 35, n. 1, p. 4–26, 2010.
- HAIDT, J. The emotional Dog and its rational Tail: A Social intuitionist Approach to Moral Judgment. **Psychological Review**, v. 108, n. 4, p. 814–834, 2001.
- HAJEK, D. Hollywood Jet Gives Fearful Fliers the Courage to Soar. **National Public Radio: Morning Edition**, 17 maio 2016.
- HARE, T. A. et al. Value Computations in Ventral Medial Prefrontal Cortex During Charitable Decision Making incorporate input from regions involved in Social Cognition. **The Journal of Neuroscience**, v. 30, n. 2, p. 583–590, 2010.
- HARE, T. A.; CAMERER, C. F.; RANGEL, A. Self-Control in Decision-Making Involves Modulation of the vmPFC Valuation System. **Science**, v. 324, n. 5927, p. 646–648, 1 maio 2009.
- HARE, T. A.; MALMAUD, J.; RANGEL, A. Focusing Attention on the Health Aspects of Foods Changes Value Signals in vmPFC and Improves Dietary Choice. **Journal of Neuroscience**, v. 31, n. 30, p. 11077–11087, 27 jul. 2011.
- HARUNO, M.; FRITH, C. D. Activity in the Amygdala elicited by Unfair Divisions Predicts Social Value Orientation. **Nature Neuroscience**, v. 13, n. 2, p. 160–161, 2010.

HORGAN, T.; TIMMONS, M. Morphological rationalism and the Psychology of Moral Judgment. **Ethical Theory and Moral Practice**, v. 10, n. 3, p. 279–295, 2007.

HUTCHERSON, C. A. et al. Cognitive Regulation during Decision Making Shifts Behavioral Control between Ventromedial and Dorsolateral Prefrontal Value Systems. **Journal of Neuroscience**, v. 32, n. 39, p. 13543–13554, 26 set. 2012.

JOYCE, RICHARD. **The Myth of Morality**. Cambridge: Cambridge University Press, 2001. v. 10

JOYCE, R. **The Evolution of Morality**. Cambridge: MIT Press, 2006.

KAHNEMAN, D. **Thinking, Fast and Slow**. New York: Farrar, Straus, and Giroux, 2011.

KAUPPINEN, A. Intuition and Belief in Moral Motivation. in Gunnar Björnsson, Caj Strandberg, ragnar Francén Olinder. In: ERIKSSON, J.; BJÖRKLUND, AND F. (Eds.). **Motivational Internalism**. Oxford: Oxford University Press, 2015. v. 10p. 237–259.

KENNET, J.; FINE, C. Could There Be an empirical Test for internalism? in Walter Sinnott-Armstrong. **Moral Psychology**, v. 3, 2008.

KOENIGS, M. et al. Damage to the Prefrontal Cortex increases Utilitarian Moral Judgments. **Nature**, v. 446, n. 7138, p. 908–911, 2007.

KOENIGS, M.; TRANEL, D. Irrational economic Decision-Making after Ventromedial Prefrontal Damage: evidence from the Ultimatum Game. **The Journal of Neuroscience**, v. 27, n. 4, p. 951–956, 2007.

KRAJBICH, IAN et al. Economic Games Quantify Diminished Sense of Guilt in Patients with Damage to the Prefrontal Cortex. **The Journal of Neuroscience**, v. 29, n. 7, p. 2188–2192, 2009.

KRIEGEL, U. Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction. **Australasian Journal of Philosophy**, v. 90, n. 3, p. 469–486, 2012.

- KUMAR, V. Moral Judgment as a Natural Kind. **Philosophical Studies**, v. 172, n. 11, p. 2887–2910, 2015.
- KUMAR, V. The empirical identity of Moral Judgment. **The Philosophical Quarterly**, v. 66, n. 265, p. 783–804, 2016.
- LEBRETON, M. et al. An Automatic Valuation System in the Human Brain: evidence from Functional Neuroimaging. **Neuron**, v. 64, n. 3, p. 431–439, 2009.
- LELAND, J. W.; GRAFMAN, J. Experimental Tests of the Somatic Marker Hypothesis. **Games and Economic Behavior**, v. 52, n. 2, p. 386–409, 2005.
- MCDOWELL, J. Values and Secondary Qualities. In: SAYRE-MCCORD, G. (Ed.). **Essays in Moral Realism**. Cornell: Cornell University Press, 1988.
- MERCIER, H.; SPERBER, D. Intuitive and reflective inferences. In: EVANS, J.; FRANKISH, K. (Eds.). **In two minds: Dual processes and beyond**. Oxford: Oxford University Press, 2009. p. 149–170.
- MERCIER, H.; SPERBER, D. Why Do Humans reason? Arguments for an Argumentative Theory. **Behavioral and Brain Sciences**, v. 34, n. 2, p. 57–74, 2011.
- OCHSNER, K.; GROSS, J. The cognitive control of emotion. **Trends in Cognitive Sciences**, v. 9, n. 5, p. 242–249, maio 2005.
- OLSON, J. **Moral Error Theory: History, Critique, Defence**. Oxford: Oxford University Press, 2014. v. 10
- PAXTON, J. M.; UNGAR, L.; GREENE, J. D. (EDS.). Reflection and reasoning in Moral Judgment. **Cognitive Science**, v. 36, n. 1, p. 163–177, 2012.
- PLASSMANN, H. et al. Marketing actions can modulate neural representations of experienced pleasantness. **Proceedings of the National Academy of Sciences**, v. 105, n. 3, p. 1050–1054, 22 jan. 2008.
- PRINZ, J. An Empirical Case for Motivational Internalism. In: BJÖRNSSON, G. et al. (Eds.). **Motivational Internalism**. Oxford: Oxford University Press, 2015. p. 61–84.

- PUJARA, M.; KOENIGS, M. Mechanisms of reward Circuit Dysfunction in Psychiatric illness: Prefrontal-Striatal interactions. **The Neuroscientist**, v. 20, n. 1, p. 82-95, 2014.
- RAILTON, P. Moral realism. **The Philosophical Review**, v. 95, n. 2, p. 163-207, 1986.
- RAILTON, P. The Affective Dog and Its Rational Tale: Intuition and Attunement. **Ethics**, v. 124, n. 4, p. 813-859, jul. 2014.
- RANGEL, A.; CAMERER, C.; MONTAGUE, AND P. READ. A Framework for Studying the Neurobiology of Value-Based Decision Making. **Nature Reviews Neuroscience**, v. 9, p. 545-556, 2008.
- RANGEL, A.; CLITHERO, J. A. The Computation of Stimulus Values in Simple Choice. In: **Neuroeconomics**. Cambridge: Academic Press, 2014. p. 125-148.
- RIDGE, M. Internalism. In: BJÖRNSSON, G. et al. (Eds.). **Motivational Internalism**. Oxford: Oxford University Press, 2015. p. 135-149.
- ROSCH, E. Principles of Categorization. In: LLOYD, B. B.; ROSCH, E. (Eds.). **Cognition and Categorization**. New Jersey: Lawrence Erlbaum, 1978. p. 27-48.
- ROSKIES, A. Are ethical Judgments intrinsically Motivational? Lessons from “Acquired Sociopathy”. **Philosophical Psychology**, v. 16, n. 1, p. 51-66, 2003.
- ROY, M.; SHOHAMY, D.; WAGER, T. D. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. **Trends in Cognitive Sciences**, v. 16, n. 3, p. 147-156, mar. 2012.
- RUDORF, S.; HARE, T. A. Interaction between Dorsolateral and Ventromedial Prefrontal Cortex Underlie Context-Dependent Stimulus Valuation in Goal-Directed Choice. **The Journal of Neuroscience**, v. 34, n. 48, p. 15988-15996, 2014.
- SANFEY, A. G. The Neural Basis of Economic Decision-Making in the Ultimatum Game. **Science**, v. 300, n. 5626, p. 1755-1758, 13 jun. 2003.

- SAUER, H. Educated intuitions: Automaticity and rationality in Moral Judgment. **Philosophical Explorations**, v. 15, n. 3, p. 255–275, 2012.
- SCHROEDER, M. **Slaves of the Passions**. Oxford: Oxford University Press, 2007. v. 10
- SCHROEDER, M. **Being For: Evaluating the Semantic Program of Expressivism**. Oxford: Oxford University Press, 2008. v. 10
- SCHROEDER, T. **Three Faces of Desire**. Oxford: Oxford University Press, 2004. v. 10
- SCHROEDER, T.; ROSKIES, A. L.; NICHOLS, S. Moral Motivation. In: DORIS, J. M. (Ed.). **The Moral Psychology Handbook**. Oxford: Oxford University Press, 2010. p. 72–110.
- SHENHAV, A.; BOTVINICK, M. M.; COHEN, J. D. The expected Value of Control: An integrative Theory of Anterior Cingulate Cortex Function. **Neuron**, v. 79, n. 2, p. 217–240, 2013.
- SHENHAV, A.; D.GREENE, J. Integrative Moral Judgment: Dissociating the roles of the Amygdala and Ventromedial Prefrontal Cortex. **Journal of Neuroscience**, v. 34, n. 13, p. 4741–4749, 2014.
- SHENHAV, A.; GREENE, J. D. Moral Judgments recruit Domain-General Valuation Mechanisms to integrate representations of Probability and Magnitude. **Neuron**, v. 67, n. 4, p. 667–677, 2010.
- SHERMAN, J. W.; GAWRONSKI, B.; TROPE, Y. **Dual-Process Theories of the Social Mind**. New York: Guilford Press, 2014.
- SINHABABU, N. **Humean Nature: How Desire Explains Action, Thought, and Feeling**. Oxford: Oxford University Press, 2017. v. 10
- SLOVIC, P. et al. Affect, risk, and Decision Making. **Health Psychology**, v. 24, n. 4, p. 35–40, 2005.
- SMITH, D. V. et al. Distinct Value Signals in Anterior and Posterior Ventromedial Prefrontal Cortex. **The Journal of Neuroscience**, v. 30, n. 7, p. 490–495, 2010.

SMITH, M. **The Moral Problem**. Hoboken, New Jersey: Blackwell Publishing, 1994.

SMITH, M. The Truth About internalism. in Walter Sinnott-Armstrong. **Moral Psychology**, v. 3, 2008.

SOKOL-HESSNER, P. et al. Decision value computation in DLPFC and VMPFC adjusts to the available decision time: Value computation adjusts to decision time. **European Journal of Neuroscience**, v. 35, n. 7, p. 1065–1074, abr. 2012.

STRANDBERG, C. The Pragmatics of Moral Motivation. **The Journal of Ethics**, v. 15, n. 4, p. 341–369, 2011.

STRANDBERG, C. Expressivism and Dispositional Desires. **American Philosophical Quarterly**, v. 49, n. 1, p. 81–91, 2012.

STRANDBERG, C.; BJÖRKLUND, G. Is Moral internalism Supported by Folk intuitions? **Philosophical Psychology**, v. 26, n. 3, p. 319–335, 2013.

STURGEON, N. L. Review: “The Moral Problem”. **The Philosophical Review**, v. 108, n. 1, p. 94–97, 1999.

TOOBY, J.; COSMIDES, L.; PRICE, M. E. Cognitive Adaptations for n-Person exchange: The evolutionary roots of Organizational Behavior. Managerial and Decision. **Economics**, v. 27, n. 2–3, p. 103–129, 2006.

TOPPINEN, T. “Pure Expressivism and Motivational Internalism”. In: BJÖRNSSON, G. et al. (Eds.). **Motivational Internalism**. [s.l.] Oxford University Press, 2015. p. 150–166.

TRESAN, J. De Dicto internalist Cognitivism. **Noûs**, v. 40, n. 1, p. 143–165, 2006.

TUSCHE, A.; BODE, S.; HAYNES, J.-D. Neural responses to Unattended Products Predict later Consumer Choices. **The Journal of Neuroscience**, v. 30, n. 23, p. 8024–8031, 2010.

WEDGWOOD, RALPH. **The Nature of Normativity**. Oxford: Oxford University Press, 2007. v. 10

WIELENBERG, ERIK J. **Robust Ethics: The Metaphysics and Epistemology of Godless Normative Realism**. Oxford: Oxford University Press, 2014. v. 10

ZAJONC, R. B. Feeling and Thinking: Preferences Need No inferences. **American Psychologist**, v. 35, n. 2, p. 466–473, 1980.

Capítulo 2

Obrigação moral e motivação moral*

*David Copp*¹

Em ética, *internalismo* é um conjunto de visões de acordo com o qual existe uma conexão *interna* entre obrigações morais e motivações ou razões para agir moralmente, enquanto *externalismo* diz que tais conexões são contingentes. Assim descrita, a disputa entre internalismo e externalismo pode parecer um debate técnico de menor interesse. Entretanto, as questões que o motivam incluem problemas profundos sobre verdade moral, realismo, normatividade e objetividade. De fato, acredito que alguns filósofos veem o externalismo como um enfraquecimento da *dignidade* da moralidade. Eles podem dizer que, se a moralidade precisa de uma "sanção externa" - se a crença que tem uma obrigação não é motivo suficiente ou razão para fazer o certo - então a moralidade é depreciada em status. Mesmo um sistema arbitrário de etiqueta poderia atrair uma sanção externa sob condições apropriadas.

Embora eu acredite que as teses internalistas mais interessantes sejam falsas, existem verdades importantes que o internalismo está tentando apreender. O mais importante deles é o fato de que os julgamentos morais são intrinsecamente 'normativo' ou 'orientador de escolha', que eles são, *grosso modo*, relevantes para ação ou escolha por causa de seu conteúdo. O internalismo tenta explicar a normatividade do julgamento moral

¹ University of California – Davis. Tradução e revisão: Evandro Barbosa; Thaís Alves Costa.

em termos de motivações ou razões, e eu acredito que este é o seu erro. Contudo, o internalismo está correto de que a normatividade é *interno* ao julgamento moral. As teorias externalistas têm negado esta normatividade intrínseca ou pelo menos não tiveram êxito em explicá-la.

Meu objetivo neste texto é apresentar uma teoria do julgamento e convicção moral que é, estritamente falando, externalista, muito embora acomode o *insight* dos internalistas sobre a normatividade da moral julgamento. Eu quero explicar como esta teoria faz isso e também explicar como ela acomoda outras intuições importantes e plausíveis que motivam o internalismo. Eu apresentei esta teoria em um recente livro, mas minha discussão sobre o internalismo no livro foi distribuída ao longo de vários capítulos.² Quero reunir os pontos-chave neste texto. Além disso, quero mostrar mais detalhadamente como minha teoria pode responder a certas objeções internalistas, prestando especial atenção a alguns novos argumentos de Michael Smith.³ Começo com algumas considerações preliminares.

1. O debate entre o internalismo e o externalismo

O termo internacionalismo é usado para nomear doutrinas sobre motivação, bem como doutrinas sobre razões. Em geral, uma doutrina internalista afirma há uma conexão necessária entre o estado de ter uma moral obrigação, ou o estado de acreditar ou reconhecer que existe uma obrigação moral, e o estado de estar motivado - ou o estado de ter razões cumprir esta obrigação. Cada tese internalista corresponde a uma versão do *externalismo*, que é

* Sou grato a Philip Clark, Earl Conee, Jeffrey C. King e Michael Smith por seus frutíferos comentários. Uma versão anterior deste texto foi apresentada no *Philosophy Colloquium* na Universidade da Carolina do Norte, Greensboro, em 1995. Sou grato a todos aqueles que contribuíram para a discussão naquela ocasião.

² David Copp, *Morality, Normativity, and Society*. New York: Oxford University Press, 1995.

³ Michael Smith, *The Moral Problem*. Oxford: Blackwell, 1994.

simplesmente a sua negação. O externalismo, como tal, nega todas as doutrinas internalistas.

Começarei com as doutrinas do internalismo motivacional. Usando terminologia introduzida por David Brink, podemos distinguir entre *internalismo de agente*, *internalismo de crença* e *internalismo híbrido*.⁴ O internalismo de agência é a doutrina de que é uma verdade necessária que, se uma pessoa tem uma obrigação moral de fazer alguma coisa, então a pessoa tem alguma motivação para fazê-lo (ou ela teria alguma motivação, sob condições relevantes). O internalismo de crença sustenta que é a crença da pessoa de que ela tem uma obrigação moral que implica nela está motivada de forma relevante. O internalismo híbrido sustenta que é a crença de uma pessoa, quando a crença é verdadeira, que garante que ela está motivada de forma relevante.⁵

O internalismo pode permitir casos em que a pessoa com a obrigação - com a crença ou com a crença verdadeira de que ela tem uma obrigação - não está realmente ou *ocasionalmente* motivado. Isso explica a qualificação nas formulações acima. A pessoa em questão tem alguma motivação fazer a ação relevante ou *teria* alguma motivação *sob condições relevantes*. Diferentes internalistas especificariam as condições relevantes de maneiras diferentes,⁶ mas é, obviamente, muito importante para um

⁴ David o. Brink, *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press, 1989, 40. Brink fala de "internalismo de avaliação" que parece ser a mesma que eu chamo de "internalismo de crença". W.K.

Frankena reconheceu as distinções em seu clássico ensaio "Obligation and Motivation in Recent Moral Philosophy", em Kenneth E. Goodpaster, ed., *Perspectives on Morality: Essays of William K. Frankena*. Notre Dame, IN: Notre Dame University Press, 1976, 60. As doutrinas que Frankena rotula em (1), (7) e (8) são, respectivamente, internalismo de agência, de crença e formas híbridas de internalismo.

⁵ Isto é, de acordo com o internalismo de crença, é uma verdade necessária que se uma pessoa acredita que ela tem uma obrigação moral de fazer alguma coisa, então ela tem alguma motivação para fazê-lo (ou teria alguma motivação sob condições relevantes). De acordo internismo híbrido, é uma verdade necessária que se uma pessoa acredita realmente que ela tem uma obrigação moral de fazer algo, então ela tem alguma motivação para fazê-lo (ou teria alguma motivação sob condições relevantes).

⁶ Por exemplo, Michael Smith defende uma versão do internalismo de crença de acordo a qual, se uma pessoa não está realmente motivada a fazer o que ela acredita que seria certo, então ela é

internalista evitar banalizar sua posição. É trivial que uma pessoa com uma obrigação moral, necessariamente, adquira a motivação relevante sob *algumas* circunstâncias, mas devemos tomar uma posição como internalista apenas se fizer uma afirmação não trivial sobre a conexão entre obrigação e motivação. Isso significa que a distinção entre internalismo e externalismo não é muito nítida por causa da imprecisão na noção de trivial, mas a literatura felizmente contém casos claros de internalismo e externalismo.

Como eu disse, *internalismo* é usado, às vezes, para nomear uma tese sobre uma conexão necessária entre obrigação e *razões*. Podemos distinguir versões do internalismo de razões que correspondem às versões de internalismo motivacional. Existe uma versão do internalismo do agente, bem como uma versão do internalismo de crença e uma versão híbrida.

Existem complicações relacionadas à noção de uma razão. Se alguém tem uma obrigação moral de fazer algo, é trivial que ele tenha uma razão *moral* para fazê-lo. Porém, eu acredito que o internalismo de razões poderiam, normalmente, levar a implicar que uma obrigação moral ou uma crença de que alguém tem uma obrigação garante uma razão-sem-qualificação - uma *Razão* que motivaria qualquer um que age conforme requerido, se fosse racional.⁷ Alguns filósofos combinam o internalismo de razões com uma doutrina internalista sobre a conexão entre *razões* e *motivação*, uma doutrina no sentido de que a motivação é *interna* ao estado de ter uma razão ou crer ou crer realmente que alguém possui uma razão.⁸ Eu ignorarei amplamente tais doutrinas e limitarei a atenção às doutrinas internalistas sobre a obrigação.

"praticamente irracional". Ou seja, suas motivações estão sujeitas a uma "influência distorcida" de alguma forma de "desrazão prática", como fraqueza de vontade. Smith, M. *The Moral Problem*, 61.

⁷ Como veremos, esta visão parece ser mantida por Smith. Smith, *The Moral Problem*, 62. Neste contexto, Smith cita Christine Korsgaard "Skepticism about Practical Reason," *Journal of Philosophy* 83 (1986) 5-25.

⁸ Tal visão é defendida por Bernard Williams em "Internal and External Reasons" em sua obra *Moral Luck*. Cambridge: Cambridge University Press, 1981.

Mais especificamente, irei me concentrar na versão do internalismo de crença da motivação até o final.

De acordo com essa doutrina, existe uma conexão *necessária* entre a crença de uma pessoa que ele possui uma obrigação e que ela está apropriadamente motivada. Considere, no entanto, o caso de Alice como um contra-exemplo putativo: Alice foi criada para acreditar que a teoria do comando divino é correta. Ou seja, como a própria Alice poderia dizer, ela foi criada para acreditar que nossas obrigações morais são determinadas pelos mandamentos de Deus. Ela também foi criada para acreditar que Deus é um governante vingativo e que Ele nos julgará [segundo o princípio] *olho por olho*. Segundo este princípio de *olho por olho*, Alice acredita que a pena capital é obrigatória em casos de assassinato, e ela acredita que tem a obrigação de apoiar a pena de morte. Entretanto, ela é profundamente piedosa, e ela é não nutre qualquer simpatia com o que ela considera ser a vingança de Deus. Por causa de sua compaixão, ela não está motivada nem um pouco a apoiar a pena de morte. Na verdade, ela é ativa em se opor a ela, mesmo que acredite que seja moralmente proibido fazer isso.

Externalistas estariam inclinados a aceitar o estado de espírito de Alice como psicológica e logicamente possível, mas uma internalista de crença teria que rejeitar o exemplo. De fato, qualquer versão plausível do internalismo de crença é uma doutrina qualificada no sentido de que uma crença na moral obrigação implica em motivação *sob condições relevantes*. Não obstante, poderíamos, presumivelmente, preencher nossa descrição do caso para que Alice esteja, claramente, sob as condições relevantes requeridas pelo internalista. Se fizéssemos isso, então o internalista teria que alegar que Alice não está genuinamente acreditando que ela é moralmente obrigada a apoiar a pena de morte ou que, se ela acredita nisso, ela está de fato *motivada* a apoiá-la em algum grau mínimo. Por exemplo, Michael Smith argumentaria que Alice não pode realmente acreditar que ela possui uma obrigação *moral* de se opor à pena de morte ou que ela não *realiza* um julgamento *moral*

quando ela diz que está *moralmente obrigado* a se opor à pena de morte.⁹ Seu argumento depende no *requerimento de praticidade* (*practicality requirement*), que discutirei mais adiante. Em geral, o internalista deve insistir que o conceito ou a natureza da obrigação moral exclui a possibilidade lógica de Alice ter a psicologia atribuída a ela no exemplo citado.¹⁰

Um internalista normalmente consideraria que sua doutrina está vinculada a uma análise adequada do conceito de obrigação moral.¹¹ Entretanto, porque este conceito é em si mesmo objeto de controvérsia um internalista não pode defender sua doutrina simplesmente derivando-a da análise proposta do conceito. Como veremos, alguns internalistas oferecem explicitamente argumentos independentes de suas doutrinas, embora tais argumentos não pareçam ser os fonte de sua crença no internalismo, pois eles, normalmente, não estão inclinados abandonar o internalismo quando se descobre que o argumento é infundado. Além disso, para defender o internalismo de crença frente ao caso de Alice, o internalista precisa argumentar que Alice não pode ter o conceito de obrigação moral ou que suas palavras não expressam o julgamento moral que parecem expressar. Consequentemente, o internalismo de crença precisa estar apoiado por uma explicação sobre o que significa ter um conceito, bem como por uma explicação da semântica do julgamento moral.

Na verdade, eu acredito que a discordância entre internalistas e externalistas é motivada por diferenças sobre as

⁹ Smith, *The Moral Problem*, 66-71. Smith está respondendo ao argumento de David Brink.

¹⁰ Desentendimentos desse tipo entre externalistas e internalistas levaram W.O. Falk a concluir, com efeito, que externalistas e internalistas têm diferentes conceitos de obrigação moral. Mas, se assim fosse, então os externalistas e os internalistas realmente não discordam. Eles estariam apenas fazendo alegações sobre diferentes tipos de obrigação. Penso, pelo contrário, que eles têm um desacordo genuíno. Confira W.O. Falk, “‘Ought’ and Motivation”, *Proceedings of the Aristotelian Society* N.S. 48 (1947-48) 137, 124-5.

¹¹ Smith diz que “nosso conceito de correção” suporta uma “conexão conceitual entre o julgamento moral e a vontade” e a “conexão conceitual entre ... os fatos morais ... e nossas razões para a ação”. Smith, *The Moral Problem*, 61, 65, 37-8. Brink vê as versões do internalismo como doutrinas sobre o conceito de obrigação moral. Brink, *Moral Realism and the Foundations of Ethics*, 40.

principais questões metaéticas do que por uma diferença que está restrita a uma questão analítica sobre o conceito compartilhado de obrigação moral. W.K. Frankena disse:

Cada teoria [intemalista e extemalista] teoria tem pontos fortes e fracos, e decidir entre eles envolve determinar seus valores totais relativos enquanto modelos de moralidade. Mas tal determinação (...) exige uma investigação muito ampla (...) sobre a natureza e função da moralidade, do discurso moral e da teoria moral (...).¹²

Devemos perceber que “nenhum tipo de filosofia moral pode ser decisivamente refutada pelo outro, e (...) devemos abandonar a busca de tais certezaa no sentido de não mais esperar por tais refutações”.¹³ Na descrição de Frankena, as questões levantadas pelo debate entre o internalismo e o externalismo se ramificam ao longo da teoria metaética. Uma posição no debate não pode ser adequadamente defendida sem defender uma teoria metaética que aborde questões sobre a moral verdade, realismo, normatividade e objetividade, entre outros.

Para ver isso, considere a ideia de que as reivindicações de obrigação moral são normativas. Alguns internalistas sustentaram que a explicação mais plausível disso construiria o julgamento moral, na medida em que é normativo, como uma expressão de uma atitude, em vez de uma expressão da crença em um proposição. Agora, considerando-se que a visão de que reivindicações morais expressam proposições é geralmente chamado de *cognitivismo* ou *descritivismo*, tais internalistas seriam não-cognitivistas.¹⁴ Eles estão argumentando que o não-cognitivismo é apoiado pela normatividade do juízo moral. Desta forma, o debate entre internalismo e externalismo estaria ligado

¹² Frankena, “Obligation and Motivation in Recent Moral Philosophy”, 73.

¹³ *Ibidem*, 69.

¹⁴ R.M. Hare e Allan Gibbard são internalista que rejeitam (ou parecem rejeitar) o cognitivismo. Confira R.M. Hare, *The Language of Morals*. Oxford: Oxford University Press 1952; Allan Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press 1990.

aos debates sobre verdade moral e realismo moral, considerando-se que o realismo moral e a ideia de que existem verdades moral pressupõem cognitivismo.¹⁵

Agora, considere a ideia de que as alegações de obrigação moral são *categóricas* ao menos neste sentido: uma pessoa conceitualmente competente, que produziu sinceramente uma reivindicação moral no sentido de que ele é obrigado a fazer algo, não poderia retratar (apropriadamente) sua alegação meramente ao perceber que ele não tem qualquer motivação para fazê-lo. Se isto estiver correto, então nenhuma forma simples de internalismo de crença pode ser verdadeiro, pois, sobre uma forma simples de internalismo de crença, uma pessoa não iria realmente acreditar que ele tem uma obrigação se ela não estivesse motivada adequadamente. Quer dizer, se uma pessoa perceber que não está motivada da forma correta, o que ela percebe implica que ela não tem a crença de que ela está obrigada a agir daquela forma e, por isso, seria apropriado para ele deixar de expressar a crença. O debate entre o internalismo e externalismo está ligado, desta forma, à ideia de que as reivindicações de obrigação são categóricas. Internistas de crença devem proteger suas reivindicações sobre a conexão entre obrigação e motivação ou, então, explicar a aparência de que as reivindicações de obrigação são categóricas.

Se aceitarmos o internalismo e o cognitivismo e também pensarmos que as afirmações de obrigação moral são categóricas no sentido que expliquei, nós estamos no caminho para aceitar as premissas do argumento da estranheza (*argument from queerness*) de J.L. Mackie, segundo o qual pensou nenhuma proposição moral é verdadeira.¹⁶ Assim, pode parecer que apenas uma posição externalista pode reivindicar o julgamento moral como cognitivo, categórico e, às vezes, correto.

¹⁵ Eu não me preocupei em formular o internalismo para ser compatível com não-cognitivismo, porque eu estou assumindo uma visão cognitivista.

¹⁶ J.L. Mackie, *Ethics: Inventing Right and Wrong*. Harmondsworth, Middlesex: Penguin 1977, 38-42.

Dadas às questões em jogo, parece claro que o debate sobre internalismo não será resolvido baseando-se em argumentos finalmente refinados e vistos como anteriores ou impondo restrições ao debate sobre a natureza da moralidade. Frankena disse que o debate “não pode ser resolvido, como muitos parecem pensar, por meio de (...) uma pequena escala lógica ou argumento semilógicos (...)”¹⁷ Da mesma forma, o debate não será resolvido através de contra-exemplos putativos. Em vez disso, devemos considerar a questão em um “plano macroscópico em vez de microscópico.”¹⁸ É o que proponho fazer neste texto, ainda que admitidamente de forma incompleta. Eu irei propor uma teoria externalista de julgamento e convicção moral e, em seguida, tentar mostrar como esta teoria pode lidar com certas objeções internalistas.

2. Uma Teoria do Julgamento Normativo Baseado-em-Padrão¹⁹

Eu não pretendo fornecer um argumento decisivo para o externalismo ou um argumento decisivo contra o internalismo. Em vez disso, nesta e na próxima seção, eu oferecerei uma explicação parcial do conteúdo das proposições sobre obrigação moral, bem como uma explicação sobre convicção moral. Eu quero mostrar que estas explicações, apesar de externalistas, preservam muitas das intuições que motivam o internalismo. Se eu estiver correto, meu argumento mina o internalismo na medida em que seu apelo é devido ao pensamento de que apenas uma teoria internalista pode preservar as intuições.

Vou simplificar as coisas de várias maneiras. Primeiro, vou restringir a atenção para reivindicações nas quais dizemos que um agente tem uma obrigação moral. Frankena chama tais alegações

¹⁷ Frankena, “Obligation and Motivation in Recent Moral Philosophy”, 69.

¹⁸ *Ibidem*, 49.

¹⁹ Nas duas seções seguintes, apresento idéias que são desenvolvidas de forma mais completa em meu livro *Morality, Normativity, and Society*. New York: Oxford University Press, 1995.

de *juulgamentos de obrigação moral*²⁰, desconsiderando os muitos outros tipos de reivindicações morais. Em segundo lugar, ignorarei todas as diferenças que pode haver entre julgamentos sobre obrigações morais, julgamentos sobre deveres morais e julgamentos no sentido de que um agente deve moralmente realizar determinada ação. Vou me referir a todos eles como julgamentos de obrigação moral.

Eu também assumirei que o *cognitivismo* ou *descriptivismo* são verdadeiros e que proposições morais expressam proposições. Quando fazemos uma reivindicação moral, nós proferimos sentença declarativa e, assim, afirmamos algo. É razoável supor que o que afirmamos é verdadeiro ou falso, que expressamos uma proposição, assim como fazemos em casos não-morais.²¹ Preciso supor que cognitivismo está correto para que eu possa avançar de forma eficiente para as questões que eu deseja abordar.

Eu proponho o que se segue como uma explicação parcial do conteúdo ou das condições de verdade das proposições expressas pelos juízos de obrigação moral:

Uma proposição de que um agente tem uma obrigação moral de fazer algo implica (em um sentido não trivial) que existe um padrão moral justificado que exige que o agente faça isso.

Várias ideias precisam ser explicadas. Para começar, deixe-me enfatizar que estou propondo apenas uma condição para a verdade de uma proposição de obrigação moral. Parece que isso vai representar um problema para o meu argumento, mas as outras condições que são necessárias para a verdade de tais proposições são irrelevantes para o que eu tentarei mostrar. Compare, por exemplo, a proposição de que a tentativa de Jimmy Carter para resgatar os reféns era moralmente obrigatória e a proposição de que o

²⁰ Frankena, "Obligation and Motivation in Recent Moral Philosophy", 71.

²¹ Confira Nathan Salmon e Scott Soames, "Introduction," in Nathan Salmon and Scott Soames (eds.), *Propositions and Attitudes*. New York: Oxford University Press, 1988, 1.

reconhecimento de Richard Nixon da China cumpriu um dever moral. Estas são proposições, obviamente, distintas e com diferentes condições de verdade. Ainda assim, a diferença entre ambas não é relevante para qualquer coisa que eu discutirei a seguir.

Deixe-me explicar agora a parte principal da minha proposta, começando com a noção de um padrão. Regras, normas e imperativos são exemplos de padrões. A noção de norma para Allan Gibbard parece ser a mesma que a minha noção de padrão. Gibbard diz que uma norma é “uma possível regra ou prescrição expressa por um imperativo.”²² Esta formulação seria adequada para explicar o que quero dizer com um padrão. Dizer que um padrão *exige* uma ação é dizer que o padrão está em conformidade apenas no caso da ação ser executada. Por exemplo, o imperativo “Bata a porta!” pede que a porta seja fechada.

Precisamos da noção de um padrão na teoria semântica para oferecer uma explicação sobre o que é expresso por sentenças imperativas – embora, obviamente, não precisamos chamar o que é expresso como *padrão*. Frege disse: “Não devemos desejar negar sentido a um comando, mas esse sentido não é tal que a questão da verdade possa surgir para ela. Portanto, eu não posso chamar o sentido de um comando de um pensamento.”²³

Frege parece afirmar que, embora as sentenças imperativas expressem algo ou possuem *sentido*, eles não expressam proposições, porque, diz ele, a questão da verdade não surge com relação ao que quer que seja expresso por uma sentença imperativa. Na minha terminologia, tais sentenças expressam *padrões*. Assim como as proposições são expressas por sentenças declarativas típicas, os padrões são expressos por sentenças imperativas típicas.

Para um padrão se qualificar como justificado não é necessário que alguém tenha apresentado qualquer argumento

²² Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment*, 46.

²³ Gottlob Frege, “Thoughts” (tradução de P. Geach e R.H. Stoothoff), em Nathan Salmon and Scott Soames (Eds.), *Propositions and Attitudes*. New York: Oxford University Press, 1988, 37.

sobre ela ou provado qualquer coisa sobre a mesma. Para um padrão estar justificado significa que ele possui um *status*, o status de ser moralmente *obrigatório*. A ideia é que, a menos que um padrão tenha esse status, as proposições correspondentes de obrigação não são verdadeiras.

O status em questão não é determinada por padrões arbitrários, inventados ou injustificados, pois não acreditamos que as proposições morais correspondentes as eles sejam verdadeiras. Por exemplo, nós não pensamos que os padrões aceitos por Nietzsche que nos chamam a nos esforçar para nos tornar o *super-homem* têm o status que eles precisariam para que seja verdade que estamos obrigados a nos esforçar para nos tornar o super-homem. Não acreditamos que tenhamos qualquer obrigação de nos esforçar para nos tornarmos tal. Por outro lado, acreditamos que temos a obrigação de nos opor à escravidão e podemos formular um padrão relevante que nos convide a nos opor à escravidão. Concordamos que esse padrão está justificado ou é moralmente obrigatório.

Em meu livro recente, forneço uma teoria substantiva do status que padrões justificados possuem.²⁴ Aqui, preciso ser menos ambicioso. Felizmente, muitas explicações diferentes da justificação de padrões morais são familiares na literatura sobre o tema; e, embora eles não são geralmente descritos como teorias da justificação dos padrões morais, ainda assim podem ser utilmente considerados como tais na estrutura que estou propondo.

Evidentemente, é possível tomar nossa compreensão do status de *estar justificado* como sendo derivativo de nossa compreensão sobre verdade moral. Nós poderíamos insistir que, se uma proposição de obrigação moral é verdadeira, então existe um padrão moral justificado a ela relacionado e não poderíamos dizer mais nada sobre o que o status de estar justificado consiste. Teríamos que oferecer alguma explicação das condições sob as

²⁴ Cf. meu livro *Morality, Normativity, and Society*.

quais uma proposição de obrigação moral seria verdadeira. Mas, novamente, poderíamos refinar essa questão simplesmente afirmando que tal proposição é verdadeira apenas no caso de a ação relevante ter a propriedade de ser moralmente obrigatória. Poderíamos deixar por isso mesmo, embora isso não seja filosoficamente esclarecedor.

Mais úteis para meus propósitos são as teorias das escolas kantiana, aristotélica ou teoria da escolha racional, incluindo os modelos contemporâneos hobbesianos, na medida em que tais teorias podem ser vistas como fornecendo explicações sobre as condições sob as quais um padrão moral seria justificado. Um kantiano pode dizer que um padrão moral é justificado apenas no caso de uma pessoa agir a partir dele e poderia “ao mesmo tempo querer que se tornasse uma lei universal.”²⁵ Ou, um kantiano poderia dizer que um padrão moral está justificado apenas no caso de que um agente totalmente racional, com incentivos apenas puramente racionais, pretenderia cumpri-lo.²⁶ Um aristotélico poderia dizer que um padrão moral está justificado apenas no caso de qualquer pessoa que vive uma vida que seria preenchida por um humano teria, por esse fato, intenções que o levariam a se conformar com a norma.²⁷ Um hobbesiano contemporâneo, como David Gauthier, poderia dizer que um padrão moral está justificado apenas no caso em que um agente, ao maximizar a satisfação de suas preferências, seria racional ao se dispor a cumpri-lo.²⁸

De minha parte, defenderei uma teoria de acordo com a qual um padrão moral está justificado em relação a uma sociedade apenas no caso em que, se esta sociedade escolhesse um código moral para observá-lo, seria racionalmente necessário para ela

²⁵ Kant, Immanuel. *Fundamentação da Metafísica dos Costumes*. Lisboa: Edições 70, Ak 421.

²⁶ *Ibidem*, Ak 449.

²⁷ Espero que isto seja, ao menos, vagamente aristotélico.

²⁸ Gauthier, David. *Morals by Agreement*. Oxford: Oxford University Press, 1986. Confira o Capítulo 06.

selecionar um código moral que contém ou implica um tipo de padrão. Obviamente, não posso defender essa abordagem aqui.

Estas são versões extremamente abreviadas e cruas das teorias em questão, mas os detalhes não importam para os propósitos atuais. Cada uma dessas teorias oferece uma explicação diferente do status possuído por um padrão moral daquilo que é moralmente obrigatório. Entretanto, se pudermos entender a natureza do debate entre essas teorias, e entendê-la como relevante para a questão sobre quais proposições morais são verdadeiras e em que base elas são verdadeiras, então a noção de justificação de um padrão moral não deveria ser um mistério.

Qualquer teoria das condições sob as quais os padrões morais estão justificados levantaria profundas questões filosóficas sobre a natureza da moralidade. De todo modo, minha proposta de condição necessária para a verdade de um julgamento de obrigação é neutra em relação a estas diferentes teorias. Eu acredito que, uma vez que está claro qual é a minha proposta, não deve parecer controverso exceto para aqueles que duvidam do cognitivismo que estou assumindo aqui.

Agora, eu afirmo que o fato de que uma proposição de obrigação moral implica (não trivialmente) que alguma norma relevante está justificada é o que explica sua normatividade, bem como a normatividade das proposições que utilizamos para expressá-la. Uma proposição normativa é intrinsecamente relevante para a ação ou a escolha. O fato de que uma proposição de obrigação moral implica que alguns normas relevante estão justificadas é o que explica a conexão entre uma proposição de obrigação e ação. Isso significa que se uma pessoa tem uma obrigação moral de fazer alguma coisa, então há um padrão que *chama* a pessoa fazer tal coisa, e esse padrão está *justificado*. Chame isso de teoria normativa “baseada em padrões”.²⁹

²⁹ Na verdade, é um esquema, haja vista que precisa ser completado por uma teoria de justificação para padrões morais. Eu forneço uma análise completa em *Morality, Normativity, and Society*.

Uma intuição muito poderosa compartilhada por internalistas é que proposições de obrigações morais são *intrinsecamente normativas*. Mackie diz que propriedades morais são *intrinsecamente orientadoras de ação*. Ele diz que uma ação errada “não seria indiferente de algo embutida nela mesma”³⁰. Jonathan Dancy argumenta que “a moralidade é essencialmente prática, de modo que seria estranho alguém dizer ‘Essa ação é errada, mas eu não a vejo de modo relevante para a minha escolha.’”³¹

A intuição de que a normatividade é interna às proposições de moral é defendida pelo modelo baseado-em-padrão. Tal defesa implica que uma proposição moral é normativa em virtude de seu conteúdo, em virtude do fato de que implica (em um sentido não trivial) que um padrão moral relevante está justificado. Além disso, se estendêssemos a conta de forma natural Assim, dir-se-ia que as propriedades morais são *intrinsecamente normativas*. Deixando de lado detalhes irrelevantes, uma ação tem a propriedade de ser obrigatória se e somente se for exigida por um padrão moral justificado, um padrão de acordo com o qual isto deve ser feito. A erradeza (*wrongness*) de uma ação é *relevante* para a escolha, pois ela implica que a ação é impedida por um padrão moral justificado. Além disso, uma extensão natural desta explicação implica que se uma pessoa tem uma obrigação moral de fazer algo, ela tem uma *razão moral* para fazê-lo. Como argumentarei, uma pessoa tem um *razão moral* para fazer algo apenas no caso em que um padrão moral justificado pede para ela fazer isso.

Tudo isso é totalmente compatível com o externalismo. Quanto à motivação, tal modelo deixa em aberto a possibilidade de uma pessoa ser moralmente obrigada a fazer alguma coisa e ainda

³⁰ Mackie, *Ethics: Inventing Right e Wrong*, 32, 40, respectivamente.

³¹ Jonathan Dancy, *Moral Reasons* (Oxford: Blackwell 1993) 4. Dancy acrescenta que “considerações morais são aquelas cuja relevância prática não pode ser perdida, dizendo: ‘Eu não me importo com esse tipo de coisa’”. ‘Mas se as afirmações morais são categóricas no sentido que eu expliquei anteriormente, então algo mais forte é verdadeiro, a saber, a ‘relevância prática’ de uma proposição de obrigação moral não podendo ser dito que “eu não estou de todo motivado por esse tipo de coisa”. Isso prejudicaria o internacionalismo em vez de ajudá-lo.

não estar motivada em qualquer grau a fazê-lo. Ele também deixa em aberto a possibilidade de uma pessoa poder acreditar que ela é moralmente obrigada a fazer alguma coisa e ainda não estar motivado em qualquer grau para fazê-lo. Possuir esta crença é simplesmente aceitar uma proposição que implica que um padrão justificado pede alguma ação, e nada sobre as motivações daquele que acredita são implicadas por sua aceitação de tal proposição. Voltando agora para o internalismo de razões, uma pessoa pode ter uma obrigação moral e, portanto, ter uma razão *moral* para agir de acordo sem ter um razão-sem-qualificação, uma Razão que motivaria qualquer agente racional. E ela pode não ter uma razão de *autointeresse* ou, para usar uma terminologia que considero mais adequada, uma *razão autofundamentada*, uma razão fundamentada em sua própria natureza. Finalmente, essa proposta deixa em aberta a questão sobre se uma pessoa que acredita que possui uma obrigação moral pode não ter uma Razão ou uma razão autofundamentada para agir adequadamente.

Falando da maneira correta, com certeza o modelo baseado-em-padrão é neutro em relação ao internalismo e ao externalismo. Quando combinado com determinadas teorias da justificação, o modelo não implica qualquer doutrina internalista. Por outro lado, quando combinado com outras teorias da justificação, isso implica também em doutrinas internalistas. Na verdade, qualquer doutrina internalista é presumivelmente implicada por alguma teoria da justificação quando é combinada com o modelo baseado-em-padrão. David Gauthier, por exemplo, pretende mostrar que “todos os deveres” de moralidade “também são verdadeiramente endossados na razão de cada indivíduo.”³² Sua descrição da moralidade implica uma versão do internalismo de razões e ainda assim é compatível com a minha abordagem baseada em padrões. No entanto, ao usar o modelo baseado-em-padrões para explicar a normatividade do julgamento moral, eu não utilizei nenhuma ideia

³² Gauthier, *Morals by Agreement*, 1

internalista. Essa explicação foi extenalista e não pressupõe nada sobre motivações ou Razões.

Eu não descrevi em detalhes minha teoria da justificação dos padrões morais, e eu dei apenas uma condição de verdade necessária às reivindicações de obrigação. Portanto, pode ser objetado que, por tudo que eu disse, internalismo pode fazer parte da explicação da normatividade através da minha teoria da justificação ou através de alguma condição necessária oculta para a verdade de uma afirmação de obrigação moral.³³ Evidentemente, eu não posso provar que isso não seja assim. No entanto, pode-se ver pela análise da minha visão centrada na sociedade que a explicação da justificação dos padrões morais é externalista. Isto não implica qualquer doutrina internalista - exceto por doutrinas que não estão seriamente em disputa, como a doutrina de que qualquer pessoa com uma obrigação moral possui uma razão moral correspondente.

Minha principal alegação é de que a normatividade de uma proposição da obrigação moral é explicada pelo fato de que ela envolve, em um sentido não trivial e em virtude de seu conteúdo, que algum padrão moral relevante seja justificado. Como eu disse, um padrão no sentido de que algo que deve ser feito *chama por* uma ação, o que dá sentido à ideia de que tal padrão *direciona* a escolha da ação.³⁴ Em resumo, o modelo baseado-em-padrões explica que proposições morais são intrinsecamente normativas e que são capazes de direcionar escolhas em virtude de seu conteúdo. Esta é uma versão cognitivista. E, combinado com uma teoria apropriada de justificação, isso implica que as afirmações morais são categóricas no sentido que expliquei acima.

³³ Tanto Earl Conee quanto Michael Smith fazem essa objeção.

³⁴ Claro, alguns padrões estão preocupados com outras coisas além de ações, tais como estados de caráter ou estrutura das instituições.

3. Crença moral e convicção

As teorias não cognitivistas de Allan Gibbard e R.M. Hare são semelhantes à teoria baseada em padrões, haja vista que elas também utilizam a ideia de um padrão³⁵. Porém, Gibbard e Hare vêem a normatividade do julgamento moral como decorrente do assentimento a um padrão, e não do modo como os padrões entram na semântica das reivindicações morais. Gibbard afirma que a aceitação de ‘normas’ envolve um elemento de endosso, para ele este é “o elemento especial que torna o pensamento normativo e a linguagem normativa”³⁶. Hare afirma que quando concordando sinceramente com um comando endereçado a nós mesmos estamos envolvidos em fazer ou resolver aquilo que foi dito para fazermos³⁷. Esse tipo de teoria sustenta que quando a pessoa faz tal julgamento, ela está expressando seu padrão de endosso, que por sua vez, representa a normatividade dos juízos de obrigação moral.

Muitas vezes tomamos o comportamento das pessoas como prova de suas crenças morais. Para não-cognitivistas como Gibbard e Hare isso é explicado observando a ‘crença’ moral como, essencialmente, uma questão de endosso de uma norma em vez de, uma questão de aceitar uma proposta. Para endossar uma norma é presumivelmente, ao menos em parte, estar disposto a agir apropriadamente. É claro que eu rejeito o não-cognitivismo e o internalismo. Mas parece plausível que uma pessoa que acredita sinceramente que tem uma obrigação moral normalmente estaria motivada a agir de acordo.

A teoria baseada em padrões pode aceitar isso apesar de insistir, ao contrário do não-cognitivismo, que a crença moral é

³⁵ Gibbard, *Wise Choices, Apt Feelings*. Hare, *The Language of Morals*. Gibbard chama de padrões “normas”, e Hare chama de “imperativos”. Hare usa o termo ‘padrões’ de uma maneira diferente, isto é, imperativos morais que têm moeda em uma sociedade ou que são subscritos por uma pessoa. (Hare, *The Language of Morals*, 7).

³⁶ Gibbard, *Wise Choices, Apt Feelings: A Theory of Normative Judgment*, 33.

³⁷ Hare, *The Language of Morals*, 19-20.

literalmente a *crença*. O padrão baseado em um ponto de vista se distingue entre a *proposição* expressa por uma pessoa em fazer uma reivindicação de obrigação moral e o *padrão* ‘correspondente’, cuja justificação é implicada por essa proposição. Claro, não faz sentido supor que alguém acredite em um padrão, mas uma pessoa que sinceramente faz uma reivindicação moral normalmente se baseia em um padrão relevante. Assim, em casos típicos, onde uma pessoa faz uma reivindicação de obrigação moral sincera, sua proposição expressa e assinala o padrão correspondente. Em tais casos, direi, sua reivindicação expressa uma ‘convicção moral’, uma combinação de crença e assinatura para o padrão correspondente.

A idéia de “assinatura” de um padrão é semelhante à ideia de Gibbard da aceitação da norma, e da ideia de Hare de concordar com um comando. O ponto importante para os propósitos atuais é que quando uma pessoa ‘assina’ um padrão para o qual ela pode seguir pretende ou segui-lo ou adequar-se a ele. Especificamente, uma pessoa ‘assina moralmente’ um padrão somente se,

- (1) Ela pretende se adaptar ou faz uma política para estar em conformidade, assumindo que é um padrão ao qual ela pode seguir, e
- (2) Ela pretende apoiar a conformidade dessa ou faz uma política para apoiar a conformidade.

Os detalhes restantes que seriam necessários em uma compreensão completa não serão assuntos para os propósitos atuais.

Eu estou dizendo, então, que em casos típicos onde uma pessoa é sincera em fazer uma reivindicação de obrigação moral, ela tem ‘convicção moral’, o que implica na assinatura do padrão correspondente. Essa pessoa é, portanto, motivada a se adequar ao padrão, pois pretende se conformar. Uma pessoa com a convicção de que ela tem a obrigação moral de fazer algo é, portanto, motivada em algum grau para fazê-lo.

Dada a maneira pela qual adquirimos atitudes e crenças morais, será dada a dissonância psicológica que seria experimentada por uma pessoa que não subscreveu padrões correspondentes à sua crença moral - comprometida em ver como justificada - é de se esperar que as pessoas normalmente assinem esses padrões. Tipicamente, crenças morais são convicções completas. Pais que ensinam as crianças sobre moralidade presumivelmente tentam criar crenças apropriadas, pois raciocinamos questões morais, nesse sentido, as proposições morais que acreditamos são os veículos para o nosso raciocínio. No entanto, é pelo menos igualmente importante ensinar sobre moralidade para fazer com que as crianças subscrevam padrões apropriados correspondentes às crenças que queremos que elas tenham. Queremos que as pessoas vivam as suas vidas de acordo com certos padrões e tenham certas políticas que regem o seu comportamento, ou seja, queremos que eles se inscrevam nos padrões relevantes. Nós queremos que eles acreditem que esses padrões são justificados ou garantidos, o que equivale a dizer que queremos que eles tenham crenças morais correspondentes. Idealmente queremos que eles se inscrevam para os padrões porque acreditamos que sejam justificados. Portanto, é de se esperar que a educação moral e os mecanismos sociais que apoiam e reforçam a educação moral visaria à inculcar crença moral, bem como a assinatura dos padrões correspondentes.

Existe, no entanto, uma possibilidade real de uma pessoa acreditar que ele tem uma obrigação moral sem subscrever o padrão correspondente ou vice-versa. Os melhores exemplos disso são casos em que uma pessoa tem uma crença metaética que está em tensão com suas atitudes morais subjacentes.

Considere novamente o caso de Alice. Alice foi criada para acreditar na teoria divina do comando e vê Deus como um governante vingativo que nos conduzirá à pena capital. Alice ainda acredita em tudo isso, mas ela é uma pessoa gentil e compassiva. Sua compaixão a leva a falhar para apoiar a pena de morte e ela

não pretende, de fato, apoiar ou torná-la uma política. Portanto, ela não se inscreve no padrão, pois ela não apoia a pena capital mesmo acreditando que esse comportamento é moralmente obrigatório.

O caso de Huckleberry Finn pode ilustrar a mesma possibilidade. Huck acredita que é obrigado a entregar Jim às autoridades porque Jim é um escravo fugitivo, mas Huck não o entrega. Claro, seu fracasso em entregar Jim não é suficiente para mostrar que ele não tem uma política de entregar escravos fugidos, pois às vezes não nos conformamos com nossas próprias políticas pessoais, mas certamente podemos imaginar que Huck não tem uma política de entregar escravos ou a intenção de entregá-los. Portanto, há uma elaboração do caso de Huck em que, embora ele acredite ser obrigado a entregar escravos fugitivos às autoridades, não se inscreve no padrão correspondente.

Finalmente, considere o caso de Bill. Bill é um niilista ingênuo que está convencido que nenhum padrão moral é justificado e, portanto, que não (basicamente) há proposição moral verdadeira. No entanto, ele foi criado com valores morais e os possui. Ele subscreve uma variedade de padrões incluindo, por exemplo, uma proibição de pena de morte. Ele iria admitir que seja moralmente contra a pena capital, mas negaria que a pena de morte está errada ou que é moralmente obrigado a opor-se, e negaria que *acredita* nessas coisas. Agora, ele pode ser autoenganoso sobre o que acredita, mas eu estipulo, neste caso, que Bill está correto. Ele não acredita que a pena capital esteja errada. Ele é comprometido por seu niilismo a acreditar que é falso que a pena de morte está errada, por isso seria inconsistente para ele acreditasse que a punição capital está errada. Ele certamente não está comprometido com essa inconsistência simplesmente em virtude de sua oposição moral à pena capital - em virtude de sua assinatura de um padrão que proíbe a pena de morte.

Mackie observa que “visões de primeira e segunda ordem ... são completamente independente Um homem poderia ter fortes visões morais ... enquanto acredita que eles eram simplesmente

atitudes e políticas em relação à conduta que ele e outras pessoas mantem”³⁸. Não é óbvio como melhor interpretar essa observação. Na minha terminologia, Mackie pode estar apontando para a possibilidade de ter crenças morais ao acreditar (falsamente) que há “simplesmente” o fato de que subscrevemos certos padrões. Alternativamente, ele pode estar apontando para a possibilidade de subscrever padrões morais, enquanto ao mesmo tempo, acredita (novamente falsamente) que não há verdades morais sem crenças morais, estritamente falando, e que existiria “simplesmente” o fato de que, por acaso, subscrevemos certos padrões. Isso me parece ser uma virtude que permitiria ver posições possíveis ou pelo menos coerentes.

Os exemplos de Alice, Huck e Bill sustentam minha ideia de que convicções morais não são simplesmente crenças. Claro, eu digo isso por estipular que ‘convicção moral’ consiste em uma crença assinada, mas acredito que essa estipulação não é artificial. Existe um complexo estado de acreditar numa proposição moral e subscrever o correspondente padrão. É este estado que normalmente estamos interessados quando tentamos determinar a natureza das convicções morais de uma pessoa. Pois é este estado, não o estado mais simples de acreditar em uma proposição moral, que implica motivação. Uma pessoa com a convicção acerca de uma obrigação moral de fazer algo é motivado a fazer isso em algum grau. Uma pessoa que meramente tem crença pode não estar motivada a fazer algo. Pois, como os exemplos sugerem, embora a crença moral seja normalmente acompanhada de assinatura da norma correspondente, não necessita ser.

Posições externas geralmente supõem que nossa motivação para obrigações é um desejo de agir moralmente, e interpretam esse desejo externo à crença moral. Eu concordo, claro, que a *crença* moral não garante, por si só, qualquer motivação para cumprir nossas obrigações. Mas a *convicção* moral é o estado mais complexo

³⁸ Mackie, *Ethics: Inventing Right and Wrong*, 16.

que combina crença com assinatura, pois convicção moral não garante motivação para cumprir as obrigações de que estamos “convencidos”. Portanto, podemos ser internalistas sobre convicções morais, mesmo se somos externalistas sobre crenças morais, e nós não precisamos aceitar a psicologia moral crua dos padrões externalistas. Talvez as pessoas tenham o desejo de agir moralmente, mas a motivação moral em casos típicos é explicada pela assinatura de um padrão que é constitutivo da convicção moral.

4. Objeções Internalistas

Seguindo Frankena, recomendei que adotássemos uma perspectiva “macroscópica” sobre o debate entre internalismo e externalismo. Idealmente, avaliá-riamos a minha proposta como um todo, comparando suas vantagens gerais e desvantagens com as dos seus concorrentes. Não posso esperar fazer isso em um ensaio, é claro. No entanto, um de meus objetos é minha proposta desse externalismo sem compará-lo completamente ao que foi desenvolvido pelos concorrentes internalistas. Para fazer isso, proveitosamente, seria necessário argumentar diretamente por uma doutrina internalista. Muitos internalistas confiam mais em apelos à intuição do que em argumentos desenvolvidos, entretanto Michael Smith apresenta argumentos para as duas doutrinas internalistas, cada uma das quais incompatíveis com a minha opinião.

A defesa de Smith dessas duas doutrinas é parte de um argumento maior pretendido para mostrar que os juízos de obrigação moral como proposições sobre o que nós quereríamos se fôssemos totalmente racionais³⁹. No entanto, vou restringir a atenção aos seus argumentos para as duas doutrinas internalistas.

A primeira é a doutrina que ele chama de “racionalismo”: “Se é [moralmente obrigatório] para os agentes [fazer A] nas circunstâncias C, então há uma razão para aqueles agentes fazerem

³⁹ Smith, *The Moral Problem*, 184.

[do A] em C”.⁴⁰ Relembrando as distinções que fizemos antes, podemos ver que o racionalismo é uma versão do agente com razões internas; alegando que há razões para o agente fazer aquilo que ele tem uma obrigação moral de fazer.

Smith chama sua segunda doutrina de “requisito de praticidade”: “e um agente julga que é [moralmente obrigatório] para ele [fazer A] em circunstâncias C, então ele está motivado a fazer A em C ou ele é praticamente irracional”.⁴¹ O requisito de praticidade é uma versão de crença de motivação-internalista. É uma doutrina sobre as motivações de um agente acerca de acreditar ter uma obrigação moral, desde que não afligido por qualquer forma de irracionalidade “prática”, como fraqueza de vontade ou depressão⁴².

Começarei com o requisito de praticidade, pois é a chave do internacionalismo de Smith. Essa é uma premissa em seu principal argumento racionalista⁴³, na qual ele busca responder a certas objeções ao internalismo⁴⁴. Através dela, ele mostra que Alice não pode genuinamente acreditar que tem uma obrigação moral.

O argumento de Smith sobre o requisito da praticidade começa com o que ele chama de “fato marcante” segundo o qual “*uma mudança na motivação* segue de forma confiável o despertar

⁴⁰ Ibid., 62. Smith formula ambas doutrinas como afirmações mais sobre retidão moral do que alegações de obrigação moral. Mas eu tenho incluído julgamentos de ação correta sob a rubrica “julgamentos de obrigação moral”. Nenhuma substância será perdida se eu expressar suas doutrinas como afirmações sobre a obrigação moral.

⁴¹ Ibid., 61. Ver 62.

⁴² Ibid., 61, 120.

⁴³ Ibid., 86. At 87-90, Smith oferece um segundo argumento para o racionalismo que parece não depender do requisito de praticidade. Eu discuto abaixo. Smith também acredita que o racionalismo implica o requisito de praticidade. (62).

⁴⁴ Smith responde ao argumento de Brink de que o internacionalismo não pode aceitar o desafio do amoralista. Veja Brink, *Moral Realism and the Foundations of Ethics*, 45-50, 57-62. Smith argumenta que o amoralista não faz realmente julgamentos morais, mas seu argumento se baseia na exigência de praticidade. Smith, *The Moral Problem*, 68-71. Smith também responde ao argumento de Foot de que os requisitos morais são análogos aos requisitos de etiqueta e, como os últimos, não são requisitos da razão. Smith argumenta contra a analogia, mas seu argumento se baseia no requisito de praticidade. Smith, *The Moral Problem*, 80-4

de uma *mudança no julgamento moral*, pelo menos para a pessoa boa e forte”⁴⁵. Suponha que você consiga convencer Carol de que valores fundamentais estão errados e, como resultado, ela passa a ter diferentes crenças morais daquelas que ela tinha antes. Ela passa a acreditar que possui a obrigação de se opor à pena capital enquanto, antes, acreditava que tinha a obrigação de apoiá-la. Se ela for uma “pessoa boa e forte”, diz Smith, ela agora será motivada a se opor à pena capital, embora antes estivesse motivada do contrário. Este exemplo ilustra a conexão entre mudança de crença moral e a mudança de motivação que Smith acredita ser bastante confiável. Ele argumenta que a confiabilidade desta conexão deve ser explicada de duas maneiras. Primeiro, em termos do “conteúdo do julgamento moral”, citando o requerimento da praticidade. Segundo, pode ser explicado em termos de “disposição motivacional que pessoas boas e obstinadas possuem”. Smith argumenta que qualquer explicação do segundo argumento seria compreender mal a natureza da pessoa boa. Portanto, o requisito da praticidade fornece a única explicação viável de conexão confiável. Diante disso, devemos aceitar o requisito de praticidade ou rejeitar a conexão confiável.

Como exatamente o requisito da praticidade explicará a conexão confiável? O requisito postula uma conexão entre crença moral e motivação em uma pessoa *racional*, mas a conexão confiável de Smith é entre crença moral e motivação em uma pessoa com vontade boa e forte. Smith deve, portanto, assumir que nenhuma pessoa boa e forte pode sofrer de irracionalidade prática de qualquer tipo. Desde que a fraqueza da vontade, que acredito ser o seu principal exemplo de irracionalidade prática, da “pessoa *obstinada*” fosse exatamente a de pessoas que não têm qualquer forma de irracionalidade prática. Nesta suposição, segue-se imediatamente a exigência de praticidade para que a pessoa

⁴⁵ *Ibid.*, 71. Neste e nos parágrafos seguintes, eu resumi e reconstruí o argumento de Smith em *The Moral Problem*, 71-6.

obstinada possa acreditar que que está certa. A obrigação também passa a ter uma nova motivação correspondente (a menos que, claro, que ela coincidentemente tenha uma motivação relevante o tempo todo). Em minha suposição, portanto, o requisito de praticidade pelo menos começa a explicar a conexão confiável de Smith⁴⁶.

Extetalistas que negam o requisito de praticidade obviamente devem explicar a conexão de alguma outra *forma*. Como vimos, Smith argumenta que eles devem explicá-lo em termos de algo na natureza da pessoa de vontade boa e forte. Em particular, argumenta, que esses devem explicar através da suposição de que uma pessoa boa, invariavelmente, deseja fazer tudo o que é obrigada a fazer. O problema, para Smith, é que esse desejo é vicioso, então uma pessoa boa *não* teria esse desejo.

Enquanto Smith vê a pessoa *boa* como preocupada diretamente com a honestidade, justiça e bem-estar de seus amigos. Por ter certeza que tudo o que ela desejar fazer será moralmente obrigatório, o que não quer dizer que, tudo o que é moralmente obrigatório ela querera fazer. Smith chama isso de “desejo *de dicta*”. Se a pessoa boa deseja ser honesta, de acordo com o desejo *de dicta*, não seria esse um desejo direto de ser honesto; mas um desejo derivado por este desejo *de dicta*, juntamente com a sua crença de que é obrigatório ser honesto. O desejo de fazer qualquer coisa que é obrigado a fazer é “um fetiche ou vício moral”, de acordo com Smith⁴⁷. Desde que a explicação extenalista da conexão confiável dependente da ideia de que a vontade da pessoa boa e forte seja caracterizada por esse desejo, seja necessariamente rejeitada.

⁴⁶ Se a minha suposição é incorreta, e ser “forte de vontade” é compatível com algumas formas de irracionalidade prática, então o requisito de praticidade não explica a conexão confiável. Uma pessoa com vontade boa e forte pode deixar de experimentar uma mudança de motivação consequente a uma mudança de crença moral, pois ele pode ser deprimido ou sofrer de alguma outra forma de “irracionalidade”.

⁴⁷ *Ibid.*, 75

Argumentarei que as principais afirmações de Smith sobre o desejo *de dicta* estão incorretas. O desejo não é vicioso; uma pessoa boa *poderia* ser motivada por isto. Além disso, a conexão confiável não existe em pessoas que são meramente boas sem serem *obstinadas*. Devido a isso, o extenalista pode apelar para o desejo *de dicta* para explicar a conexão confiável em pessoas que são boas e obstinadas ao mesmo tempo em que negam que o desejo *de dicta* seja essencial para uma pessoa ser boa. Finalmente, existem explicações externalistas da conexão confiável que não faz referência ao desejo *de dicta*.

Deixe-me começar com o último ponto. Os externalistas podem explicar a conexão confiável de Smith no sentido da natureza da pessoa *obstinada*. Para Smith, uma pessoa de vontade fraca é aquela que, pelo menos, às vezes não consegue desejar fazer o que acredita ser moralmente obrigatório⁴⁸. Presumivelmente, então, a pessoa obstinada não deixa de querer fazer o que acredita ser obrigatório. Portanto, quando uma pessoa *boa e forte* sofre uma mudança em suas crenças sobre suas obrigações, passando a corresponder suas motivações (a menos que, é claro, que ele coincidentemente tenha as motivações relevantes o tempo todo). Esta explicação de confiabilidade da conexão de Smith pode ser aceita por um externalista. Isso não requer que o externalista pense que a chamada “fraqueza de vontade” de Smith seja irracional ou uma espécie de fraqueza propriamente dita. Confiando apenas na ideia de que a força da vontade exclui (o que Smith chama) fraqueza da vontade.

O segundo ponto é relacionado a não ter uma conexão confiável entre mudança de crença moral e mudança de motivação em pessoas que são boas, mas não são obstinadas. Suponha que um demagogo convença Carol que ela é, de fato, obrigada a apoiar

⁴⁸ Caso contrário, a fraqueza da vontade não seria uma forma de irracionalidade prática que poderia “quebrar a conexão” entre o julgamento de uma obrigação e a motivação para agir de acordo com ela. *Ibid.*, 120.

a pena de morte. Enquanto, você a persuadi de que ela deveria se opor a isso, e você consegue vincular a natureza dela fundamentalmente compassiva à sua oposição à pena capital. Este link é cimentado com tanta firmeza, que vamos supor que o demagogo seja incapaz de superar a repulsa de Carol à pena de morte. Ele convence com sucesso que ela é obrigada a apoiá-lo, mas continua a se opor a isso. Agora, Carol pode ser uma pessoa boa, apesar do fato de ser “fraca de vontade”; sua crença sobre a pena capital mudou sem acompanhamento da mudança motivacional, ainda que (suponho) suas motivações fundamentais ou valores sejam apropriados. De fato, sua bondade parece depender dela não ter o desejo de fazer o que quer que seja obrigada a fazer, pois se tivesse esse desejo, suas crenças vacilantes sobre suas obrigações trariam motivações oscilantes. O exemplo mostra que uma pessoa boa pode experimentar uma mudança de crença moral sem passar por uma mudança correspondente de motivação. Isso significa que a conexão confiável de Smith não é encontrada em pessoas sem restrições. Encontra-se apenas em pessoas boas que são possuem força de vontade. Para explicar isso, um externalista deve, portanto, fazer uso da ideia de que força de vontade exclui a fraqueza da vontade.

Smith pode negar que Carol seja uma boa pessoa, pois ele afirma que é “constitutivo de ser uma pessoa moralmente boa que você tenha a preocupação com o que você acha que é certo”⁴⁹. Ou seja, se Carol é uma boa pessoa, então, por qualquer coisa que ela acredita que é moralmente obrigada a fazer, ela terá uma preocupação direta em fazê-lo. Mas o exemplo é para desafiar essa ideia. Caso eu esteja certo, Carol, no exemplo, é uma boa pessoa, embora acredite que seja obrigada a apoiar a pena de morte e não esteja preocupada em apoiá-la. Da maneira que eu vejo as coisas, a bondade é fundamentalmente uma questão de motivações e não de crenças ou conexão entre as crenças e as motivações da pessoa.

⁴⁹ Note que “isto é lido como *de re* e não *de dicta*.” *Ibid.*, 76.

Uma boa pessoa se inscreve nos padrões morais apropriados. Smith pode responder que Carol não é boa no sentido relevante, pois ela não tem princípios. No entanto, isso não está claro, uma vez que ela subscreve um padrão moral que proíbe a pena de morte, então não são apenas os sentimentos dela que a leva a se opor à pena de morte, mas a atitude dela em relação a um princípio. Se assim for, acho que intuitivamente a contaríamos como principialista. Em qualquer caso, o argumento de Smith não é convincente se a consideramos “principialista” ao invés de ‘boa’. Como vimos, ele argumenta que a boa pessoa se preocupa diretamente com coisas como honestidade, justiça e o bem-estar de seus amigos. O que quer que pensemos destas alegações sobre a pessoa boa, será certamente menos plausível se lidas como sendo de uma pessoa ‘principista’⁵⁰. Suponho que uma pessoa de princípios age de acordo com regras ou princípios ao invés de certas preocupações diretas. Por conseguinte, continuarei a investigar a ideia de pessoa boa.

Deixe-me agora seguir com a afirmação de Smith de que o desejo *de dicta* é “um fetiche ou vício moral”⁵¹. Se ele está correto, então *nenhuma* pessoa (inteiramente) boa tem esse desejo, nem mesmo uma boa pessoa que também tem força de vontade. O ponto chave em seu raciocínio é a alegação de que uma característica de boa pessoa é um conjunto de preocupações diretas. Isso pode estar correto, mas não se segue que nenhuma pessoa boa possa ter o desejo de dicta. Para ver isso, considere o caso de Dena. Dena não tem o desejo *de dicta*. Ela tem uma variedade de preocupações diretas por coisas como justiça, honestidade e bem-estar de seus amigos e, por qualquer coisa que ela acha ser moralmente obrigada a fazer, ela é motivada a fazê-lo, ou seja, é uma boa pessoa. Mas agora suponha que Dena *adquira* o desejo *de dicta*. Smith deve dizer que ela não é mais uma boa

⁵⁰ Smith sugere a resposta em correspondência.

⁵¹ Ibid., 75.

pessoa, mas ele certamente está errado sobre isso. Tampouco, Dena é menos boa do que costumava ser. Ela ainda se preocupa diretamente com as coisas que ela costumava cuidar. A única mudança é que ela tem o desejo de – fazer - o – que – é – obrigada – a - fazer. Não havendo razão para pensar que esse desejo seja cruel para ela, uma vez que ainda se importa com em fazer todas as coisas corretamente e só se importa adicionalmente em fazer a coisa obrigatória.

A preocupação de Smith é presumivelmente sobre o que aconteceria com Dena se ela adquirisse uma nova crença acerca do que é obrigada a fazer. Ele deve pensar que de alguma forma seria vicioso para ela derivar o desejo de fazer a coisa nova de sua crença recém-adquirida em conjunto com o desejo *de dicta*. Suponha, então, que depois disso Dena adquira um desejo de fazer coisas novas. Isso seria vicioso? Dena inicialmente se importaria apenas de maneira derivada sobre a coisa nova. No entanto, isso não seria suficiente para mostrar que ela é uma pessoa menos boa do que costumava ser, pois ainda se importaria diretamente sobre tudo (quase) o que ela costumava se importar, e se ela continuasse a se importar com a coisa nova, ela poderia vir a se importar sobre isso diretamente e não apenas porque ela acredita que é obrigatório. Nem parece que Dena teria sido uma pessoa melhor se ela tivesse conseguido imediatamente adquirir um desejo direto e não - derivativo de fazer a coisa nova. Suponho que ela *poderia* ter sido uma pessoa melhor naquele caso, mas parece-me que esta é uma questão contingente.

Por tudo o que eu disse, Smith pode estar correto que uma pessoa cuja motivação moral é *inteiramente* derivada do desejo *de dicta* não seja uma pessoa boa. Isso, no entanto, não é uma objeção ao tentar usar o desejo *de dicta* para explicar a conexão confiável de Smith. Pois mesmo que algumas pessoas boas não tenham o desejo *de dicta*, como é ilustrado pelo exemplo de Carol, e mesmo que qualquer pessoa boa tenha alguma preocupação direta, pode

ser que qualquer pessoa *boa e forte* tenha o desejo *de dicta* além de suas preocupações diretas.

Concluo, portanto, que a tentativa de Smith em apoiar o requisito da praticidade falha. Um externalista pode explicar a confiabilidade da conexão entre mudança de crença moral e mudança de motivação em uma pessoa de vontade boa e forte com base nas motivações confiáveis para fazer o que se acredita ser obrigatório. Tais pessoas são, afinal de contas, obstinadas.

A combinação da minha teoria baseada em padrões de convicção moral tem um número de vantagens sobre o padrão externalista relado que invocam o desejo *de dicta* de explicar a motivação moral. Primeiro, minha proposta pode fornecer uma explicação mais plausível da psicologia da pessoa boa do que a baseada no desejo *de dita*. *Podemos* ter o desejo de fazer – o – que – é – obrigatório, mas também normalmente nos inscrevemos em um padrão de ‘segunda ordem’, que nos chama para subscrever padrões morais que correspondem às proposições morais que aceitamos. Para subscrever neste padrão, temos que ter uma intenção ou política com respeito às coisas que acreditamos serem corretas, formando uma intenção ou tornando-a uma política para fazer tais coisas. Ou seja, inscrevemo-nos num padrão que converte crenças em convicções morais.

Em segundo lugar, a assinatura deste padrão de segunda ordem pode ajudar a fornecer uma explicação externalista da conexão confiável de Smith. Para isso, explica por que, normalmente, uma pessoa boa que venha a ter uma nova crença moral chega a subscrever o padrão moral correspondente. E o fato de que ele venha a se inscrever para este padrão explica como chega a ser motivado a fazer o que acredita ser obrigatório.

Dado que minha proposta abre espaço para uma distinção entre crença e convicção moral, sugiro a possibilidade de Smith ter descrito erradamente nossa intuição sobre a conexão confiável. Talvez nossa intuição seja de fato melhor entendida no sentido de que uma mudança de *convicção* moral que resulta numa mudança

de motivação, em vez de uma intuição sobre mudança de crença moral, falando estritamente. Não é mistério que uma nova *convicção* de que alguém tenha uma obrigação moral implica em nova motivação para agir adequadamente, essa é uma *convicção* moral, uma crença moral com assinatura de um padrão moral correspondente, implicando em uma motivação apropriada.

Finalmente, podemos explicar porque o requisito de praticidade pode parecer plausível. Afinal, se um agente tem a *convicção* moral de que é obrigatório fazer A nas circunstâncias C, ele é motivado a fazer A em C. Algo muito parecido com o requisito de praticidade é, portanto, o acordo de verdade que forneci anteriormente neste texto. Não obstante, não é verdade que um agente que acredita ser moralmente obrigado a fazer algo esteja necessariamente motivado para fazê-lo, a menos que seja praticamente irracional. Lembre-se do caso de Alice. Alice *acredita* ser moralmente obrigada a apoiar a pena capital com base nos mandamentos de Deus, mas é compassiva de uma forma que não está motivada a apoiar isso. Sua falta dessa motivação se deve à sua compaixão, não para qualquer irracionalidade que eu possa ver⁵².

Deixe-me agora seguir para o racionalismo de Smith, segundo a qual, se é moralmente obrigatório aos agentes fazerem A nas circunstâncias C, então existe uma razão para fazerem A em C. É trivial, é claro, que se um agente tiver uma obrigação moral de fazer algo, ele terá uma razão *moral* para fazê-lo. Como explicado anteriormente, isto está implícito na teoria baseada em padrões. Porém, de acordo com o racionalismo, o agente ter uma obrigação moral de fazer algo implica em haver uma razão para fazê-lo, uma razão tal que qualquer um “estaria motivado a agir igualmente se

⁵² Smith veria Alice como fraca de vontade fraca, considerando tal fraqueza como uma forma de irracionalidade (61). Mas se uma pessoa que deixa de estar motivada a fazer o que pensa ser obrigatório for considerada irracional apenas com base nisso, então a exigência de praticidade é tautológica. Eu suponho que, apesar do fato de Alice ser o que Smith chamaria de “fraca de vontade”, ela pode ser racional em outros aspectos.

fosse racional”. Uma pessoa é “praticamente irracional se não está motivada a agir”⁵³. Anteriormente, por simplicidade, referi-me às Razões como “razões sem qualificação”. Smith afirma que está é uma verdade conceitual que a existência de uma obrigação moral implica a existência de uma razão-sem-qualificação. Isso não é uma afirmação trivial.

Por analogia, considere as regras de etiqueta. A etiqueta requer a configuração da mesa com o garfo no lado esquerdo do lugar sentado. Segue que há uma razão de etiqueta para colocar a mesa dessa forma. Mas isso não significa que há uma razão para colocar a mesa desta forma - uma razão tal que alguém estaria motivado a configurar a mesa desta forma, se ele fosse racional. Há uma questão substantiva se somos necessariamente racionais ao agir tal como somos obrigados pela etiqueta a agir, ademais muitos pensariam que a resposta para a pergunta é negativa. Da mesma forma, há pelo menos uma tradição na filosofia que pensa haver uma pergunta substantiva, isto é, se somos racionalmente exigidos a agir assim como somos moralmente obrigado a agir. A resposta racionalista exigirá uma defesa independente.

O argumento racionalista de Smith inicia com a premissa de que “esperamos” que agentes racionais façam o que eles são moralmente obrigados a fazer⁵⁴. Isto é, diz ele, acreditamos que os agentes racionais *farão* o que são moralmente exigidos⁵⁵. “Ser racional, como tal, deve, portanto, ser suficiente para fundamentar nossa [crença] de que os agentes racionais farão o que são moralmente obrigados a fazer”. Mas se esta crença é bem fundamentada para todos os agentes apenas pelo fato de serem racionais, deve haver *Razões* para os agentes racionais fazerem o que eles são moralmente obrigados a fazer. Por isso, conclui ele, se

⁵³ *Ibid.*, 62. Aqui, Smith cita Korsgaard, “Ceticismo sobre razões práticas”.

⁵⁴ *Ibid.*, 85-6. Em sequência resumi e reconstruí o argumento de Smith.

⁵⁵ *Ibid.*, 89, 85-6.

é moralmente obrigatório para os agentes fazerem algo, então há uma razão para fazerem esse algo⁵⁶.

Smith não explica o que quer dizer com “agente racional”, mas isso decorre da definição de que uma ‘Razão’, necessariamente, implica que se há uma Razão para fazer alguma coisa, um agente racional será motivado a fazê-la. Para simplificar a discussão vou assumir que, qualquer Razão é uma Razão para que “todas as coisas sejam consideradas”, e qualquer agente racional faz o que for uma Razão para fazer. Para tal, será útil distinguir entre agentes invariavelmente movidos por Razões e agentes que às vezes não conseguem ser movidos por Razões, ainda que tenham a *capacidade* de serem movidos por tais. Escreverei que os primeiros são *Racionais*, enquanto os últimos são meramente *c-racionais*.

A premissa chave desse argumento de Smith é a proposição de que se somos razoavelmente ‘agentes racionais’ *faremos* o que é moralmente requerido que se faça. Claro, percebemos que agentes meramente *c-racionais* podem deixar de agir moralmente, mesmo que haja motivos para agir moralmente. A premissa de Smith deve, portanto, ser a de que somos razoáveis em acreditar que a agentes *Racionais* *agirão* moralmente. Smith oferece dois argumentos para apoiar esta premissa. Seu argumento principal baseia-se no requisito de praticidade, que pode ser ignorado aqui⁵⁷. Seu segundo argumento está relacionado com a adequação das atitudes de aprovação e desaprovação!⁵⁸

Este segundo argumento depende da doutrina de nossa desaprovação moral do comportamento ilícito das pessoas que pressupõe a legitimidade da nossa expectativa de que agirão corretamente; isto é, Smith afirma que nossa desaprovação moral pressupõe a justificabilidade da crença de que os agentes *farão* o

⁵⁶ *Ibid.*, 85.

⁵⁷ Para este argumento veja Smith, *The Moral Problem*, 86-7.

⁵⁸ *Ibid.*, 87-91.

que são moralmente obrigados a fazer⁵⁹. Todavia, a desaprovação moral seria apropriada em qualquer caso no qual um agente Racional não fazer o que é moralmente obrigado a fazer⁶⁰, e isso pressupõe a justificabilidade da crença de que qualquer agente Racional fará o que é moralmente obrigado a fazer. Isto é, “Ser suficientemente racional para aterrar a expectativa de que as pessoas façam o que são moralmente obrigadas a fazerem”⁶¹. Assumindo que as precondições da desaprovação moral são satisfeitas, então, podemos razoavelmente acreditar que as pessoas racionais farão o que serão obrigadas a fazerem simplesmente porque são Racionais⁶².

Entretanto, Smith está completamente errado em pensar que a legitimidade da desaprovação pressuporia a razoabilidade de acreditar que as pessoas agirão corretamente. Considere que não está totalmente claro o que Smith quer dizer. No entanto, primeiro a legitimidade de desaprovar uma pessoa ou sua ação não pressupõe que acreditamos razoavelmente ou que teríamos sido razoáveis em acreditar que o agente *por ele mesmo* agiria corretamente. Mesmo se soubéssemos muito bem que alguém fosse fazer algo errado, não significa que seria inadequado desaprovarmos isso. Suponha que haja alguém que sabemos ser um mentiroso inveterado, que mentirá se tiver algo a ganhar. Isso não torna inadequado desaprová-lo por sua mentira. Não precisa haver nada moralmente questionável na desaprovação do comportamento de amoralistas ou de pessoas que, por algum outro motivo, não reconhecem suas obrigações. Muito menos a

⁵⁹ *Ibid.*, 89. Como Smith observa, “dizer que esperamos que alguém faça alguma coisa pode significar ou que acreditamos que eles *farão*, ou que acreditamos que eles *deveriam* fazer”. Para Smith, a primeira interpretação passa por todas as questões. (*Ibid.*, 85-6. Smith enfatizou a correspondência pessoal também.) O primeiro é a interpretação para a proposição que ele está se apoiando, segundo a qual acreditamos que os agentes racionais *farão* o que eles são moralmente obrigados a fazer.

⁶⁰ *Ibid.*, 89. Veja 85.

⁶¹ *Ibid.*, 90.

⁶² *Ibid.*, 90.

legitimidade da desaprovação de uma pessoa ou ação pressuporia que *as pessoas em geral* irão agir corretamente. Suponha que uma pessoa má mate todos e uma pessoa boa, então depois disso, as pessoas em geral não agiriam mais corretamente. Isso não tornaria inapropriado que essa pessoa boa desaprove a pessoa má.

Pode-se propor que seja legítimo desaprovar apenas pessoas que tenham capacidade de *saber* o que é moralmente obrigado a fazer. Caso fosse legítimo desaprovar os agentes Racionais que agissem errado, a única coisa que esses teriam em comum seria a racionalidade que permitiria a eles saberem o que são moralmente obrigados a fazer, considerando apenas o exercício da razão comum. Portanto, seria o caso de conhecimento moral *a priori*. Infelizmente, mesmo se aceitarmos este raciocínio, o argumento de Smith continuará insuficiente, pois não demonstra que o conhecimento moral é conhecimento de Razões. Não demonstrando a existência de uma Razão para fazer o que for moralmente obrigatório.

Pode-se objetar que, se não há razões para fazer o que somos obrigados a fazer - se razões morais não são razões - então, um agente Racional não *poderia* fazer o que é obrigado a fazer. Este resultado pode parecer um conflito com a máxima de “dever” implicar “poder”. Mas esta inferência é enganosa; mesmo que razões morais não sejam Razões, não significa que um agente racional não pode fazer o que é obrigado a fazer. Para ver isso, suponha que Earl seja Racional. É uma verdade necessária que, enquanto Earl é racional, ele age segundo Razões. Mas isso não significa que Earl só possa agir corretamente se houver uma razão para agir corretamente. Não é uma necessidade para Earl ser Racional. Ele pode agir corretamente mesmo sem razões para agir corretamente, pois ele pode deixar de ser racional. Portanto, a ideia de que razões morais não são Razões não implica que os agentes Racionais não têm nenhuma obrigação moral.

Em resumo, não vejo como gerar um argumento de sucesso para o argumento racionalista das observações de Smith.

A teoria baseada em padrões suporta uma doutrina superficialmente semelhante ao racionalismo, pois, como eu disse antes, implica que se um agente tem uma obrigação moral de fazer algo, ele tem uma razão *moral* para fazer isso. Consequentemente, se existe uma obrigação moral de fazer algo, há uma razão (de algum tipo) para fazer isso. Mas o racionalismo é uma tese mais forte do que se um agente tem ou não uma obrigação moral de fazer alguma coisa, diante disso, há uma razão ou uma ‘razão-sem-qualificação’ para ele fazer algo. A existência de uma razão – sem - qualificação para cumprir nossas obrigações morais não se segue da existência de uma razão *de qualquer tipo* para cumprir nossas obrigações morais. Nós já não estamos comprometidos com o racionalismo simplesmente devido ao reconhecimento de que as obrigações morais implicam razões morais.

Pode-se pensar que as únicas razões verdadeiras ou genuínas são razões sem qualificação. Isso significaria que as chamadas ‘razões morais’ não são *razões* a menos que sejam razões – sem - qualificação. Claro, não tenho objeções à política de reservar a palavra ‘razão’ para as razões –sem - qualificação, mas normalmente não fazemos isso. Uma pessoa racional responde apropriadamente às razões de todos os tipos, tomando sua verdadeira medida de acordo com o ponto de vista da razão; ela não está necessariamente motivada por razões de todo tipo. Uma pessoa irracional é uma pessoa que falha em responder às razões de maneira apropriada. Então, para estabelecer que há uma razão para fazer algo de qualquer tipo ainda não significa que uma pessoa racional estaria motivada a fazer tal coisa ou que seria irracional não estar motivado a fazê-lo. Ainda não é preciso estabelecer se há uma razão – sem - qualificação para fazer isso.

Christine Korsgaard propôs um “requisito de internalismo” para o efeito de que razões devem ser *capazes* de motivar a nós, na

medida em que somos racionais⁶³. Mas a proposta de Korsgaard não minou meu ponto de vista. Minha opinião é que há uma razão moral para fazer algo apenas quando o padrão moral justificado exige que façamos algo. Uma pessoa racional é capaz de subscrever tal padrão e se ela assina, está motivada para se adequar a esse padrão. Portanto, razões morais são capazes de nos motivando mesmo que não sejam Razões.

Obviamente, estaria além da minha capacidade, contrapor cada argumento internalista frente ao tipo de externalismo proposto. A minha questão foi apenas tentar apresentar o que é o julgamento moral baseado em padrões, associado à convicção moral de bem contra os argumentos de Michael Smith. Além disso, de alguma maneira, o balanço realizado tenta explicar porque muitos filósofos estão convencidos de que o internalismo está correto, inferindo a verdade das doutrinas que são muito semelhante às internalistas apresentadas. Para eles, as proposições morais são intrinsecamente normativas. O que por sua vez, implica que as considerações morais são fontes de razões morais. Dessa forma, eles concluem através do fato de que pessoas boas não são geralmente motivadas pelo desejo – de – fazer – o – que – é – obrigatório. Por tudo isso, consideram que a convicção moral é parcialmente constituída pela motivação moral. Por todos esses aspectos, sou simpático ao internalismo ainda que, estritamente falando, isso seja externalista.

⁶³ Korsgaard, “Skepticism about Practical Reason”. Em algumas partes, ela fala de “alegações de razão” ou de “considerações racionais”, e não de razões, quando esta considerando o “sucesso” em nos motivar, e não de ser capaz de nos motivar. Veja pp. 11, 15, 23. Para uma visão superficialmente similar, veja Williams, “Internal and External Reasons.”

Capítulo 3

Dando razões práticas

*David Enoch*¹

1. Introdução

Estou escrevendo um artigo medíocre sobre um tópico que você não está particularmente interessado. Você não tem, parece seguro assumir, uma razão (normativa) para ler meu rascunho. Então, pergunto se você estaria disposto a dar uma olhada e me dizer o que pensa. Subitamente, você tem uma razão (normativa) para ler o meu rascunho. O que exatamente aconteceu aqui? Você tendo a razão para ler o meu rascunho – certamente, o próprio fato de que existe tal razão – depende, aparentemente, da minha parte ter pedido a você para o ler. Por ter pedido, consegui fazer com que seja o caso de você ter tal razão, ou dei uma razão para que você leia o meu rascunho. No que este “dar razões” consiste? E como é que podemos fazê-lo? Seria um tipo de magia normativa? Especialmente, se razões estão em um sentido objetivo importante e autônomo, como é que por puro ato da vontade nós podemos trazê-las à existência, mudar sua força e, talvez, até mesmo eliminá-las?² Se, como parece provável, “a razão deve restringir e guiar a vontade”, como é que podemos criar razões à vontade, por exemplo, fazendo uma requisição?³

¹ The Hebrew University. Tradução: Gustavo Teles Jr. - Revisão: Mariana Burkle

² Para o *locus classicus* de reflexões sobre promessas e mágica, veja o Tratado de Hume 3.2.5.

³ Essa forma de coloca-las em um contexto intimamente relacionado, veja Watson (2009, p. 158)

Requisições não esgotam, claramente, o espaço de dar razões. Algo bastante similar parece acontecer, por exemplo, quando uma autoridade emite um comando, dando assim, ao endereçado, uma razão (talvez de um tipo especial) para agir de uma certa forma, uma razão que não estava lá antes do comando. Quando prometo, uma das coisas que pareço fazer é dar a mim mesmo uma razão (talvez de um tipo especial) para agir de determinada forma, uma razão que não estava lá antes da promessa.⁴ Agora, mais controversamente, talvez formando uma intenção, dou a mim uma razão que não estava lá antes para seguir esta intenção. Talvez haja também outros casos.⁵ Todos esses casos parecem – bem pré-teoricamente – ter algo em comum a estes. (Na seção 3, argumentarei que essa suposição pré teorética é a que devemos aceitar em valor nominal [*face-value*].) Será isso que tentarei entender – e desmistificar – nesse texto. Se fizermos progresso neste fenômeno geral em que todos estes casos são instâncias, parece uma forma promissora de então fazermos progresso nas perspectivas filosóficas da autoridade, das promessas, e das intenções, e talvez, em outros fenômenos relacionados (trabalho que deixarei para ocasiões futuras).

Na próxima seção, distinguirei entre o que podemos chamar de dar razões puramente epistêmicas, dar razões meramente acionadas, e o tipo de dar razões em que estou primariamente interessado, o tipo que presumivelmente envolve requisições, que chamo de dar razões robustas. Então, na Seção 3, tento caracterizar em detalhes os dados intuitivos ou fenomenológicos. Tento, em outras palavras, clarificar o que queremos acomodar com a perspectiva de *dar razões robustas*. Entretanto, no final da Seção 3 permanece completamente em aberto caso se qualquer

⁴No contexto de uma discussão sobre promessas, Watson (2009, p. 160) observa-se que elas são um exemplo particular de um fenômeno muito mais amplo.

⁵ Talvez, por exemplo, pelo desculpar, alguém dá razões para o perdão. (Eu agradeço a David Sosa por essa sugestão.) E talvez por tomar responsabilidade alguém dá aos outros razões para considerar alguém responsável. Veja o meu “Being Responsible, Taking Responsibility, and Penumbral Agency”.

perspectiva possível, de fato, satisfaça essas desideratas. Na Seção 4, então, prossigo investigando se tal perspectiva pode ser encontrada. Argumento que a única forma possível de compreender *dar razões robustas* é em uma instância única e particular de *dar razões acionadas*. Então, caracterizo essa instância única e particular, antes, em termos de intenções complicadas do concessor de razões e do fundo normativo.

Antes de prosseguir, duas preliminares são necessárias.

Primeiramente, o papel das requisições neste texto – como no parágrafo de abertura – é o de um exemplo paradigmático de *dar razões* do tipo em que estou interessado, o tipo que chamo de *dar razões robustas*. Apesar do fenômeno de *dar razões robustas* ser – se eu estou certo, pelo menos – muito mais amplo do que no sentido das requisições, ainda assim, há boas razões metodológicas para focar em requisições no estudo de *dar razões robustas*. Desta forma, o caso da autoridade e do comando é mais complicado, naquilo em que, indiscutivelmente, pelo menos, a razão ali dada é de um tipo especial – isto é, um dever talvez, ou obrigação, ou tal como estes. No caso das promessas é mais complicado que este *dar requisições*, tanto porque a razão dada é indiscutivelmente um dever ou uma obrigação, quanto porque promessas são um caso de *dar razões reflexivas*, e tal reflexividade pode resultar em maiores complicações.⁶ Então, o caso das requisições é um caso mais simples de *dar razões robustas* que estes dois, e por esse motivo parece metodologicamente aconselhável começar com requisições. Além disso, comandos e promessas são tópicos de outras grandes controvérsias filosóficas. E, por esse motivo, também parece que há maior esperança de progresso se começarmos por outro lugar, onde implicações políticas e comprometimentos filosóficos são menos prováveis de enviesar a caracterização dos dados fenomenológicos. Requisições parecem idealmente adequadas para

⁶ Por exemplo, esta reflexibilidade dá origem às preocupações espontâneas que não parecem surgir para *dar razões robustas* em geral, ou para *requisições* em particular.

este papel. Contudo— apesar do papel central das requisições no que está a seguir—este texto é primariamente sobre dar razões robustas, e não sobre requisições. Portanto, enquanto muito da discussão se focará em requisições, não tentarei oferecer uma análise completa delas: eu faço uso requisições somente em ordem de realçar as características cruciais do dar razões robustas. Não terei nada a dizer sobre outras questões que talvez sejam relevantes para uma discussão completa sobre requisições (como, por exemplo, quando é e quando não é apropriado fazer uma requisição).

Em segundo, minha discussão sobre dar razões robustas é obviamente intimamente relacionada – inspirada, pelo menos – à recente ênfase de Darwall sobre o ponto de vista da segunda pessoa, no livro intitulado com este nome (2006) [*The Second-person Standpoint*], e certamente estou em débito por este trabalho. Porém, precisamente por causa da similaridade entre minha discussão e a de Darwall, talvez valha a pena realçar algumas importantes diferenças entre os dois já neste estado inicial. Uma diferença deve ter sido notada: obrigações, que são centrais para a discussão de Darwall, são sobretudo irrelevantes na minha. Em partes do livro próximas ao meu tópico, no entanto, Darwall discute as *razões* de segunda pessoa, que ele oficialmente⁷ define assim:

Uma razão de segunda pessoa é uma cuja validade depende na pressuposta relação de autoridade e responsabilização entre pessoas, e, portanto, na possibilidade da razão ser endereçada de pessoa à pessoa. [2006, p. 8]

Enquanto há claras similaridades entre as razões de segunda pessoa e o que pode ser chamado de razões *dadas* (como a razão que você tem para ler o meu rascunho), ainda assim há

⁷ Enquanto essa é a única definição oficial de razões de segunda pessoa que pude encontrar no livro de Darwall, há momentos em que penso que Darwall usa essa locução de maneira mais folgada.

importantes diferenças entre elas. Sua razão para ler o meu rascunho não precisa – por nada que foi dito até então, neste caso – depender na relação de autoridade e responsabilização (embora talvez dependa de algo análogo, ou em alguma relação generalizada em que as relações de autoridade de Darwall sejam um exemplo). O possível endereço é irrelevante para razões dadas. Em vez disso, o que importa aqui é que essa razão *era na verdade* (ao invés de possível) *dada* (ao invés de endereçada). Embora haja, em um sentido intuitivo, algo de segunda pessoa sobre razões dadas (como essas criadas por requisições); elas ainda devem ser distintas, pelo menos, das razões de segunda pessoa de Darwall, como foi oficialmente caracterizado. Além disso, não estou convencido de que razões de segunda pessoa existem (como oficialmente definido por Darwall), e enquanto penso que razões dadas existem, não tenho certeza se elas abrangem um tipo interessante de razões, um “tipo normativo”, por assim dizer. Meu foco aqui não é sobre razões dadas, mas em vez disso, sobre dar razões. Ademais, enquanto Darwall está interessado em fundamentar a moralidade em “sua forma de” razões de segunda pessoa, isso não faz parte da minha ambição aqui (exceto por uma sugestão nesta direção no último parágrafo deste texto). Mantenho as coisas muito mais simples. Deixando a moralidade para outra ocasião, eu foco, ao invés disso, nos inegáveis dados intuitivos – frequentemente damos uns aos outros razões para agir – e tento apresentar uma análise destes dados. Se algo de maior interesse geral – nomeadamente, a moralidade – vai se seguir, não é minha maior preocupação aqui.

Essas diferenças, não obstante a discussão aqui, parecem a mim que são de relevância para uma avaliação do projeto de Darwall. Uma vez que, se não pudermos entender o fenômeno de dar razões, então, *a fortiori*, as razões de segunda pessoa de Darwall estarão em perigo. Além do mais, se o fenômeno de dar razões pode ser entendido redutivamente – de uma maneira que não é essencialmente e irredutivelmente em segunda pessoa em

nenhum sentido interessante – então a afirmação de Darwall sobre a irredutibilidade do círculo de conceitos de segunda pessoa(e.g., p. 11) também recai sobre sérias dúvidas. Não é completamente claro para mim em que uma caracterização de razões robustas que acabo oferecendo (na seção 4.4) é suficientemente de segunda pessoa para satisfazer Darwall.⁸

Nesse sentido, então, me parece que a discussão deste artigo é mais geral que, e talvez filosoficamente anterior que, a de Darwall. Mesmo que o projeto de Darwall falhe, ainda assim nós precisamos de uma caracterização do dar razões, e sem a viabilidade de tal caracterização, e assim também de razões dadas, o projeto de Darwall não pode prosseguir.

2. Algumas Distinções Relevantes

Tanto aqui como em outros lugares, a linguagem natural é enganosa. É importante distinguir entre inúmeras leituras possíveis de tais locuções como ‘dar a alguém razões para Φ ’. Isso será útil para limitar o sentido sobre o sentido que estamos perseguindo, o sentido que tento capturar nas próximas seções.

Decidi hesitantemente contar a um colega exatamente o que penso sobre ele, e não será uma cena encantadora. Você insiste para não o fazer. Posso então dizer algo como “Me dê uma razão para não o fazer!” Suponhamos que você responda apontando para os efeitos ruins que tal cena teria sobre a atmosfera intelectual do nosso departamento. Aparentemente você conseguiu me dar uma razão para não prosseguir com meu plano insensato. Não há nada de errado então em descrever a situação. No entanto, claramente, não é diante disto que estamos aqui. O que você fez – a coisa naturalmente descrita em termos de me dar uma razão para me calar –foi *indicar* para mim, ou me *mostrar*, uma razão que existia o tempo todo, independentemente de você tê-la dado para

⁸ Voltarei a este ponto em uma nota de rodapé na seção de conclusão.

mim. Talvez, em meu furor, eu não tenha prestado atenção a isso e, portanto, sua intervenção pôde fazer a diferença. Contudo, isso não fez uma diferença por criar uma nova razão⁹, como ocorre em minha requisição para ler o meu rascunho. Podemos chamar este dar razões de *puramente epistêmico*, pois o papel dar a razão aqui não tem nada a ver com a existência da razão, e tudo a ver com meu conhecimento de que ela está ali, apreciá-la, e agir através dela. Agora, requisições certamente possuem uma dimensão epistêmica. Frequentemente, uma das coisas que uma requisição faz é informar ao endereçado que o falante possui certas necessidades, ou ânsias, ou preferências, ou que os desejos que o endereçado já tinha conhecimento são particularmente fortes, ou algo do tipo. Então, não quero alegar que as requisições não envolvem dar razões epistêmicas. Não obstante, parece claro a mim que o dar razões envolvido em requisições não é *esgotado* por este dar razões epistêmicas¹⁰. Para ver isso, devemos imaginar casos que são epistemicamente “transparentes”, então: ambos, falante e endereçado, sabem tudo que há para saber sobre as preferências e afins do falante, e além disso ambos sabem que ambos sabem essas coisas, e por aí vai (assim, estes fatos relevantes são matéria de conhecimento comum). Todavia, em tais casos também as requisições parecem fazer uma diferença normativa – o endereçado parece ter dado uma razão através de um pedido, apesar de que nada de puramente epistêmico mude. Isso significa que o tipo de dar razões envolvido em pedidos não é simplesmente este do dar razões puramente epistêmicas¹¹. O

⁹ É possível que enquanto a razão estava lá independentemente de sua intervenção, eu tê-la, depende de sua intenção. (Mas veja Schroeder [2008] para uma crítica à ideia aparentemente pressuposta por esta forma de falar.) Eu posso seguramente ignorar tais dificuldade aqui. Aos meus propósitos, é suficiente que a existência de uma razão é aqui independe de sua concessão.

¹⁰ Para uma distinção similar entre razões puramente epistêmicas e o tipo que é envolvido em requisições, veja Cupit(1994, p. 449). E Wallace (2007, p. 24) enfatiza que a noção de Darwall de razões de segunda pessoa não é fundada em considerações epistêmicas.

¹¹ Por me pressionar neste e em pontos relacionados, eu agradeço à Hagit Benbajui, Yuval Eylon, David Heyd e Adi Koplovitz.

mesmo parece verdadeiro – talvez ainda mais claro – sobre comandos, de modo que sua importância também não é esgotada por seu valor epistêmico. É *muito* difícil pensar em promessas como meramente dar razões epistemicamente, vendo que com promessas o concessor de razões e o recebedor de razões são um e o mesmo. Então, aparentemente o contraste entre dar razões puramente epistêmicas é uma característica do dar razões robustas no geral (e não somente em requisições).

Agora suponhamos que o merceeiro do seu bairro aumentou o preço do leite. É natural dizer que ele, deste modo, deu a você uma razão para reduzir seu consumo de leite. É verdadeiro, afinal, que você não tinha essa razão antes da ação relevante dele, e que você a tem depois dela, e além disso, você tem essa razão *por causa* dele ter aumentado o preço. Então, em um sentido perfeitamente ordinário, ele criou essa razão: o merceeiro deu a você uma razão para comprar menos leite. Entretanto, não há nada de misterioso – sem magia normativa aqui. A coisa óbvia a se dizer sobre esse caso é que o dar razões aqui, é um dar ou criar razões muito menos radical do que no caso das requisições, comandos e promessas. O que o merceeiro fez, parece natural dizer, foi meramente manipular as circunstâncias não normativas de tal forma que acionou uma razão dormente que estava lá o tempo todo, independentemente das ações do merceeiro. Indiscutivelmente, você tem uma razão geral (grosseiramente) para poupar dinheiro. Essa razão não depende de o merceeiro aumentar o preço do leite. Aumentando o preço do leite, este acionou essa razão geral, assim, fazendo com que seja o caso de que você tenha uma razão para reduzir o seu consumo de leite. De fato, talvez, você já tivesse o tempo todo a razão condicional para-comprar-menos-leite-se-o-preço-do-leite-subir. Novamente, a existência dessa razão condicional não depende das ações do merceeiro. Contudo, o merceeiro pode tornar a razão condicional em incondicional, simplesmente manipulando as circunstâncias não normativas. E isso é o que ele fez aumentando o preço do leite.

Exemplos destes casos de acionamento estão por toda nossa volta. Colocando seu pé na rua, o pedestre pode dar ao motorista uma razão para parar, mas somente porque o motorista tinha o tempo todo, e independentemente das ações do pedestre, a razão condicional para-parar-se-um-pedestre-começar-a-atravesar. Colocando seu pé na rua, o pedestre, portanto, aciona essa razão pré existente, assim dando ao motorista uma razão para parar.¹² Talvez todos nós tenhamos razões (algumas) para fazer o que podemos para ajudar os famintos. Então, se eu dar a você muito dinheiro, em algum sentido, eu estou dando a você uma razão para doar mais para o alívio da fome; devido a manipular as circunstâncias não normativas habilitando a condição para a razão relevante, que não fora satisfeita, ser então satisfeita, fazendo com que seja o caso em que você tenha uma razão para dar mais dinheiro para o alívio da fome. Talvez haja outras maneiras pelas quais, manipulando as circunstâncias não normativas, se pode fazer com que seja o caso de que uma razão pré-existente se aplique, maneiras que não precisam envolver a satisfação da condição pela qual uma razão condicional é condicionada, ou pela habilitação da condição para uma razão (talvez, por exemplo, se poderia anular [*defeat*] uma anuladora [*defeater*] por essa razão). Chamarei todos esses casos, casos de dar razões *no sentido acionador*, e usarei o acionamento de razões condicionais como o paradigma de um fenômeno mais geral.¹³

Quando peço a você leia meu rascunho, parece que algo mais está acontecendo. É verdade que o que faço é manipular as

¹² Veja Estlund (2008, p. 143) para esse exemplo.

¹³ Mark van Roojen apontou minha atenção para a relevância aqui da locução “Eu vou lhe dar uma razão”, declarada em um tom de voz ameaçador. (Voltarei às ameaças mais adiante.) De fato, tais declarações transitam sobre uma ambiguidade em “dar uma razão”. Tipicamente, alguém perguntará por uma razão para ser dada em um sentido puramente epistêmico, isto é, para a indicação da existência de uma razão da qual não depende sobre o ato de dar a razão. E a razão “Você quer uma razão? Eu vou lhe dar uma razão!”expressará uma intenção para dar uma razão no sentido acionado. Nenhum desses sentidos, então, é aquele que estou procurando aqui, aquele que está indiscutivelmente no lugar quando eu lhe dou uma razão para ler o meu rascunho ao requisitar que você faça isso.

circunstâncias não normativas, mas não parece ser que *meramente* faça isso. Em vez disso, parece que estou dando uma razão mais robusta, em um sentido ainda a ser especificado.¹⁴ Esse sentido ainda a ser especificado será o tópico principal para o restante deste texto. Por enquanto, no entanto, deixe-me enfatizar o seguinte: requisitar que você leia o meu artigo parece significativamente diferente de informar que o fato de você ler meu artigo me dará prazer (um caso de dar razões no sentido puramente epistêmico) ou de a não colegialidade fazer um fundamento para negar a sua efetivação [no departamento] (um caso de dar razões no sentido acionado). E o mesmo é verdadeiro – nesse nível pré-teorético, nível fenomenológico – para comandos e promessas. Os casos de dar razões em que estou interessado, como o caso das requisições, não são *meramente* aqueles em que uma razão é dada no sentido puramente epistêmico ou acionado. As *razões dadas*, neste sentido, podem não ser tão únicas (por exemplo, elas não precisam ser mais fortes que outras razões dadas de forma não robusta). No caso mencionado o *dar razões* é um fenômeno distinto, ou – se isto é uma instância de um dos outros tipos de dar razões – isto é especialmente uma instância interessantemente particular, uma cujas características especiais valem a pena uma discussão separada. Sem exigir qualquer questionamento, então, vamos chamar esse tipo de dar razões – o tipo presumivelmente presente no caso dos pedidos, e aqueles em que vou focar – de *dar razões robustas*.

3. O Que Mais Podemos Querer?

¹⁴Para uma distinção similar no contexto de comandos, veja, Estlund (2008, pg. 143). Estlund o chama os casos (análogos de) que chamo de razões meramente acionadas “casos de efeito colateral”. Por razões que ficarão claras mais tarde, não acredito que essa é uma boa sugestão terminológica. Estlund não desenvolve uma perspectiva da diferença entre razões acionadas e razões robustas. E veja Postema(2001, pg.482-3) para uma pesquisa histórica de uma distinção similar daquela em texto (também no contexto dos comandos).

Com essas distinções em mente, então, o que mais nós vamos querer do dar razões robustas, além do que pode ser acomodado por dar razões puramente epistêmicas ou por dar razões meramente acionadas? Nesta seção, realçarei alguns dos dados fenomenológicos pré-teoréticos, tentando cuidadosamente caracterizar algumas características únicas do dar razões robustas. Estarei, primeiramente, usando o caso das requisições, mas, os dados fenomenológicos que coletarei aqui se aplicam igualmente, parece-me, aos comandos e promessas. Então, podemos estar moderadamente confiantes de que o que irei em efeito caracterizar é o fenômeno de dar razões robustas (ao invés do caso particular das requisições).

3.1 Ameaças

Quando emitimos (condicionalmente) ameaças nós (parecemos) dar às pessoas razões para ação. Ameaças parecem a mim – pré-teoricamente – ser um caso particular de dar razões no sentido de acionamento. Ameaçando votar contra você assumir uma efetivação caso você não leia meu artigo, eu estou simplesmente manipulando as circunstâncias não normativas com o acionamento de uma razão pré-existente, ao saber de sua razão para promover suas chances de ser efetivado. Eu estou também, claramente, deixando você saber que manipulei as circunstâncias.

Ameaças são uma instância particular singularmente interessante de dar razões acionadas, pois elas são em segunda pessoa em um sentido que em muitas outras instâncias não são. Compare o caso da ameaça e o caso de aumentar-o-preço-do-leite. Uma das características distintivas no caso da ameaça é que o dar razões é uma parte maior do ponto predominante da ameaça. Há um sentido em que a ameaça também aumenta o preço de alguma coisa (a saber, o de não ler o meu artigo). Todavia, isto é um caso de aumentar o preço precisamente *a fim de* dar a você uma razão para ler o meu artigo. Não é este o caso do preço do leite, onde não

era a intenção do merceeiro quando aumentou o preço, ou sua razão para fazer isso, dar-lhe uma razão para comprar menos leite. Ela sabe como o mercado funciona, é claro, e então ela prevê que isso seria uma consequência de sua ação, mas ainda assim, isto não é o que ela pretende que aconteça. (No entanto, ele talvez pense nesta razão como contando *contra* o aumento do preço, uma razão que é contraposta nas circunstâncias por razões mais fortes para aumentá-lo)

Isto, tomarei, no sentido que ameaças são mais em segunda pessoa: o dar razões envolvido nestas é parte de seu ponto. Mas ainda assim, o dar razões envolvido quando ameaço você para que leia meu artigo é muito diferente daquele envolvido quando peço a você que o leia.¹⁵ Embora ambos sejam casos de dar razões, e apesar de que algumas consequências desagradáveis sejam previstas em ambos os casos (talvez, por exemplo, você preveja que caso negue meu pedido e falhe em ler meu rascunho, eu falharei em comentar no seu próximo artigo), ainda assim, uma ameaça parece meramente servir para acionar uma razão condicional¹⁶, e uma requisição parece fazer algo a mais.¹⁷ Um entendimento sobre o dar razões robustas deve tanto explicar porque ameaças são tão próximas do caso do preço-do-leite para o

¹⁵ Comandos talvez sejam um caso híbrido interessante: eles certamente envolvem dar razões robustas, mas talvez haja algo semelhante às ameaças sobre comandos também. Espero discutir comandos em outra ocasião.

¹⁶ Aqui está outra consideração mostrando que é de tal forma, (agradeço a Janice Dowell por chamar minha atenção para isso). Ameaças vazias – ameaças em que o receptor não pensa no conteúdo da ameaça como algo mal – não obtêm sucesso em dar razões. Isso mostra que ameaças tomam toda a sua força normativa por acionar razões que já estão lá. Não há fenômeno paralelo, quão longe posso ver, para requisições.

¹⁷ De fato, ameaças combinam também elementos de razões dadas puramente no sentido epistêmico. Considere o diálogo estilo Hollywood: “Isso é uma ameaça? Não, isso é uma advertência”. Uma advertência genuína é simplesmente uma tentativa de chamar a atenção de alguém para uma razão pré-existente, uma que é independente da advertência em si mesma. Uma ameaça é uma tentativa de simultaneamente criar uma razão (no sentido acionado) e alertar o endereçado sobre a existência de uma ameaça. (Afinal, uma ameaça não pode obter sucesso como uma ameaça se o endereçado não sabe sobre ela.) Para algumas discussões iniciais sobre a distinção entre ameaças e advertências, veja Darwall (2006, p. 50-2).

dar razões robustas, quanto porque ameaças, no entanto, não contam como dar razões robustas.

3.2 Pessoalidade

No dar razões robustas, parece que pessoas relevantes estão envolvidas de uma maneira muito mais pessoal, comparada com seu envolvimento nos casos puramente epistêmicos ou acionados de dar razões.

Considere um concessor de razões primeiro. No caso de dar razões puramente epistêmicas, o papel do concessor de razões é o de uma indicação, chamando a atenção para uma razão que já está lá. Quase qualquer coisa pode desempenhar este papel. A leitura promovida por um barômetro pode, em um certo sentido, dar a você uma razão para pegar seu guarda-chuva. No sentido puramente epistêmico de dar razões, então, o concessor de razões não precisa ser, sequer, uma pessoa.¹⁸ Similarmente, para dar razões no sentido de acionamento: a tempestade prestes a acontecer pode dar a você uma razão para não sair. Dar razões robustas, o tipo de dar razões envolvidos em pedidos (e também, é claro, na emissão de comandos e em fazer promessas) é diferente: tais razões podem somente ser dadas por pessoas, ou ao menos, por agentes como pessoas. Talvez, por exemplo, Deus possa dar razões dessa forma. Talvez o estado, ou talvez também (mesmo que eu duvide) o seu cachorro. Porém o barômetro não, nem a tempestade iminente.

Seguindo em frente para o fim recebido agora, há alguma restrição comparável sobre a quem podemos robustamente dar uma razão? Bem, indiscutivelmente, podemos dar razões somente para pessoas, mas por razões mais gerais que não tem a ver com o

¹⁸ Talvez haja alguns tipos de dar razões puramente epistêmicas que possam somente ser feitas por pessoas. Talvez – não estou certo – conselhos sejam dessa natureza. Mas dar razões puramente epistêmicas, em geral, não requer pessoalidade.

nosso assunto aqui. Talvez, no entanto, somente pessoas (ou somente agentes) podem *ter* razões. E é claro: você não pode dar uma razão para alguém que não pode ter uma razão. Se então somente a pessoas (ou a agentes) podem ser dadas razões, isso não nos diz nada interessante sobre dar razões.

A pessoalidade do recebedor de razões não irá nos ajudar, então, a focar nossa atenção sobre o dar razões robustas. Porém, há algo mais útil nas vizinhanças aqui. A fim de (intencionalmente) dar a alguém uma razão para ação no sentido puramente epistêmico, parece que devo acreditar que ela é capaz de responder à uma razão relevante pré-existente. Com a finalidade de (intencionalmente) dar a alguém uma razão no sentido de acionamento, parece que devo acreditar que ele é capaz de responder a uma razão relevante (e através dela, ir em direção a razão geral ou condicional que estava lá o tempo todo). E com o intuito de dar a alguém uma razão robusta, parece que devo acreditar que ela possa responder a uma razão *então dada*. Este ponto se aplica às ameaças tanto quanto se aplica às requisições: se eu penso que você não pode responder às ameaças, eu não posso, sinceramente, emitir uma ameaça endereçada a você. Similarmente, se penso que você não responderá às requisições, eu não posso sinceramente requisitar a você que leia meu artigo, assim tentando dar a você uma razão no sentido usual em que se fazem requisições. (Eu talvez tenha outras razões para declarar palavras relevantes – eu retornarei neste ponto a seguir – mas isto é irrelevante aqui.)

Portanto, a habilidade do recebedor de razões de responder através de uma razão dada parece ser assumida em algum sentido pelo próprio ato de dar razões. A habilidade parece ser uma “condição de conveniência normativa”¹⁹ do dar razões relevante,

¹⁹ Um termo que Darwall usa extensivamente em um contexto intimamente relacionado. Veja Darwall (2006, p. 24). Para uma ideia de alguma forma similar no contexto das requisições, veja Cupit (1994, p. 450). E para uma pesquisa histórica sobre essa ideia na tradição do pensamento sobre leis como comando, veja Postema (2001, p. 477-9).

uma em que a ausência do dar razões falha, ou não funciona, ou sem a crença em que a requisição é insincera, ou algo assim. Talvez diferentes habilidades sejam necessárias a fim de responder a diferentes tipos de dar razões. Em particular, parece que poderia haver um agente relativamente simples que poderia responder apenas às razões epistemicamente dadas, e às razões dadas no sentido de acionamento, mas não às requisições e comandos. Eu voltarei a este ponto adiante.²⁰

Por conta da centralidade das pessoas para o dar razões robustas, não é surpresa que relações pessoais sejam relevantes aqui. Ambos os sentidos de dar razões puramente epistêmicos e acionadas podem frequentemente ser exaustivamente entendidos em termos relativamente finos (*thin*): não precisamos saber nada sobre a relação entre o merceeiro e o consumidor em ordem de entender a maneira pela qual o aumentar o preço do leite do primeiro dá razões para este último. Contudo, com o dar razões robustas nós frequentemente precisamos uma descrição mais grossa (*thicker*) da relação entre as pessoas envolvidas. Afinal, não é como se as requisições criassem (ainda que fracas) razões para ações. É plausível supor que um dos fatores determinantes em que uma dada requisição dá uma razão para ação (e certamente um dos fatores determinantes da força de tais razões dadas onde elas existem) é a natureza do relacionamento relevante. Isto também é verdadeiro, certamente, em comandos, que são plausivelmente parasíticos em alguma relação de autoridade entre duas pessoas relevantes.²¹ Dar razões robustas, então, é em um sentido, sustentada pela natureza do relacionamento relevante de uma

²⁰ Acredito que é essa intuição que subjaz a fala de Darwall da posição para emitir razões de segunda pessoa, e sobre a suposição de que há condições de conveniência normativa aqui sobre a habilidade tanto do concesso quanto do receptor de razões para considerar a razão relevante como uma razão.

²¹ Promessas são mais difíceis aqui, por causa da natureza reflexiva do dar razões que elas envolvem. Mas é difícil negar que as pessoas frequentemente mantêm uma relação normativamente significativa com elas mesmas.

maneira em que geralmente o dar razões puramente epistêmicas e acionadas não são.²²

3.3 Dando Razões Epistêmicas

Até agora falei somente sobre razões *para ação* que damos uns aos outros. Todavia, nós também damos razões uns aos outros *para crer*. Em qual dos três sentidos distinguidos acima nós nos damos razões para crer?

Obviamente, há exemplos de dar razões puramente epistêmicas deste tipo. Por direcionar sua atenção para uma digital, eu posso epistemicamente dar-lhe uma razão para crer que o mordomo fez isso. Ainda há casos bem diretos de dar razões epistêmicas no sentido acionado também. Possivelmente, por exemplo, por conduzir um certo experimento original, eu posso fazer com que seja o caso em que você tenha uma razão para acreditar na verdade da minha teoria, uma razão que não existia antes do experimento.²³ A questão mais interessante no nosso

²² O ponto desse texto é consistente com a observação (da qual eu agradeço à Hagit Benbaji e Yuval Eylon) de que no contexto de relacionamentos pessoais, algumas vezes requisições existam em um sentido redundante. Isso talvez seja assim, primeiramente, porque, como já foi dito, requisições algumas vezes tenham uma função epistêmica – apontando a atenção do endereçado para as necessidades ou desejos relevantes do falante – mas dentro de relacionamentos próximos aqui, algumas vezes talvez não haja necessidade para requisições desempenhar este papel epistêmico, pois aqueles em relacionamentos pessoais por vezes sabem muito mais sobre as necessidades e desejos uns dos outros do que os outros sabem. E segundo, se preocupar pelo bem-estar um do outro é presumivelmente uma parte constitutiva de muitas formas de relacionamentos próximos, e então o endereçado de uma requisição em potencial talvez tenha uma razão para desempenhar uma ação relevante independentemente de uma requisição. Além disso, essa razão não dependente de requisição talvez seja esmagadoramente mais forte que aquelas presumivelmente fornecidas pela requisição. Então, fazendo a requisição será, embora normativamente não inerte, ainda assim praticamente redundante. São tais considerações que explicam como, dentro de relacionamentos próximos, reclamações como “você não deveria ter pedido” talvez façam total sentido. Mesmo em tais contextos, entretanto, requisições ainda farão diferença, como talvez possa ser observado na adequação de locuções tais como “Eu preciso mesmo pedir”, uma locução que simultaneamente indica tanto a (presumida) presença de uma razão que é independente da possivelmente futura requisição, e o ato que a possivelmente futura requisição irá, ainda assim, ser realizada, fazer uma diferença normativa.

²³ Mas as coisas são enganosas aqui. No caso de razões epistêmicas, a distinção entre dar razões puramente epistêmicas e meramente acionadas as vezes se torna problemática. Pode ser talvez

contexto é, claramente, em que nós podemos dar uns aos outros *robustamente* razões epistêmicas. Aqui parece para mim que a resposta é claramente um não.

Suponha que a minha requisição não seja para você leia meu artigo, mas sim, para que você acredite que é um bom artigo. Parece ter algo profundamente errado com está requisição. Não parece o tipo de coisa que pode ser candidata a uma requisição genuína. Tanto faz, que haja uma pressão significativa neste caso para interpretar a minha requisição (ou minha “requisição”) de alguma forma não padrão: talvez eu esteja confuso sobre a natureza de requisições, ou talvez eu esteja realmente pedindo que você *diga* que ele é bom, ou de outra forma que você haja *como se* você acreditasse que ele é bom, ou que você de alguma forma *prepare-se* para ter esta crença. Pontos similares parecem se aplicar às tentativas em dar razões epistêmicas empregando comandos (“Eu, por este meio, ordeno você a acreditar que meu artigo é bom! ”) ou promessas (“Deixe-me ler o seu artigo. Eu prometo que eu vou acreditar que ele é bom”).

A diferença entre razões epistêmicas e práticas podem ser colocadas também em termos de uma discussão da subseção anterior. Aqui, observo como a natureza do relacionamento pessoal relevante é relevante para o dar razões robustas, mas não para outros tipos de dar razões. Parece a mim que este resultado coere satisfatoriamente com o ponto dos últimos dois parágrafos, de acordo com o qual razões epistêmicas não podem ser dadas robustamente, pois, parece a mim, que relacionamentos pessoais são completamente irrelevantes quando vamos dar razões epistêmicas.

argumentado, por exemplo, que o caso de experimento é um em que a razão para acreditar na verdade da teoria estava lá o tempo todo, e por conduzir o experimento eu meramente aponte sua atenção para ela, então que é um caso de dar razões puramente epistêmicas ao invés de acionadas. Ou talvez pode ser argumentado que enquanto a razão estava lá o tempo todo, você não a tinha antes do experimento. Essas complicações, embora sejam interessantes, não são aquelas às quais preciso me direcionar em detalhes para meus propósitos aqui.

Razões epistêmicas, então, não podem ser robustamente dadas.²⁴ Uma consideração sobre dar razões robustas deve explicar o porquê.²⁵

4. Podemos Conseguir o Que Queremos?

Casos de dar razões robustas – como em requisições, comandos, promessas – são diferentes de casos de dar razões puramente epistêmicas ou meramente acionadas. Dar razões robustas parece envolver uma dimensão pessoal que tipicamente não é uma parte de dar razões puramente epistêmicas ou meramente acionadas. Enquanto razões epistêmicas podem ser certamente dadas, elas não podem ser robustamente dadas. Como para ameaças: enquanto elas dão razões de uma maneira próxima àquelas das requisições (ou para dar razões robustas mais comumente) em que elas envolvem uma intenção para (desta forma) dar a alguém uma razão, ainda assim, ameaças não contam como dar razões robustas.

A discussão até então, resumida no último parágrafo, estava num sentido não ambicioso, por enquanto foi estabelecido a caracterização do que parece ser os nossos pensamentos pré-teóricos sobre o tipo de dar razões envolvidos (por exemplo) em requisições, reservando-lhe o nome de ‘dar razões robustas’. Mas agora é hora de teoria. Por nada até então dito, exclui-se a possibilidade que dar razões robustas como caracterizado, envolveria muito mais uma forma de magia normativa, e é,

²⁴ Todavia veja nota 57 adiante para uma complicação relevante.

²⁵ Darwall (2006, p. 253) também discute a diferença relevante entre raciocínio teórico e prático, sugerindo que reside no fato de que enquanto as visões dos outros podem ser completamente descartadas epistemicamente se equivocadas, isso não ocorre no caso prático. Se aplicarmos o ponto das requisições, se torna um ponto interessante em que requisições que não devem ser feitas não são necessariamente normativamente vazias (veja Cupit [1994, pg; 543]). Mas de fato acredito que aqui Darwall não está apreciando a complexidade do caso epistêmico. As visões dos outros, mesmo quando equivocadas, podem servir como evidência. Para alguma discussão relevante no contexto dos fenômenos de “desacordo entre pares”, veja o meu “Not Just a Thuthometer: Taking Oneself Seriously (But Not Too Seriously) in Cases of Peer Disagreement” (forthcoming).

portanto, simplesmente impossível, então, que ninguém pudesse dar para alguém uma razão para Φ ou qualquer coisa como o sentido que tentei capturar na seção anterior.²⁶ Nesta seção, então, avançarei alguns passos na direção de desenvolver uma consideração do que consiste em dar razões robustas, uma consideração que, se bem-sucedida, vai acomodar os dados da seção anterior.²⁷ Nas Seções 4.1 e 4.2, mostrarei porquê dar razões robustas “deve ser” – se existe tal coisa, e aparências ao contrário, não obstante – uma instância particular do dar razões acionadas. Na seção 4.3, brevemente considerarei opções mais radicais de uma teoria do erro e uma visão da irredutibilidade do dar razões robustas. Na Seção 4.4, finalmente apresentarei minha consideração sugerida de dar razões robustas, e na seção 4.5, mostrarei como esta consideração satisfaz a desiderata da seção anterior.

4.1 Razões condicionais e Condicionais

Não há espaço plausível no espaço lógico para dar razões robustas ao que não é uma instância de dar razões acionadas. Aqui, a história iniciada do porquê (que será refinada adiante).

Pense sobre a minha requisição para ler o meu rascunho. Estamos assumindo que, antes da requisição, você não tinha razões para ler o meu rascunho, e que depois você tinha. Isso significa que o condicional “Se eu pedir para você ler o meu rascunho, você terá uma razão para o ler” era verdadeiro o tempo todo, ou de qualquer

²⁶ A preocupação análoga no que diz respeito à teoria de Darwall é que razões de segunda pessoa como ele define então simplesmente não existem. Não encontrei no livro de Darwall uma réplica convincente para a preocupação. Para sua tentativa, veja, por exemplo, Darwall (2006, p. 299).

²⁷ É dito algumas vezes que cada um de nós é uma “fonte auto originadora de reivindicações válidas”. (Rawls [1980, p. 546], ou algum desses. Não é de todo claro, naturalmente, por que essa metáfora aparece. Mas primeiro, é importante ver que essa metáfora é no melhor um nome cativante para o nosso problema (no que o dar razões robustas consiste, e como é que é possível), não uma solução para ele. E segundo, a discussão que se segue talvez seja vista como uma forma de pesar essa metáfora.

forma concisamente anterior, – e independentemente de – eu realmente fiz seu antecedente verdadeiro (requisitando que você leia o meu rascunho). É muito tentador pensar neste caso como ainda outro caso de dar razões acionadas; pois tudo o que fiz aqui foi manipular as circunstâncias não normativas de forma que acionei sua razão condicional de-ler-o-rascunho-se-eu-te-pedir.

Entretanto, seria muito rápido concluir imediatamente que todos os casos de dar razões robustas são instâncias do dar razões acionadas. Como está, o argumento no último parágrafo move também rapidamente a verdade do condicional (“Se eu pedir para você ler meu rascunho, você terá uma razão para o ler”) e sua independência da minha requisição, para a existência de uma razão condicional que é independente da minha requisição. Como é bem conhecido, tais condicionais podem ser lidas de mais de uma maneira, e nem todas as leituras autorizam tal inferência.

O condicional pode ser lido pelo menos das duas formas: o operador normativo (“você tem uma razão para”) pode ser entendido como tendo um escopo largo, alcançando todo o condicional, resultando em, grosseiramente:

Escopo largo: você tem uma razão para ler o meu rascunho se eu pedir para você o ler.

Ou ele pode ser lido como tendo um escopo limitado, alcançando somente o consequente, com o antecedente permanecendo inteiramente não normativo, fora do escopo de qualquer operador normativo relevante:

Escopo limitado: Se eu pedir para você ler meu rascunho você tem uma razão para o ler.

Nós não temos que nos comprometer aqui com a literatura recente sobre escopos largos ou limitados.²⁸ Para nossos

²⁸ Veja, por exemplo, Schroeder (2004), e as referências ali

propósitos, será suficiente que o escopo limitado é *uma* leitura possível do condicional, e de fato uma que é talvez mais próxima do que uma alternativa à sua formulação na linguagem natural. Além disso, a linha intuitiva de pensamento apresentada acima – a saber, aquela pela qual você não tinha uma razão antes de eu pedir-lhe para que lesse meu artigo, e possui mais tarde, mostra que o condicional é verdadeiro – não suporta o condicional de escopo largo sobre o de escopo limitado. A viabilidade desta leitura de escopo limitado do condicional, como também outros dados²⁹, mostra que a verdade do condicional não precisa implicar na existência de uma razão condicional. A verdade do condicional em si mesma, então, não estabelece a afirmação de que dar razões robustas pode somente ser uma instância do dar razões acionadas.

No entanto, uma preocupação permanece.³⁰ Pois é um pensamento natural que enquanto a leitura do escopo limitado não implica a leitura do escopo largo, ainda assim a única *explicação* plausível para a primeira seria em termos da última. Afinal, requisições são especiais em alguma maneira normativamente relevante. Se eu tivesse declarado palavras bem diferentes, teria me comprometido com atos de fala bem diferentes (ou se tivesse evitado de me comprometer a quaisquer atos de fala relevantes), eu não teria lhe dado uma razão para ler o meu artigo. Além disso, nosso relacionamento pessoal é, como enfatizei acima, relevante aqui. Talvez, por exemplo, se uma completa estranha pedisse para você ler o rascunho dela, ela não conseguiria deste modo dar-lhe uma razão para fazê-lo. Minha requisição é especial, então, não apenas comparada a outras coisas que eu poderia fazer (ou falhado fazer) mas também comparado à (algumas) outras requisições. O

²⁹ Como notado mais cedo, o modelo acionador encompassa mais que simplesmente o acionar de razões condicionais. Ele inclui também o acionar de uma razão por assegurar que uma condição habilitada está no lugar, por anular o anulador, etc. Nestes casos também, o condicional relevante é verdadeiro, mas não há razão condicional relevante.

³⁰ A discussão no resto da subseção e na próxima depende pesadamente sobre o artigo de Mark Schroeder “Cudworth e Normative Explanations” 2005, no qual estou muito endividado.

que é isso, então, que explica o porquê minha requisição para que você leia o meu rascunho consegue dar-lhe uma razão, mas todas essas outras coisas não conseguem? A resposta natural parece ser em termos de algo como o condicional de escopo largo: a unicidade normativamente relevante das requisições, e realmente da minha requisição, é precisamente devido à verdade de algo como o condicional de escopo largo. A única coisa que pode explicar porque minha requisição criou razões aqui, em minha exclamação “O rascunho que estou trabalhando é muito legal!”, não é precisamente porque você tem uma razão condicional prévia para-ler-o-meu-rascunho-se-eu-lhe-pedir, mas porque você não tem uma razão condicional prévia para-ler-meu-rascunho-se-eu-dizer-que-ele-é-muito-legal. A preocupação, então, é que a única forma que o condicional de escopo limitado pode ser *não misteriosamente* verdadeiro é se o condicional de escopo largo explicar sua verdade. E nós sabemos que sempre que uma razão pode ser robustamente dada, pelo menos alguma coisa, como o condicional de escopo limitado, deve ser verdadeira. Então em ordem de evitar a misteriosidade de um condicional de escopo limitado bruto deste tipo, nós devemos concluir que sempre que uma razão pode ser robustamente dada, o condicional de escopo largo é verdadeiro, independentemente do ato de dar razões (como, na realização de uma requisição). Então, nós devemos concluir que qualquer caso de dar razões robustas (ou aparentemente robustas) é realmente um caso de (grosseiramente falando) uma razão condicional.³¹

4.2 O Modelo Constitutivo

Schroeder argumenta, bem convincentemente, acredito, que o modelo explicativo subjacente a este pequeno argumento – o que

³¹ Esse argumento é uma instância particular do “Argumento Cudworthy” de Schroeder, em sua generalização do argumento contra a teoria do comando divino ele encontra em Cudworth, ainda que aplicado às razões, não às obrigações. E observe especialmente a citação de Price na página 12 e a discussão relevante de Schroeder.

ele chama de “O Modelo Padrão”, de acordo com o qual a verdade de qualquer condicional normativo de escopo limitado é explicada pela verdade de alguma declaração normativamente categórica (por exemplo, os condicionais de escopo largo da seção anterior) – é *senão um* modelo explicativo, em que há, pelo menos, uma alternativa, ao que ele chama de “O Modelo Constitutivo”.

Se você for um teórico do comando divino, por exemplo, você acredita que para qualquer Φ , se Deus comandou que você Φ , então você está sobre uma obrigação para Φ . Isso não significa que você está comprometido com a afirmação de que há outra obrigação mais geral, uma da qual não dependa do comando de Deus, isto é a obrigação geral para fazer como Deus comanda, ou a obrigação condicional para- Φ -se-Deus-comanda-que- Φ , ou qualquer coisa assim. É verdade que, você ainda nos deve uma perspectiva de como é que quando Deus comanda que você Φ , você repentinamente está sobre uma obrigação para Φ , mas você frequentemente não está, quando eu comando (ou “comando”) que você Φ . A forma de explicar isso é apontar para o fato que obrigações são (talvez parcialmente) *constituídas* pelos comandos de Deus (mas infelizmente, não pelos meus), que ser comandado por Deus para Φ (talvez de forma grosseira) *é somente o que é* estar sobre uma obrigação.³² O mesmo acontece, argumenta Schroeder, para qualquer outra teoria perfeitamente geral da obrigação moral: o condicional capturando o núcleo de qualquer que seja a teoria – sempre que Deus comanda que você Φ . Você está sobre a obrigação para Φ : sempre que Φ -zar irá maximizar a utilidade, você está sobre a obrigação para Φ ; sempre que as partes em alguma situação privilegiada de escolha requeira que você Φ , você está sobre uma obrigação de Φ ; etc – deve ser explicado de acordo com o Modelo Constitutivo ao invés do Modelo Padrão (na

³² Como John Deigh lembrou-me, este rascunho de uma teoria do comando divino talvez seja uma caricatura. Talvez haja versões da teoria do comando divino de acordo com a qual o poder de dar razões de deus não é tão diferente daquele de um dar razões robustas humano. Esses detalhes não nos preocupam aqui, entretanto.

pena de ser vítima do “Argumento Cudworthy” de Schroeder, uma instância que concluiu na última subseção).

Voltando, então, para dar razões, o argumento que tenta mostrar que qualquer dar razões robustas é, na verdade, meramente o acionamento de (grosseiramente) uma razão condicional pré-existente que pode ser resistido se pudermos oferecer uma explicação das verdades dos condicionais relevantes (como “Se requisito que você leia o meu rascunho, você vai ter uma razão para fazê-lo”) junto às linhas do Modelo Constitutivo. Isso pode ser feito?

Bem, em ordem de fazer isso, alguém teria que argumentar que ter uma razão para ler um rascunho *consiste em* eu ter-lhe pedido para o lesse, que a requisição *constitui* haver uma razão para o ler, que ter sido requisitado para ler o rascunho é (talvez parcialmente) *simplesmente o que é* ter uma razão para o ler. Contudo, isso parece, somente, completamente implausível.³³ Quaisquer que sejam os problemas da teoria do comando divino, pelo menos o lugar de Deus no universo (e na teoria) dá alguma plausibilidade para a afirmação de que estar sob uma obrigação é *simplesmente* ser comandado por Deus. Nenhuma plausibilidade é carregada para o caso em mãos. A sugestão de que a relação entre requisições – ainda pior, *minhas* requisições – e de que você tenha razões que são constitutivas é muito para acreditarmos. Parece ainda mais difícil de acreditar que qualquer coisa como isso está acontecendo quando lembramos que requisições dão razões somente algumas vezes, dependendo de muitos fatores contextuais. O teórico do comando divino tem um condicional simples que precisa explicar usando uma explicação Modelo-

³³ Veja a distinção entre dizer que uma requisição constitui a razão, é dizer que ter sido requisitado é somente o que é ter uma razão (aqui). No texto atribuo implausibilidade não ao primeiro, mas somente ao último, e é o último que é um análogo apropriado da reivindicação de que ser comandado por Deus é somente o que é estar sobre uma obrigação para o exemplo da teoria do comando divino. Então, isto é essa reivindicação (implausível) que é necessária para a aplicação do modelo constitutivo para dar razões robustas.

Constitutiva (sempre que Deus comandar que você Φ você está sob uma obrigação para Φ), e então a afirmação constitutiva que ela determina é similarmente simples (estar sobre uma obrigação para Φ simplesmente é ter sido comando por Deus para Φ). O condicional que teríamos que explicar em ordem de aplicar o Modelo Constitutivo para o caso de dar razões teria de ser muito mais complicado: quando alguém pede para que você Φ , e quando Φ -zar não é tão difícil, e quando um relacionamento pessoal entre você e a pessoa emitindo a requisição é tal para suportar tais requisições, e quando Φ -zar não é muito imoral, e ... então você tem uma razão para Φ . Consequentemente, e também por que nem todas as razões dependem de requisitos, a afirmação constitutiva aqui necessária seria terrivelmente complicada, e muito menos plausível do que a afirmação constitutiva da qual o teórico do comando divino necessita.

O Modelo Constitutivo talvez seja aplicado aqui também de uma maneira um pouco diferente.³⁴ Suponha que pensamos que haja algum outro condicional que é explicado na maneira constitutiva apontada. Talvez, por exemplo, pensemos que a verdade do condicional “se você quer X, e acredita que Φ -zando você pod conseguir X, então você tem uma razão para Φ ” é explicada pelo fato de que ter uma crença relevante e um desejo é simplesmente o que é ter uma razão. Suponha que você acredita que com Φ -zando você pode conseguir X, mas você (ainda) não quer X. Eu posso então dar-lhe – robustamente, aparentemente – uma razão para Φ fazendo com que você queira X. Se eu posso fazer com que isso lhe dê um desejo, irei assim estar fazendo com que seja o caso de você ter uma razão para Φ (por ter uma crença de que você já tem o desejo, que vou fazer com que você adquira, é somente o que é preciso para ter uma razão), e não por acionar uma razão condicional pré-existente (para a explicação do condicional “se você quer X, e acredita que Φ -zando pode fazer

³⁴ Eu agradeço à Mark Van Roojen por me fazer ver isso.

“você conseguir X, então você tem uma razão para Φ ” é uma instância do Modelo Constitutivo ao invés do Modelo Padrão).

Concordo que há lugar no espaço lógico para essa forma de aplicar o Modelo Constitutivo. No entanto, acredito que não precisamos persegui-lo mais aqui, pelas duas seguintes razões: primeira, não é claro para mim que essa maneira de pensar o dar razões robustas pode satisfazer a desiderata sublinhada na Seção 3 (não é óbvio, por exemplo, porque tal dar razões não pode ocorrer no que diz respeito às razões epistêmicas). Segunda razão, e mais importante, eu simplesmente não vejo como os detalhes podem ser preenchidos aqui de uma maneira plausível. Eis as condições que devem ser satisfeitas para algo, ao longo dessas linhas, trabalhar para nosso caso das requisições: primeiro, deve haver alguma característica constitutiva de ter uma razão (prática); e segundo, deve ser o caso em que, por requisitar que você leia o meu artigo, eu faça com que seja o caso em que aquela coisa que é constitutiva de ter uma razão – qualquer que seja ela – de fato se sustente. Eu não possuo um argumento mostrando que não há forma de preencher isto em detalhes que possam satisfazer essas condições. Entretanto, não posso pensar em nenhuma forma remotamente plausível de fazer isso. Então presumo que não há nenhuma forma.

Vamos recapitular. A preocupação era que, dada a verdade dos condicionais relevantes (“Se eu pedir para você ler o meu rascunho, você vai ter uma razão para o ler”), e a necessidade para explicá-lo, seríamos forçados a reconhecer algo como uma razão condicional (para ler-se-eu-pedir-para-você). Seguindo Schroeder, aponte que talvez haja outras formas de explicar os condicionais relevantes, por exemplo, de acordo com o Modelo Constitutivo. Porém, agora reivindiquei que para o Modelo Constitutivo se aplicar para os casos em mãos, algumas proposições altamente implausíveis seriam verdadeiras. Então o Modelo Constitutivo – quaisquer que sejam seus méritos em geral – não pode nos ajudar aqui.

Existem, assim, outros tipos de explicações dos condicionais de tipo relevante? Schroeder introduz (12) “A Conjectura Padrão-

Constitutiva”, de acordo com a qual *somente* explicações de tais condicionais são tanto alinhadas com o Modelo Padrão, quanto alinhadas com o Modelo Constitutivo. Ele as introduz como uma conjectura, não oferecendo nenhum argumento para a afirmação de que é de fato verdadeira. Não possuo tal argumento para oferecer, também. Contudo, não posso pensar em outro possível modelo explicativo aqui³⁵, e então vou prosseguir na suposição de que nenhum poderá ser encontrado.³⁶

4.3 A Cila da Teoria do Erro e o Caríbdis da Irredutibilidade

Na seção 4.4 vou oferecer uma perspectiva que compreende o dar razões robustas como uma instância particularmente interessante do dar razões acionadas³⁷, as características especiais das quais (comparado aos outros casos de dar razões acionadas) acomodam a desiderata da seção 3. Antes de fazer isso, no entanto, deixe-me brevemente endereçar mais duas estratégias extremas, que a esta altura, já se sugerem.

A primeira é uma teoria do erro sobre dar razões robustas. Através dos argumentos até agora estabelecidos, pode-se demonstrar que simplesmente nada pode se qualificar como dar razões robustas genuínas. Nada, afinal, pode ser uma instância de dar razões acionadas (como a discussão nesta seção sugere que dar razões robustas deve ser) e falhar em ser uma instância de dar

³⁵ Em conversa, Mark Schroedder deixou claro que mesmo ele não pode.

³⁶ John Gardner sugeriu-me a seguinte forma heurísticamente útil de pôr o meu ponto aqui. Se eu penso em razões tendo em mente um clássico silogismo prático, então dar razões epistêmicas equivale a apontar em um completo silogismo prático que era disponível ao endereçado o tempo todo; dar razões acionadas equivale a acarreta uma mudança na premissa menor; instâncias do Modelo Constitutivo são aquelas onde a nova premissa maior é acarretada; e a reivindicação de que o Modelo Constitutivo não se aplica às requisições equivale à reivindicação que requisições podem somente acarretar em mudanças na premissa menor, embora em silogismos com, interessantemente, somente premissas maiores.

³⁷ Raz (1986, p. 84) parece dizer explicitamente que (nos meus termos) dar razões robustas deve ser uma instância de dar razões acionadas, quando ele escreve (no contexto de uma discussão sobre consentimento): “Embora a [vontade] possa, dentro dos limites, criar razões, ela pode fazer isso somente quando há uma razão não baseada na vontade.” Para um ponto semelhante, veja Watson (2009, p. 159)

razões acionadas (como a discussão na seção anterior parece sugerir que deve ser verdadeira sobre o dar razões robustas). Dar razões robustas, então, é simplesmente impossível.

No outro extremo, nós talvez sejamos conduzidos para uma visão da irreducibilidade, de acordo com a qual o fenômeno de dar razões robustas é *sui generis*, não melhor entendido ou explicado em outros termos. O erro na discussão, até agora, embora possa prosseguir, foi uma tentativa de oferecer uma perspectiva de dar razões robustas em outros termos, não tão proximamente relacionados. Todavia é esta tentativa que foi reduzida ao absurdo, e então deve ser abandonada. Nós devemos simplesmente reconhecer que a esfera normativa é mais rica do que pensávamos, e ela contém – no topo de tudo mais que ela contém – também a relação irreducível de dar razões robustas.

Ambas estas sugestões são de fato opções, mas são aquelas opções somente para um último (ou penúltimo) recurso, pois ambas são altamente implausíveis. Começando com a teoria do erro. Tornar-se um teórico do erro sobre dar razões robustas implica em haver algum erro sistemático na nossa prática de dar razões (por exemplo, na nossa prática de fazer requisições), portanto, que nós nunca, de fato, conseguimos dar uns aos outros razões da forma caracterizada na Seção 3. Como é comum com teorias do erro, para a teoria ser de todo plausível, uma explicação deve ser oferecida sobre o erro relevante. Entretanto, tal teoria do erro, aparentemente, é suficientemente implausível³⁸ sendo evitável se possível.

Assim como para a irreducibilidade, aqui como em outro lugar, há algo menos que teoricamente satisfatório sobre este movimento. Nós já vimos que, em que dar razões robustas é possível, um certo tipo de condicional é verdadeiro, tal como “se eu

³⁸ Talvez pudesse ser um pouco (mas não muito) menos implausível se suavizado por algum tipo de instrumentalismo ou ficcionalismo sobre dar razões robustas. Watson (2009, p. 157, nota de rodapé 9) sugere algo sobre essas linhas como uma possibilidade de ler Hume sobre promessas.

pedir para você ler o meu rascunho, você terá uma razão para fazê-lo”. Discutimos dois tipos possíveis de explicação para a verdade de tais condicionais. Endossando dar razões robustas como um fenômeno normativo, irreduzível, *sui generis*, seria o equivalente então a admitir a verdade de tais condicionais como brutas, inexplicadas ou inexplicáveis. E isso parece um movimento desesperado (embora talvez não tão desesperado *como* endossar uma teoria do erro).³⁹

De todo possível, então, nós devemos dirigir rumo a algum lugar entre os dois extremos de uma teoria do erro e irreduzibilismo. Agora, dada a discussão anterior nesta seção, dar razões robustas (se for possível) não pode ser senão uma instância de dar razões acionadas. E nós já sabemos – da seção anterior – que dar razões robustas têm características que não são compartilhadas por outros casos, talvez, mais paradigmáticos de dar razões meramente acionadas. Então, se dar razões robustas é possível, deve ser porque são uma instância especialmente interessante do acionamento de razões⁴⁰, um dos elementos especiais (esses não compartilhadas por outras instâncias do dar razões acionadas) que explicam as características únicas de dar razões robustas.

4.4 Minha Sugestão

Um começo promissor é pensar sobre as intenções características do concessor de razões (e talvez também do

³⁹Darwall (2006) enfatiza a irreduzibilidade da segunda pessoa. E isso já é algo insatisfatório. Mas Darwall em nenhuma parte discute algo como os condicionais do tipo mencionado no texto, e então, não reconhece o ainda mais problemático resultado – admitindo a verdade de tais condicionais como brutos. Lavin (2008) também enfatiza um semelhante desafio explicativo deixado não respondido por Darwall.

⁴⁰No contexto da discussão da autoridade – novamente, plausivelmente, uma instância particular de dar razões robustas – Raz (2005, p. 1012-3, 1020) claramente pensa que o dar razões envolvido é uma instância particular do (que eu chamo) de dar razões acionadas. Não é claro a mim se ele pensa que isso é meramente tal instância particular, ou uma que pode ser especial de forma grosseira, que sugiro a seguir.

recebedor de razões) em casos de dar razões robustas, comparado com casos de dar razões puramente epistêmicas ou meramente acionadas. Parece uma condição necessária para que alguma coisa seja qualificada como uma requisição, que a pessoa fazendo a requisição tenha uma intenção para, deste modo, dar uma razão para o endereçado. Algo como essa condição necessária também se sustenta para o caso de comandos, como a literatura sobre autoridade deixa claro: ninguém pode emitir um comando inadvertidamente, sem ter a intenção de emitir um comando, isto é, (talvez misturado com outras coisas) dar ao endereçado uma razão forte para performar a ação relevante. Note que a condição não é necessária para muitos casos de dar razões puramente epistêmicas ou meramente acionadas. Alguém pode, por exemplo, indicar que há uma razão inadvertidamente, em todo caso, sem pretender fazer isso. E alguém pode certamente acionar uma razão condicional sem pretender fazer isso, como o exemplo do merceiro mostra claramente. Então, insistir na intenção, parece certamente um passo na direção correta.

Entretanto, já sabemos que a intenção não é uma condição suficiente para dar razões robustas, pois há casos onde a intenção está presente, mas a razão dada relevante não é robusta. Um tipo de caso que já foi mencionado – o das ameaças. Mais comumente, tome um caso de dar razões acionadas (como esse do merceiro e o preço do leite), então adicione uma intenção para dar uma razão, e ainda tudo que você possui é um caso de dar razões acionadas, não robustas. Ou considere um exemplo que tirei de David Estlund(2008, p.188).⁴¹ O filho de um ditador brutal “ordena” que você realize uma ação. Ele não é, claramente, autorizado a emitir tal ordem, nem mesmo de acordo com as regras que seu pai ditador aceita. Entretanto, se você não fizer como a criança diz, seu pai ficará muito irritado, de fato irritado, ao ponto de brutalizar algumas pessoas inocentes. Neste caso, parece que você agora tem

⁴¹Estlund mesmo, deu crédito por esse exemplo para John Deigh.

uma razão (de fato, uma obrigação; mas para o meu propósito uma razão adianta) para fazer o que quer que seja que o filho do ditador lhe ordenou. O filho do ditador conseguiu lhe dar uma razão para ação. De fato, isso é exatamente o que ele tinha intenção de fazer. Claramente, este não é um caso de dar razões robustas. É meramente um caso de dar razões acionadas: você tem uma razão que é sustentada, a fim de prevenir desastres horríveis de acontecer com pessoas inocentes, e o filho do ditador manipulou com sucesso as circunstâncias não normativas de modo que essa razão geral irá implicar em alguém específico fazer o que ele diz.⁴²

Para fazer algum progresso, compare ameaças com advertências. Quando todas as outras coisas são iguais, se uma ameaça (“se você não fizer Φ , eu vou te matar!”) lhe dá uma razão para Φ , então a advertência relevantemente similar seria (“se você não fizer Φ , O Cara Mal vai te matar!”). A intenção daquele emitindo a ameaça, é lhe dar uma razão para ação – embora seja uma parte do que faz uma ameaça – ainda neste sentido, desliga-se da imagem normativa.⁴³ Uma pessoa benevolente emitindo uma advertência, ou de fato um indicador que não é uma pessoa de um perigo iminente, faria o mesmo nos termos de lhe dar uma razão para Φ .

Agora considere requisições novamente, e considere os dois seguintes casos. No primeiro, pedirei para você ler o meu artigo, e você vai em frente e o lê porque, bem, eu lhe pedi para fazê-lo. No segundo caso, você não pode se importar menos comigo e com minha requisição. Porém, você percebeu que nossa chefe de departamento me ouviu pedindo para que você leia o meu artigo, e você sabe o que ela pensa que você deve fazer enquanto eu lhe

⁴²Estlund distingue o caso do filho do ditador do caso de uma autoridade real. Mas ele não oferece uma explicação satisfatória da distinção. Veja que o próprio fato de que esse caso pareça tão diferente do caso de uma autoridade genuína, mostra que seria errado pensar sobre casos de autoridade como instâncias não interessantes particulares de dar razões acionadas.

⁴³ Talvez isso é grosseiramente o que Estlund tinha em mente quando ele disse (sobre o exemplo do filho do ditador): “o comando em si mesmo sai fora do conjunto de razões para ação” (2008, p. 118).

peço, e você reconhece sua razão geral, razão permanente de mantê-la feliz. Você aceita ler meu artigo meramente para evitar um conflito com a chefe de departamento. Parece haver uma diferença importante entre os dois casos. Embora você tenha feito o que lhe pedi em ambos os casos, no último a requisição não funcionou da maneira que intencionei que funcionasse. Pense desta maneira: se você vier a acreditar que a minha requisição não era de fato sincera (isto é, eu estava simplesmente fingindo fazer uma requisição), neste primeiro caso você concluiria que você não tem nenhuma razão para ler o meu rascunho. Não é assim no segundo caso: desde que a chefe de departamento acredite que a requisição foi sincera, sua razão para ler o rascunho permanece. Então, o que é importante no caso das requisições é que o conessor de razões não só tenha a intenção de dar uma razão, mas também que ele tenha a intenção de que conceder a razão dependa do reconhecimento do recebedor da razão, de que esta é sua intenção, e de fato que este reconhecimento desempenhe um papel apropriado no raciocínio prático do recebedor de razões. Não há nada, então, que permanece na mesma relação para requisições como advertências estão para ameaças: uma ameaça tratada pelo endereçado meramente como uma advertência é uma ameaça completamente bem-sucedida. Uma requisição tratada pelo endereçado meramente como um incentivo não é uma requisição completamente bem-sucedida.

Podemos ganhar mais suporte para isso pensando sobre a variante do caso de manter-a-chefe-de-departamento-feliz. Suponha que eu sabia o tempo inteiro quão pouco você se importava comigo e minhas requisições, e então eu nunca iria ter lhe incomodado ao pedir para você ler o meu artigo se não fosse pela presença da chefe de departamento. Dado sua presença, eu sei que ao declarar “você se importaria em dar uma olhada no meu rascunho?” eu conseguiria dar-lhe uma razão para ler o meu rascunho. Parece a mim que este caso – no qual eu tenho a intenção de dar a você uma razão e consigo fazer isso ao declarar

palavras que soam como fazer uma requisição – não é uma requisição genuína afinal, precisamente porque o dar razões envolvido não é robusto. Antes, isto é um caso de dar razões acionadas intencionalmente. O que parece estar faltando aqui é a intenção complicada mencionada anteriormente: apesar de a intenção de lhe dar uma razão, eu não tive intenção de lhe dar a razão dependendo do seu reconhecimento da minha intenção.

Essas considerações, suportam, então, a seguinte perspectiva que será conveniente para apresentar como uma perspectiva de tentativa para dar razões robustas, então, adicionando as condições de sucesso relevante:

Uma pessoa A tenta dar robustamente à outra pessoa B uma razão para Φ apenas no caso (e por que):

- (i) A tem intenção de dar para B uma razão para Φ , e A comunica essa intenção para B;
- (ii) A tem intenção de B reconhecer sua intenção;
- (iii) A tem intenção de que a razão dada a B Φ dependa de uma forma apropriada sob o reconhecimento de B da intenção comunicada de A para dar a B uma razão para Φ .⁴⁴

⁴⁴ Há uma óbvia similaridade estrutural entre essa perspectiva e (uma versão de) uma perspectiva de Grice de sentença-significado. Não é claro a mim se a similaridade é meramente superficial. Seja como for, mesmo que existam algumas objeções que façam uma perspectiva Griceana do significado inaceitável, elas não parecem se aplicar, a mim, para o caso de dar razões robustas. Pelo contrário, elas são instâncias de correntes causais desviantes do tipo que menciono no texto abaixo, ou sua força como objeções a perspectiva griceana depende inteiramente do último possuir uma perspectiva do significado. Este ponto está mais claro em Schiffer (1972, p. 42-3), no qual ele distingue o fenômeno do falante-significado e de contar (que ou para), relevando que (nos meus termos) dar razões puramente epistêmicas é consistente com o primeiro, mas não com o último. E para uma investigação dos problemas enfrentados amplamente pela perspectiva Griceana do significado, veja Schiffer (1987, Capítulo 9). Para um entendimento semelhante de requisições, desta vez inspirado por Searle ao invés de Grice, veja Cupit (1994, p. 450): “Requisitar é uma tentativa de afetar as ações do outro, não fazendo mais que apresentar aqueles desejos de uma forma que constitua uma tentativa para afetar a ação.” Similarmente, Raz (1975, p. 83) escreve: “Uma pessoa que faz uma requisição tem a intenção de que fazer a requisição seja uma razão para o endereçado cumprir com ela [a intenção]. E há algo na mesma direção geral em Robertson (manuscrito). E para discussões Griceanas explícitas dos assuntos intimamente relacionados, veja Hart (1982), e Sciaraffa (2009).

A terceira condição pode ser entendida como uma generalização de tal pensamento natural, como quando peço a você Φ , tenho a intenção de que sua razão para Φ -zar seja aquela *que eu pedi para você*; quando eu comando que você Φ , eu tenho a intenção de que a sua razão para Φ -zar *seja essa que disse*, etc.

Então, quando eu peço para você ler o meu rascunho, (i) eu tenho a intenção de lhe dar uma razão para ler meu rascunho, e comunico essa intenção a você pela forma de dizer algo como “você se importaria em dar uma olhada?”⁴⁵, (ii) eu tenho a intenção de que você reconheça essa intenção; e (iii) eu tenho a intenção de que seu reconhecimento desempenhe um papel apropriado no seu raciocínio prático, como pode ser visto através do fato de que meu pedido não funciona quando você lê o meu rascunho (como lhe pedi) simplesmente pela razão de que isso irá manter a chefe de departamento feliz. E uma análise similar parece plausível para, pelo menos, alguns outros casos de dar razões robustas (por exemplo, aquele de comandos).

Há dois tipos de condições de sucesso necessárias para o dar razões robustas. O primeiro tipo é a não normativa: Para a tentativa de A em dar robustamente para B uma razão para Φ com sucesso, B deve reconhecer anteriormente as intenções especificadas de A, e além disso, B deve permitir que essas intenções desempenhem um papel apropriado no seu raciocínio prático.⁴⁶ Note que para essa condição, não é necessária para a tentativa ter sucesso resultando em equivaler a um dar razões

⁴⁵ A comunicação é importante aqui. Não é suficiente para dar razões robustas que eu tenha a intenção que você saiba das minhas intenções relevantes: é importante (como pode ser visto em [iii]) que o que desempenhe o papel no seu raciocínio prático seja que eu atualmente tenha comunicado para você a intenção de que você tome essa mesma intenção comunicada como uma razão. (Eu agradeço a Joseph Raz por enfatizar este ponto para mim.)

⁴⁶ Veja que o “comunicar” na condição (i) não é entendido como um termo de sucesso, requerendo absorção de outro tipo. Requer meramente a tentativa de comunicar a intenção relevante. (Eu agradeço Hanoch Sheinman por um ponto relacionado.)

robustas, mas, ao invés disso, em ter sucesso com o tipo de efeito intencionado no mundo.

Essa condição não é suficiente para uma tentativa em dar razões robustas ter sucesso, nem mesmo para ter sucesso equivalente a um dar razões robustas. Para ver isso, pense novamente sobre o filho do ditador, e agora assuma, primeiramente, que ele tem todas as intenções especificadas acima, de modo que ele genuinamente tenta dar uma razão robusta; e segundo, que a pessoa a quem ele endereça a razão trata-o como genuinamente autorizado ou, em alguma outra forma, capaz de dar razões robustas dessa maneira, então, que essa pessoa permite que as intenções de dar razões da criança desempenhem o papel apropriado em seu raciocínio prático. Ainda assim, é claro que a criança não conseguiu dar razões robustas para a ação. O que está faltando neste caso é a condição de sucesso normativa, à saber, a tentativa deve fazer com que seja o caso de que a razão para Φ realmente surja (da forma apropriada). Já sabemos que se este procedimento vai resultar em haver uma razão para Φ , irá depender de haver uma razão independente, que é acionada por este procedimento⁴⁷ – grosseiramente, uma razão (para B) para fazer com que A tenha a intenção de que B tenha uma razão para agir. No exemplo da criança do ditador, não há razão geral que faça com que a criança “comande”, e então suas intenções no dar razões não aciona tal razão geral, portanto, ele não tem sucesso em dar robustamente uma razão, mesmo que sua audiência acredite que ele consiga. Não obstante, quando eu peço a você que leia meu artigo, presumivelmente há uma razão geral (para fazer como eu peço, dentro dos limites, em um certo contexto, etc.), uma que eu presumivelmente tenho sucesso em acionar fazendo a requisição. Isto é, desta forma, então, que a perspectiva sugerida de dar razões robustas é uma instância particular (mas uma importante e única)

⁴⁷ Ao menos o Modelo Constitutivo se aplica. Em tais casos – como Dimitrios Kyritsis notou – a condição de sucesso normativo é trivialmente satisfeita.

de dar razões acionadas.⁴⁸ Agora, eu não tenho (aqui) mais nada a dizer sobre quando tais razões acionadas existem e quando não existem, que algumas vezes elas não existem, que a força da razão (quando ela existe) irá variar de caso para caso, e que uma resposta geral para a questão “quando potencialmente uma razão acionada existe, e quão forte ela é?” é a parte de uma teoria substantiva mais geral sobre quais razões nós temos, uma teoria que eu infelizmente não tenho embaixo da manga.⁴⁹ Então, terei que decidir por afirmar a condição normativa em si mesma: para ocorrer o dar razões robustas, deve haver, independentemente da tentativa de dar razões robustas, uma razão acionada pela tentativa.⁵⁰

Não tenho certeza sobre o que há mais a dizer sobre a qualificação de “forma apropriada” em (iii). É importante eliminar correntes causais desviantes (e talvez outras).⁵¹ Teria sido

⁴⁸ Espero discutir o caso da autoridade – plausivelmente uma instância particular de dar razões robustas – em outra ocasião. Deixe-me rapidamente notar aqui que o ponto no texto é no qual discussões de justificação da autoridade aparecem. Na estrutura da perspectiva sugerida de dar razões robustas, por exemplo, a concepção de serviço de autoridade de Raz (e.g. Raz, 2005) é melhor vista como uma perspectiva da qual razões antecedentes do condicional existem, e então, quem tenta dar razões robustas (na maneira especial da autoridade) pode ter sucesso.

⁴⁹ Na introdução eu insisti que o fenômeno em que estou interessado – dar razões robustas – não, diferente do fenômeno que Darwall está interessado, necessariamente inclui nada como relações de autoridade assumidas. Uma forma de descrever o ponto no texto, aqui, é no termo relacionado (mas, claramente, não necessariamente institucional) de ponto de vista. Talvez, por exemplo, a questão de que se há tal como uma razão acionada pode ser reescrita como a questão: se o falante tem um ponto de vista para robustamente dar certas razões (como, por exemplo, se ele tem um ponto de vista para fazer as requisições relevantes).

⁵⁰ Acima, eu rejeitei uma teoria do erro sobre dar razões robustas. O fato de que eu não forneço aqui uma teoria geral sobre quais razões nós temos, ou em uma resposta particular para a questão quando existem razões que são acionadas na forma descrita no texto, significa que há outro tipo de teoria do erro que eu ainda não argumentei contra. Pois talvez, embora não haja nada incoerente em dar razões robustas, ainda assim, não há razões que são acionáveis dessa forma. Sendo assim então, não há casos de dar razões robustas, mas isso é dessa forma por causa de considerações normativas substantivas, ao invés daquelas que tem a ver com a natureza de dar razões robustas. Agora, para repetir, acredito que nós algumas vezes conseguimos dar razões robustas uns aos outros, e então, rejeito essa versão de uma teoria do erro também. Mas eu não argumento por essa reivindicação neste artigo. Eu agradeço Nadeem Hussain e Joseph Raz por me pressionar neste ponto.

⁵¹ Para a discussão de alguns casos de correntes causais desviantes no contexto da discussão da perspectiva de Grice sobre significado, veja Schiffer (1972, p. 17 e ff.)

agradável ter uma perspectiva explícita sobre como exatamente fazer isso. Entretanto, terei de me contentar em notar que, geralmente, conhecemos uma cadeia causal desviante quando a vemos, e reivindicar companheiros de culpa – visto que quase todo mundo precisa de uma consideração sobre correntes causais desviantes. A qualificação em (iii), então, não faz (iii) (ou a perspectiva da qual é uma parte) vazia, nem levanta qualquer novo problema que seja peculiar a minha perspectiva de dar razões robustas. (Se, no entanto, você pensar que não há maneira para a qualificação de “forma apropriada” tornar-se respeitável, então você talvez tenha que reconsiderar a “saída” da irreduzibilidade da seção 4.3 acima).

Também devo observar algo: *não* é preciso que o papel desempenhado pela razão dada no raciocínio prático do receptor seja apropriado. Não é requerido que o papel seja, por assim dizer, definitivo.⁵² Em outras palavras, é perfeitamente consistente com dar razões robustas então, entendido que haja uma última, completa, e talvez uma história mais básica do porquê é que B aceita e deve aceitar as intenções relevantes de A ao dar razões. Talvez, por exemplo, B é um simples utilitarista, e vamos supor ainda que um utilitarismo simples é de fato a verdadeira história fundamental sobre todas as razões para ação. Se, então, B tomará a requisição de A como uma razão para Φ se e somente se, e porque, fazendo isso, irá maximizar a utilidade. Isso não significa que ele não toma, nestes casos, a requisição de A sendo uma (não definitiva) razão. A questão crucial é onde a história definitiva (ou talvez, apenas a mais básica) será uma que atravessa as intenções especiais do concessor de razões identificadas acima (e o reconhecimento do recebedor disso), como no caso do recebedor de requisições utilitarista, no qual, neste caso, talvez tenhamos um caso de dar razões robustas: ou, queiraque a história mais básica funcione aqui diretamente, não deixando lugar para as intenções

⁵²Agradeço a Cia Dorr por me pressionar sobre um ponto relacionado.

específicas que fazem o dar razões robustas (como no caso do exemplo do filho do ditador). Casos do último tipo não são, na perspectiva que sugiro aqui, casos de dar razões robustas. E isso me parece o resultado independentemente plausível aqui.

Note que as intenções mencionadas acima não incluem algo como a intenção que B efetivamente realize Φ . Isso é, porque A pode dar a B uma razão para fazer Φ sabendo bem que outras razões possam ser relevantes, incluindo razões possivelmente mais fortes para não fazer Φ .⁵³ De fato, parece a mim que A pode fazer uma requisição genuína para que B faça Φ , sabendo o tempo inteiro que, se certas considerações tiverem peso neste caso, B não irá (todas as coisas consideradas) fazer Φ . Não queremos restringir o dar razões robustas a *somente* os casos em que o concessor de razões tem a intenção de que as razões dadas sobreponham todas as outras. Por razões similares, A não precisa ter a intenção de que a razão dada seja a única razão para B fazer Φ .

Você talvez preocupe-se sobre circularidade: as intenções aparecendo na análise sugerida de dar razões robustas são elas mesmas selecionadas em termos de dar razões. Observe que o analysandum é dar razões *robustas*. O dar razões aparecendo no analysans não precisa ser robusto: pode ser, e talvez geralmente seja, dar razões acionadas.⁵⁴ Portanto, quando eu peço a você que leia o meu rascunho, eu tenho a intenção de lhe dar uma razão para ler o meu rascunho e para que essa razão desempenhe o papel apropriado no seu raciocínio prático. Contudo, não preciso ter a intenção de lhe dar *robustamente* uma razão para que leia o meu rascunho. Certamente, não preciso ter a intenção de lhe dar uma razão *sob essa descrição*. Eu talvez tenha a intenção de acionar uma razão para você ler o meu rascunho com minhas intenções relevantes complexas, e este dar razões talvez conte

⁵³ Veja Raz (1975, p. 83).

⁵⁴ Não penso que uma intenção para epistemicamente dar razões seria suficiente aqui. É por isso que advertências não se qualificam como casos de dar razões robustas. (Agradeço a Tom Hurka por este ponto.)

como um caso de dar razões robustas, mas não preciso ter a intenção *como tal*. A perspectiva sugerida, então, explica a natureza do dar razões robustas em termos de intenção para (não necessariamente de forma robusta) dar razões. No pior – quando as intenções por si mesmas são para dar razões robustas – há algo benignamente recursivo na perspectiva sugerida.⁵⁵

4.5 Avaliando a Sugestão: A Desiderata

Novamente, a perspectiva sugerida de dar razões robustas satisfaz a desiderata da Seção 3? Quão bem ela se encaixa com – e quão bem explica – o fenômeno de dar razões robustas aqui caracterizado?

Primeiramente, essa perspectiva parece atingir os casos centrais corretamente, e pelo que parece ser, pelas razões corretas. Ameaças, no entanto, não são casos de dar razões robustas (assim são os casos de dar razões *meramente* acionadas), porque, embora a intenção (i) esteja presente em ameaças, em (ii) e (iii) tipicamente não. Claramente, as intenções (ii) e (iii) *talvez* estejam presentes em algumas ameaças, talvez, aquele que está emitindo a ameaça também tente dar uma razão robustamente. Todavia, estas intenções adicionais não são de forma alguma *necessárias*, ou mesmo *típicas*, para ameaças. O merceiro que aumenta o preço do leite intencionando dar ao seu cliente uma razão para comprar menos leite também carece dessas intenções adicionais. Isso é, porque, ele tendo a intenção (i) não faz com que o seu dar razões seja robusto. E já vimos que se eu emito qualquer som como uma requisição, mas eu não tenho as intenções (ii) e (iii), então há uma considerável pressão para não interpretar minhas palavras como uma requisição genuína. Como para o filho do ditador, em ordem

⁵⁵ Se você não está convencido, compare a definição recursiva, mas não circular de valor instrumental: “Algo é de valor instrumental se, e somente se, ela é instrumental para algo que é de valor” (no que a última ocorrência de “valor” fica para “valor não necessariamente instrumental”).

de saber se ele tenta dar uma razão robusta, precisamos saber mais sobre sua psicologia. A história é consistente, no entanto, com ele possuindo as intenções (i) – (iii), e então, como declarado acima, é bem possível que ele tente fazer isso. Porém, ele não consegue por causa da condição de sucesso normativo realçada acima: enquanto o filho do ditador talvez tenha a intenção de que sua razão dada dependa de uma forma apropriada no seu reconhecimento da intenção dele para lhe dar uma razão, sua razão não depende disso ao todo (mesmo que você pense de outra forma).

Na Seção 3, eu enfatizei a relevância da pessoalidade e relacionamentos pessoais para dar razões robustas. Parece a mim, que a perspectiva sugerida pode explicar este fato. Primeiro, e antes de tudo obviamente, por causa das intenções complexas que alguém deve ter em ordem de dar robustamente razões, dar razões robustas é restringido a, preferencialmente, criaturas complicadas, talvez apenas criaturas que são pessoas ou semelhante a pessoas. Um cachorro, por exemplo, pode certamente pedir sua recompensa, e em alguns contextos será natural dizer que ele pediu. Entretanto, parece que seria uma grande extensão dizer que o cachorro faz um pedido genuíno, que ele robustamente lhe dá uma razão para dar a ele sua recompensa. E a perspectiva sugerida explica porque cachorros não podem e talvez não possam ter as intenções necessárias para dar razões robustas. No entanto, por causa de ter uma intenção, que p é inconsistente com a crença na impossibilidade de p ⁵⁶, se A intenciona que B reconheça sua intenção ao dar razões e responda a ela, A deve conceber B como pelo menos em princípio capaz de reconhecer e de responder a sua intenção. Então, a perspectiva sugerida corretamente prevê que há restrições sobre a concepção de A sobre B se A está tentando dar

⁵⁶ Isto é muito, eu tomo como, incontestável ao longo daqueles escritos sobre intenções. As questões controversas são uma intenção p que implica em uma crença que p , ou algo do tipo. Note que a assunção no texto é muito fraca a uma afirmativa para essa questão.

robustamente para B uma razão para ação.⁵⁷ Além do mais, mesmo que A esteja confuso e tente dar robustamente ao seu cachorro uma razão para ação, as condições de sucesso normativo garantem que ele não pode conseguir, porque não há razão geral que se aplique para cachorros e que seja acionada pelas intenções complexas envolvida na confusa tentativa.

Na seção 3, também notei que há um sentido no qual dar razões robustas – ao contrário de dar razões puramente epistêmicas e meramente acionadas – é aproximadamente mais apoiada sobre a relação pessoal entre o concesso de razões e o recebedor de razões. Isto também pode ser explicado pela perspectiva sugerida no dar razões robustas. Por isso, segue dessa perspectiva que em ordem de A tentar dar robustamente uma razão para B, A deve (de alguma forma, grosseiramente) ter a intenção de que B tome as intenções de A para dar uma razão como uma razão para ação. Todavia, não há nada óbvio sobre isso. No que B toma – ou mesmo, se B deve tomar – as intenções complexas de A como uma razão, é algo que é altamente dependente de contexto. Uma das características do contexto dependente é precisamente a natureza da relação entre A e B.

Outra característica de dar razões robustas, que eu enfatizei na seção 3, era a de que razões epistêmicas não parecem ser concedidas robustamente. Isso também é algo que a perspectiva sugerida explica satisfatoriamente. Para ver isso, suponha (por *reductio*) que A tenta dar para B uma razão para acreditar que p; na caracterização sugerida, para isto ser o caso, A deve intencional:

- (i) dar uma razão para B acreditar que p;

⁵⁷Darwall (2006, p. 22 e ff.) enfatiza a distinção entre dar a alguém um incentivo, endereçar uma demanda de uma forma que ao endereçado é esperado reconhecer e internalizar. Aqui, também, a perspectiva no texto difere da de Darwall nos seus detalhes, mas talvez seja similar a sua em espírito. Darwall (2006, p. 269) também insiste que endereçar razões de segunda pessoa sempre pressupõe que o endereçado pode livremente e racionalmente determinar a si mesmo pelas razões endereçadas. Aqui também há similaridade ao ponto no texto é evidente.

- (ii) B reconhecer esta intenção;
- (iii) A razão dada para B acreditar que *p* depende de uma forma apropriada sobre o reconhecimento de B da intenção comunicada de A em dar uma razão para B acreditar que *P*.

Eu não vejo nenhum problema interessante com (i) e (ii). Porém, (iii) é altamente problemático. A intenção do dar razões de A – e o reconhecimento de B disso – parece do tipo errado para uma razão epistêmica depender dela. Afinal, eles não têm nada a ver com a verdade de *p*, ou com a evidencia suportando a crença de que *o*. Prontamente, a caracterização sugerida, em ordem de *mesmo tentar* dar robustamente uma razão epistêmica, o concessor de razões deve estar confuso sobre a distinção entre razões epistêmicas e razões pragmáticas (por exemplo, para crer). E claro, mesmo que tal tentativa seja possível (por que o concessor de razões está confuso desta forma), ainda assim nenhuma tentativa pode ter sucesso, por conta da condição de sucesso normativo: simplesmente não há razões *epistêmicas* pré-existentes que possam ser acionadas dessa forma.⁵⁸

⁵⁸ Há agora uma visão na epistemologia da testemunha – algumas vezes chamada de Visão da Confiança – de acordo com entendimento da qual o status epistemológico do testemunho é pensado de uma forma intimamente semelhante ao dar razões robustamente epistêmicas. Veja, por exemplo, Moran (2005). Eu não posso discutir a visão longamente aqui (e acredito que farei isso em algum lugar). Deixe-me simplesmente fazer os dois pontos, então. Primeiro, parcialmente para as razões rascunhadas no texto aqui, para a extensão da Visão da Confiança para ser entendida como reivindicando que razões epistêmicas podem ser robustamente dadas, e por essa mesma razão é altamente implausível. E segundo, nós devemos distinguir entre a fenomenologia do dar razões (como pode ser visto nas leituras do artigo de Moran ao lado deste). Mas disso nada segue em relação à epistemologia do testemunho.

5. Conclusão

Eu tentativamente concluo, então, que quando peço a você que leia meu rascunho, eu, dessa forma, com sucesso, lhe dou uma razão para o ler, eu aciono uma razão que estava lá o tempo todo independentemente da minha requisição. Porém, isso não mostra que dar razões robustas – no caso de requisições, ou mais generalizadamente – é uma ilusão, ou que não há nada de mágico ou misterioso sobre isso. Dar razões robustas é uma instância particular, particularmente interessante, do que eu chamei de dar razões acionadas. Ele é único no complexo de intenções que envolve, e assim sendo, também algo de suas características fenomenológicas.

Deixe-me fingir que essa perspectiva é sem problemas. Para algo, e como eu já observei, é em algum respeito incompleta (faltando uma perspectiva de quando há e quando não há razões acionáveis do tipo relevante, e faltando uma perspectiva de correntes causais desviantes). Além disso, apesar de eu pensar que a perspectiva sugerida toque nos casos centrais pelo menos grosseiramente corretamente, talvez tenha alguns resultados moderadamente contra intuitivos. Por exemplo, crianças razoavelmente jovens – talvez jovens demais para ter um complexo de intenções que minha caracterização utiliza – parecem ser capazes de emitir requisições de sangue puro, assim como o resto de nós. Talvez tenhamos que classificar tais requisições como menos que requisições de sangue puro, afinal, (na outra direção, me parece que a “requisição” endereçada *para* crianças muito jovens não são constantemente requisições de puro sangue). Isso pode ser um resultado um pouco estranho, mas não, penso, muito estranho, dada as vantagens da perspectiva sugerida.

Deixe-me mencionar outra possível objeção. As requisições em que essa perspectiva, grosseiramente, se encaixa, talvez seja argumentado, são requisições grossas – requisições que se sustentam em algumas relações pessoais relativamente grossas,

que são de importância não trivial, e assim por diante. Muitas, talvez a maioria, das requisições não são em nada como essas. Se fôssemos dois completos estranhos, e você me pedisse para lhe dizer as horas, não consegui dessa forma emitir uma requisição de sangue puro? Com tal requisição trivial (por exemplo, entre estranhos) não seria demais incluir como requisições de intenções complexas o que minha perspectiva de dar razões robustas utiliza?

Posso pensar em duas formas de refletir sobre requisições triviais em nosso contexto. Uma seria subjugar-las como menores que requisições de sangue puro, ou talvez permitir que elas contem como requisições de sangue puro pelo preço de romper o nó entre requisições e dar razões robustas. Isso não é completamente implausível: talvez, quando a requisição é trivial o suficiente, a distinção entre dar razões robustas e dar razões meramente acionadas perde alguma da sua força (certamente perde muita de sua importância).

Uma resposta muito mais interessante, no entanto, seria insistir que mesmo em tais requisições triviais entre completos estranhos, há casos de dar razões robustas. Tal insistência envolveria insistir que nós temos as intenções complexas relevantes quando fazemos requisições triviais, e então, também, que nós assumimos que completos estranhos podem e devem tomar nossas intenções para dar a eles razões como o genuíno dar razões. Talvez essa forma de pensamento sobre as coisas nos coloque de volta ao espírito da insistência de Darwall sobre as condições de felicidade normativa do endereço de segunda pessoa, o tipo de endereço que todos experienciamos, tanto endereçantes como endereçados.⁵⁹ Talvez nossa tendência de fazer requisições

⁵⁹ Há agora uma visão na epistemologia da testemunha – algumas vezes chamada de Visão da Confiança – de acordo com entendimento da qual o status epistemológico do testemunho é pensado de uma forma intimamente semelhante ao dar razões robustamente epistêmicas. Veja, por exemplo, Moran (2005). Eu não posso discutir a visão longamente aqui (e acredito que farei isso em algum lugar). Deixe-me simplesmente fazer os dois pontos, então. Primeiro, parcialmente para as razões rascunhadas no texto aqui, para a extensão da Visão da Confiança é para ser entendida como reivindicando que razões epistêmicas podem ser robustamente dadas, e por essa mesma razão é

triviais, e em respondê-las, mostra que em um sentido, assumimos que quase todos nós podemos responder às razões, e de fato, às razões robustamente dadas. (Isso contaria como outra forma em que requisições são um melhor ponto de partida no nosso entendimento de dar razões robustas, comparado a comandos, por exemplo, como claramente muito mais é necessário para comandos do que para requisições serem possíveis.) E talvez isso mostre não que dar razões robustas seja possível mesmo entre pessoas que continuam sem relação significativa normativamente umas com as outras, mas antes, a possibilidade de dar razões robustas entre estranhos mostra que mesmo completo estranhos continuam, afinal, em alguma importante relação normativa uns com os outros, e de fato, que todos nós sabemos o suficiente.⁶⁰

Referências

CUPIT, G. How Requests (and Promises) Create Obligations. **The Philosophical Quarterly**, v. 44, p. 439–455, 1994.

DARWALL, S. **The Second-Person Standpoint: Morality, Respect, and Accountability**. Cambridge, MA: Harvard University Press, 2006.

DARWALL, S. Reply to Korsgaard, Wallace, and Watson. **Ethics**, v. 118, p. 52–69, 2007.

altamente implausível. E segundo, nós devemos distinguir entre a fenomenologia do dar razões (como pode ser visto das leituras do artigo de Moran ao lado deste). Mas disso nada segue em relação à epistemologia do testemunho.

⁶⁰ Pelas discussões e conversas úteis, agradeço a Hagit Benbaji, Mitch Berman, Anne Burkard, Terence Cuneo, John Deigh, Cian Dorri, Janice Dowell, Luís Duarte d'Almeida, Yuval Eylon, Alon Harel, Scott Hershovits, Tom Hurka, Arnon Keren, Serena Olsaretti, Joseph Raz, Hili Razinsky, Gideon Rosen, Mark Schroeder, Stefan Sciaraffa, Russ Shafer-Landau, Hanoch Sheinman, David Sobel, David Sosa, Nick Southwood, Sigrún Svavarsdóttir, Mark van Rooijen; audiências na Universidade Hebraica, Oxford, Rice; e dois peritos extremamente úteis para o *The Philosopher's Imprint*. Enquanto trabalhei neste artigo, fui pesquisador do Instituto de Estudos Avançados da Universidade Hebraica de Jerusalém e agradeço o apoio do Instituto. Além disso, esta pesquisa foi apoiada pela Fundação de Ciência Israelense (concessão nº 136/09).

ENOCH, D. **Not Just a Truthometer: Taking Oneself Seriously (But Not Too Seriously) in Cases of Peer Disagreement**. Mind, No Prelo.

ENOCH, D. Being Responsible, Taking Responsibility, and Penumbral Agency. In: HEUER, U. (Ed.). . **Luck, Value and Commitment: Themes from the Ethics of Bernard Williams**. [s.l.] Oxford University Press, 2011.

ESTLUND, D. M. **Democratic Authority: A Philosophical Framework**. Princeton e Oxford: Princeton University Press, 2008.

HART, H. L. A. Commands and Authoritative Legal Reasons. In: **his Essays on Bentham**. Oxford: Oxford University Press, 1982. p. 243–268.

LAVIN, D. Stephen Darwall, The Second Person Standpoint: Morality, Respect, and Accountability. **Notre Dame Philosophical Reviews**, [s.d.].

MORAN, R. Getting Told and Being Believed. **Philosophers' Imprint**, v. 5, n. 5, 2005.

POSTEMA, G. Law as Command: The Model of Command in Modern Jurisprudence. **Philosophical Issues**, v. 11, p. 470–501, [s.d.].

RAWLS, J. Kantian Constructivism in Moral Theory. **The Journal of Philosophy**, v. 77, p. 515–572, 1980.

RAZ, J. **Practical Reason and Norms**. London: Hutchinson & Co, 1975.

RAZ, J. **The Morality of Freedom**. Oxford: Oxford University Press, 1986.

RAZ, J. The Problem of Authority: Revisiting the Service Conception. **90 Minnesota Law Review**, p. 1003–1044, 2005.

ROBERTSON, S. Creating Reasons. **Manuscrito**, [s.d.].

SCHIFFER, S. **Meaning**. Oxford and New York: Oxford University Press, 1972.

SCHIFFER, S. **Remnants of Meaning**. Cambridge, MA: The MIT Press, 1987.

SCHROEDER, M. The Scope of Instrumental Reason. **Philosophical Perspectives (Ethics)**, v. 18, p. 337–364, 2004.

SCHROEDER, M. Cudworth and Normative Explanations. **Journal of Ethics and Social Philosophy**, v. 1, n. 3, 2005.

SCIARAFFA, S. On Content-Independent Reasons: It's Not in the Name. **Law and Philosophy**, p. 233–260, 2009.

SCROEDER, M. Having Reasons. **Philosophical Studies**, v. 139, n. 1, p. 57–71, abr. 2008.

WALLACE, R. J. Reasons, Relations, and Commands: Reflections on Darwall. **Ethics**, v. 118, p. 24–36, 2007.

WATSON, G. Promises, Reasons, and Normative Powers. In: WALL, S.; SOBEL, D. (Eds.). . **Reasons for Action**. Cambridge: Cambridge University Press, 2009. p. 155–178.

Capítulo 4

O ponto de vista da segunda-pessoa, parte II: obrigação moral e o problema da responsabilidade

*Evandro Barbosa*¹

– “Se todos aspiram a Lei, como é que, durante todos esses anos, ninguém mais, senão eu, pediu para entrar?”, disse o homem. O guarda da porta, apercebendo-se de que o homem estava no fim, grita-lhe ao ouvido quase inerte: – “Aqui ninguém mais, senão tu, podia entrar, porque só para ti era feita esta porta. Agora vou-me embora e fecho-a.” (Kafka, O processo)

Apresentação

Na primeira parte deste trabalho, vimos que Darwall utiliza o chamado ponto de vista da segunda-pessoa para oferecer um tipo de justificação metaético para o problema moral. Partindo desse pressuposto teórico, o objetivo desta segunda parte do texto é distinguir entre os usos de obrigação moral como requerimento moral (*moral requirement*) e obrigação morais diretivas ou relacionais (*directed or relational obligations*), destacando a análise das chamadas obrigações bipolares (*bipolar obligations*) a partir do texto *Morality, Authority, & Law – “Essays in Second-Personal Ethics I”* (2013)². O texto está dividido em três partes. Primeiro, ofereceremos uma definição de moralidade enquanto

¹ Professor dos Programas de Graduação e Pós-Graduação da Universidade Federal de Pelotas – UFPel. Contato: evandrobarbosa2001@yahoo.com.br / webpage: <https://wp.ufpel.edu.br/ebarbosa/>

² Para fins de citação, esta obra será abreviada para *MAL*.

um sistema normativo. Em segundo lugar, analisaremos e distinguiremos obrigação moral como tal (*Obligation Period*) e obrigação bipolar (*Bipolar Obligation*), bem como o tipo de autoridade a elas associadas. Por fim, ofereceremos argumentos para defender que ambos os tipos de obrigação, como tal e bipolar, fornecem uma noção de responsabilidade baseada na noção de culpa moral.

I

A. Moralidade como um sistema normativo.

Corriqueiramente, tomamos diferentes normas daquilo que consideramos correto a partir de diferentes pontos de vistas. Por exemplo, dizemos *torturar crianças é errado!* porque afirmamos que isso é moralmente condenável; consideramos que *devemos cuidar de nossos pais na velhice* como um tipo de convenção social a ser seguida; da mesma forma, admitimos que legalmente nós *não podemos nos apropriar de bens públicos*. O que esses diferentes exemplos demonstram é que, ao explicar por que devemos agir de forma correta, de um ponto de vista prático amplo, reconhecemos que existe um conjunto de regras que consideramos válidas, sejam elas morais, sociais ou jurídicas. Em outras palavras, diríamos que as normas criam um universo de justificação que indicam o que podemos fazer, bem como o que podemos exigir uns dos outros em termos de dever. É nesses termos que coloco que existem contextos normativos, os quais pressupomos oferecer uma resposta correta para cada caso. Obviamente, com a moralidade não parece ser diferente.

Considere o seguinte exemplo: Pouco tempo atrás, Pedro prestou vestibular em determinada universidade para o curso de medicina, assinando uma autodeclaração afirmando que era negro e que gostaria de ingressar na universidade via sistema de cotas. Passados dois anos, a universidade abriu um processo interno para

investigar uma possível falha no sistema de autodeclaração após uma denúncia de que muitos desses alunos cotistas pretensamente negros, dentre eles Pedro, não atendiam aos requisitos necessários para serem considerados como tal. A acusação era de que Pedro e outros não poderiam ser considerados negros e, logo, não poderia estar ali pelo sistema de cotas. Após reuniões infundáveis e longas discussões, o reitor decidiu convocar uma reunião para que todos pudessem apresentar seus argumentos a favor ou contra a contemplação desses alunos pelo sistema de cotas, e tomar uma decisão sobre o caso. Pedro pediu a palavra e disse o seguinte: “O meu argumento é extremamente simples: eu tenho todo o direito de estar aqui, afinal eu passei no vestibular de forma justa e, se o único critério exigido era a autodeclaração, qual o problema em afirmar que sou negro, posto que, muito embora meus pais tenham ascendência escandinava, eu me identifico com a cultura afro-brasileira? O que eu fiz de errado? As regras foram postas e eu segui elas à risca.” Após certo silêncio, muitos na sala passaram a considerar que o argumento de Pedro parecia ser plausível e, talvez, devesse ser levado em consideração. Embora o exemplo levante questões relevantes sobre justiça social, eu gostaria de tratar a questão a partir da visão sobre diferentes sistemas ou contextos normativos como se fossem *circuitos fechados*, que admitem ou não a possibilidade de mudança a partir de considerações externas. Em comum, esses sistemas (moral, jurídico ou social) oferecem uma resposta correta para situações distintas e necessitam de uma justificação correspondente.

B. A distintividade da moralidade enquanto sistema normativo

O problema acima exposto ecoa na teoria moral de Darwall, uma vez que os conceitos morais centrais – incluindo aqueles de obrigação moral, certo, errado e demais direitos morais que possuem uma estrutura irredutivelmente em segunda-pessoa – se

referem às reivindicações e demandas que passa pela definição de moralidade como um sistema normativo:

Moralidade, como os filósofos usam atualmente o termo, refere-se a algo diferente do que quando alguém fala da moralidade da antiga Atenas ou das tribos Pashtuns do Afeganistão contemporâneo. Tampouco é o mesmo que a moralidade de qualquer indivíduo. O que nós temos em mente nesses últimos usos são costumes, hábitos sociais, normas e valores aos quais algum indivíduo se submete. Moralidades são identificadas social ou psicologicamente por meio da sua aceitação por algum indivíduo ou grupo. São, portanto, itens que podem ser estudados pelas ciências sociais empíricas, entendidas em termos amplos, talvez, para incluir princípios empáticos ou hermenêuticos. Já a moralidade como nós, os filósofos, a entendemos é um **conceito essencialmente normativo** e não empírico. [...] Além disso, *moralidade*, nesse sentido, admite apenas um uso singular. Existem muitas morais reais e possíveis (encarnadas psicologicamente e socialmente), mas há apenas uma moralidade [...] Um antropólogo, historiador ou sociólogo histórico pode nos informar sobre alguns moralidade da sociedade ou do grupo, mas a perícia desse tipo não constituiria competência sobre a questão normativa do que a moralidade exige ou mesmo no que a moralidade requer de pessoas que pertencem a esta sociedade ou grupo.³

Temos, então, que a moralidade possui essa visão multifacetada (sociológica, psicológica, normativa, etc.), o que dificulta sua compreensão. Entretanto, a moral possui a condição distinta da normatividade, por isso “se a moralidade é inerentemente normativa, ela não pode ser constituída por qualquer padrão normativo.”⁴ Isso constitui um traço distintivo à

³ MAL, 03, grifo meu. Bernard Gert explica a polissemia do termo moralidade, distinguindo-o por seus campos. Segundo ele, podemos utilizar o termo: “(A) *Descritivamente*, para se referir a certos códigos de conduta levados em consideração por uma sociedade, grupo (como uma religião) ou mesmo aceitos por um indivíduo para seu próprio comportamento; (B) *Normativamente*, para se referir a um código de conduta que, dadas as condições especificadas, seria levando em consideração por todas as pessoas racionais.” (Gert, 2016, Introdução). Confira também sua obra *The Definition of Morality* (2005), especialmente “Part I – Conceptual Foundations”.

⁴ *Idem*, 04.

moralidade ao mesmo tempo em que oferece uma definição à mesma. Nesse ponto, Darwall segue a caracterização de Gertrude E. M. Anscombe (1998) e de Henry Sidgwick (1964) acerca da “filosofia moral moderna”⁵, afirmando que o que é distintivo da moralidade na modernidade é o modo como a centralidade conceitual das noções irreduzíveis de segunda-pessoa de obrigação e prestação de contas/responsabilização são tratados.

Vamos considerar rapidamente a distinção entre ética antiga e moralidade moderna. Em *Modern Moral Philosophy* (1958), Elizabeth Anscombe argumenta que os conceitos modernos de *obrigação* e *dever moral*, bem como aquilo que consideramos em termos de moralmente certos ou errados, não vão ser encontrados

⁵ Confira *Modern Moral Philosophy* (1998) de G. E. M. Anscombe e *Outlines of the History of Ethics for English Readers* (1964) de H. Sidgwick. Charles Larmore, em *The Morals of Modernity* (1996) faz uma distinção entre o que chama de visões imperativas e atrativas da ética. A visão atrativa parte da pressuposição de que é ético aquilo que consideramos como verdadeiramente *bom* para o próprio agente. Por isso, a ação, antes de ser obrigatória, é *boa* para a florescimento do agente; na mesma medida em que, para determinar o justo (conteúdo do dever), o agente toma como ponto de partida aquilo que considera *bom* do ponto de vista moral. Por sua vez, a visão imperativa parte do ponto de vista de impessoalidade e imparcialidade, cuja origem remonta à visão deontológica moderna de uma teoria do dever. Segundo o autor, a origem da deontologia se deve “desgaste” da visão teleológica clássica e tem dois pontos centrais. Primeiro, a transformação da visão ordinária de bem pessoal, ou seja, não estamos tratando tão somente de uma visão sobre o bem, e temos uma gama de opções morais à mão, tais como o hedonismo (Epicuro), egoísmo ético (Hobbes), o perfeccionismo (Platão), etc. Some-se a isso a tentativa de superar a visão *aretéica* de mundo por uma visão do dever, dado a consideração sobre a pluralidade de “bens pessoais” presentes no mundo. Nesses termos, a clássica pergunta “por que ser moral?” é recolocada pela pergunta “o que é correto?”. Um dos problemas derivados dessa visão é que a filosofia moral moderna promoveu ampla discussão sobre as normas (*éthos*), mas relegou a um segundo plano a discussão sobre o caráter do agente moral (*êthos*). De forma semelhante, Frankena estabelece uma distinção interessante entre aquilo que chama de moralidade de imperativos *versus* moralidade de qualidades morais com base na distinção entre: “1. *Princípios* ou tipos de ação corretos ou obrigatórios, *e. g.*, os princípios da justiça e da beneficência; 2. *Valores morais* ou coisas que são moralmente boas, *e. g.*, as virtudes e certos motivos; 3. *Valores* ou coisas que são não-moralmente boas, quer *espirituais*, como o conhecimento ou comunhão com Deus, quer *materiais*, como automóveis e casas, quer *intrínsecas* ou *extrínsecamente* boas, etc.” (1975, 73-74) As visões imperativas endossam um tipo de ‘moralidade’ centrada nos imperativos (dever) que nos obrigam a realizar (ou não) determinada ação em termos de certo/errado. Por sua vez, as visões atrativas de moralidade sugerem centrar a discussão no agente e seu caráter. Com isso, temos que, do ponto de vista imperativo, moralidade requer a normativa do correto/errado para distinguir entre os diferentes padrões morais que estabelecemos; por sua vez, do ponto de vista atrativo, a moralidade parece ser uma condição do caráter do agente.

entre os antigos. Entre suas críticas, Anscombe argumenta que muitos teóricos morais, inclinados a concordar com a visão moderna, discutem o problema da moralidade de forma negligente. Diz ela:

Se alguém pretende expor as teses de Aristóteles e fala à maneira moderna sobre isto e aquilo ser *moral*, deve se tratar de alguém muito obtuso se não se sentir constantemente como alguém cuja mandíbula, de alguma forma, está fora do alinhamento: os dentes não se juntam em um mordida adequada.⁶

Isso se deve ao fato de que, segunda ela, o conceito moderno de moralidade está baseado na ideia de norma ou lei como uma noção jurídica de obrigação, culpabilidade ou culpa (*guilt*). Ao que parece, então, a visão moderna não faz uma distinção adequada entre moral e legal. Neste sentido, os autores modernos não estariam preocupados em identificar um tipo peculiar de norma, cujas mudanças conceituais do termo se orientariam pelas mudanças de costumes ou de sociedade, o que pareceria endossar uma tese relativista à moral. Em vez disso, nós deveríamos nos imaginar como seres humanos que estão sujeitos à moralidade em um sentido amplo, sentindo-nos obrigados a uma lei moral não por habitar qualquer jurisdição particular ou por fazer parte de qualquer contexto, mas apenas porque somos agentes morais na companhia de outros. Para Anscombe, esta visão racionalista e universalista reduziria a moralidade a condição quase-jurídica, o que seria problemático no seu entender.

Da mesma forma, Sidgwick desenha um contraste semelhante em *The Methods of Ethics* (1874) ao afirmar que, na filosofia grega em geral, existia uma faculdade reguladora reconhecida sob o nome de Razão. Por outro lado, na modernidade encontramos um tipo de dualismo entre razão universal e razão egoísta ou amor-próprio que evidenciam um tipo de *antagonismo de interesses*. De acordo com Sidgwick, existem dois ditames racionais fundamentalmente

⁶ Anscombe, 1998, 27.

independentes. O primeiro é o ditame da prudência racional, o qual indica que devemos visar nosso próprio interesse ou bem. Por outro lado, possuímos também um tipo de benevolência racional, posto que, na condição de ser racional, “sou obrigado a visar o bem em geral; *logo* sou obrigado a visar o bem de qualquer outro indivíduo tanto quanto o [meu].”⁷ Segundo Sidgwick, apenas este último – também conhecido como o “Axioma da Benevolência Universal” (*Axiom of Universal Benevolence*) – é uma obrigação moral distintiva, ou seja, nos coloca na condição de estarmos moralmente conectados a outros agentes.

Entretanto, o que é que Sidgwick considera ser distintivo sobre a moralidade e seu dever, tornando-a distinta da força normativa da prudência? Segundo ele, a prudência é desejável e racional do ponto de vista do próprio agente, ou seja, diz respeito a um desejo de realização egoísta ou de amor-próprio. Por sua vez, a moralidade diz respeito ao que é desejável ou racional a partir de um contexto mais inclusivo ou, na memorável expressão de Sidgwick, “o ponto de vista do universo” (*the point of view of the universe*).⁸ Os ditames da consciência ou moralidade seriam essa voz da razão de um ponto de vista universal. Faz parte do projeto de Sidgwick superar o método do hedonismo egoísta, cujo princípio afirma que devemos maximizar nossa própria felicidade, via hedonismo universalista, segundo o qual devemos maximizar a felicidade geral. Some-se a isso sua tese intuicionista, que o autor deseja combinar com sua defesa utilitarista, de que existem três princípios morais básicos que passam no teste da autoevidência. Primeiro, o princípio da justiça enquanto um princípio de universalização, através do qual ele endossa a tese utilitarista. Segundo, o princípio do amor-próprio racional, o qual exprime a visão prudencial de nossas ações e está atrelado ao método do hedonismo egoísta. Terceiro, o princípio da benevolência racional,

⁷ Sidgwick, 1964, 381-2.

⁸ Cf. Sidgwick, 1967, 382.

que exprime a ideia de que devemos considerar o bem dos indivíduos em geral para a realização de nossas ações.⁹

Temos, assim, uma visão geral sobre a dimensão da moralidade. Tanto Anscombe quanto Sigdwick concordam que a consciência no pensamento moderno ético é essencialmente concebida com noções *quase-jurídicas* de obrigação, dever e direito. Entretanto, enquanto Anscombe estabelece uma crítica a essa visão, Sigdwick pode ser tratado com um reformista da visão moderna universalista e imperativa. A questão que resta é saber como Darwall se posiciona diante do quadro apresentado. O autor contemporâneo concorda que as ideias de obrigação moral, certo e errado, não podem ser capturadas em termos daquilo que é desejável ou racionalmente aconselhável de um ponto moral impessoal ou imparcial. Nesse caso, se essa é uma possibilidade conceitual à moralidade, então os conceitos de obrigação, certo e errado, poderiam ser tomados como conceitos distintivamente diferentes daqueles sobre o que é desejável ou aconselhável, ou indesejável ou desaconselhável do ponto de vista moral.

Darwall irá concordar com Anscombe ao afirmar que o conceito de obrigação moral está conceitualmente vinculado ao de responsabilidade moral ou prestação de contas, considerando-se duas razões principais. Primeiro, é uma verdade conceitual que o que nós somos moralmente obrigados fazer, ou o que seria errado não fazer, é o que nos mantém responsáveis por fazer ou culpados por não fazer se omitimos a ação sem uma desculpa válida.¹⁰ Segundo, embora seja certamente verdade que os filósofos, no período moderno, começaram a levar a sério a possibilidade não-eudemonista de que considerações impessoais ou imparcial possam fornecer razões normativas independentemente de

⁹ Confira a tradução portuguesa do obra *Os métodos da ética* (2013), especialmente a excelente introdução à obra feita por Pedro Galvão (ULisboa).

¹⁰ Confira a distinção feita por Grotius entre fazer demandas legítimas e oferecer *conselhos racionais*, considerando-se que as obrigações morais estão diretamente atreladas à ideia de prestar contas (*accountability*) moderna.

benefícios ao próprio agente (Francis Hutcheson é um exemplo disso), é importante ressaltar que esse pensamento é diferente do tipo de autoridade racional suprema que os filósofos morais modernos, como Butler e Kant, vieram reivindicar à moralidade.

Ao mesmo tempo, o autor também concorda com o contraste de Sidgwick de que, enquanto a filosofia grega antiga e o pensamento pré-moderno derivaram da visão eudemonista, levando o bom para o agente como a única fonte das razões normativas para a ação, muitos pensadores do período moderno passaram a acreditar em uma ordem normativa jurídica, distintamente obrigatória e moral que se propõe fornecer uma fonte de motivos distintos e potencialmente conflitantes com a prudência ou autointeresse racional, ainda que estes sejam amplamente concebidos. Sua proposta decorre da pressuposição de que obrigação moral está conceitualmente conectada à ideia de responsabilização ou prestar contas (*accountability*) nestes termos quase-jurídicos, o que torna os agentes, em seu conjunto de atitudes, mutuamente responsáveis. É nesse sentido que a distintividade da moralidade “[...] consiste em normas que qualquer agente responsável pode legitimamente manter (por ele mesmo e outros), enquanto um agente mutuamente responsável.”¹¹

Como complemento à sua teoria, ele irá afirmar, seguindo Strawson, que estar *amarrado* à responsabilidade significa dizer que a obrigação moral é diferente de estar sujeito a qualquer norma ou padrão de excelência ou virtude, assim como não pode ser reduzido a uma visão prudencial da razão. Somos moralmente responsáveis por fazer parte de uma comunidade moral (*moral community*)¹², uma vez que são as obrigações que nos tornam responsáveis enquanto

¹¹ MAL, 07.

¹² Darwall cita a comunidade moral como o espaço onde os membros em condições iguais interagem, é o universo normativo em que as relações interpessoais ocorrem. Nela, “[...] a obrigação moral está relacionada à moral responsabilidade [...] temos que ver isso como envolvendo demandas que estão em vigor do ponto de vista moral, isto é, da perspectiva (em primeira pessoa do plural) da comunidade moral.” (2006, 09)

agentes morais diante dos outros. Em outras palavras, a comunidade moral é o elo que estabelecemos e exige que sejamos capazes de nos responsabilizarmos mutuamente. Em suma:

[...] tenho argumentado que nosso conceito (moderno) de moralidade é jurídico e inclui uma conexão conceitual entre obrigação e responsabilidade moral [...]. O que torna a moralidade distinta entre as noções normativas é a sua rede de ideias jurídicas, incluindo formas de responsabilização que assumem, na ideia de moralidade, uma normatividade fundamental *de jure*, uma normatividade que está faltando em qualquer sistema de direito *de fato* [...] ¹³

Se assim for, a distintividade da moralidade decorre de seu caráter normativo e do reconhecimento do ponto de vista da segunda-pessoa como a base teórica para defendermos que somos obrigados e responsáveis moralmente uns com os outros. Neste caso, obrigações e razões morais são sempre em segunda-pessoa, posto que elas são endereçadas a um agente, o qual se encontra moralmente sob uma condição de responsabilização.

O que torna o conceito de obrigação moral em segunda-pessoa, tal como argumento, é a sua conexão conceitual com a responsabilidade ou *justificação* moral, o que implica uma posição para endereçar (e ser endereçado) reivindicações e demandas legítimas. ¹⁴

Além disso, a estratégia de Darwall parte da pressuposição de que, para existir responsabilização em segunda-pessoa, não é necessário apenas que exista uma regra moral ou mesmo razões neutras em relação ao agente, ¹⁵ mas também condição *relacional*. É isso que tornar os conceitos morais centrais – obrigação, dever,

¹³ MAL, 19. Para uma discussão sobre autoridade *de jure* e *de facto*, confira a primeira parte deste trabalho em Barbosa, 2018, I, Item A.

¹⁴ MAL, 21.

¹⁵ Sobre este ponto, confira a primeira parte deste trabalho em Barbosa, 2018, especialmente II, Item B.

certo e errado –conceitualmente ligados a uma noção de responsabilidade. Esta é uma pressuposição do ponto de vista da segunda-pessoa da sua teoria, segundo a qual o caráter eminentemente prático de nossas relações – uma fenomenologia da moral, nos termos de Strawson – determinam o tipo de obrigação decorrente. Ademais, obrigação decorreria não apenas das razões em segunda-pessoa, mas da censura de determinada ação por parte dos demais agentes.

O que é moralmente obrigatório não é apenas dizer que há boas razões morais para fazer, por mais importante que essas razões possam ser. Pelo contrário, é o que seria moralmente *errado* não fazer e o erro moral não é apenas qualquer tipo de falha moral. Um ato é moralmente errado se e somente se seria condenável (*blameworthy*) se feito sem desculpas.¹⁶

II

A. Normatividades Distintas: Obrigação Moral Period e Obrigação Bipolar

Entretanto, mesmo que reconheçamos o que significa estar obrigado, ainda precisamos definir o tipo de obrigação à qual nos submetemos. Nesse ponto, o texto de Darwall é esclarecedor:

Obrigação bipolar investiga o conceito de um chamado *dever relacional* que é devido a alguém. Ao contrário da ideia de obrigação moral como tal (*period*), as obrigações bipolares não são apenas de devedores (*obligors*), agentes que estão sujeitos à obrigação, mas também de obrigados (*obligees*), aqueles a quem a obrigação é devida. Elas implicam, portanto, uma reclamação correlativa de que o obrigante deve reclamar do devedor o comportamento que o devedor é obrigado a realizar em relação ao obrigante. Embora as obrigações bipolares tenham figurado de forma proeminente na teoria moral e política recente, pouca atenção tem sido dada à sua

¹⁶ MAL, 21.

metaética. Em que consiste uma obrigação bipolar? Eu forneço uma teoria do conceito de obrigação bipolar que está conectado ao modelo de obrigação moral *como tal* apresentada em *Second-Person Standpoint*. Ambos os conceitos são mostrados como pessoais, porém de modos distintos. Alguém está moralmente obrigado a fazer algo *como tal* apenas no caso em que seu ato é algo que qualquer um (incluindo ele mesmo) justificadamente requer dele na condição de *pessoa representativa*. E alguém é obrigado em relação a outro a fazer algo apenas no caso de o credor ter a autoridade discricionária individual para fazer exigências do devedor como obrigado e responsabilizar pessoalmente o devedor. Obrigações morais *como tal* implicam implicitamente *autoridade representativa*, enquanto as obrigações bipolares envolvem *autoridade individual* do obrigado. Ambas as formas de autoridade são em segunda-pessoa (*second-personal*); no entanto, uma vez que implicitamente elas envolvem endereçar, não importa se essa autoridade é exercida pela primeira, segunda ou terceira parte envolvida.¹⁷

O primeiro ponto a ser considerado é a distinção entre obrigação bipolar e obrigação *como tal* (*period*). Uma forma de observarmos isso é através da distinção que Frankena estabelece entre juízos de obrigação moral particulares e juízos de obrigação moral gerais.¹⁸ Podemos reconhecer como um juízo de obrigação moral geral que “devemos respeitar as promessas feitas”,

¹⁷ MAL, Introdução, xiii. Segundo Darwall: “Originalmente, o termo *obrigação* referia-se a algo essencialmente relacional ou bipolar, quer dizer, *obrigação* era usada para se referir ao resultado de um ato de *compelir*. Se pensava que este *compelir* incluía ações como conceder um favor, celebrar um acordo ou um contrato, fazer um juramento e dentre outros. De acordo com este uso, ao realizar um bom serviço, por exemplo, um benfeitor obriga seu beneficiário a ele como uma dívida de gratidão. Ao selar um acordo ou celebrar um contrato, as partes obrigam-se mutuamente a seguir conforme o acordado. Ao juramentar ou fazer um juramento a alguém, o juramentador obriga-se e faz-se responsável perante a pessoa a quem jura ou para quem jura pelo que jura. E assim por diante. Obrigações no sentido original da palavra eram sempre devidas por um obrigado (devedor) a um obrigante (credor).” (*Idem*, 25)

¹⁸ Frankena distingue, ainda, juízos de obrigação e juízos de valor moral, ambos *juízos normativos*, como ele mesmo coloca. Enquanto tomamos os juízos de obrigação moral como cumprimento ou não de um dever, os juízos de valor não se referem às ações, mas às pessoas, motivos, intenções, traços de caráter, etc. Por fim, ela ainda fala sobre juízos de valor não-moral, os quais são dirigidos a outras espécies de coisas. Para mais detalhes, confira Frankena, 1975, 23-25.

aplicando-o na sua forma mais ampla, ou seja, tratando dele como um tipo de obrigação como tal. É uma obrigação cumprir uma promessa e seria errado não fazê-lo e *ponto*, não estando especificado seu contexto de aplicação.

Por sua vez, as obrigações bipolares são um tipo de obrigação que envolve necessariamente dois polos:

- i. De um lado, um agente que é obrigado (o *devedor*);
- ii. De outro, um indivíduo ou grupo a quem ele é obrigado, ligado ou vinculado (o *credor*).

Considere que João fez a promessa a Manoel de que iria assistir ao jogo do time de seu amigo caso ele chegasse na final do campeonato estadual. Embora João não torça para o mesmo time que seu amigo, ainda assim ele se considera em um contexto relacional que o obriga a realizar essa ação, pois quebrar uma promessa implica prejudicar o promitente (Manoel) – lembre-se de que algo é condenável (*blameworthy*) se e somente feito sem desculpas. Afinal, no entender de Darwall, “um promitente tem uma posição ou autoridade normativa distinta em relação ao promissor que terceiros não têm”.¹⁹ Isso significa que existe um tipo de relação entre devedor e credor que implica um tipo de obrigação moral bipolar entre aquele que faz a promessa (promitente) e aquele a quem ela é feita (promissor). Neste caso, não estamos falando simplesmente do juízo geral de que devemos respeitar as promessas feitas: João seria responsável por manter a promessa com base em um juízo de obrigação moral particular do tipo “devo cumprir minha promessa feita a meu amigo Manoel”.

Isso oferece uma distinção entre obrigação bipolar e obrigação como tal:

¹⁹ MAL, 20.

Obrigações bipolares exigem a relação entre *devedor e credor moral*, ou seja, elas pressupõem a relação entre os agentes (devedor e credor) para que a promessa tenha “peso” de obrigação. Em outras palavras, “[...] para que uma obrigação bipolar exista alguma ação deve *prejudicar um promissor*; deve constituir um *erro* e não apenas um ponto equivocado. [Por outro lado] [...] a existência de um credor faz parte do conceito de uma obrigação bipolar, embora não seja para obrigação moral *como tal*.²⁰

Outra forma de observarmos essa distinção pode ser feita a partir do texto de Michael Thompson (2004), no qual Darwall se apoia para reforçar seu argumento. Nele, Thompson distingue entre o que chama de ordem deontológica e ordem *dikaiológicas*. As categorias deontológicas monádicas definiriam a *ordem deontológica*:

- (1) Fazer A é errado (inadmissível).
- (2) Não fazer A é moralmente obrigatório.
- (3) Fazer B é moralmente admissível (não errado).
- (4) Não fazer B não é moralmente obrigatório.
- [A ordem *dikaiológica* faz um complemento:]
- (5) X prejudicou Y fazendo A.
- (6) X tem um dever para Y não fazer A.
- (7) Y tem um direito contra X que X não faça A.²¹

Darwall transcreve essa distinção para os seus próprios termos ao afirmar que existem duas opções em termos de obrigação:

1. Obrigação como tal: X é moralmente obrigado a fazer A.
Fazer A – Condição do dever pelo dever (*ordem deontológica*).
2. Obrigação bipolar: X é moralmente obrigado a fazer A para Y.
Para Y – Condição relacional entre devedor e credor (*ordem dikaiológica*)

Considerando-se a obrigação como tal, diríamos que João (X) é obrigado a fazer A (ir ao jogo com seu amigo) porque existe um

²⁰ *Idem, ibidem.*

²¹ Thompson, 2004, 335-338 *apud* MAL, 28.

juízo moral geral que o torna responsável nos termos da ordem deontológica, especialmente (1) e (2). Por sua condição geral, o juízo “devemos respeitadas as promessas feitas” poderia se aplicar não apenas a João e Manoel, mas a outros indivíduos em situações semelhantes. Por sua vez, a obrigação bipolar insere a condição *dikaiológica* (5), (6) e (7) para a realização da ação. Quer dizer, João está moralmente obrigado a fazer A (ir ao jogo com seu amigo), especialmente por sua condição relacional à Manoel (Y). Nestas condições, a obrigação é bipolar exatamente porque existe um promissor e um promitente, enquanto a obrigação como tal desconsidera essa condição relacional mais imediata, o que interferiria diretamente no tipo de obrigação decorrente.

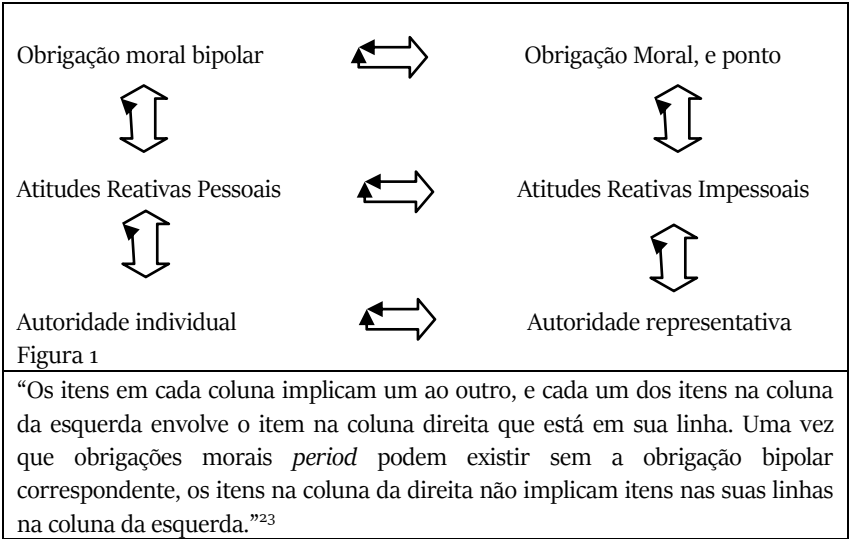
B. Autoridade individual e autoridade representativa.

Nesse universo moral normativo, Darwall entende que as obrigações bipolares pressupõem um tipo de autoridade individual do obrigante, ao passo que as obrigações morais como tal pressupõe uma autoridade representativa de qualquer pessoa como membro da comunidade moral. Por exemplo, a possibilidade de responsabilizar diretamente alguém por meio de uma atitude reativa, tal como o ressentimento, pressupõe um tipo de *autoridade individual*. Logo, o indivíduo está a ela vinculado em um “nexo prático” que permite fazer uma reivindicação com razões em segunda-pessoa a partir do reconhecimento de sua autoridade legítima. Por sua vez, a autoridade representativa estaria associada à obrigação moral como tal e envolveria um tipo de atitude reativa impessoal como a censura. Considerando o exemplo acima, Manoel ficará ressentido se João quebrar sua promessa de ir ao jogo com ele, e nós iremos censurá-lo – à *distância* – por tal atitude. Darwall entende que, mesmo que o conteúdo moral seja idêntico para obrigação moral como tal e obrigação moral bipolar, isso não implica que a autoridade *individual*, oriunda da condição bipolar,

permite responsabilizar do ponto de vista pessoal outros por violações sobre a obrigação moral como tal.

Do ponto de vista normativo, precisamos também considerar se as obrigações morais como tal são *exigências* morais genuínas e perguntar quem possui autoridade para fazer tais exigências. Segundo Darwall, “obrigações morais como tal implicam analiticamente uma autoridade em segunda-pessoa não-discrecionária que todos *compartilhamos* como pessoas representativas ou membros da comunidade moral para nos responsabilizar uns aos outros e exigir o cumprimento desse tipo de obrigação.”²² De todo modo, os aspectos normativos distintivos do conceito de obrigação bipolar e obrigação moral como tal indicam que tais obrigações envolvem autoridades distintas. Respectivamente, autoridade individual e autoridade representativa.

Darwall apresenta um quadro explicativo destes termos:



²² MAL, 32. Seguindo a pressuposição quase-jurídica da moral, autoridade não-discrecionária indica que tal autoridade não é passível interpretação ou ponderação de um ponto de vista particular, *i. e.*, seguimos tal e qual ela se apresenta. Por sua vez, a possibilidade discrecionária da autoridade indica que a mesma permita uma interpretação extensiva.

²³ *Idem*, 35.

Se assim for, podemos dizer que um credor ou titular do direito não precisa realmente reivindicar seu direito ou exigir a sua conformidade para o direito de existir, ou que o direito consiste inteiramente em qualquer exigência de que eletenha a autoridade individual para fazer.¹⁹ Ademais, nenhuma obrigação moral como tal, obrigação moral bipolar ou direito de reivindicação moral podem existir a menos que existam exigências *não-discricionárias* que não dependam de serem feitas por qualquer pessoa com a autoridade individual para fazê-las ou não.

Vale destacar também que autoridade individual vai *ao* (e não *de*) encontro à autoridade representativa na teoria de Darwall. Pelo contrário, pelo quadro acima exposto, autoridade representativa é condição *sine qua non* para a autoridade bipolar.

Se isto for correto, segue-se que a autoridade individual que está envolvida na bipolaridade obrigações não podem existir sem a autoridade representativa que está envolvida nas obrigações morais como tal. Uma vez que tanto a autoridade individual quanto a representativa são noções pessoais, segue-se que a normatividade bipolar não pode ser explicada sem as ideias gerais de autoridade e razões de segunda-pessoa.²⁴

Isso ocorre porque, para falarmos em obrigação no sentido moral, devemos considerar o contexto específico da moralidade como sistema normativo. Antes de tudo, enquanto representantes desta comunidade moral, a qual está justificada a partir do ponto de vista da segunda-pessoa, estamos em uma condição de responsabilização, o que significa reconhecer, antes de tudo, a obrigação moral como tal da comunidade moral como legítima e como sendo minha, enquanto agente moral. Após essa consideração como membro desta comunidade moral, então eu considerar minha autoridade individual e estabelecer uma condição relacional bipolar. Dessa forma, a

²⁴ *Idem*, 24.

distintividade normativa da moralidade é apresentada. De um lado, a normatividade bipolar envolve um tipo de autoridade individual (discricionária), baseada em atitudes reativas pessoais na condição de um obrigante (credor) em relação ao obrigado (devedor). De outro, a normatividade *period* desta moralidade implica em um tipo de autoridade representativa (não-discricionária), cujas atitudes reativas são impessoais por parte dos membros desta comunidade moral.

C. Responsabilidade e culpa em segunda-pessoa (Second-Personal)

Como vimos, obrigação moral (certo e errado) está ligada conceitualmente à responsabilidade e à culpa moral conceitualmente, por isso afirmamos que responsabilidade ou justificação moral implica relação interpessoal (*Interpersonal* – Strawson) ou em segunda-pessoa (*Second-Personal* – Darwall). Aqui, Darwall segue a fenomenologia da moral pressuposta por Strawson²⁵, afirmando que nós nos colocamos como moralmente responsáveis por meio da atitude moral distinta da culpa,²⁶ através da qual implicitamente tratamos demandas supostamente legítimas:

O **moralmente obrigatório é o que a moral exige**. Isso significa [...] o **conceito de obrigação moral é normativo** em primeira instância para as atitudes distintivas através das quais julgamos uns aos outros e nós mesmos moralmente responsáveis.²⁷ **[grifos meus]**

²⁵ Darwall extrai esta tese de Strawson, embora reconheça que ele não a tenha definido muito claramente: “Strawson não deu uma definição formal das atitudes reativas, mas suas características centrais são claras do papel que desempenham em seu argumento sobre a responsabilidade moral e a liberdade da vontade. A ideia central de Strawson é que as atitudes reativas envolvem uma maneira característica de considerar os indivíduos que são seus objetos que comprometem o detentor da atitude a certas suposições sobre o objeto isolado e suas capacidades de regular sua vontade.” (MAL, 21)

²⁶ Isso seria o equivalente à atitude reativa do ressentimento de Strawson. Cf. *Freedom and Resentment* (1968) de Strawson.

²⁷ *Idem*, 33.

Darwall define que as atitudes objetivas tais como desdém, aversão e aborrecimento estão em oposição às chamadas atitudes reativas, estas reconhecidas como atitudes em segunda-pessoa, posto que elas “exigem envolvimento ou participação com outras pessoas em relacionamentos humanos interpessoais”²⁸. Por isso, uma atitude reativa pessoal como o ressentimento é sentida da perspectiva de uma parte envolvida da relação, enquanto as atitudes reativas impessoais são sentidas a partir de um ponto de vista de terceiros, ou seja, daqueles não *diretamente* envolvidos. O ponto comum é que tanto atitudes reativas pessoais, quando as atitudes reativas impessoais são essencialmente interpessoais (Strawson) ou em segunda-pessoa (Darwall).

Assim, as atitudes reativas de *primeiro segmento*, como a culpa e as atitudes do *segundo segmento*, como o ressentimento e as atitudes de *terceiro segmento*, como a indignação ou a culpa moral são todas igualmente “interpessoais” no sentido de Strawson e, logo, em segunda-pessoa no meu sentido. “Segunda pessoa” não significa “segundo segmento”.²⁹

Em resumo, podemos afirmar que a normatividade da obrigação moral para a ação decorre, portanto, da sua normatividade para a culpa. Primeiro, a normatividade da obrigação moral como tal diz respeito às atitudes reativas impessoais, enquanto a normatividade das obrigações bipolares (e a correlativa reivindicação de direitos) diz respeito às atitudes reativas pessoais como o ressentimento, embora ambas sejam atitudes em segunda-pessoa.

Com base nisso, Jay Wallace oferece uma objeção à posição de Darwall afirmando que, na verdade, são os elementos em segunda-pessoa que resultam como aspectos da obrigação e normatividade bipolares. Se assim fosse, as obrigações bipolares

²⁸ Strawson, 1968, 79.

²⁹ *Idem*, 22.

seriam o *locus* das noções de segunda-pessoa e não o contrário. Como réplica, Darwall afirma que mesmo as atitudes reativas impessoais devem ser consideradas como essencialmente em segunda-pessoa – ou “interpessoais”, nos termos de Strawson. A culpa moral é um exemplo disso, pois ela não pode ser entendida como meramente conceitual: se a culpa “endereça uma exigência *como legítima*, então assume-se, necessariamente, uma autoridade para fazer tal exigência. E isso é o que a torna segunda-pessoa, segundo compreendo.”³⁰ Ao mesmo tempo, a culpa, enquanto fenômeno moral, indica que devemos atender a determinada exigência que não provém tão somente de uma autoridade individual, mas também da condição de pessoa representativa da comunidade moral. Por isso, “as atitudes reativas impessoais como a culpa moral pressupõem uma *autoridade representativa*, a qual é complementar à autoridade individual e está pressuposta às obrigações bipolares.”³¹

Se faço uma censura a João por quebrar sua promessa de ir ao jogo com Manoel, eu pressuponho (a) ter razões em segunda-pessoa para que ele cumpra a demanda que estou exigindo sobre cumprir sua promessa em relação a seu amigo. Entretanto, esta autoridade que pressuponho possuir não provém da minha condição individual, mas de (b) uma condição de autoridade representativa enquanto membro da comunidade moral. Se alguém possui tal autoridade (c), certamente é Manoel quem a possui, o qual está em uma condição privilegiada para demandar que é errado – feito sem desculpa do um ponto de vista moral – quebrar a promessa de ir ao jogo.³² Disso tiramos algumas conclusões:

³⁰ *Idem*, 36-37.

³¹ *Idem*, *ibidem*.

³² A distinção entre razão interessada (amor-próprio) a razão benevolente de Sigdwick parece fazer sentido neste ponto.

- (a) Minha censura sobre João decorre de uma autoridade representativa oriunda da comunidade moral.
- (b) As razões que ofereço representam um tipo de obrigação como tal.
- (c) Apenas Manoel pode fazer exigências ou demandas legítimas a partir de uma autoridade individual, tendo em vista que Manoel é o obrigante e João é o obrigado nesta relação bipolar.

Podemos verificar, assim, que a culpa, a qual expressa uma violação à obrigação moral *period*, depende da existência de uma autoridade para exigir que qualquer agente moral, Manoel, João ou qualquer outro, compartilhe tal condição como pessoa representativa ou membro da comunidade moral. Quando culpamos alguém estamos pressupondo, antes de tudo, que esta é uma exigência feita por todos da comunidade moral. Para além disso, Darwall reconhece ainda o fato de que o credor (neste caso, Manoel) possui uma “base privilegiada para reclamação” (*privileged basis for complaint*) em relação a João que outros não possuem.³³ Essa base privilegiada é a condição bipolar da relação.

O que podemos verificar é que existe uma conexão conceitual entre a obrigação bipolar e a obrigação moral *period*. Como vimos: “Se X tiver uma obrigação de Y para fazer A, e Y, portanto, um direito de reivindicação contra X para que X faça A, então a falha de X para fazer A não prejudicaria apenas Y, seria errado enquanto obrigação como tal em relação aos outros [membros].”³⁴

D. O que podemos dizer sobre Pedro e sua autodeclaração? Algumas conclusões.

Temos, agora, alguns elementos importantes para determinar se Pedro agiu ou não de forma aceitável. Um primeiro ponto a ser considerado é o alcance da moralidade. Chegamos,

³³ Cf. *MAL*, 38-39.

³⁴ *Idem*, 38-39.

assim, a algumas conclusões sobre esse ponto. Primeiro, a metaética da obrigação bipolar é paralela à obrigação moral *period*. Ambas são pessoais, mas de maneiras diferentes. Por um lado, a normatividade bipolar envolve uma autoridade distintiva individual que os credores têm para fazer exigências e responsabilizar os devedores. Por outro lado, a obrigação moral *period* implica uma autoridade representativa que alguém compartilha como pessoa representativa ou membro da comunidade moral. “Segue-se [então] que o conceito de obrigação bipolar não é mais fundamental do que os de autoridade e razão em segunda-pessoa. [...] A ideia da autoridade em segunda-pessoa é uma noção mais geral, e é necessário compreender a metaética das obrigações bipolares e do período de obrigações morais.”³⁵ A “erradeza” da ação, como no exemplo de Pedro, indica uma quebra na autoridade legítima que alguém possui para fazer qualquer tipo de demanda (obrigação). A questão é que, se tomarmos o contexto normativo onde a situação ocorre – no exemplo de Pedro, o contexto jurídico – como um sistema *fechado*, parece haver uma incoerência sobre a legitimidade da reivindicação vinda de outro contexto normativo. Seria como se estivéssemos transferindo um elemento normativo moral para esse contexto de forma arbitrária. Contra isso, a teoria moral de Darwall parece oferecer alguns argumentos promissores.

Primeiro, se compreendermos moralidade a partir do ponto de vista da segunda-pessoa, poderíamos tomá-la em um sentido *lato*, ou seja, a dimensão jurídica deveria estar pressuposta por esta condição normativa mais ampla da moral. Nestes termos, não trataríamos a questão como mera observação em termos jurídicos, pois existiria tanto sua condição de membro da comunidade moral, quanto a relação de dois polos entre Pedro e seus concorrentes) que é determinante para a definirmos a normatividade bipolar de sua obrigação.

³⁵ *Idem*, 39.

Segundo, podemos imaginar o sistema moral em um sentido estrito da comunidade moral, então Pedro estaria obrigado, nos termos de Darwall, não apenas no sentido impessoal de uma obrigação deontológica, mas também naquele sentido *dikaialógico*. Quer dizer, ele se encontrar duplamente obrigado: há uma obrigação moral como tal dada pela autoridade representativa dos agentes da comunidade moral, assim como há obrigação bipolar oriunda da autoridade individual daqueles que estão na posição de credor em relação a ele, o devedor. Se Pedro está sob uma obrigação moral em relação Y (seus concorrentes) de fazer A (ser honesto em sua autodeclaração), então Pedro estaria, pelo menos dentro de outras variáveis, sob um ponto de obrigação moral para fazer A. Ações que levam equivocadamente alguém a violar uma obrigação bipolar também estão erradas enquanto obrigação *como tal*.

Por fim, de modo semelhante ao que acontece com o exemplo entre João e Manoel, a censura sobre Pedro pelo fato dele não ser sincero em sua autodeclaração, pressupõe um tipo de razão em segunda-pessoa para que ele cumpra a minha demanda. Isso requer autoridade legítima do tipo representativo. Por sua vez, seus concorrentes se encontram em uma base privilegiada, ou seja, possuem autoridade individual para demandar que é errado – feito sem desculpas de um ponto de vista moral – não ser honesto em sua afirmação apenas para ganhar vantagem em determinada situação.³⁶ Admitir a culpa da atitude de Pedro envolve reconhecer a legitimidade da queixa dos concorrentes, pressupor que eles possuem uma autoridade especial (*individual*) para fazê-lo e, por fim, a aceitação interna da legitimidade da queixa, pois esta foi endereçada a partir da condição em segunda-pessoa. Logo, Pedro não possui uma base legítima para fazer sua demanda, sendo a

³⁶ A distinção entre razão interessada (amor-próprio) a razão benevolente de Sigdwick parece fazer sentido neste ponto.

mesma é suplantada pela exigência de um prestar contas que, necessariamente, pressupõe a normatividade do sistema moral.

Referências

ANSCOMBE, G. E. M. “Modern Moral Philosophy”. In: CRISP, R.; SLOTE, M. (Eds.). . **Virtue Ethics**. Oxford: Oxford University Press, 1998. p. 26–44.

ANSCOMBE, G. E. M. “Reasons, Relations, and Commands”. **Ethics**, v. 118, p. 24–36, 2007.

AUDI, R. “Reasons for action”(Parte II, 24). In: SKORUPSKI, J. (Ed.). . **The Routledge Companion to Ethics**. New York: Routledge, 2010.

BARBOSA, E. “Obrigações deontológicas no contratualismo moral de Thomas Scanlon.” In: BRESOLIN, K.; BARBOSA, E. (Eds.). Caxias do Sul, RS: Educs, 2017. p. 08–30.

DARWALL, S. **The Second-Person Standpoint: Morality, Respect, and Accountability**. Cambridge: Harvard Press, 2006.

DARWALL, S. **Morality, Authority, and Law (I)**. Oxford: Oxford Press, 2013a.

DARWALL, S. “Respeito e a perspectiva da segunda pessoa”. In: DARWALL, S.; GIBBARD, A.; RAILTON, P. (Eds.). Florianópolis: Editora da UFSC, 2013b.

GERT, B. **Morality: Its Nature and Justification**. Oxford: Oxford University Press, 2005.

GERT, B.; GERT, J. The Definition of Morality. In: ZALTA, E. N. (Ed.). . **The Stanford Encyclopedia of Philosophy**. Fall 2017 ed. [s.l: s.n.].

MILLER, A. **An Introduction to Contemporary Metaethics**. Malden, M.A: Polity Press, 2003.

NAGEL, T. **The View from Nowhere**. New York: Oxford University Press, 1986.

NAGEL, T. **Visão a partir de lugar nenhum**. São Paulo: Martins Fontes, 2004.

- NUNES, C. DE M. “Atitudes reativas e responsabilidade moral”. In: CONTE, J.; GELAIN, I. L. (Eds.). . **Ensaio sobre a filosofia de Strawson**. Florianópolis: Editora da UFSC, 2015. p. 207-219.
- PUFENDORF, S. **On the Law of Nature and Nations**. Tradução C.H. Oldfather and W. A. Oldfather. Oxford: Clarendon Press, 1934.
- SCANLON, T. M. **What We Owe to Each Other**. Cambridge, MA: Harvard University Press, 1998.
- SIDGWICK, H. **Outlines of the History of Ethics for English Readers**. Boston, MA: Beacon Press, 1964.
- STRAWSON, P. F. “Freedom and Resentment”,. In: **Studies in the Philosophy of Thought and Action**. London: Oxford University Press, 1968.
- STRAWSON, P. F. “Liberdade e ressentimento.” In: In: CONTE, J.; GELAIN, I. L. (Eds.). . **Ensaio sobre a filosofia de Strawson**. Florianópolis: Editora da UFSC, 2015. p. 245-269.
- THOMPSON, M. “What Is It to Wrong Someone? A Puzzle about Justice.” In: WALLACE, R. J. (Ed.). . **Reason and Value: Themes from the Moral Philosophy of Joseph Raz**. New York: Oxford University Press, 2004. p. 333-84.
- VALLIER, K. L. **Politics and Public Faith: Beyond Separation**. New York: Routledge, 2014.
- WALLACE; JAY, R. **Responsibility and the Moral Sentiments**. Cambridge, MA: Harvard University Press, 1994.

Capítulo 5

Pessoas virtuosas podem fazer coisas más? Aristóteles entre a idealização e o realismo

João Hobuss¹

Howard Curzer, em um artigo intitulado “How good people do bad things: Aristotle on the misdeeds of the virtuous”, apresenta uma visão original, embora sujeita a inevitáveis questionamentos, que busca apresentar a figura do virtuoso - em variadas tipologias -, como despidido da aura de perfectibilidade geralmente associada ao mesmo como geralmente interpretado pela, digamos, maioria dos comentadores e dos comentadores da ética aristotélica, mais especialmente da *Ética Nicomaqueia*.

Este texto procurará empreender uma primeira análise de alguns pontos e passagens levantadas por Curzer, buscando entender se a argumentação levada a cabo por ele se sustenta, mais especificamente no que concerne à figura do virtuoso em sentido forte, ou seja, aquele que possuiria a virtude de pleno direito por cumprir os requisitos próprios para tal, o que não impede uma apresentação inicial mais ampla, contudo um pouco mais breve, de sua posição geral. Uma análise mais extensa e vertical será realizada posteriormente, em outra oportunidade.

No sentido de mostrar as falhas apresentadas pela tese da idealização do papel do virtuoso, Curzer tenta demonstrar que

¹ Universidade Federal de Pelotas – CNPq

existem sete modos de agir mau no que tange aos virtuosos (2005, p. 255):

- (1) os virtuosos podem não ter consciência de determinados fatos cruciais quando age;
- (2) ele pode sucumbir à fortes pressões;
- (3) eventualmente podem agir independentemente do seu caráter;
- (4) o virtuoso pode possuir pequenas falhas em suas virtudes, o que deixaria umas poucas situações descobertas;
ou
- (5) levá-lo a alguns poucos atos extremos;
- (6) vícios podem regularmente mascarar virtudes;
- (7) dilemas morais podem requerer atos viciosos, de modo que agir em conformidade com a virtude em dada situação é agir incorretamente.

Passemos à descrição de cada um desses sete modos de agir mau do virtuoso, que, *a priori*, pelo fato de ter a reta razão e o reto desejo jamais poderia “escolher ações viciosas” (p. 234), na medida em que é o critério último da ação moral, ou como afirmou Aristóteles na *EN* III, o homem de bem, o virtuoso deve ser entendido como regra e medida (*kanôn kai metron*, cf. 1113a33), pois percebem, nas coisas, a verdade que elas carregam (a32). Esse tipo de interpretação, segundo Curzer, é perigosa, pois a idealização é sempre perigosa,

“pois descreve caracteres perfeitos, caracteres que são, no melhor dos casos, aproximados, mas nunca realizados por agentes humanos, concepções idealizadas podem nos dar um marco para atingir” (p. 234).

Tal visão também pode ser encontrada em Destrée (2011, p. 302 e n. 13), que sustenta que o prudente não existe realmente, sendo apenas um ideal a ser buscado. Logo, uma pessoa virtuosa “não age perfeitamente em toda as situações” (Curzer, 2005, p. 234), essa não seria a posição verdadeira de Aristóteles, que não

teria em mente em sua ética construir o virtuoso como se fosse uma pessoa ideal (p. 234), na medida em que ele pode admitir distintas formas de imperfeição:

“A pessoa que Aristóteles esboça através da *Ética Nicomaqueia* não é uma pessoa que sempre faz as coisas de modo correto, mas antes é a que possui um caráter que é bom o suficiente para contar como virtuoso” (idem).

Assim, é possível que pessoas virtuosas possam agir incorretamente, e ainda permanecerem virtuosos, pois Aristóteles não teria em mente uma pessoa que tenha de agir sempre virtuosamente, na medida em que é o caráter de alguém, se suficientemente bom, que faz dela alguém virtuoso (p. 234). Por exemplo, na passagem citada por Curzer (1121a1-7), há a afirmação que, no caso do liberal, “[ele] ficaria mais aborrecido se não gastasse o que deve do que aflito se tivesse gasto o que não deve”. Ora, a segunda parte da frase indicaria que o liberal poderia agir incorretamente se gastasse o que não deve, e isso ocorreria pelo fato de que ao liberal, em determinada circunstância, poderia faltar a consciência de alguns fatos cruciais, como aponta o primeiro modo do virtuoso agir mal (modo 1 do virtuoso agir mau, p. 237). No modo 2, temos a possibilidade de, embora virtuosos, agirmos viciosamente, pois uma pressão demasiada, e que supera a natureza humana, não poderia evitar a realização de uma ação má (cf. 1110a23-26, 1115b7-8). Torna-se inexequível para uma pessoa de bem não ceder à uma dada e extrema ação incorreta, mas ainda assim a ação deveria ser perdoada (p. 237), pois há coisas que não podem ser demandadas às pessoas pela sua própria condição, pela sua própria natureza (p. 238).

Um outro modo de agir mau do virtuoso é quando ele age independentemente do seu caráter (modo 3), o que pareceria ser o caso de 1132a2-4. Ali, Aristóteles menciona o fato de que não importa se um homem bom (*epieikês*) defraudou um homem vil (*phaulos*), ou vice-versa, ou se um homem bom (*epieikês*) cometeu

adultério (p. 239). Uma questão que Curzer não menciona é se ele pode tomar o *epieikês* na passagem como sinônimo de virtuoso, pois há comentadores que sustentam que apenas o *spoudaios* pode ser entendido como o virtuoso em sentido pleno (cf. Bodéüs, 2004, p. 460, n. 2).

O modo 4 indica que em certas situações que ocorrem raramente, uma pessoa virtuosa pode agir viciosamente, embora de modo atípico, o que não retira o fato dela permanecer virtuosa (p. 241). Curzer nos dá o seguinte exemplo: “Quando Carl apresenta sua esposa Dara para sua amante, não seria surpreendente se Dara exibisse rudez excessiva” (idem), embora em todas as outras situações ele mantém um padrão de ação virtuoso. Seria o caso também da passagem da *Política* 1287a31-32: “As paixões, pervertem as mentes dos governantes, até mesmo quando são o melhor dos homens”. Se um dado herdeiro assumisse de chofre o trono depois da morte repentina de vários aspirantes que o antecederiam, ele poderia vir a se corromper, mas a virtude, por ser uma disposição seria dificilmente perdível, embora concretamente pudesse ser. Assim, haveria uma janela de tempo que permitiria que a corrupção engendrada pelas paixões fosse detida, e o caráter virtuoso resgatado. Nesses exemplos eles agem mal, segundo Curzer, a despeito de suas virtudes (p. 242).

O modo 5 de agir mal para um virtuoso parte do fato que muitas vezes o virtuoso age mal em função de suas virtudes, e não a despeito delas. Três passagens são mencionadas: 1120b4-6, 1125b35-a3, 1136b19-21. Todas elas possuem como propriedade a indagação de se a virtude consiste em uma mediedade, “como uma ação excessiva ou deficiente podem ser *características* de uma pessoa virtuosa?” (p. 243). Um liberal pode ser considerado ainda um liberal se agir no limite da prodigalidade? Podemos lembrar, contra Curzer que, quando Aristóteles discorre sobre a liberalidade, ele afirma que existem coisas que podem ser bem ou mal utilizadas, em outras palavras, podem ser utilizadas de modo correto ou incorreto, como é o caso da riqueza, e quem utiliza algo

de modo correto, possuirá a virtude concernente (EN 1120a4-6). A questão não estaria, então, em estar mais perto do excesso ou da deficiência, mas antes responder do ponto de vista da qualidade da ação, responder o que demanda a situação, e esta pode demandar ir em direção mais do excesso ou da deficiência. O que deve ser preservado é a ‘essência’ da virtude, qual seja, agir em conformidade com a reta razão, tendo como fim o nobre (*tou kalou heneka*, EN 1115b 12-13. Cf. 1120a23-24: *Ai de kat’ aretên praxeis kalai kai tou kalou heneka*).

No que se refere ao modo 6, Curzer tem em mente pessoas que agem frequentemente mau são virtuosas (p. 246). Como isso poderia se dar? Acenando para a *aretê phusikê* (virtude [moral] natural) no livro VI da *Ética Nicomaqueia*. Ele afirma que uma pessoa pode possuir algumas, mas não todas, virtudes [morais] naturais: “Uma pessoa pode ser naturalmente virtuosa a respeito de alguma esfera da vida humana, enquanto permanece continente, incontinente ou até mesmo vicioso a respeito de alguma outra esfera” (p. 246-247). Desse modo, alguém pode ser um “liberal covarde”, ou, talvez, um “temperante injusto” e assim por diante, ou seja, “Aristóteles concede que pode haver um período de tempo no qual uma pessoa age frequentemente de modo não-liberal sem cessar de ser liberal” (cf. 1115a20-22), até mesmo porque “a erosão da virtude toma tempo” (p. 248). A argumentação de Curzer é bastante criativa, se bem que não explique como dar-se-ia a erosão do virtuoso, mas causa certo desconforto no sentido de “forçar” uma interpretação do texto aristotélico que parece não estar devidamente respaldada no próprio texto, como a afirmação que “em geral, a pessoa virtuosa age, regularmente, corretamente, mas ‘regularmente’ não significa ‘sempre’”. Se uma pessoa age corretamente 99,97 por cento do tempo, mas previsivelmente age incorretamente 0,03 por cento do tempo, então aquela pessoa devia ser considerada virtuosa” (p. 245). Sim, mas é isso que Aristóteles afirma peremptoriamente? Por que previsivelmente?

Seguindo a tipologia esboçada no início do texto, passamos ao modo 7 de como o virtuoso pode agir viciosamente. Aqui, Curzer entra no problema dos dilemas morais que aparecem no livro III da *EN*, sobretudo no que concerne às passagens acerca das ações mistas, ou ações que são ao mesmo tempo as corretas a fazer e más, “onde o ato correto é um ato vicioso. É moralmente requerido e moralmente repugnante” (p. 249), o que opera como um caso de *dirty hands*. Vejamos a passagem:

“Há discussão para saber se são voluntárias ou involuntárias as ações praticadas por medo de males maiores antes que com vistas a algo belo; por exemplo, se um tirano ordenasse a alguém fazer algo ignóbil retendo em seu poder pais e filhos que seriam salvos se o fizesse, mas morreriam se não o fizesse. Algo semelhante ocorre também a propósito do lançamento ao mar da carga de um navio durante uma tempestade: ninguém lança, sem mais, a carga ao mar voluntariamente, mas, para salvação de si mesmo e dos restantes, todas as pessoas sensatas agem assim. Portanto, tais ações são, de um lado, mistas; de outro, assemelham-se mais às voluntárias, pois são escolhidas no momento em que são praticadas, e o fim da ação se dá conforme à ocasião.” (1110a4-13, tradução de Marco Zingano)

O caso mais chocante, evidentemente, é a do tirano que obriga o virtuoso a cometer um ato ignóbil, vicioso, que acaba por estabelecer um paradoxo, pois não agir em conformidade com o que o tirano lhe pede é um ato mau do ponto de vista moral, pois colocaria em risco a vida dos seus. Logo são casos de dilemas morais, casos de *dirty hands*, onde o conflito entre o que fazer e o que não fazer aflora em toda a sua dimensão (uma outra passagem citada por Curzer sobre dilemas morais é 1164b31-11654).

É difícil perceber aqui uma distinção fundamental em relação em atos que ocorrem *out of character*, independentemente do caráter virtuoso do agente. Saliento isso por uma passagem bastante conhecida do *Organon*, onde há a afirmação de que o homem poderia agir de modo oposto ao seu caráter. Algumas

passagens poderiam dar guarida a este tipo de leitura, como é o caso dos *Tópicos* IV, 5, 126a 34-36 em que aparece a afirmação de que até mesmo Deus e o homem são capazes de fazer o mal, embora isto não constitua o seu caráter:

“(...) pois até mesmo Deus e o homem bom (*spoudaíōs*) são capazes de fazer coisas más, mas isto *não é o seu caráter*; pois é sempre a respeito de sua escolha (*proairesin*) que os homens maus são assim chamados.”

Tal relação poderia ser melhor explorada, mas não é o intuito desse texto, mas parece haver, sim, uma relação entre o modo 7 e o modo 3, se não totalmente, talvez de maneira complementar.

Existe no artigo de Curzer um pano de fundo metodológico, que faz sentido, mas não parece ser razão suficiente para permitir uma leitura tão larga no que se refere às variadas possibilidades levantadas com o propósito de listar as condições em que o virtuoso age viciosamente (eventualmente, raramente, frequentemente, ligeiramente etc., e *de forma alguma* sempre - pois o virtuoso, na concepção de Curzer, não pode agir virtuosamente sempre). A razão metodológica aventada por Curzer é o fato de que os assuntos práticos, próprios do campo da ética ocorrem *no mais das vezes* (p. 234-235), o que não é sem dúvida uma surpresa, já que é clara a visão aristotélica da extensão da *hexis*, se não incontornável como pensam alguns, seguramente difícil de ser contornada. Nesse último caso, a não inflexibilidade da mesma permite trabalhar em termos de regularidade, do usual, do frequente, no mais das vezes, mas não sempre. Isso abre espaço para o entendimento dos variados tipos de imperfeição a que estamos submetidos, o que permite, em dadas circunstâncias, que ajamos diferentemente de nossa disposição virtuosa. É uma posição razoável, mas insuficiente para assentar uma tese tão forte quanto a que Curzer defende em seu artigo. Uma coisa é afirmar que no mais das vezes fazemos x, mas em algumas vezes fazemos

~x, outra coisa é sustentar que ~x representa a o antípoda extremo de x, ~x parece ser tão somente diferente de x. No mais das vezes comemos pão com manteiga no café da manhã, eventualmente preferimos pão com geleia: ~x aqui é o pão com geleia. Como Aristóteles observou, a mediedade (*mesotês*) é relativa a nós, agentes e situação, e existirão situações em que a minha ação aproximar-se-á mais de um extremo ou do outro, já que a doutrina da mediedade não é uma estanke doutrina da moderação.

Mas talvez a debilidade da argumentação, me parece, é dar uma extensão, uma amplitude, demasiada alargada para a noção de virtuoso, e, ao mesmo tempo, não se ocupar mais profundamente da razão prática, e da força atribuída a ela por Aristóteles.

Por isso é necessário tratar do livro VI 13 da *EN*, pois é um *locus* interessante para testar a referida tese, e restringir a um personagem quem é verdadeiramente virtuoso, pois há uma distância tangível entre quem possui a virtude própria e a virtude [moral] natural.

No livro VI 13, Aristóteles propõe-se a reexaminar a virtude (1144b1), ressaltando que a virtude é similar de certo modo à prudência, e esta está relacionada com a habilidade, não sendo a mesma relação, mas uma relação semelhante, assim a virtude [moral] natural está relacionada com a virtude própria, pois estamos desde o nascimento inclinados a sermos justos, temperantes ou corajosos, “entretanto buscamos o bem propriamente dito como sendo outra coisa, isto é, procuramos fazer que essas qualidades ocorram de um outro modo” (1144b6-8, tradução de Lucas Angioni), pois, “se se adquire inteligência (*ean de labei noun*), o que faz diferença no agir e, assim, a disposição, sendo semelhante, será virtude propriamente dita” (1144b12-14, tradução de Lucas Angioni, conservadoramente adaptada). Assim, se temos em relação a parte opinativa a habilidade e a *phronêsis*, no concerne ao caráter possuímos dois tipos de virtude, a *kuria aretê* e a *aretê phusikê*, a virtude própria e a virtude [moral]

natural, sendo que a primeira não se dá sem a prudência (1144b14-17), já que, como salienta Aristóteles, não é possível ser prudente sem ser bom (1144a36-b1), ou melhor, não podemos ser prudentes sem sermos virtuosos, em função de que a virtude garante a retitude do fim, enquanto a prudência assegura a retitude do que conduz ao fim (1145a2-5).

A virtude em questão é uma disposição de agir em conformidade com a reta razão, e “correta é a razão conforme à prudência”. Logo, “a virtude é uma tal disposição, a que é conforme à prudência. Mas é preciso fazer uma pequena mudança. Virtude não é apenas a disposição conforme à razão correta (*kata ton orthon logon*), mas a disposição com a razão correta” (1144b21-23, tradução de Lucas Angioni conservadoramente adaptada). E a razão correta a respeito desses assuntos é a prudência, com que faz com que ocorra o fato de que as virtudes sejam acompanhadas pela razão (*meta tou orthou logou*, 1144b24-28). Por essa razão, pode-se afirmar que (1144b30-1145a1):

“Pelo que foi dito, é evidente que não é possível ser propriamente bom sem prudência, tampouco é possível ser prudente sem a virtude do caráter. É também por essa via que se pode refutar o argumento pelo qual se defenderia que essas virtudes são separadas uma da outra: não é o mesmo indivíduo, pois, que é naturalmente bem disposto a todas elas, de modo que já seria possuidor de uma, mas não teria ainda adquirido a outra. *Ora, isso é possível para as virtudes naturais*, mas não é possível para as virtudes segundo as quais alguém é reconhecido como “bom” sem mais. *De fato, todas ocorrem junto com a prudência, que é uma só.*” (tradução de Lucas Angioni, conservadoramente adaptada)

Relembremos uma afirmação de Curzer mencionada anteriormente: “Uma pessoa pode ser naturalmente virtuosa a respeito de alguma esfera da vida humana, enquanto permanece continente, incontinente ou até mesmo vicioso a respeito de alguma outra esfera” (p. 246-247). Sim, isso pode servir para as

virtudes [morais] naturais como o primeiro itálico acima destaca. Isso pode ocorrer tão somente na esfera da virtude [moral] natural, pois na esfera da prudência tal situação não é factível, como mostra o segundo itálico no interior da citação: a posse da prudência implica a posse das outras virtudes (1145a1-2, 1146a7-9).

Claro, poderia ser alegado que essa é uma tese forte da conexão que não se mostra realista, pois se além a uma composição idealizante do virtuoso, sobre sua perfeição inerente indiscutível. Mas isso não se dá em uma tese mitigada acerca da conexão das virtudes, se entendermos que a presença da prudência, em dada situação, requer algumas virtudes, não todas, mas que, em outra situação, outras virtudes devem acompanhar a prudência, o que não abre uma brecha para uma leitura fundamentada em ‘duas esferas’, pois há uma só esfera, a do prudente, da razão prática e da ação virtuosa sem a possibilidade de se agir viciosamente em nenhuma circunstância distinta, afinal, como vimos, o prudente é *kanôn kai metron*, mas ele também é o responsável por dar razões, pesar razões para a ação em vista de um fim moralmente bom: ele segue a reta razão, seu desejo é reto, o que faz com que, na hora da ação, independentemente da situação, ele aja virtuosamente sem fricção interna, sem conflito interno. Não há evidência textual aparente que permita a debacle do prudente, ao menos nos termos de Curzer.

Ou Curzer pode se refugiar na negação da tese idealizante, ou buscar outras evidências para além da perspectiva realista que apresentou em seu artigo - até mesmo recuar e reconhecer a vaidade da tese idealizante, o que parece improvável -, que dá a impressão de tão somente residir em agentes que ainda não desenvolveram a virtude plena, a virtude própria, mas apenas em figuras tais quais os continentes, incontinentes ou detentores de virtudes [morais] naturais, ressalvando que essas últimas estão desacompanhadas de prudência.

Referências

- ARISTÓTELES. **Ethica Nicomachea** (I. Bywater, ed.). Oxford: Oxford Classical Texts, 1942.
- _____. **Política** (D. Ross, ed.). Oxford: Oxford Classical Texts, 1988.
- _____. **Éthique à Nocamaque** (traduction et présentation para R. Bodéüss). Paris: GF Flammarion, 2004..
- CURZER, H. “How good peoople do bad things: Aristotle on the misdeeds of the virtuous”. **Oxford Studies in Ancient Philosophy**, 28 (1), 2005, p.233-256.
- DESTRÉE, P. 2011. “Aristotle on responsibility for one’s character”. In: PAKALUK, M.; PEARSON, G. (eds). **Moral psychology and human action in Aristotle**. Oxford: Oxford University Press.
- ZINGANO, M. **Estudos de ética Antiga**. São Paulo: Discurso Editorial, 2007.
- _____. **Aristóteles. Ethica Nicomachea I 13 – III 8. Tratado de Filosofia Moral**. São Paulo: Odysseus, 2008.

Capítulo 6

A questão do realismo

*Kit Fine*¹

Meu objetivo neste texto é ajudar a estabelecer a fundamentação conceitual e metodológica para o estudo do realismo. Cheguei a duas principais conclusões: primeira, que há uma concepção metafísica primitiva da realidade, a qual não pode ser entendida em termos fundamentalmente diferentes; e segunda, que as questões sobre o que é real devem ser resolvidas com base em considerações de fundamentação. As duas conclusões estão um tanto em tensão uma com a outra, pois a falta de uma definição do conceito de realidade parece estar no caminho do desenvolvimento de uma metodologia sólida para determinar sua aplicação, e uma das minhas principais preocupações têm sido mostrar como a tensão entre as duas pode ser resolvida.

O texto é dividido em duas partes principais. Na primeira, aponto para as dificuldades na elaboração de uma concepção metafísica da realidade. Começo por distinguir esta concepção da concepção ordinária da realidade (§1) e depois, mostrar como os dois principais concorrentes para a concepção metafísica – o factual e o irreduzível – parecem, ambos, resistir a formulação em outros termos. Isso leva ao desafio quietista, no qual as questões do realismo são sem sentido ou inúteis (§4); e a segunda parte do

¹ New York University. Tradução: Marina M. Burkle.

texto (§§ 5-10) é amplamente dedicada a mostrar como este desafio pode ser vencido. Começo por introduzir a noção de fundação (§ 5) e, em seguida, mostrar como pode ser usada de base para resolver questões tanto de factuality (§§6-7) quanto de irredutibilidade (§§8-9). Concluo com algumas observações sobre a unidade essencial destas duas questões e dos meios pelos quais elas devem ser respondidas (§10).

1. Realidade

Entre as questões mais importantes em filosofia estão aquelas sobre a realidade desta ou daquela característica do mundo. Existem números ou outros objetos abstratos? É tudo mental ou tudo físico? Existem fatos morais? É através da tentativa de resolver tais questões que a filosofia mantém a promessa de nos apresentar uma imagem de mundo, uma imagem de como o mundo é e do nosso lugar dentro dele.

No entanto, como é muitas vezes verdade em filosofia, as dificuldades começam com a formulação da questão, em vez da tentativa de resposta. O antirrealista sobre números alega:

Não há números.

Mas a maioria de nós, em nossos momentos não-filosóficos, estamos inclinados a pensar:

Ha números primos entre 2 e 6.

E a segunda dessas afirmações ainda implica que existem números, o que é incompatível com a primeira afirmação. Da mesma forma, o antirrealista sobre moralidade mantém:

Não há fatos morais.

Mas ele também pensa:

Matar bebês por diversão é errado.

E, no entanto, a segunda afirmação implica ser um fato que matar bebês por diversão é errado e, como esse é um fato moral, sua existência é incompatível com a primeira afirmação.

Como, à luz de tais possíveis conflitos, as afirmações realistas e antirrealistas devem ser construídas? Devemos tomar o conflito entre antirrealismo e opinião dada não-filosófica como um conflito genuíno ou não? E se não, então como é que o aparente conflito entre eles é dissolvido?

Se tomarmos o conflito como genuíno, obtemos o que foi chamado de concepção “eliminativa” ou “cética” de antirrealismo. O antirrealista será levado a colocar em discussão o que nós ordinariamente aceitamos, o que o realista, por sua vez, será levado a endossar. Assim, o antirrealista sobre os números será levado a negar, ou duvidar, que há números primos entre 2 e 6; e da mesma forma, o antirrealista moral será levado a negar, ou duvidar, que matar bebês por diversão é errado.

É claro que a mera rejeição do que ordinariamente acreditamos é negativa e, portanto, presumivelmente, o interesse do antirrealismo nesta concepção, deve derivar do pressuposto de que a filosofia é capaz de nos fornecer algumas razões especiais para duvidar do que normalmente acreditamos. Assim, o antirrealista pode tentar nos convencer de que não temos nenhuma boa razão para acreditar em um reino de objetos abstratos com os quais não podemos ter contato causal ou que, em questões morais, não podemos ter justificativa para ir além da mera expressão de aprovação ou desaprovação. Nossa visão de mundo será, portanto, o produto de lidar com essas dúvidas, seja por colocá-las de lado ou por recuar para o ceticismo.

O antirrealismo, assim entendido, tem uma longa e ilustre história; e certamente seu interesse não deve ser negado. Contudo, nesta época de modéstia pós-mooreana, muitos de nós estão

inclinados a duvidar que a filosofia está em posse de argumentos que podem genuinamente servir para minar o que nós ordinariamente acreditamos. Talvez se admita que os argumentos do cético parecem ser totalmente convincentes, contudo, os Mooreanos entre nós sustentarão que a própria plausibilidade de nossas crenças ordinárias é razão suficiente para supor que *deve* haver algo de errado nos argumentos do cético, ainda que sejamos incapazes de dizer o que. Até agora, então, já que as pretensões da filosofia de fornecer uma visão de mundo se apoiam em sua pretensão de estar em posse de um terreno epistemológico elevado, seria melhor desistir de tais pretensões.

Existe espaço para outra forma de antirrealismo – e outra perspectiva das pretensões da filosofia – que não a coloque em conflito com a opinião dada advinda do senso comum? Se houver, então requer que possamos consistentemente afirmar que algo é o caso e ainda assim negar que é realmente o caso.² Requer, em outras palavras, uma concepção *metafísica* da realidade, que nos permite distinguir, dentro da esfera do que é o caso, entre o que é realmente o caso e o que é apenas aparentemente o caso.

Mas o que poderia ser essa concepção metafísica da realidade? Duas principais respostas a esta questão foram propostas. De acordo com a primeira, a realidade metafísica deve ser identificada com o que é “objetivo” ou “factual”. O antirrealista, nesta concepção, nega que existam quaisquer fatos “lá fora” em virtude dos quais as proposições de um dado domínio podem ser verdadeiras. As proposições do domínio não estão no “negócio” de declarar tais fatos; elas servem apenas para indicar nosso engajamento com o mundo sem afirmar, de maneira objetiva, como o mundo é. Como exemplos familiares de tal posição, temos

2 Eu aqui ignoro a possibilidade de que a reconciliação seja alcançada modificando nossa visão da opinião dada advinda do senso comum, seja por não considerá-la como aquilo que acreditamos ou supondo que seu conteúdo é diferente do que nós naturalmente acreditamos que seja. Essas tentativas de reconciliação, na minha opinião, apenas mudam o conflito com a opinião dada para outro lugar.

o expressivismo em ética, segundo o qual juízos éticos são meras expressões de uma atitude; o formalismo em matemática, segundo o qual juízos matemáticos são meras mudanças dentro de um sistema de regras formais; e o instrumentalismo em ciência, segundo o qual teorias científicas são meros dispositivos para a previsão e controle de nosso ambiente. De acordo com a segunda concepção, a realidade metafísica deve ser identificada como o que é “irredutível” ou “fundamental”. Nesta perspectiva, a realidade é constituída por certos fatos irredutíveis ou fundamentais; e ao negar a realidade em um determinado domínio, o antirrealista está afirmando que os fatos são todos redutíveis a fatos de algum outro tipo. Assim, o naturalista ético alegará que todo fato ético é redutível a fatos naturais, o logicista alegará que todo fato matemático é redutível a fatos da lógica, e o fenomenalista alegará que todo o fato sobre o mundo exterior é redutível a fatos sobre nossos dados dos sentidos.

Podemos ver o antifactualista e o reducionista como apontando duas maneiras diferentes em que uma proposição pode falhar em “corresponder” aos fatos. Pois a proposição pode falhar até mesmo ao apontar a direção dos fatos, por assim dizer; ou pode-se deixar de indicar, em um nível mais fundamental, como são os fatos. No primeiro caso, as proposições de um determinado domínio nem mesmo *representarão* os fatos, enquanto no outro, as proposições não representarão os fatos de forma *perspicaz* — haverá algumas divergências entre como os fatos são “em si” e como eles são representados como sendo. Se alguma dessas concepções metafísicas da realidade é viável, então parece nos fornecer uma maneira de defender uma forma não-cética de antirrealismo, pois será perfeitamente compatível com afirmar qualquer proposição que negue que a realidade é genuinamente factual ou genuinamente fundamental. O expressivista, por exemplo, pode afirmar que matar bebês por diversão é errado e ainda negar que, ao afirmar tal sentença, ele está fazendo uma afirmação factual; e o logicista pode afirmar que $5 + 7 = 12$ e ainda

negar que ele está afirmando assim algo fundamental. A verdade é uma coisa, o *status* metafísico, outra.

Porém o problema agora não é defender a posição antirrealista, mas ver como ela poderia ser inteligível. Considere o antifactualista em ética. Como ele é considerado não cético, ele presumivelmente estará disposto a afirmar que matar bebês por diversão é errado. Sendo assim, então ele não deveria estar preparado para admitir que ele está, dessa forma, fazendo uma afirmação sobre como as coisas são? E isso não é uma afirmação sobre como as coisas estão no mundo — o único mundo que conhecemos, que inclui tudo que é o caso e exclui qualquer coisa que não for o caso? Então ele não é comprometido com proposições serem factuais?

É claro que o antirrealista insistirá que ele foi mal interpretado. Ele sustentará que a proposição de que matar bebês por diversão é errado não faz uma afirmação sobre o mundo real como ele o concebe e que, embora possa ser correto afirmar que matar bebês por diversão é errado, ainda não há um fato “lá fora” no mundo real ao qual seja correspondível. Todavia, a dificuldade então é entender o contraste entre o seu mundo —o mundo real “lá fora” — e o mundo dos fatos mundanos comuns. Pois que espaço existe, em nossa concepção comum da realidade, para qualquer distinção maior entre o que é genuinamente um fato e o que meramente tem a aparência de um fato?

Similarmente, o reducionista em ética afirmará que os fatos éticos são redutíveis a fatos de outro tipo e, deste ponto de vista, nega que eles sejam reais. Agora, pode ser concedido que existe um sentido no qual certos fatos são mais fundamentais do que outros; eles podem servir para explicar outros fatos, ou talvez, de algum outro modo, serem constitutivos deles. Porém, como isso fornece um fundamento para negar a realidade dos outros fatos? Certamente, eles tinham uma explicação ou constituição em termos dos fatos reais que pareceria indicar que eles mesmos eram reais.

Qual é então a concepção de redução para a qual o redutível não será real?³ Assim como havia uma dificuldade em compreender uma concepção metafísica dos fatos, uma concepção que possa servir para sustentar uma forma metafísica de antirrealismo, também há uma dificuldade em entender uma concepção metafísica de redução. Em qualquer um dos casos, parecemos evitar os absurdos do ceticismo, mas apenas comprando as obscuridades da metafísica. Um tipo de problema de fundamento superior foi simplesmente trocado por outro.

2. Factualidade

Existe alguma maneira de sair das dificuldades anteriores? Uma forma de antifactualismo ou reducionismo pode ser sustentada? Vamos discutir uma questão por vez.

No caso do antifactualismo, tem sido comumente suposto que há alguma característica dos pronunciamentos não declarativos — como “Ai!” ou “Saia daqui!” — que obviamente torna-os não factuais, e que também é possuída, embora não tão obviamente, pelas proposições declarativas de um determinado domínio.⁴ Assim, apesar dessas proposições serem declarativas em forma, devem ser classificadas como pronunciamentos não declarativos, como não factuais. Então, por exemplo, na versão tradicional deste tipo de pronunciamento, uma proposição “não-cognitiva” ou “não-factual” não é considerada candidata à verdade ou falsidade, e o antirrealista é levado a negar que as proposições de um determinado domínio em discussão são candidatas à verdade ou falsidade.

Mas o problema com esta abordagem é que o que se considera como uma característica não óbvia da proposição em

3 Muitos filósofos não aceitam a redução por ter importação antirrealista. Seu conceito de redução parece corresponder de maneira mais próxima ao que mais tarde chamarei de “fundamento”.

4 Eu aqui relevo sobre a dificuldade de saber se os portadores da não-factalidade podem propriamente ser proposições.

discussão, é na verdade, uma característica que obviamente lhe falta. Portanto, dado que matar bebês por diversão é errado, se segue —no sentido simples e comum de “verdade” — que é verdade que matar bebês por diversão é errado; e assim a proposição de que matar bebês por diversão é errado é um candidato à verdade ou falsidade, afinal. O não-cognitivista tradicional deve, portanto, ou estar usando os termos “verdadeiro” e “falso” em algum sentido metafisicamente inflacionado que ainda precisa ser explicado, ou, ele deve considerar-se um completo cético que não está disposto a afirmar qualquer proposição do domínio em discussão.

Também não ajuda apelar para outras características factuais óbvias de proposições no lugar da aptidão à verdade ou falsidade. Pode-se, talvez, sugerir que uma proposição factual é aquela capaz de ser acreditada ou asserida, figurada em inferências, ou incorporada em contextos linguísticos maiores. No entanto, o mesmo ponto se aplica. Pois no sentido comum de “acreditar”, “asserir”, etc., nós temos crenças morais, fazemos asserções morais, tiramos conclusões morais, e incorporamos proposições morais em contextos linguísticos maiores; e similarmente para as proposições da matemática, da ciência ou de outras áreas de discussão. De fato, uma vez dada uma destas características, o resto parece se seguir — sua posse é, na maior parte, um “pacote”. Parece, portanto, não haver esperança razoável de identificar uma forma não-cética de factualidade em termos de posse de algumas dessas características em oposição às outras.

Qualquer visão razoável e não cética deve, portanto, conceder que proposições do tipo que figuram na discussão realista, possuirão todas as armadilhas óbvias da factualidade: elas serão capazes de ser verdadeiras ou falsas, acreditadas ou asseridas, incorporadas em contextos linguísticos maiores, e assim por diante. O antifactualista deve, portanto, ser um *quase-realista* e atribuir ao não factual todas as características que foram tradicionalmente pensadas pertencerem ao factual. Mas se o não factual não deve ser distinguido do factual em termos das

armadilhas óbvias da factualidade, então como é que ele será distinguido? O que, em uma palavra, é a diferença entre *quase-realismo* e *genuíno* realismo?

Vários critérios mais sofisticados que parecem evitar essas dificuldades foram propostos.⁵ Vou me concentrar no critério proposto por Dummett, segundo o qual o realismo para uma dada área de discurso é principalmente constituído em questão de sua conformidade com o Princípio da Bivalência, o princípio de que toda declaração do discurso deve ser verdadeira ou falsa.⁶ Vou argumentar que o critério, mesmo quando suplementado, é insatisfatório, e depois tentar tirar algumas conclusões mais amplas.

Devemos notar de imediato que a proposta, mesmo que de outro modo mais aceitável, não atende às nossas necessidades. Uma vez que estamos atrás de uma forma não cética de antirrealismo, que não esteja em desacordo com a opinião dada advinda do senso comum. Porém, em relação às muitas áreas do discurso, a opinião dada advinda do senso comum afirma que as declarações são sujeitas a bivalência em dada dimensão e assim, qualquer forma de antirrealismo de Dummett deve, portanto, ser cética. Consequentemente, até o momento, como os filósofos desejaram defender uma forma completamente não-cética de antirrealismo, o critério de Dummett deve ser considerado insatisfatório.

Outro problema diz respeito à aplicação do critério para declarações particulares. Certamente desejamos ser capazes de afirmar ou negar que uma declaração particular—como $7 + 5 = 12$ — responde ou não por uma “realidade exterior”, ou seja, possui

5 Uma pesquisa útil de tais critérios é dada em Wright [1992] e algumas críticas gerais a elas devem ser encontradas em Rosen [1994], Dworkin [1996] e Stroud [2000], capítulos 1-2.

6 Uma proposta semelhante foi feita por Gaïfman [1975]. Deve-se notar que Dummett [1993b], p. 467, não está inclinado, como eu, a assimilar o não-factualismo de uma forma sofisticada de expressivismo com o não-factualismo de uma posição construtivista. No entanto, nenhuma das minhas críticas se voltará a fazer essa assimilação.

“importação” realista. No entanto, qual é o critério a ser aplicado em tal caso? Presumivelmente, associando a declaração com uma área particular do discurso. Mas qual? A resposta pode muito bem depender do que dizemos. No caso de ' $7 + 5 = 12$ ', por exemplo, a área do discurso poderia ser a linguagem da aritmética equacional (sem variáveis ou quantificadores), a linguagem de primeira ordem para adição ou a linguagem de primeira ordem para adição e multiplicação. Contudo, o finitista pode muito bem aceitar a bivalência para a primeira linguagem, embora não para as outras, enquanto o construtivista pode aceitar a bivalência para as duas primeiras linguagens, embora não para a terceira. Portanto, nenhuma resposta estável é assegurada. (O problema aqui é análogo ao problema em determinar uma classe de referência para “casos únicos” de probabilidades, e tolhe vários outros critérios para o realismo também.)

Mesmo na aplicação em áreas do discurso, a bivalência não é, por si só, suficiente para o realismo. Um simples contraexemplo é proposto da seguinte maneira.⁷ Suponha que um antirrealista venha a ser completamente teimoso sobre o discurso dado: ele adquire uma visão (por razões internas ao discurso, embora talvez muito ruins) acerca da resposta verdadeiro/falso para cada questão particular que possa surgir. Ele então seria comprometido com cada instância da bivalência, e contanto que ele estivesse ciente de ter se tornado completamente teimoso, ele também estaria comprometido com manter a bivalência em toda a declaração do discurso. Não obstante, parece absurdo supor que, naquele dia fatídico em que a última questão se enquadraria na oscilação de sua opinião, ele estaria destinado a se tornar um realista. Como ser um antirrealista pode impedi-lo de formar uma opinião sobre o assunto?

⁷ Ver Edgington [1980-81] e Winkler [1985] para críticas adicionais da bivalência como uma condição suficiente. Eu também acredito que há problemas com a bivalência como uma condição necessária para o realismo, mesmo quando as fontes óbvias de lacunas de valores de verdade (tais como vagueza ou falha de referência) são removidas, mas isso não é algo que eu discutirei.

Claramente, a aceitação da bivalência para uma determinada área do discurso não é suficiente para garantir o realismo. Esta tem de ser aceita pelo motivo certo. Mas existe alguma maneira de suplementar a bivalência, de modo a assegurar que tal aceitação será feita pelo motivo certo? Duas propostas para suplementar a bivalência foram consideradas (frequentemente combinadas) - uma epistêmica e a outra semântica.

De acordo com a proposta epistêmica, a noção de verdade que ocorre na bivalência deve ser tal que seja possível uma declaração do discurso dado ser verdadeira e ainda assim incognoscível — ou mesmo sem qualquer evidência possível de sua verdade. De fato, se uma declaração dada fosse verdadeira e ainda assim incognoscível, então isso parece fornecer uma forte razão para tomar a declaração como tendo “importação” realista, além de qualquer discurso em que possa ser colocada.

Entretanto, acredito que mesmo essa plausível proposta esteja sujeita a um contraexemplo, embora de um tipo mais sofisticado. Considere uma semântica na qual o significado de uma sentença é dado pelas situações evidenciais em que sua asserção é garantida. O princípio orientador da semântica é que alguém está garantido a asserir uma sentença em uma situação evidencial dada se, e somente se, para qualquer melhoramento na situação evidencial de alguém, há um melhoramento posterior no qual sua asserção seria garantida.⁸ Deve então ficar claro que a Lei do Terceiro Excluído e, portanto, o Princípio da Bivalência, são

8 Vamos usar ' $t \models A$ ' para ' t garante A ' e ' $t \geq s$ ' para ' t melhora s ' (suponho que qualquer situação é uma melhoria em si mesma). Fórmulas moleculares dentro da semântica proposta podem então estar sujeitas às seguintes cláusulas:

$$(i) \ t \models B \text{ e } C \text{ se } t \models B \text{ e } t \models C$$

$$(ii) \ t \models B \text{ ou } C \text{ se } (\forall t^* \geq t) (\exists u \geq t^*) (u \models B \text{ ou } u \models C)$$

$$(iii) \ t \models \neg B \text{ se } (\forall t^* \geq t) (\text{não-}(t^* \models B)).$$

Dado que as fórmulas atômicas p satisfazem a condição de não encerramento:

$$(*) \ t \models p \text{ se } (\forall t^* \geq t) (\exists u \geq t^*) (u \models p)$$

então assim será toda fórmula A .

válidos, isto é, são garantidos em qualquer situação evidencial.⁹ Pois tome qualquer melhoria nessa situação. Nesse caso ou o melhoramento garante $\neg A$ e, conseqüentemente, tem um melhoramento (ou seja, em si) que garante $(A \vee \neg A)$, ou tem um melhoramento que garante A e, portanto, garante $(A \vee \neg A)$.

Claramente, tal semântica pode ser adotada por um antirrealista. E assim resta mostrar como tal antirrealista pode estar justificado em manter que não há evidência possível a favor ou contra a verdade de uma determinada declaração A . Não está claro como isso pode se dar, pois para qualquer situação evidencial dada, ou irá garantir não- A , ou permitirá um melhoramento que garantirá A . Nosso antirrealista, no entanto, pode estar trabalhando com uma noção “objetiva” de garantia. Pode ser uma questão objetiva — uma para a qual ele não tem necessariamente acesso epistêmico — o que as situações evidenciais possíveis são e, conseqüentemente, o que elas garantem. As declarações em questão podem dizer respeito ao mundo exterior, por exemplo, e as situações evidenciais podem ser dadas pelos cursos de experiência que alguém pode realmente sofrer. Nosso antirrealista pode então argumentar que, embora eu esteja objetivamente garantido em acreditar que MacCavity não estava aqui, na cena do crime, já que nenhuma contra evidência jamais seria apresentada, mesmo assim ainda não posso ter evidência de que ele não tenha estado aqui, já que posso não ter base para excluir as possibilidades da contra-evidência. Claro, nosso antirrealista é, em certa medida, um realista — pois ele é realista sobre as possibilidades objetivas da experiência, mas ele não é um realista sobre o mundo exterior, que é o que está aqui em questão.

A segunda estratégia é fornecer um suplemento semântico para a bivalência: não só manter a bivalência, mas manter uma questão semântica. Conquanto, o nosso contraexemplo anterior

9 Dummett [1978], pp. 365-67, tem ressalvas em fazer a transição da Lei para o Princípio, mas não acredito que se apliquem no presente caso.

permanece, uma vez que a validade da bivalência é uma questão semântica sob a semântica antirrealista que propus (dado que existe uma sustentação semântica para aplicar a Lei do Terceiro Excluído às declarações de garantia.) Parece, portanto, necessário aprofundar-se no mecanismo pelo qual a linguagem deve ser interpretada. Mas o que poderia ser? Uma sugestão é que precisamos da “aceitação da semântica clássica de dois valores [...] em sua totalidade” (Dummett [1993], p. 468). Isso exigiria não só a bivalência, mas a exclusão de termos vazios, das cláusulas padrão para os conectivos e assim por diante. O problema agora é que, dado que nosso antirrealista está disposto a aceitar a bivalência, não está claro por que ele deveria estar disposto a aceitar o resto da semântica clássica — embora, é claro, sob sua própria compreensão do que isso significa. Outra sugestão, é que exigimos que nossa compreensão da linguagem seja verocondicional, ou seja, que nossa apreensão do significado de uma declaração deva consistir no conhecimento das suas condições de verdade. Embora os filósofos usem essa fraseologia como se soubessem o que isso significa, não é de todo claro que possa ser explicada de maneira que implique o realismo e, ainda assim, não pressuponha que as condições de verdade já sejam entendidas de forma realista. Pode-se, claro, insistir que a noção relevante de verdade deve estar em conformidade com a bivalência ou ser transcendente em evidências. Mas isso nos leva de volta as nossas dificuldades anteriores.

O que, acredito, fez esses vários critérios tão atraentes é que muitas vezes é difícil ver como se poderia plausivelmente manter que um dado critério é satisfeito (ou não satisfeito) e ainda assim ser realista (ou antirrealista). Realismo sobre matemática, por exemplo, é uma razão — e talvez a única boa razão — para sustentar que toda declaração matemática é verdadeira ou falsa ou que pode haver verdades matemáticas que estão além da nossa compreensão. Mesmo assim, deve ser reconhecido que, mesmo que a existência de uma realidade exterior talvez torne plausível que o

nosso contato linguístico e epistêmico com aquela realidade é de um certo tipo, não é nisso que a externalidade da realidade consiste. Ao pensar sobre esses assuntos, precisamos nos restaurar a um estado de inocência em que as afirmações metafísicas são vistas como tratando do assunto em questão — seja matemática, moralidade ou ciência — e não do nosso relacionamento com esse assunto.

De fato, uma conclusão mais ampla pode ser justificada. Pois nós vimos que, mesmo quando nos empilhamos nas condições, estávamos incapazes de encontrar uma condição suficiente para o factualismo (e, posso acrescentar, para o não-factualismo, não que este seja um caso que eu tenha considerado). Isto sugere, portanto, que pode não haver nenhuma condição de todo suficiente para algo ser factual (ou não-factual) — a menos, é claro, pela razão trivial de que as condições não podem ser satisfeitas devido a petição de princípio, segundo a qual uma concepção problemática de factualidade já foi pressuposta.

Se isto está certo, então significa que nem mesmo será possível fornecer uma formulação adequada de qualquer posição factualista ou antifactualista, ou seja, uma posição que implicará que a posição é de fato comprometida com o factualismo ou com o antifactualismo no domínio em questão; e a análise das formulações reais de tais posições confirma isso. Considere o expressivismo, por exemplo. O expressivista deseja sustentar que as afirmações morais são expressivas da mesma maneira que expressões como "Ai". Mas, é claro, a mera afirmação de que as afirmações morais são expressivas não serve para distinguir sua posição da posição do realista moral, já que até este pode estar disposto a manter que afirmações morais são usadas para expressar nossas atitudes morais bem como para relatar os fatos morais. O expressivista deve, portanto, estar alegando que as afirmações morais são *meramente* expressivas, que elas não possuem outra característica que sirva para fazer delas factuais. Agora, no caso de "Ai", podemos ver porque deve ser assim, já que

a expressão “Ai” não é usada para dizer qualquer coisa verdadeira ou falsa. É neste exato ponto que as afirmações morais diferem de expressões como “Ai”. Assim, permanece completamente opaco o que há exatamente na posição do expressivista que o obriga a abraçar a não-factualidade do discurso moral; e, similarmente, para as outras formas particulares de antifactualismo que foram propostas.

3. Redutibilidade

Nós nos voltamos para a segunda das duas concepções metafísicas da realidade, a concepção da realidade como irreduzível. Esta concepção não pode estar em melhor situação que o conceito de redução com a qual está associada; e então podemos perguntar: o que é para uma proposição (ou declaração ou sentença) ser redutível às outras?

Três principais linhas de resposta a esta questão foram propostas. De acordo com a primeira, a redução é uma questão de análise lógica. Dizer que uma sentença se reduz a, ou é analisável em termos de outra, é dizer que elas expressam a mesma proposição, mas que a forma gramatical da segunda está mais próxima da forma lógica da proposição do que a forma gramatical da primeira.¹⁰ Assim, a redução revela uma discrepância entre a forma gramatical “aparente” da sentença e a forma lógica “genuína” da proposição e serve para aproximar as duas. Para dar um exemplo paradigmático, a sentença ‘O americano médio tem 1,54m’ reduziria ‘A soma das alturas de todos os americanos dividido pelo número de americanos é de 1,54m’, uma vez que a última nos traz para mais perto da forma lógica da proposição que é expressa.

Esta abordagem sofre de ao menos dois pormenores. Primeiro, é incapaz de lidar com reduções um-para-muitos. O

¹⁰ Às vezes a ênfase está em uma correspondência com “fatos” em vez de proposições.

filósofo que não acredita em fatos conjuntivos irá querer dizer que a verdade da conjunção $S \& T$ se reduz a verdade de seus conjuntos S e T . Mas aqui não há questões sobre uma única proposição ser expressa à esquerda e à direita. Também é incapaz de lidar com reduções de um-para-um em que a sentença redutora é apenas suficiente para a sentença a ser reduzida. Por exemplo, quando S é verdadeiro e T é falso, podemos querer dizer que a verdade do disjuntivo $S \vee T$ é redutível a verdade do disjuntivo S . Todavia, novamente, nenhuma proposição única é expressada.

A segunda dificuldade diz respeito a reduções *de re*. Assim como há afirmações *de re* modais que devem ser distinguidas das suas homólogas *de dicto*, também há afirmações de redução *de re* que devem ser distinguidas das correspondentes afirmações de redução *de dicto*. Assim, podemos querer afirmar não apenas que a sentença ‘O casal Jack e Jill é casado’ é redutível a sentença ‘Jack é casado com Jill’, mas também que a satisfação da frase aberta ‘ z é casado’ pelo casal Jack e Jill é redutível à satisfação da frase aberta ‘ x é casado com y ’ por Jack e Jill. No entanto, é difícil enxergar, na visão proposta, como isso poderia ser um caso de análise lógica, pois a proposição expressa pela frase aberta ‘ z é casado’ sob a atribuição do casal para z é presumivelmente a proposição singular de que esse casal é casado, e proposição expressa pela frase aberta ‘ x é casado com y ’ sob a designação de Jack para x e Jill para y é presumivelmente a proposição singular que Jack é casado com Jill; e ainda assim estas duas proposições devem ser tomadas como distintas, já que, sem uma noção refinada de identidade proposicional, nós estaremos perdidos para explicar como uma forma gramatical pode estar mais próxima da forma lógica genuína do que outra.¹¹

11 A importância das reivindicações redutivas *de re* não foi devidamente avaliada. Eles permitem que se obtenha uma enorme simplificação na formulação de muitas reduções e, em certos casos—como a teoria do agrupamento de particulares —, eles são essenciais para entender o próprio ponto da redução.

Pode parecer estranho que expressemos uma redução de casais aos seus membros, fazendo referência a casais, já que não é o ponto da redução mostrar que casais são uma “ficção lógica” e, portanto, não são realmente existentes? Entretanto esta linha de pensamento representa uma confusão entre as formas céticas e não-céticas de antirrealismo. Há uma ficção lógica, no sentido relevante, que não nos impede de fazer afirmações *não-filosóficas* sobre casais, como por exemplo, de que todos os casais no quarto são casados; e não mais nos impediria de fazer alegações *filosóficas* sobre casais do tipo “tipificado” pelas reduções. É meramente essa consistência que demanda que essas afirmações, em algum momento, devam ser reduzidas.¹²

A dificuldade mais séria com a presente abordagem é que esta repousa sobre uma concepção problemática de forma lógica. Para sustentar que uma sentença concernente a nações, digamos, expressa a mesma proposição, ou afirma o mesmo fato, como uma concernente a indivíduos e ainda é menos próxima em sua forma a essa proposição ou fato, já é comprar uma concepção metafísica carregada de forma lógica. Não há nada além de uma base metafísica para fazer tal afirmação. Pode-se tentar, é claro, fornecer uma explicação da forma lógica em termos metafisicamente neutros - talvez em termos do que é necessário para uma explicação satisfatória das "condições de verdade" ou da inferência válida. Mas na medida em que a perspectiva é bem-sucedida nesse aspecto, sua importância metafísica não é clara. Por que deveria a explicação mais satisfatória da inferência válida ou de condições de verdade ter quaisquer implicações sobre como o mundo realmente é, a menos que isso seja algo que já é requerido em tal explicação?

12 Da mesma forma, é perfeitamente plausível para o antifactualista ético expressar sua opinião de que não há propriedades morais genuínas na forma “Para qualquer propriedade moral e qualquer possível portador da propriedade, a posse do portador da propriedade é um assunto não-factual”. Quantificamos sobre todas as propriedades morais para expressar a opinião de que nenhuma delas é “real”.

De acordo com a segunda abordagem, a redução é um assunto *semântico*. É tomado como uma relação que se mantém em virtude do significado das frases às quais se aplica; e o que é mais distintivo sobre esta relação é que, dado que uma sentença se reduz a outras, deveria ser possível adquirir uma compreensão da sentença reduzida com base em uma compreensão das frases às quais ela se reduz.

Tal abordagem evita as dificuldades anteriores das reduções de um-para-muitos e um-para-um, uma vez que não há nada, em geral, para prevenir que uma sentença reduza simultaneamente várias sentenças que, em conjunto, fornecem condições suficientes — embora talvez não seja uma condição necessária — para a verdade da sentença dada.¹³ Mas ainda tropeça no problema das reduções *de re*. Para onde *c* é o casal Jack e Jill, *a* é Jack e *b* é Jill, não há conexão semântica entre *c* ser um casal e *a* ser casado com *b*,¹⁴ uma vez que não há nada semântico que possa servir para indicar que *a* e *b* são indivíduos que compõem *c*. Há sim, é claro, uma implicação semântica de *a* ser casado com *b* para *a-e-b* ser um casal. Porém, isso não é o que estamos atrás, já que queremos referência direta ao casal *c* e não referência indireta, através dos componentes *a* e *b*. Ou, novamente, pode ser uma verdade analítica que, para qualquer casal *c* há indivíduos *a* e *b* que compõem *c*, e são de tal modo que a propriedade de *c* ser casado reduz-se a *a* ser casado com *b*. Esta é uma afirmação geral e ainda deixa inexplicado o que as reduções particulares sustentam. Nós também podemos afirmar que a formulação geral da redução de tais casos nem sempre é uma verdade analítica, pois podemos querer dizer que, para qualquer quantidade *q* de água, existem moléculas de H₂O *m*₁, *m*₂, ... tal que a existência de *q* num determinado momento se reduz à existência de *m*₁, *m*₂, ... naquele

13 Dummett propõe tal concepção em [1993], pp. 56-7.

14 Ou, para colocá-lo mais linguisticamente, entre a satisfação da condição '*z* é casado' por *c* e a satisfação da condição '*x* é casada com *y*' por *a* e *b*.

momento; e esta generalização não é *a priori* e, portanto, presumivelmente não é analítica. Assim, não é como se reduções particulares pudessem sempre ter um suporte semântico.

Nem mesmo é claro que a existência de uma análise forneça uma condição *suficiente* para a redutibilidade no sentido metafísico. Assim sendo, nos deixe conceder que o que é para alguém ser solteiro, é para ele ser solteiro e ser um homem. Nós somos então obrigados a dizer que o fato de que alguém é solteiro é redutível ao fato de que ele é não-casado e o fato de que ele é homem e que, portanto, não há fato real dele ser solteiro? Talvez não. Podemos acreditar que existem realmente atributos complexos no mundo e que a sua atribuição não pode, por esse motivo, ser reduzida à atribuição dos atributos mais simples de que são compostos. Assim, embora possamos explicar sua identidade em termos dos atributos mais simples, nós não redutivamente descrevemos a sua atribuição nesses termos.

A lição a ser aprendida das críticas anteriores, eu acredito, é que a redução deve ser construída como metafísica mais do que como uma relação linguística ou semântica. Fazendo afirmações de redução, queremos falar não sobre nossa representação dos fatos, mas sobre os fatos em si. Assim, ao afirmar que duas nações em guerra se reduzem a tal e tal atividade militar por parte de seus cidadãos, nós não estamos fazendo uma afirmação sobre a nossa linguagem para descrever nações e cidadãos, ou mesmo sobre os nossos conceitos de nação ou cidadão, mas sobre as próprias nações e os próprios cidadãos e a conexão entre eles. Mais uma vez, precisamos restaurar nós mesmos a um estado de inocência metafísica no qual a redução está relacionada com o assunto em si e não com os meios pelos quais ele pode ser representado ou percebido.

De acordo com a terceira abordagem, mais recente, a redução é um assunto *modal*. Uma classe de proposições se *reduz* a — ou *sobrevem* (*supervene*) a — outra se, necessariamente, alguma

verdade de uma é acarretada pelas verdades da outra.¹⁵ Essa abordagem evita as dificuldades anteriores quanto à possibilidade de reduções *de re*, uma vez que as próprias proposições podem ser *de re*; todavia, sofre de dificuldades próprias. Por um lado, enfrenta o problema anterior da “classe de referência”, para saber quando uma proposição é redutível a outras, pois dependerá da classe de proposições com a qual elas estão associadas. Também não se aplica adequadamente aos domínios necessários, como a matemática, já que é sempre uma questão trivial que uma verdade essencial é implicada por quaisquer proposições que sejam.

Mas mesmo se limitarmos sua aplicação a domínios contingentes, existem outras duas graves deficiências em tal abordagem. Em primeiro lugar, ela não é capaz de capturar a ideia de que a verdade de uma proposição deve ser reduzida a algo mais básico. Por exemplo, velocidade em um instante, sobrevém a velocidade ao longo de um intervalo e vice-versa, e ainda assim nós não podemos dizer, sem circularidade, que cada uma reduz a outra. Tampouco ajuda insistir que a sobreveniência (*supervenience*) seja de mão única. Para supor que existem três parâmetros e que o valor de qualquer parâmetro sobrevém nos valores do dois outros, mas não no valor de um deles sozinho, então o valor de cada parâmetro será unidirecionalmente sobreveniente aos valores dos outros e, no entanto, não podemos, sem circularidade, dizer que o valor de cada parâmetro é redutível ao valor dos outros. Como um exemplo particular, podemos tomar como parâmetro a massa, volume e densidade de um dado corpo.¹⁶

Finalmente, a abordagem não é mais capacitada do que a primeira a capturar a “importação” antirrealista de afirmações redutivas. Uma vez que, enquanto a redução for considerada como

15 Os defensores dessa abordagem incluem Armstrong [1997], p. 12, Chalmers [1996], p. 48 e Jackson [1998], p.5. Muitos filósofos, devo observar, não tomam a sobreveniência para capturar uma noção metafisicamente significativa de redução.

16 Para uma discussão mais aprofundada destes pontos, ver Kim [1993], pp. 144-46, e as referências nele contidas.

nos aproximando do que é real, vamos querer negar a realidade de qualquer fato que seja redutível a outra coisa. Então como poderia a mera existência de certas conexões modais entre uma classe de proposições e outra servir para estabelecer a irrealidade (ou a realidade) dos fatos de qualquer classe?

De fato, uma conclusão mais ampla pode ser justificada, assim como no caso da factualidade. Pois é difícil ver como poderia haver *qualquer* condição suficiente para que uma proposição seja redutível a outras (casos triviais e circulares à parte). Uma vez que qualquer que seja talvez a condição suficiente, sua satisfação pareceria ser compatível com a adoção de uma verdadeira estratégia de posição realista, que tomou cada fato para ser real, e ainda compatível com a rejeição de qualquer afirmação redutiva dada.¹⁷

4. O Desafio Quietista

Vimos na discussão anterior que as perspectivas de definir as noções de factualidade e redutibilidade em termos fundamentalmente diferentes, ou mesmo fornecer conceitualmente condições suficientes e não problemáticas para a sua aplicação, não parecem boas. Apesar de termos chegado a esta conclusão negativa como resultado de uma investigação detalhada, tal conclusão poderia, talvez, ter sido antecipada desde o início. Estávamos atrás de uma forma de antirrealismo que não fosse necessariamente cética, em desacordo com a opinião dada advinda do senso comum. Agora, presumivelmente, não há nada especial sobre a opinião a ser dada a este respeito. Na medida em que uma posição antirrealista é independente de tal opinião, deve ser independente de todas as opiniões similares, dadas ou não. De fato, o central

¹⁷ No entanto, devemos observar uma assimetria nos dois casos. Pois mesmo que não haja garantia de não-factalidade, haverá garantia de não-redutibilidade. Pois se R é redutível a S então S implicará necessariamente R e assim a possível verdade de S e $\neg R$ (um fato puramente modal) garantirá que R não é redutível a S .

para a nossa atual compreensão do antirrealismo é fazer uma distinção entre o que poderíamos chamar de proposições de “primeira ordem”, que meramente dizem como as coisas estão sem levar em conta o seu *status* metafísico, e as correspondentes afirmações de “segunda ordem”, que apenas comentam sobre o *status* metafísico das proposições de “primeira ordem”. O que pareceria então garantir a possibilidade de uma forma não-cética de antirrealismo é a independência *geral* das afirmações de segunda ordem, neste sentido, a partir das proposições de primeira ordem com as quais elas lidam. Então, enquanto o critério de factualidade ou redutibilidade é declarado em termos de primeira ordem, sua inadequação simplesmente se seguirá desta forma geral de independência.¹⁸ O ceticismo tem um alcance muito amplo; e se estamos buscando escapar completamente de seu alcance, devemos adotar doutrinas de não-factualidade ou redutibilidade que estão livres de qualquer ônus de primeira ordem.

No entanto, estes resultados parecem ser profundamente perturbadores. Pois, se factualidade ou redutibilidade não devem ser entendidas em termos ordinários de primeira ordem, então como elas devem ser entendidas? Como podemos conferir sentido à ideia de que por detrás de cada fato putativo existe ou não existe algo real no mundo que o corresponde? Os resultados parecem especialmente preocupantes no caso da factualidade. Não existe nenhuma diferença de primeira ordem entre proposições factuais e não-factuais. As proposições não-factuais serão de natureza impostora — com todas as armadilhas usuais de factualidade, mas nenhuma relevante. Elas serão como “zumbis” que exibem todos os sinais externos de consciência mesmo sem ser conscientes. Isso

¹⁸ Ronald Dworkin apontou para mim que a moralidade de primeira ordem é naturalmente levada a incluir a alegação de que, se não há fatos morais objetivos, então vale tudo (ou alguma outra conclusão moral) e, dado que assim é, a segunda afirmação de que não há fatos morais objetivos não será independente da moralidade. Entretanto, para os propósitos atuais, eu não tomaria o critério acima como estritamente de primeira ordem.

posto, como vamos distinguir entre os dois? Pelo menos no caso do zumbi, talvez possamos dizer, por nossa própria experiência, o que é ser consciente. Contudo, no presente caso, pareceria não haver um ponto de vista especial a partir do qual pudéssemos fazer uma distinção entre o que é real e o que é meramente “sombra” lançada por nossa linguagem ou pensamento. Não há passo recuo aos fatos putativos para ver o que realmente está lá.

São considerações como essas que levaram vários filósofos atuais — os “quietistas” — a concluir que as noções metafísicas de facticidade e redutibilidade são desprovidas de conteúdo.¹⁹ E, é claro, uma vez que essas noções ocorrem, o mesmo então ocorre com o empreendimento metafísico associado a elas. A filosofia, neste modo de pensar, deveria abandonar sua pretensão de nos apresentar uma visão de ordem superior sobre como o mundo realmente é. Ou melhor: se existe uma visão, é a de que não há tal visão a ser obtida.

Contudo, por mais tentadora que seja essa conclusão, esta não é garantida evidencialmente. Pois a dificuldade em definir as noções pode derivar não da falta de conteúdo, mas de seu caráter distintivo. De fato, é difícil evitar essa impressão, uma vez que quando analisamos as várias tentativas de definição existentes, temos, na concepção da realidade como objetiva ou fundamental, uma ideia distintamente metafísica. Deste ponto de vista, a tentativa de definir essas noções em outros termos seria semelhante à falácia naturalista; e assim como seria um erro inferir a ininteligibilidade de noções normativas da dificuldade de defini-las em termos naturalistas, então seria um erro, no presente caso, inferir a ininteligibilidade das noções de facticidade e redutibilidade da dificuldade de defini-las em termos não-metafísicos.

19 Seu número inclui Blackburn [1992], pp. 7, 34, 168 e [1998], p. 319; Dworkin [1996]; A. Fine [1984], pp. 97-100; Putnam [1987], p. 19. Outros filósofos, como Rosen [1994], flertaram com o quietismo sem realmente abraçá-lo.

É claro que o quietista pode ter uma hostilidade geral para conceitos metafísicos, mas ele, entre todos os filósofos, não está em boa posição para justificar tal hostilidade de maneira principialista, pois, a base usual para rejeitar a inteligibilidade de toda uma esfera de conceitos é de que eles não podem ser tornados inteligíveis dentro de alguma visão de mundo escolhida — uma que vê apenas o físico ou apenas o psicológico, por exemplo, como real. Assim, pode-se tentar motivar a rejeição de conceitos metafísicos adotando uma visão de mundo que vê apenas os fatos de primeira ordem — como dados em ética, matemática ou ciência, etc.— como reais. Mas a adoção de tal visão já pressupõe a inteligibilidade de uma concepção metafísica da realidade. Deste modo, há um perigo real, pois, a posição quietista exclui como ininteligível o único fundamento que poderia torná-la plausível.

Há também fortes evidências intuitivas a favor da inteligibilidade; o fato de que uma noção parece fazer sentido é *prima facie* uma forte evidência de que ela faz sentido. Deveras, a indispensabilidade das noções na formulação de certas questões metafísicas pareceria tornar sua inteligibilidade quase impossível de se negar. Considere a questão dividindo “teóricos-A” e “teóricos-B” quanto à realidade temporal ser intrinsecamente estática (*intrinsically tensed*). Este é um assunto que não pode ser tornado inteligível sem invocar a concepção metafísica de “fato”. Pois o teórico-A vai querer afirmar, e o teórico-B vai querer negar, que há fatos estáticos no mundo; e é apenas a noção metafísica e não a noção ordinária de “fato” que pode servir adequadamente para representar o que está aqui em questão. Também podemos notar que apelar para a metáfora de um ponto de vista arquimediano é quase irresistível neste contexto. Pois o teórico-B vai querer adotar um ponto de vista arquimediano no qual a realidade temporal é descrita *sub specie aeternitatis*, enquanto o teórico-A negará que existe tal ponto de vista. Conferida inteligibilidade para a questão, deveríamos conferir inteligibilidade para a noção e para a metáfora

em aplicação neste caso particular. E se conferimos neste caso particular, então por que não conferimos em geral?

Mesmo que a acusação de ininteligibilidade não possa razoavelmente ser sustentada, há outra objeção mais moderada que pode ser, e que é igualmente devastadora nas suas implicações para a busca da metafísica.²⁰ Este quietismo é mais metodológico do que conceitual em sua orientação. A acusação não é de que não há noções significativas de facticidade ou redutibilidade, mas de que não há como determinar o que é ou não é factual ou o que reduz ou não a que.²¹ Dado que proposições não-factuais têm natureza impostora, como podemos distingui-las da “coisa real”? E dado que as reduções têm importação antirrealista, como podemos estabelecer que qualquer conexão proposta entre proposições terá essa importação? Juízos concernentes à facticidade e à redutibilidade parecem ser metafísicos no sentido pejorativo de flutuarem livres de quaisquer considerações que possam ser ditas a favor ou contra a sua verdade.

O quietista metodológico talvez possa conceder que existe uma presunção geral a favor de uma proposição ser factual. Ele também pode admitir que parecemos ter um preconceito metafísico contra certos tipos de proposições factuais — por exemplo, aquelas relativas a questões de gosto; e ele pode estar disposto a conceder a plausibilidade de certos juízos condicionais, tais como 'Se é uma questão factual se *P*, então é factual se não-*P*? Ou 'Se é matéria factual se a neve é branca, então é matéria factual se grama é verde'. No entanto, será acordado em todos os lados que tais considerações não nos levarão muito longe. O que é necessário são considerações detalhadas a favor ou contra uma dada visão realista; e o que não está claro é o que essas considerações podem ser.

20 Rorty [1979], p. 311, é alguém que adotou uma forma mais moderada de quietismo.

21 É uma esquisitice na crítica do positivista lógico à metafísica que essas duas acusações estejam ligadas. Pois, se não há como resolver questões metafísicas, quem se importa se fazem ou não sentido?

Claro, a literatura filosófica parece estar cheia de argumentos a favor ou contra essa ou aquela forma de realismo. É frequentemente mantido, por exemplo, que é mais fácil para o antifactualista acerca da matemática explicar a possibilidade do conhecimento matemático, já que o factualista enfrenta o problema de explicar como podemos estar em contato apropriado com um reino externo de fatos matemáticos; e é frequentemente pensado ser uma vantagem da visão expressivista em ética poder explicar o papel motivacional da crença moral. Mas todos esses argumentos, na medida em que são levados a suportar a questão não cética, parecem estar sujeitos a uma crítica devastadora, pois não está claro como eles retornam a adoção de um conceito metafísico, em oposição a um conceito ordinário da realidade e, portanto, por que eles estabelecem um método não-cético, em oposição a uma forma cética de antirrealismo. Assim, mesmo se empregarmos a noção fina (*thin*) e comum de fato, ainda pareceria ser um problema explicar como poderíamos ter conhecimento de fatos matemáticos; e permanece obscuro como esse problema se torna maior quando substituímos a noção espessa (*thick*) metafisicamente inflada para a noção ordinária e fina. Novamente, não está claro o que é a crença nos FATOS MORAIS (a coisa real) em oposição à crença nos fatos morais (a coisa ordinária) que faz com que seja menos plausível supor que as crenças morais são um tipo de atitude.

Acredito que este é o problema verdadeiramente sério levantado pelo quietismo. Não é que a noção de factualidade seja sem sentido, mas que é inútil; e a metafísica realista deve ser abandonada, não porque suas perguntas não possam ser levantadas, mas porque suas respostas não podem ser encontradas. O mundo real da metafísica é semelhante ao mundo numênico de Kant, uma coisa-que-nós-não-sabemos-o-que, e nenhum progresso é feito ao buscar a sua constituição. Podemos acrescentar que se o problema metodológico pudesse ser resolvido, o problema conceitual então perderia muito de sua força. Como podemos duvidar seriamente da inteligibilidade de um discurso

dado, quando o seu emprego na resolução de disputas não está, de outro modo, em dúvida?

O que eu gostaria de fazer no restante do texto é mostrar como essas preocupações podem ser satisfeitas. Quero deixar claro o papel dos conceitos de factuality e redutibilidade na discussão realista e, assim, mostrar como podemos progredir na resolução de tal discussão. Deste modo, não é meu objetivo defender a coerência desses conceitos. De fato, para o propósito de dissipar dúvidas metodológicas, é melhor jogar a cautela conceitual aos ventos e adotar modelos ou metáforas que podem nos ajudar a entender como conceitos devem ser empregados. Nem é meu objetivo mostrar como podemos *realmente* resolver questões realistas. Afinal, isso não é algo que normalmente nos consideramos capazes, mesmo nas áreas menos problemáticas da filosofia. Em vez disso, eu gostaria de mostrar como podemos proceder. Nós precisamos saber o que seria necessário para resolver as questões, mesmo que não possamos ter nenhuma garantia de resolvê-las em qualquer caso dado.

A maioria dos anti-quietistas tentou dissipar as visões quietistas ao produzir critérios de factuality ou redutibilidade em termos pelos quais questões de realismo podem, então, ser colocadas. Uma questão foi simplesmente substituída por outra. Nossa estratégia para lidar com o quietista é bem diferente. Nós tentamos ver como as questões do realismo podem se transformar em perguntas mais tratáveis, sem pressupor que elas devam ser tornadas inteligíveis em termos destas outras questões. Não nos comprometemos, assim, com a visão de que os principais conceitos-chave metafísicos *não podem* ser definidos em termos fundamentalmente diferentes, mas, claramente, na medida em que podemos permanecer neutros sobre esta questão, nossa defesa da metafísica realista provavelmente seria muito menos controversa, e diminuiríamos o perigo para o qual toda a filosofia é propensa, isto é, fazer um problema claro apenas por deturpar o que ele é.

5 Fundamento

As questões metafísicas do realismo se ligam a outras questões de natureza menos problemática. Mas quais são essas questões? Sugiro que elas dizem respeito a relações de *fundamentação* e, portanto, antes de continuarmos, vamos tentar explicar o que elas são.

Eu recomendo que uma declaração de fundamentação seja feita na seguinte forma “canônica”:

Sendo o caso que *S* consiste em nada mais do que ser o caso que *T*, *U*, ...

Em que *S*, *T*, *U*,... são sentenças particulares. Como exemplos específicos de tais declarações, temos:

Sendo o caso que o casal Jack e Jill é casado consiste em nada mais do que ser o caso que Jack é casado com Jill.

Sendo o caso que a Grã-Bretanha e Alemanha estavam em guerra em 1940 consiste em nada mais do que ...,

Em que ‘...’ é uma descrição resumida da atividade de guerra de vários indivíduos.

Em tais casos, dizemos que as proposições à direita (*conjuntamente*) *fundamentam* a proposição à esquerda e que cada uma delas *fundamenta parcialmente* essa proposição. Normalmente, suponho que a proposição fundamentada e seus fundamentos são verdadeiros, embora também se possa falar de “fundamento” quando as proposições de fundamento *serviriam de base* à proposição fundamentada, *se* fossem verdadeiras.

A noção de fundamento deve ser distinguida da noção estrita de redução. Uma declaração de redução implica à irrealidade do que é reduzido, mas uma declaração de fundamento, não. Dessa forma, ao dizer que o fato de que *P&Q* se reduz ao fato de que *P* e o fato de *Q*, estamos implicando que o fato conjuntivo é irreal; mas

ao dizer que o fato de que $P \& Q$ está fundamentado, ou consiste no fato de que P e o fato de que Q , não estamos implicando tal coisa. Estamos adotando uma posição metafisicamente neutra de que realmente existem fatos (ou verdades) conjuntivos. Assim, nossa visão é a de que existe um sentido no qual mesmo um realista sobre os fatos conjuntivos pode estar disposto a admitir que o fato de $P \& Q$ consiste no fato de que P e o fato de que Q ; há uma posição aqui que pode ser adotada pelo realista e pelo antirrealista igualmente.²²

A noção de fundamento, como a noção de redução, também deve ser distinguida da análise lógica. De fato, o paradigma da análise lógica (“o americano médio”) não é para nós um caso de fundamento, já que as proposições expressas em ambos os lados da análise são presumivelmente as mesmas e, ainda assim, nenhuma proposição pode ser apropriadamente tomada para fundamentar a si mesma. Para nós, a aparência superficial potencialmente enganadora da gramática é inteiramente irrelevante para as questões de fundamento, uma vez que estamos olhando para as proposições expressas pelas sentenças, e não para as próprias sentenças. Assim, distinguimos entre a questão essencialmente linguística de determinar qual proposição é expressa por uma dada sentença (se, por exemplo, um termo é uma expressão genuinamente referente) e a questão essencialmente metafísica de determinar o que fundamenta o que.

Tomemos a fundamentação como uma relação explicativa: se a verdade de que P está fundamentada em outras verdades, então elas *respondem* por sua verdade; P é o caso *em virtude das* outras verdades serem o caso. Existem, é claro, muitas outras conexões explicativas entre as verdades. Mas a relação de fundamentação distingue-se por ser a conexão mais estreita.

22 Alguns filósofos têm pensado que a sobreveniência é uma contrapartida metafisicamente neutra da noção de redução e, nesse sentido, pelo menos, o que eles têm em mente pode corresponder à nossa noção de fundamento.

Portanto, quando a verdade de *P* explica causalmente a verdade de *Q*, ainda podemos afirmar que a verdade de *Q* consiste em algo mais (ou outra coisa) que a verdade de *P*. Ou, novamente, o fato de alguém ter quebrado uma promessa pode “normativamente” explicar que ele fez algo errado, mas que ainda é compatível com o seu delito, consistindo em algo mais do que ele ter quebrado a promessa. Não há, no entanto, nenhuma conexão explicativa que resida à base das outras formas de explicação. É a forma final de explicação; e é talvez por essa razão que não estamos inclinados a pensar na verdade de uma proposição fundamentada como um *fato adicional* para além de seus fundamentos, embora possa ser distinta de seus fundamentos mesmo que talvez ela mesma possa ser um fato real.²³

Conquanto tenhamos falado da verdade de uma proposição como sendo fundamentada na verdade das outras, isso não é estritamente necessário. Podemos expressar declarações de fundamentos na forma '*S* porque *T*, *U*, ...', contanto que o "porque" seja tomado em um sentido adequadamente forte, e assim evite qualquer referência a proposições, fatos ou conceito de verdade. A palavra “fundamento” seria, na verdade, um operador sentencial, da mesma maneira que “se então” ou “a menos que”.²⁴ Esse ponto é de algum interesse filosófico, pois mostra que não há necessidade de supor que um fundamento seja algum fato ou entidade no mundo ou que a noção de fundamento esteja indissociavelmente conectada com o conceito de verdade. As questões de fundamentos, sobre as quais se colocam as questões realistas, não precisam ser vistas como engajadas nem com o conceito de verdade, nem com a ontologia dos fatos.

23 Eu devo notar que eu não tomo todos os juízos de fundamento para ser *a priori*. Assim, a investigação filosófica da realidade deve basear-se apenas nos juízos que são *a priori* ou que podem receber algum tipo de apoio *a priori*.

24 Eu tenho uma visão semelhante sobre as noções de factualidade e irredutibilidade.

6 Solucionando Questões de Factualidade

Nesta e na próxima seção, examinaremos como as questões do factualismo podem ser resolvidas com base em considerações de fundamentos. Nossa abordagem é um pouco indireta. Começamos por apresentar um argumento abstrato, no sentido de que qualquer desacordo razoável sobre o *status* factual de uma determinada proposição levará a um desacordo sobre o que fundamenta o quê; e então tentamos mostrar como as questões de fundamentos que se seguem podem ser resolvidas. Será visto que estas últimas questões se voltam para saber se o factualista ou o antifactualista é capaz de fornecer o melhor relato de nossa “prática”.

Sugerimos que não pode haver nenhuma garantia conceitual do factualismo ou não-factualismo com base em considerações essencialmente diferentes, e assim, ao tentar traçar as possíveis diferenças entre o factualista e o antifactualista, será necessário em algum ponto apelar para o que é plausível e não para o que é conceitualmente exigido. Vamos, portanto, supor que dois filósofos diferem quanto ao *status* factual de uma determinada proposição verdadeira. Queremos então mostrar que a divergência deles, desde que suas respectivas posições sejam plausíveis, levará à discordância em alguma questão de fundamento.

Podemos ilustrar a ideia por trás de nosso argumento com a proposição de que o aborto é errado. Embora o antifactualista tome essa proposição como não-factual, ele concordará com o factualista na factuality da proposição de que fulano disse que o aborto é errado. Entretanto, enquanto o antifactualista desejará dizer em que consiste a verdade dessa proposição sem fazer qualquer referência à propriedade de ser errado, o factualista sustentará que tal explicação não pode ser dada. Assim, eles diferirão no que pode fundamentar essa nova proposição.

Vamos agora tentar colocar o argumento em forma geral. Ele prosseguirá em etapas; e, em cada etapa, explicitaremos as suposições possivelmente problemáticas que são empregadas. Articular essas suposições ajudará a esclarecer o “espaço dialético”, ou lógico, dentro do qual as questões de factualidade devem ser resolvidas, e uma vez que tenhamos completado a exposição do argumento, tentaremos mostrar que essas suposições são, de fato, defensáveis.

Estágio 1. Tomemos P como uma proposição verdadeira cujo *status* factual o factualista e o antifactualista discordam.²⁵ Digamos que uma proposição verdadeira é *básica* se não estiver fundamentada em outras proposições. Nós agora perguntamos ao factualista: “A proposição dada P é básica?” Se ele disser “sim”, prosseguiremos para o segundo estágio. Se ele disser “não”, perguntamos a ele: “quais proposições básicas fundamentam conjuntamente a proposição P ?” Concedemos que:

(a) qualquer proposição factual verdadeira não-básica é fundamentada por proposições básicas.

Haverá algumas proposições básicas que fundamentam P ; e então suponhamos que sejam $Q, R, S, \dots$

Se o antifactualista nega que essas proposições fundamentaram P , então já temos um desentendimento quanto à fundamentação. Então, suponhamos que ele concorda que elas fundamentam P . Então ele deve pegar uma delas, digamos Q , para não ser factual, já que

(b) nenhuma proposição não-factual tem um fundamento que consiste inteiramente de proposições factuais.

25 Por simplicidade, assumimos que P (e seu sucessor $P +$) são verdadeiros, embora tudo que seja estritamente necessário para o argumento seja que eles sejam possivelmente verdadeiros.

O factualista, por outro lado, levará todas elas a serem factuais, uma vez que

(c) nenhuma proposição factual é parcialmente fundamentada em uma proposição não-factual.

Desse modo, os dois filósofos diferirão quanto ao *status* factual da proposição *Q*; e podemos substituir o presente *Q* pelo *P* anterior e prosseguir para o próximo estágio.

Estágio 2. Nosso factualista e antifactualista diferirão no *status* factual da proposição *P*, que o factualista considera básico. Dado que o antifactualista considera que a proposição *P* não é factual, ele deve reconhecer que ao menos um dos seus constituintes é não-factual, já que²⁶

(d) qualquer proposição não-factual conterá um constituinte não-factual.

Assim, no caso da proposição de que o aborto é errado, o constituinte não factual presumivelmente seria o atributo *errado*.

Tomemos *C*, *D*, ... como os constituintes da proposição que nosso antifactualista considera não-factual. Assim, é concebível que nosso factualista também possa tomar alguns dos constituintes *C*, *D*, ... para não serem factuais, apesar de acreditar que a proposição dada é factual. Não obstante, qualquer posição factualista e antifactualista plausível irá certamente concordar com a questão acerca de quando uma proposição dada *seria* factual se os constituintes *C*, *D*, ... não fossem factuais. Deste modo, mesmo que

26 Intuitivamente, um constituinte não-factual é aquele que pode ser uma fonte de não-factalidade em uma proposição à qual ele pertence. Talvez a factalidade dos constituintes possa ser definida em termos de factalidade proposicional da seguinte maneira: com qualquer constituinte *c* pode-se associar a classe de proposições *Pc* em que *c* tem seu emprego primário. Um constituinte *c* é, então, não factual se qualquer proposta de *Pc* não é factual.

Eu pressupus que as proposições são entidades estruturadas construídas a partir de seus constituintes. Aqueles que não gostam deste pressuposto podem conduzir um argumento paralelo com sentenças e seus termos no lugar de proposições e seus constituintes.

eles possam discordar sobre a factualidade da proposição de que o aborto é errado, eles concordarão que a proposição seria não-factual *se* o atributo *errado* fosse não-factual. Dado que o factualista sustenta que a proposição dada é factual, e desde que ele concorde com o antifactualista de que os constituintes *C*, *D*, ... se não factuais, tornariam a proposição não factual, ele deve tomar um desses constituintes para ser factual.

Estágio 3. Nossos factualista e antifactualista discordam sobre o *status* factual de algum componente, digamos *C*, da proposição dada *P*. Digamos que uma proposição contém *essencialmente* um dado constituinte, que a sua substituição por algum outro constituinte induziria uma mudança no valor de verdade. Assim, Sócrates é um constituinte essencial na proposição de que Sócrates é um filósofo, embora não na proposição de que Sócrates é idêntico a si mesmo. Agora, qualquer visão antifactualista plausível presumivelmente sustentará que

- (e) qualquer constituinte não-factual *C* está essencialmente contido em alguma proposição factual verdadeira *P*₊.

No caso do antifactualista ético, *P*₊ pode ser a proposição de que fulano disse que o aborto é errado (ou atribuiu a *propriedade de ser errado* ao aborto) ou a proposição de que a palavra “errado” se refere à *propriedade de ser errado*. O factualista, além disso, é plausivelmente levado a concordar com o antifactualista nessa questão. Não precisa haver discordância acerca da proposição essencialmente conter o constituinte dado ou sua verdade; e o factualista pareceria ter ainda menos razão do que o antifactualista para tomar a proposição *P*₊ como não factual. De fato, suponhamos que os únicos outros constituintes da proposição, além de *C*, sejam aqueles que ambos concordam serem factuais. Então a proposição *P*₊, para o factualista, conterá apenas constituintes factuais e deve, portanto, ser factual.

Estágio 4. Nossos factualista e antifactualista concordam com a factualidade da proposição P_+ , mas ainda discordam com o *status* factual do seu componente C . Digamos que uma proposição é *imperfeitamente factual* se é factual, mas contém um constituinte não-factual, e que é *perfeitamente factual* se é factual e contém apenas constituintes factuais. Assim, a proposição P_+ , para o antifactualista, é imperfeitamente factual, mas então ele acreditará que tem um fundamento perfeitamente factual, ou seja, que consiste inteiramente em proposições perfeitamente factuais, uma vez que

(f) qualquer proposição verdadeira e imperfeitamente factual tem um fundamento factual

O factualista, por outro lado, é plausivelmente levado a crer que a proposição não tem fundamento em uma daquelas proposições que não envolva o constituinte C , uma vez que

(g) sempre que um constituinte ocorre em uma proposição factual básica verdadeira e também ocorre essencialmente em alguma proposição factual verdadeira, então qualquer fundamento para a última proposição deve conter o constituinte.

Assim sendo, nosso antifactualista manterá que existe algum fundamento – $R, S, T, \dots$ – para P_+ que não envolva o constituinte C , enquanto nosso factualista negará que $R, S, T, \dots$ é um fundamento para P_+ . O desacordo acerca de uma questão de fundamento é assim assegurado.

Vamos agora tentar defender as suposições (a) – (g) das quais o argumento depende. Acerca disso, é importante ter em mente que nosso argumento não é à prova de enganos (e, portanto, tampouco à prova de filósofos). Não tentamos mostrar que quaisquer posições factualistas e antifactualistas numa determinada proposição levarão, apenas por motivos conceituais, a um desacordo sobre alguma questão de fundamento, mas apenas

que quaisquer posições factualistas e antifactualistas *plausíveis* conduzirão a tal desacordo. De fato, se um factualista, digamos, fosse conceitualmente compelido a aceitar alguma declaração de fundamento *S* que o antifactualista não fosse obrigado a aceitar, então não-*S* implicaria que a proposição dada não era factual sem também implicar que era factual e assim nossa posição sobre a “independência” da metafísica realista não poderia mais ser mantida. Logo, há certos pontos no argumento em que devemos apelar para o que é plausível para o factualista ou antifactualista aceitar em um dado caso, e não para o que eles são obrigados a aceitar. Também devo salientar que, mesmo que minha linha específica de argumento falhe, pode haver outras semelhantes que se sucederão.

A suposição (a), de que o não-básico é fundamentado em sua base, é controversa, mas também dispensável. Pois, como ficará claro, a suposição permanecerá como “fundamental” (ou “irredutível”) no lugar de “básica”, mesmo na presença de um regresso infinito de fundamentos.²⁷

A verdade de (b), de que o factual só pode fundamentar o factual, parece clara. Como pode a verdade de uma proposição não-factual consistir inteiramente na verdade das proposições factuais? Isso não seria suficiente para tornar a proposição factual? É claro que isto não é para descartar sentidos de “fundamento” ou “dependência” nos quais o não-factual pode ser fundamentado ou depender do factual. Assim, mesmo um expressivista pode concordar que a verdade de qualquer afirmação moral é “normativamente” fundamentada na verdade de certas afirmações naturais. Entretanto, esse não é o sentido relevante de “fundamento”; não está sendo afirmado que a verdade da afirmação moral *consiste em não mais* do que a verdade das

27 Torna-se então menos claro se o factualista precisa levar o constituinte C a ser “ineliminável” no estágio 4. Mas, como veremos, os únicos casos plausíveis em que é eliminável são aqueles em que ele dá lugar a constituintes que o factualista e o antifactualista reconhecerão como equivalentes, em seu *status* de factualidade, ao constituinte dado.

afirmações naturais. De fato, este último ponto de vista comprometeria alguém a uma forma de naturalismo e, portanto, a uma negação de que as afirmações morais e naturais poderiam diferir em seus *status* factual.²⁸

A verdade de (c), que o não-factual não pode fundamentar parcialmente o factual, também parece clara. Pois como pode a verdade de algo factual parcialmente consistir em algo não-factual? Isso não seria suficiente para tornar a proposição original não factual? Um contraexemplo possível para a conjunção de (b) e (c) é a disjunção P de uma verdade factual P_1 e uma verdade não factual P_2 . Pois dado que P_1 fundamenta P , deve ser factual por (b); e dado que P_2 fundamenta P , deve ser não-factual por (c). No entanto, a nossa opinião em tal caso é de que são as duas proposições P_1 e P_2 juntas que conjuntamente fundamentam P .²⁹

A verdade de (d), que qualquer proposição não-factual contém um constituinte não-factual, é igualmente aparente, visto que, se uma proposição não factual contivesse apenas constituintes factuais, sua não-factuality não teria nenhuma fonte; não haveria nada que pudesse sensatamente ser dito para torná-la não-factual.³⁰

28 Quando P é a disjunção $P_1 \vee P_2$ de uma verdadeira proposição factual P_1 e uma falsa proposição não factual P_2 , será fundamentada apenas por P_1 e nós então gostaríamos de dizer, em conformidade com (b), que P é factual. Assim, a factuality, em nosso entendimento atual, é uma questão contingente. (Em um entendimento alternativo, uma proposição pode ser considerada não factual quando poderia ser não-factual em nosso sentido atual.) Devo acrescentar que não pensamos em imprecisão como uma fonte, por si só, de não-factuality, uma vez que uma proposição vaga pode ainda ser voltada para o mundo real. Assim, casos em que uma verdade vaga se baseia em uma verdade precisa também não são contraexemplos para (b).

29 A verdade de $P \vee \neg P$ quando P é não-factual também não é um contraexemplo para (c), pois $P \vee \neg P$, nesse caso, deve ser considerado não-factual.

30 Considere a proposição de que a última coisa que o Papa disse é verdade e suponha que a última coisa que ele disse foi que o aborto é errado. Então essa proposição não é factual (se a ética não é factual) e por isso somos obrigados, por essa suposição, a tratar algum constituinte da proposição — presumivelmente “coisa” ou “verdade” — como não factual. Isso é estranho, já que significa que a proposição é imperfeitamente factual mesmo quando a última coisa que o papa disse foi factual. Há uma dificuldade relacionada com a tese de que nenhum “deveria” pode ser derivado de um “é”. Pois a partir das suposições de que a última coisa que o Papa disse é verdade e que a última coisa que o Papa disse é que o aborto é errado, podemos concluir que o aborto é errado. Assim, a primeira

A suposição (e), asserindo a existência de uma proposição factual perfeitamente adequada P_+ , é o ponto principal do argumento no qual considerações de plausibilidade entram, já que não há nada que force o antifactualista a reconhecer a factualidade de qualquer proposição dada. Todavia, é certamente muito plausível que ele esteja disposto a reconhecer a factualidade de algumas proposições do tipo requerido. Talvez ele tenha dúvidas sobre atribuições de crença. Ele não pode, então, considerar atribuições semânticas? Talvez ele tenha dúvidas sobre estas. Então ele não pode considerar a questão de quando é *apropriado* fazer tais atribuições? De fato, sugerirei mais tarde que há considerações factuais concernentes a qualquer domínio que esteja quase prestes a surgir, qualquer que seja a visão da própria factualidade do domínio.

Pode-se, é claro, declarar categoricamente todas as proposições como não-factuais, mas essa forma global de antifactualismo não é, no atual cenário dialético, uma opção viável. Visto que, tentando invocar a não-factualidade de um dado conjunto de proposições circunscritas, o antifactualista não deve dar como certo a não-factualidade de outras proposições, assim como ao argumentar pela não-veridicalidade de um dado conjunto de experiências perceptivas, não se deve dar como certo a não-veridicalidade de outras experiências perceptivas. Desse modo, o antifactualista deve admitir — se somente temporariamente, por uma questão de argumento — que proposições fora do conjunto dado são factuais.

Nesse modo de pensar, há uma presunção geral em favor da factualidade e do estabelecimento de todo antifactualismo global, porém, de forma fragmentada e não por uma linha geral de argumentação. Deve-se, sucessivamente, desbastar o aparente edifício da factualidade; e é só então, quando cada parte for

suposição deve ser tomada como um “dever”, embora aparentemente não contenha termos morais. Talvez haja outra maneira de lidar com esses casos.

removida, que uma forma global de antifactualismo pode emergir como uma alternativa viável.

De acordo com a suposição (f), qualquer verdade imperfeitamente factual deve ter uma base perfeitamente factual, isto é, uma que possa ser declarada apenas em termos factuais. Pois considera qualquer verdade contendo um constituinte não-factual. Se alguém perguntar ao antifactualista por que ele considera factual, não obstante o constituinte não-factual, a única resposta completamente satisfatória que ele pode fornecer é que ele tem um fundamento perfeitamente factual. O pensamento metafísico subjacente aqui é a não-essencialidade do não-factual em descrever o factual. Mesmo se o não-factual fosse completamente expurgado do mundo ordinário, ainda poderíamos fornecer um relato completo da realidade factual em termos do que restava; e isso, então, forneceria um fundamento para todas as verdades factuais, formuladas em termos factuais ou não.

É importante, nesta conexão, não se deixar enganar pelo nosso exemplo de alguém que diz que o aborto é errado. Alguém poderia pensar que um antifactualista em ética poderia levar essa proposição a ser ultimamente fundamentada em alguns fatos que relacionam a pessoa ao conceito *errado*, mas, na medida em que isso é plausível, a proposição original deve ser levada a considerar o conceito *errado*, que o antifactualista pode legitimamente considerar como um elemento factual, em vez da própria *propriedade de ser errado*. Para esclarecer qualquer confusão possível sobre este ponto, tome a proposição como a pessoa atribuir a *propriedade de ser errado* ao aborto. Seria então extremamente bizarro para o antifactualista supor que essa proposição foi, em última análise, fundamentada em alguns fatos que relacionam a pessoa a *propriedade de ser errado*. Pois, como poderia haver fatos reais no mundo relacionando a pessoa à *propriedade de ser errado* e, ainda assim, nenhum fato real relacionando a *propriedade de ser errado* com as coisas erradas?

Podemos argumentar pela suposição (g), sobre a ineliminabilidade de constituintes fundamentais, da seguinte maneira: se um dado constituinte *C* ocorre em uma proposição factual verdadeira básica, então ele deve ser um elemento fundamental da realidade, mas, se alguma proposição factual verdadeira essencialmente contém *C*, deve ser verdadeira em virtude de algum aspecto de *C*. Dado que *C* é um elemento fundamental da realidade, este aspecto de *C* não pode ser fundamentado em algo que não envolva *C*.³¹

Um tipo possível de contraexemplo a essa suposição é ilustrado pela proposição de que '5' se refere a 5. Não poderia um realista concernente a números levar sua verdade a consistir simplesmente em '5' sendo o quinto termo de contagem, onde isso era algo que não envolve o número 5? No entanto, uma visão muito mais plausível para o realista é que sua verdade consiste tanto em '5' sendo o quinto termo de contagem e em 5 sendo o quinto número. É a conexão entre os dois sendo o quinto nas respectivas séries, o que ajuda a fundamentar o fato de que um se refere ao outro. Da mesma forma, um realista ético poderia supor que o fato de alguém atribuir a *propriedade de ser errado* ao aborto é de alguma forma fundamentado em seu comportamento. Novamente, o fundamento não é plausivelmente considerado completo até que a conexão entre seu comportamento e o atributo *errado* seja explicitamente dada.³²

Vale a pena observar, em conclusão, o papel crítico desempenhado pela noção de fundamento no argumento acima. Se

31 Uma suposição relacionada (que ele considera uma falácia) foi esboçada por Horwich [1998], p. 21. Isto é que “sempre que um fato tem um certo componente, então o que quer que constitua este fato deve conter o mesmo componente ou alternativamente algo que o constitui.” A suposição (g) diz que o fato constitutivo deve conter o mesmo componente quando esse componente é fundamental, isto é, tal como ocorrer em um fato não constituído. Mas se o fato constitutivo não contém algo que constitua o componente em seu sentido, não há razão para supor que seja fundamental em meu sentido. Assim, aceitar (g) é perfeitamente compatível com a rejeição de sua falácia.

32 Em certas visões deflacionárias, do tipo proposto por Field [2001], capítulos 4 e 5, essas proposições nem seriam tomadas para envolver uma relação entre um termo ou conceito e uma entidade.

tivéssemos usado uma noção explanatória mais fraca, então não haveria razão para supor que os vários princípios dos quais o argumento depende seriam sustentados. Não haveria objeção, por exemplo, a uma proposição não-factual ter um fundamento inteiramente factual, embora normativo. Este ponto é importante para entender como o factualista e o antifactualista devem ser vistos como enfrentando diferentes demandas explicativas. Pois, a menos que essas demandas sejam entendidas em termos da noção metafísica estrita de fundamento, não há nada que impeça os factualistas e antifactualistas de o verem exatamente da mesma maneira.³³

7 Solucionando Questões de Fundamento

Mostramos como elaborar uma “experiência crítica” para testar quando aceitar ou rejeitar a hipótese de que uma dada proposição é factual. A alegação de que a proposição é factual implicará que uma certa proposição relacionada tem um tipo de fundamento, enquanto a afirmação de que não é factual implicará que a proposição não tem tal fundamento. Ao determinar a resposta correta à questão do fundamento, podemos, assim, determinar a resposta correta à questão da factualidade.

Então, como as questões de fundamento são resolvidas? A noção pode não ser conceitualmente problemática da mesma maneira que as noções de factualidade ou redução, pois sua aplicação não tem nenhuma “importação” realista ou antirrealista. Vários dos nossos argumentos anteriores contra uma definição ou garantia de redução aplicam-se igualmente bem à noção de

33 Assim, o argumento de Putnam para o realismo científico (1978, p.100) e o argumento que Harman [1977], capítulo 1, considera contra o realismo moral, ambos se voltam para quando a melhor explicação de algum fenômeno (o sucesso da ciência, nossas respostas morais) envolve ou não uma referência aos fatos que estão em disputa. Mas, na medida em que a noção relevante de explicação não é metafísica, não está claro por que o factualista ou antifactualista deve diferir nessa questão. Eu também duvido, embora seja um assunto separado, que esses argumentos possam plausivelmente ser usados na forma cética do antirrealismo.

fundamento, e na ausência de uma definição ou garantia, podemos ter dúvidas metodológicas semelhantes sobre como a noção deve ser aplicada.

Existem, acredito, duas fontes principais de evidência para fazer juízos de fundamento. A primeira é intuitiva. Parecemos estar de posse de uma grande quantidade de intuições sobre o que fundamenta ou não o quê. Alguns exemplos já foram dados, mas há muitos outros. Em vista disso, o que fundamenta a verdade de uma disjunção é a verdade dos seus disjuntos que são verdadeiros, e o que fundamenta a ocorrência de um evento composto em um determinado momento é a ocorrência de seus eventos componentes. Também temos intuições sobre uma ampla gama de juízos negativos (um pouco a parte das considerações modais). É implausível, por exemplo, que o que fundamenta os fatos sobre o volume sejam fatos sobre densidade e massa, ou que o que fundamenta a verdade de que um dado objeto é vermelho é o fato de ser vermelho ou redondo e o fato de não ser redondo, mesmo embora o primeiro logicamente siga-se dos outros.

A outra principal fonte de evidências é de caráter explicativo. Como já mencionamos, a relação de fundamento é uma forma de explicação; ao fornecer o fundamento de uma dada proposição, a pessoa está explicando, da maneira mais metafisicamente satisfatória, o que é que a torna verdadeira. Assim sendo, um sistema de fundamentos pode ser avaliado, da mesma forma que qualquer outro esquema explicativo, com base em considerações como simplicidade, amplitude, coerência ou não-circularidade. Talvez a virtude mais importante a esse respeito seja a força explicativa, a capacidade de explicar aquilo que necessita de explicação e que, de outro modo, seria deixado sem explicação. E aqui não é simplesmente relevante que se fundamente e, portanto, responda por certas verdades, mas também que, ao fazê-lo, pode-se explicar a presença ou ausência de certa conexão necessária entre as proposições que são fundamentadas.

Consequentemente, questões de fundamento não são simplesmente resolvidas com base em caso a caso, mas sim em quão bem suas respostas se encaixam em um padrão geral de explicação. Nosso experimento crítico poderia parecer ter oferecido a esperança de que uma questão de factualidade pudesse ser decidida com base em uma única questão de fundamento. Contudo, na medida em que nos falta intuição sobre essa questão, qualquer resposta proposta deve ser colocada dentro de um contexto maior de respostas e avaliado com base em considerações em grande medida holísticas.

Porém, qual é esse contexto maior e quais são as considerações pelas quais suas respostas devem ser avaliadas? Visto que a factualidade de uma dada proposição está em dúvida, pensamos que ela contém certos constituintes não-factuais que ocorrem de tal maneira a tornar a proposição resultante não-factual. Vamos chamar a classe de proposições cuja factualidade é semelhantemente duvidosa de domínio *dado*, e os constituintes que se acredita serem os responsáveis pela falta de factualidade, de constituintes *contestados* (ou elementos).

Associado a um dado domínio e uma classe de elementos contestados, haverá outra classe, que chamamos de domínio *estendido*, consistindo de todas aquelas proposições que (essencialmente) contêm os elementos dados, mas que são concordadas como factuais. Já vimos dois tipos de exemplos de tais proposições — a proposição de que fulano disse que o aborto é errado e a proposição de que "errado" se refere a *propriedade de ser errado*. Na medida em que o domínio dado é usado para descrever os "fatos" de uma determinada área, pode-se pensar que o domínio estendido descreve nossa "prática" de lidar com esses fatos. Assim, quando um domínio diz respeito à moralidade, à ciência ou à matemática, digamos, o outro domínio diz respeito à nossa prática moral, científica ou matemática.

Duas partes significativas de nossa prática estão relacionadas à nossa representação e cognição dos fatos dados.³⁴ Assim, no caso da moralidade, o domínio estendido pode incluir proposições representativas no sentido de que acreditamos tal qual em um princípio moral, ou que tenhamos feito tal qual uma reivindicação moral, ou que um termo moral significa e que faz, e pode também incluir proposições cognitivas no sentido de que sabemos ou somos justificados em manter uma determinada crença moral, ou que somos moralmente sensíveis, ou que somos tendenciosos em nossas visões morais. O domínio estendido também pode incluir proposições peculiares à área em questão. Então, pode incluir proposições para o efeito de que fomos motivados por tal e tal crença moral no caso moral, ou proposições envolvendo a aplicação da matemática no caso matemático.

Para certas formas radicais de antifactualismo, muitos desses aspectos de nossa prática podem ser considerados não-factuais. Desse modo, um antifactualista sobre o significado poderia levar todas as proposições relativas ao significado, referência e justificação a serem não-factuais. Entretanto, suspeito que mesmo nesses casos, será possível encontrar aspectos de nossa prática cuja factualidade e partes em disputa possam concordar. Em primeiro lugar, parece-me que o antifactualista — em comum com o factualista — deveria estar disposto a reconhecer que há um *fato* padrão de correção. Naturalmente, o padrão óbvio de correção será não-factual; para a correção do juízo de que o aborto é errado, digamos, simplesmente equivale ao fato de o aborto ser errado — o que, para o antifactualista, é um assunto não-factual. Esse padrão não-factual de correção vive na sombra, por assim dizer, de um

34 Outro aspecto de nossa “prática” diz respeito ao status metafísico das proposições do domínio dado. Assim, mesmo que um antifactualista tenha uma proposição não factual, ele ainda pode considerar que é factual que não é factual, ou que é contingente, ou que está fundamentada em outras proposições. A questão do *status* metafísico dessas afirmações metafísicas, embora de enorme interesse filosófico, normalmente não é tão relevante para as disputas judiciais sobre a factualidade de um dado domínio de primeira ordem.

padrão factual, pois a exatidão de nossos julgamentos deve de alguma forma se envolver com o mundo real; deve haver algo que almejamos na crença e cuja realização é uma questão factual. Assim, para o expressivista, por exemplo, o padrão factual de correção para um juízo pode ser que ele reflita fielmente os compromissos (possivelmente implícitos); enquanto para o formalista matemático, pode ser que o juízo esteja de acordo com as regras do jogo. Assim, mesmo nesses casos, há algo de tipo factual que conta como entender as coisas corretamente.

Da mesma forma, parece haver um senso factual de que alguém pode ser considerado como responsivo aos fatos. Novamente, o sentido mais óbvio de ser responsivo aos fatos pelo antifactualista é aquele que é em si não factual, uma vez que repousará sobre o que leva os fatos a serem. Todavia, à sombra desse sentido, existe outro sentido que é factual. Nossa atividade epistêmica deve de alguma forma se envolver com o mundo real; deve haver algo que almejamos — no objetivo de sermos agentes epistêmicos bem colocados — cuja realização é uma questão factual. Dessa maneira, para o expressivista, isso pode consistir em ser apropriadamente sensível aos compromissos (implícitos) de alguém na formação de suas crenças éticas, enquanto, para o formalista, pode consistir em ser apropriadamente responsável pelas regras do jogo ao executar uma prova.

Admitindo que eles podem concordar com uma prática comum, o factualista e o antifactualista, cada um nos deve uma consideração do que é e daquilo em que consiste. Ou, para ser mais preciso, cada um deve nos fornecer uma descrição do que pode fundamentar as proposições do domínio estendido. Suas considerações estão sujeitas a restrições muito diferentes. O antifactualista deve fornecer uma descrição da prática sem fazer qualquer referência aos constituintes contestados. O expressivista, por exemplo, deve ser capaz de dizer em que consiste uma crença moral sem fazer uso de vocabulário moral, e o formalista deve ser capaz de dizer em que consiste uma prova matemática sem fazer

uso de vocabulário matemático. É essa restrição que explica por que um antifactualista deve ser capaz de fornecer alguma alternativa a uma explicação condicional da verdade sobre nossa compreensão da linguagem, pois a verdade, em sua aplicação às sentenças de um dado domínio não-factual, não é factual e deve, portanto, ser eliminável. Também explica por que as formulações padrão de posições antifactualistas — expressivismo, construtivismo, formalismo etc. — são comumente tomadas como antifactualistas, apesar de não estritamente implicadas pelas formulações delas mesmas, pois eles fornecem os meios gerais pelos quais a restrição pode ser vista como satisfatória. Fica imediatamente claro na posição do expressivista, por exemplo, como ele desejaria dar conta da crença ética sem fazer uso de termos éticos.

O factualista, pelo contrário, é obrigado a fazer referência aos elementos contestados (pelo menos na medida em que são considerados básicos ou fundamentais). O platonista moral, por exemplo, não pode dar conta do que é sinceramente atribuir a *propriedade de ser errado* a um dado ato em termos de ter uma atitude para com esse ato, uma vez que a conexão com o atributo é assim perdida. O platonista aritmético não pode dar conta do que é se referir aos números naturais sem fazer apelo aos números naturais. É também essa restrição que explica por que o factualista pode achar uma conotação condicional da verdade tão agradável, já que conectará nossa compreensão da linguagem da maneira necessária com os elementos da realidade com os quais ela lida.

A consideração do factualista deve, nesse sentido, ser *representacional*: deve ligar a prática com os fatos ou assuntos subjacentes, enquanto a consideração do antifactualista será *não-representacional*. No primeiro caso, a prática deve ser vista como envolvida com os possíveis fatos e deve ser entendida — pelo menos em parte — em termos de *como* ela se conecta com esses fatos. No outro caso, a prática é considerada desconectada dos fatos e, em vez de compreender a prática em termos de como ela

representa os fatos possíveis, os próprios fatos devem ser entendidos em termos de como são “projetados” pela prática. É nesse sentido que eles são subjetivos ou não estão “lá fora”. Uma proposição não-factual é, em última análise, entendida não em termos de seus fundamentos, do que no mundo a torna verdadeira, mas em termos de seu papel dentro de uma dada prática. Eles são proposições metafisicamente incompletas, por assim dizer, e devem ser entendidas da mesma maneira que os “símbolos incompletos” de Russell, por meio do contexto de seu uso, e não por meio de sua aplicação isolada ao mundo.

A questão de ser ou não um factualista é, portanto, a questão de adotar ou não considerações representativas do que fundamenta nossa prática. Esta questão, por sua vez, é em grande parte uma questão de determinar qual das perspectivas rivais é mais capaz de atender às demandas explicativas que podem ser colocadas sobre ela. Pode o factualista moral explicar o papel motivacional da crença moral ou o antifactualista moral explicar seu papel inferencial? O factualista matemático pode explicar a capacidade referencial da linguagem matemática, ou o construtivista, para sua aplicação à ciência? Pode o factualista sobre o significado fornecer uma explicação adequada dos fundamentos das declarações de significado que fazemos, ou o antifactualista uma explicação adequada dos padrões factuais de correção pelos quais eles parecem ser governados? É na sua resposta a estas e muitas outras questões que a exatidão de uma posição factualista ou antifactualista será resolvida.³⁵

Dado que as considerações estão sujeitas a diferentes restrições, o factualista e o antifactualista atenderão a essas demandas de maneiras distintas. Contudo, podemos ter certeza de que essas diferenças nos permitirão julgar entre suas perspectivas?

35 Outros filósofos (principalmente Dummett, *passim* e Blackburn [1984], p. 169 e [1998], p. 50) também enfatizaram o papel da prática na adjudicação de disputas realistas. O que é distintivo em minha opinião é a maneira exata com a qual ele articula o que é uma prática e o que está envolvido em contabilizá-la.

O que é que descarta a possibilidade de um impasse em que as vantagens e desvantagens das duas perspectivas parecem ser mais ou menos iguais? Ou ainda, se uma perspectiva parece preferível a outra, talvez não seja tão plausível a ponto de ser digna de crença. Suponha, por exemplo, que a consideração representacional mais plausível de nossa prática matemática seja epifenomenal: consiste de uma parte não representativa e uma parte matemática “paralela” (por exemplo, a referência de ‘5’ a 5 consistindo em ‘5’ sendo o quinto termo de contagem e 5 o quinto número). É difícil ver como devemos escolher entre essa perspectiva e a correspondente perspectiva não representacional.

Deve-se admitir que não temos base *a priori* para excluir tais casos. Mas também não temos boas razões para esperá-la. Alguns filósofos, é verdade, ficaram impressionados com nosso fracasso repetido em resolver os problemas do realismo no passado e tornaram-se completamente pessimistas em relação à nossa capacidade de fazer algum progresso sobre eles no futuro. Suspeito que esses filósofos não tenham apreciado completamente o quanto precisa ser feito antes que esses problemas possam ser resolvidos adequadamente, pois, ao fornecer uma explicação de uma determinada prática, devemos apresentar o que é, com efeito, uma epistemologia completa, uma psicologia filosófica e uma teoria da linguagem para a área em questão; e, ao avaliar tal posição, devemos, de fato, resolver todos os principais problemas filosóficos que a área dá origem. Até que tenhamos resolvido a questão acerca das crenças morais necessariamente terem força motivacional, por exemplo, não estamos em posição de dizer se é um ponto a favor de uma dada posição quanto à nossa prática moral que as dota com tal força e, até que tenhamos decidido se as crenças matemáticas podem ser conhecidas *a priori*, seremos incapazes de dizer se é um ponto a favor de uma perspectiva de nossa prática matemática que se permita que eles tenham tal *status*. Uma conclusão realista ou antirrealista, portanto, representa o término da investigação filosófica em uma determinada área, em vez de seu

ponto de partida, e assim, não é de se surpreender que tamanho pequeno progresso tenha sido feito dentro da metafísica realista, mesmo em comparação com outros ramos da filosofia.

8 Realidade como fundamental

Distinguimos entre duas concepções da realidade — como factual e como fundamental. Agora nos voltamos para a segunda e, depois de esclarecer o conceito na seção atual, tentamos mostrar na próxima seção como as questões relativas à sua aplicação podem ser resolvidas.

É natural entender o conceito de realidade fundamental em termos do conceito relativo de uma coisa ser *menos fundamental do que*, ou *reduzível* a outra — o ser fundamental, o que não se reduz a nada (mas ao qual outras coisas se reduzirão). Desse modo, parece que estamos nos jogando nas mãos do quietista. Como pode uma conexão explicativa ser determinante do que é e do que não é real? Podemos admitir que algumas coisas são explicitamente mais básicas que outras. Porém, por que isso os tornaria mais reais?

O que eu gostaria de sugerir, diante dessa dificuldade, é que rejeitemos a ideia de que a noção absoluta de realidade fundamental precisa de uma sustentação relacional. A concepção de realidade que estamos buscando é simplesmente a concepção da realidade como ela é em si mesma. Assim, mesmo que duas nações possam estar em guerra, podemos negar que é assim que as coisas são real ou fundamentalmente, porque as entidades em questão, as nações e a relação entre elas, não fazem parte da realidade em si mesma. Pode-se pensar no mundo e nas proposições pelas quais o mundo é descrito, cada um tendo sua própria estrutura intrínseca; e uma proposição então descreverá como as coisas são em si

quando sua estrutura corresponde à estrutura do mundo.³⁶ Isto posto, é essa ideia positiva da estrutura intrínseca da realidade, e não a ideia comparativa de redução, que deve ser tomada para informar a concepção relevante do que é fundamental ou real.

Também é importante distinguir a noção de Realidade em si de certas outras noções de realidade intrínseca. Ao falar da natureza intrínseca do mundo físico, por exemplo, pode-se ter em mente suas características não-relacionais ou não-disposicionais, mas essas características podem não fazer mais parte da Realidade em si mesma, no meu entendimento, do que as características relacionais e disposicionais. Ou ainda, ao falar da natureza intrínseca do mundo físico, pode-se estar atrás de uma descrição que é intrínseca ao mundo no sentido de ser não-perspectiva.³⁷ Logo, termos de cores podem ser excluídos com base em que nossa compreensão deles é baseada em uma forma peculiar de percepção sensorial. Nada obstante, desde que esses termos identifiquem as propriedades físicas fundamentais, eu gostaria de aceitar seu uso na descrição da Realidade em si, independentemente do modo que eles possam ter sido compreendidos.

Dada a noção de realidade como primitiva, é então possível definir a noção de redução. Intuitivamente, uma proposição se reduzirá a outras se elas nos aproximarem do que é real. Portanto, uma condição necessária para a proposição *P* se reduzir às proposições *Q*, *R*, ... é que ela seja fundamentada naquelas outras proposições; e uma condição necessária e suficiente para *Q*, *R*, ... estar mais próximo da realidade do que *P* em tal caso é que *P* seja irreal e cada um de *Q*, *R*, ... seja real ou “a caminho” do que é real. Entretanto, o último é presumivelmente apenas uma questão de a proposição estar fundamentada no que é real. Assim chegamos à seguinte definição:

36 Isto é meramente uma imagem. Ele não precisa se comprometer com a visão de que existem fatos no mundo cuja estrutura pode corresponder à estrutura das proposições ou sentenças pelas quais eles podem ser descritos, como na exposição padrão do atomismo lógico (J. Wisdom [1969]).]

37 Ver Williams [1978], p. 66, e [1985], p. 241.

a proposição verdadeira P se reduz para as proposições $Q, R, \dots$
 se (i) P não é real; (ii) P é fundamentada em $Q, R, \dots$; e
 (iii) cada $Q, R, \dots$ é ou real ou fundamentada no que é real.³⁸

Nessa abordagem, a redução deve ser entendida em termos de realidade fundamental, e não o inverso, e não há mistério sobre por que a redução tem “importação” antirrealista, uma vez que essa “importação” é incorporada à própria noção. Também é capaz de evitar os problemas que afligiam as perspectivas anteriores de redução, uma vez que compartilha, com a noção subjacente de fundamento, a capacidade de desfrutar de casos um-para-muitos, de aplicar-se ao *de re*, e de ter “importação explicativa”.

9 Estabelecendo o que é fundamental

Naturalmente, o quietista extremo não estará feliz com o conceito de Realidade em si mesma. Mas deve ser lembrado que nosso alvo é o quietista moderado. Nós jogamos a cautela conceitual ao vento e nossa única pergunta é: dado que o conceito é inteligível, então como é que sua aplicação deve ser estabelecida? Como devemos determinar, a partir de todas as verdades possíveis, quais são descritivas — ou possivelmente descritivas — de tal realidade, e quais não são?

Ao tentar determinar o que é real dessa maneira, não podemos simplesmente apelar para o fato de que uma determinada proposição é básica, visto que uma proposição básica pode ser não factual e claramente nenhuma verdade não factual é descritiva da realidade fundamental. No entanto, qualquer proposição *factual* básica será real, pois qualquer proposição

³⁸ Pode-se expressar a noção de realidade por meio de um conectivo “é constitutivo da realidade que ...”, assim como no caso das noções de factualidade e fundamento. Um interesse considerável seria então ligado ao desenvolvimento da lógica de tais noções.

factual verdadeira é real ou fundamentada no que é real, e assim a proposição, se básica, será real.

Nossa metodologia anterior para determinar o que é factual pode, portanto, ajudar a determinar o que é real. Pela primeira vez, é resolvido que um dado domínio é factual, então quaisquer proposições básicas desse domínio podem ser consideradas reais. Mesmo sem qualquer ajuda dessa metodologia, talvez haja uma suposição geral a favor de que qualquer proposição dada seja factual, e desse modo, o mero fato de que uma proposição é básica nos dará razão para acreditar que ela é real, na ausência de qualquer razão para o contrário.

Assim, como não podemos distinguir o que é real do que é básico, não podemos distinguir o que é irreal do que não é básico. De fato, é possível imaginar cenários metafísicos nos quais o não-básico, ou fundamentado, é plausivelmente considerado real. Suponhamos, para tomar um tipo de caso, que Aristóteles esteja certo sobre a natureza da água e que ela seja ao mesmo tempo indefinidamente divisível e totalmente cheia de água. Então é plausível que qualquer proposição sobre a localização de um dado corpo de água seja fundamentada em algumas proposições sobre a localização de corpos d'água menores (e em nada mais). A proposição de que esse corpo de água esteja aqui, à minha frente, por exemplo, será fundamentada na proposição de que a metade está aqui, à esquerda, e a outra metade ali, à direita. Mas qual dessas várias proposições que descrevem a localização da água é real? Não podemos dizer que algumas são reais e outras não, já que não há base sobre a qual tal distinção possa ser feita. Assim, devemos dizer que elas são todas reais ou que não são. Dado que a localização da água é uma questão factual, devemos tomar todas elas como reais, não obstante, o fato de que cada uma delas está fundamentada em proposições do mesmo tipo.³⁹ Outro tipo de

39 Ao decifrar este caso, é essencial distinguir entre o fundamento e a redução. Se for tomado um fundamento para aproximar o que é real, então é difícil ver como poderia haver uma regressão infinita de fundamentos (sem nada no fundo). Pois como alguém pode se aproximar do que não

caso envolve considerações “horizontais” em vez de “verticais”. Imagine uma ontologia que leve certos eventos simples e os relacionamentos causais entre eles como reais. Suponha agora que um evento simples causa um composto de eventos simples. Então isto presumivelmente consiste em causar um componente do composto e em causar o outro componente. Agora suponha que um composto de eventos simples cause um evento básico simples, mesmo que nenhum componente do composto cause o evento. Então não está claro o que o fundamento poderia ser. Ainda assim, se esse relacionamento causal é considerado básico e, portanto, real, então os eventos compostos também devem ser considerados reais, e assim a causação do composto — que é uma relação real entre os reais — também deve ser considerada real, apesar de estar fundamentada em outras relações causais.

Então, dado que uma proposição é fundamentada em outras, como devemos verificar se ela é real ou não? O que eu gostaria de sugerir é que existe uma suposição geral a favor do fato do fundamento não ser real. Na ausência de qualquer razão do contrário, como aquelas ilustradas pelos casos acima, devemos supor que qualquer proposição fundamentada dada é irreal.

A suposição pode ser justificada por referência aos objetivos gerais da metafísica realista. A distinção entre o que é e o que não é real representa uma estratégia geral para fazer sentido metafísico do mundo factual. Pois, de toda a estrutura que o mundo exhibe, alguns podem ser considerados reais, pertencendo ao próprio mundo e outros, apenas aparentes e entendidos por referência ao que é real. Vamos chamar uma divisão de todas as proposições àquelas que são reais e àquelas que são irrealis de *visão de mundo*. Assim, uma visão de mundo corresponderá a uma tentativa

existe? Mas, uma vez que os fundamentos são considerados metafisicamente neutros, não há mais dificuldade do que no caso da causa em concebê-los para formar uma regressão infinita. (Cf. a discussão da posição de Leibniz sobre essa questão em Adams [1994], pp. 333-8).

particular de ver o mundo como inteligível em termos da distinção entre o que é e o que não é real.

Vamos agora chamar uma proposição factual de *discutível* se ela estiver fundamentada e se não houver uma razão especial para pensar que é real. Existem então três possíveis visões de mundo que uma pessoa pode adotar: a minimalista, qual toma cada proposição discutível como irreal; o maximalista, que toma cada proposição discutível como real; e a mediana, que toma algumas proposições discutíveis como irreais e outras como reais. (Não precisamos considerar proposições não-factuais, pois é claro que elas são irreais.) Agora destas três alternativas, a terceira deve ser excluída pelo fundamento de que ela coloca uma distinção arbitrária entre proposições discutíveis, levando algumas a serem reais e outras não, quando não há base para o fazer. Assim, não corresponde a uma concepção metafísica inteligível do mundo. A segunda também deve ser excluída. Não é arbitrária em seu tratamento de proposições discutíveis, mas ao tratá-las todas como reais, ela efetivamente abandona a estratégia explanatória para a qual a distinção entre o que é e o que não é real foi planejada. Em termos dessa estratégia, ela é efetivamente equivalente a adotar uma posição que se recusa a reconhecer a distinção entre o que é e o que não é real. Assim, a única alternativa razoável é a primeira, e é isso que então nos justifica a levar toda proposição discutível a ser irreal.

10 A Unidade da Metafísica Realista

Vimos que questões de factualidade e realidade devem ser respondidas essencialmente pelos mesmos meios. Não é meramente que a determinação do que é factual é relevante para a determinação do que é real, mas que, em ambos os casos, as questões são em grande parte resolvidas por considerações de fundamento. Em um dos casos, devemos olhar para as proposições do domínio ampliado para ver se a consideração de seus

fundamentos é melhor dada em termos representacionais ou não-representacionais. No outro caso, devemos olhar para as proposições do domínio dado, assumindo-as como factuais, e tentar averiguar, a partir da estrutura geral de seus fundamentos, como a divisão acerca do que é e o que não é real é melhor efetuada. Assim, uma vez decididas todas as questões de fundamento, todas as questões do que é real — seja no sentido do que é factual ou o que é fundamental — podem ser resolvidas. Subjacente a essa unidade metodológica, está, acredito, uma significativa unidade conceitual. Até agora, tratamos o factual e o fundamental como concepções independentes da realidade metafísica, mas estes estão intimamente relacionados. É claro que qualquer proposição (fundamentalmente) real é factual e que qualquer proposição embasada no factual e, portanto, qualquer proposição fundamentada no real, é factual. Contudo, também se uma proposição é factual, então ela deve ser verdadeira pelo mundo real, e se não for ela mesma real, deve ser fundamentada no real. Portanto, chegamos à seguinte definição:

uma proposição é factual, se é real ou é fundamentada no que é real.⁴⁰

A metafísica realista, nessa visão, tem um único foco — o fundamentalmente real — e nosso interesse por outras categorias de realidade derivará de sua conexão com essa categoria mais fundamental. É o eixo explicativo, por assim dizer, sobre o qual uma consideração sobre mundo girará. Uma dada proposição pode ser idêntica ao real (o real em si) ou ser redutível ao real (o irreal) ou não ser nem idêntica nem redutível ao real (o não-factual ou

40 É possível imaginar uma posição semi-quietista que aceita o conceito de factualidade, mas rejeita o conceito de fundamento. O mundo se dividiria em uma parte objetiva e não-objetiva, sob esse ponto de vista, mas a parte objetiva seria uma massa indiferenciada e não poderia ser levada a possuir qualquer estrutura intrínseca em particular. A definição acima não estaria disponível e o estudo do que é factual teria que proceder de maneira independente a qualquer consideração sobre o que é fundamental.

irreal). Corresponder a cada tipo de proposição será uma perspectiva característica da “importação” metafísica da proposição – de *como* relaciona-se com a realidade. Pode até ser real, no caso em que não há mais nada a ser dito, já que a proposição tem sua importação “na cara”; ou pode ser irreal, caso em que sua “importação metafísica” é dada por seus fundamentos; ou pode ser irreal, no caso em que sua importação metafísica é dada por aquelas proposições factuais que refletem seu uso. O objetivo da metafísica realista é tornar o mundo inteligível em termos da distinção entre o que é e o que não é real, e sua tarefa é completa quando fica claro como o que é aparente, ou irreal, deve ser inteligível em termos do que é real.⁴¹

Referências:

- ADAMS, R. M. **Leibniz: Determinist, Theist, Idealist**. Oxford: Oxford University Press, 1994.
- ARMSTRONG, D. M. **A World of States of Affairs**. Cambridge: Cambridge University Press, 1997.
- BLACKBURN, S. **Spreading the Word**. Oxford: Oxford University Press, 1984.
- BLACKBURN, S. **Essays in Quasi-Realism**. Oxford: Oxford University Press, 1992.
- BLACKBURN, S. **Ruling Passions: A Theory of Practical Reasoning**. Oxford: Clarendon Press, 1998.
- CHALMERS, D. J. **The Conscious Mind**. Oxford: Oxford University Press, 1996.

⁴¹ Eu tenho muita sorte pela assistência que recebi ao escrever este artigo. Participantes de seminários na UCLA e Princeton e de palestras nas Universidades de Columbia, Michigan, Pensilvânia e Rutgers fizeram muitos comentários excelentes; tive conversas valiosas ou correspondências sobre os tópicos do artigo com Rogers Albritton, Paul Boghossian, Ruth Chang, Harry Field, Mark Johnston, David Kaplan, David Lewis, Mark Moyer, Gideon Rosen, Stephen Schiffer e Bartosz Wieckowski; e, acima de tudo, Ronald Dworkin, André Gallois, Gil Harman, Paul Horwich, Joshua Schechter e dois pareceristas do Imprint forneceram-me uma extensa lista de comentários escritos, que foram de grande ajuda na revisão do artigo.

- DUMMETT, M. **Frege: Philosophy of Language**. London: Duckworth, 1973.
 DUMMETT, M. What is a Theory of Meaning? (II). In: MCDOWELL, J.; EVANS, G. (Eds.). . **Truth and Meaning: Essays in Semantics**. Oxford: Oxford University Press, 1976. p. 67–137.
- DUMMETT, M. **Truth and Other Enigmas**. London: Duckworth, 1978.
- DUMMETT, M. **Frege: Philosophy of Mathematics**. London: Duckworth, 1991.
- DUMMETT, M. **The Seas of Language**. Oxford: Clarendon Press, 1993.
- DWORKIN, R. Objectivity and Truth: You'd Better Believe It. **Philosophy and Public Affairs**, v. 25, p. 87–139, 1996.
- EDGINGTON, D. Meaning, Bivalence and Realism. **Proceedings of the Aristotelian Society**, v. 81, p. 153–73, 1980.
- FIELD, H. **Truth and The Absence of Fact**. Oxford: Clarendon Press, 2001.
- FINE, A. The Natural Ontological Attitude,. In: LEPLIN, J. (Ed.). . **Scientific Realism**. Berkeley: University of California Press, 1984. p. 149–77.
- GAIFMAN, H. Ontology and Conceptual Frameworks,. **Erkenntnis**, v. 9, p. 329–353, 1975.
- HARMAN, G. **The Nature of Morality**. Oxford: Clarendon Press, 1977.
- HORWICH, P. **Meaning**. Oxford: Clarendon Press, 1998.
- JACKSON, F. **From Metaphysics to Ethics**. Oxford: Clarendon Press, 1998.
- KIM, J. **Supervenience and Mind**. Cambridge: Cambridge University Press, 1993.
- PUTNAM, H. **Meaning and the Moral Sciences**. London: Routledge & Kegan Paul, 1978.
- PUTNAM, H. **The Many Faces of Realism**. La Salle. Illinois: Open Court, 1987.

RORTY, R. **Philosophy and the Mirror of Nature**. Princeton: Princeton University Press, 1979.

ROSEN, G. Objectivity and Modern Idealism: What is the Question? In: MICHAELIS, M. (Ed.). . **Philosophy in Mind**. Amsterdam: Kluwer, 1994. p. 277-319.

STROUD, B. **The Quest for Reality: Subjectivism and the Metaphysics of Color**. New York: Oxford University Press, 2000.

WILLIAMS, B. **Descartes: The Project of Pure Enquiry**. Harmondsworth: Penguin, 1978.

WILLIAMS, B. **Ethics and the Limits of Philosophy**. Cambridge: Harvard University Press, 1985.

WINKLER, K. Skepticism and Antirealism,. **Mind**, v. 94, p. 36-52, 1985.

WISDOM, J. **Logical Constructions**. New York: Random House, 1969.

WRIGHT, C. **Truth and Objectivity**. Cambridge: Harvard University Press, 1992.

Capítulo 7

Correção dos sentimentos em Hume: Intersubjetividade sem objetividade.

Max Barkhausen¹

Este texto mostra que, a fim de entender o sentimentalismo de Hume e suas implicações metaéticas, devemos nos voltar para a sua explicação da correção dos sentimentos exigidos para o julgamento moral. Começo discutindo uma questão interpretativa que ainda não recebeu a atenção que merece: como devemos entender o padrão de imparcialidade para o qual seu recurso de correção apela? Ao responder esta pergunta, percebemos que Hume endossou não apenas o sentimentalismo causal e a visão de que os juízos morais típicos são formados em resposta a sentimentos morais, mas também tipos de sentimentalismos constitutivos e epistêmicos. Os sentimentos morais constituem correção moral e podem servir como um guia restrito para corrigir juízos morais. Como argumentarei, a visão sentimentalista resultante implica em uma forma de relativismo moral.

1. Introdução

A fama de Hume como filósofo moral deriva, em grande parte, de sua visão sentimentalista de que a moralidade é "mais propriamente sentida do que julgada" (T3.2.2.1; SBN 470). Em

¹ New York University. Tradução: Celina A. Brod

contrapartida, sua visão de que nossos sentimentos morais não devem servir como um guia irrestrito para a verdade moral, e que de fato eles precisam de correção, tem sido apreciada por estudiosos, mas não atingiu o mesmo grau de notoriedade.

O presente texto parte da convicção de que, para entender o sentimentalismo de Hume corretamente, devemos nos voltar para sua explicação da correção dos sentimentos. Começo discutindo uma questão interpretativa e que ainda não recebeu a atenção que merece: como devemos entender o padrão de imparcialidade ao qual sua doutrina de correção apela?

O padrão que está em jogo na correção dos sentimentos está enraizado na função social da moralidade. A grosso modo, este habita na função de realizar a coordenação sobre regras de conduta mutuamente benéficas e assegurar que nós ajamos de acordo com tais regras até mesmo quando elas conflitam com nossas paixões e preferências. Na medida em que a correção é necessária para os julgamentos tanto sobre as virtudes naturais como artificiais, verifica-se que ambas estão intimamente ligadas à função coordenadora da moralidade, embora de maneiras ligeiramente diferentes.

O entendimento da correção dos sentimentos desenvolvidos aqui gera uma interpretação do sentimentalismo de Hume. Como veremos, Hume endossou não apenas o sentimentalismo causal, a visão de que os juízos morais típicos são formados em resposta aos sentimentos morais, mas também tipos constitutivos e epistêmicos de sentimentalismo, segundo os quais nossos sentimentos morais constituem verdade moral e sentimentos morais podem nos servir como uma heurística, no entanto, não como uma autoridade irrestrita para verdade moral.

A interpretação presente serve para esclarecer não apenas o sentimentalismo de Hume, mas também sua visão anti-racionalista de que a razão não pode discernir verdade moral. Hume pode conceber que a razão é capaz de discernir verdade moral, porém, apenas a verdade relativa ao conjunto de disposições emocionais

compartilhadas em respostas às nossas muitas convicções morais formadas.

Por fim, o sentimentalismo de Hume tem uma consequência relativista importante. Na interpretação desenvolvida no presente trabalho, juízos morais corretos tomam como base sentimentos morais, contudo, estes estão sujeitos a um padrão de imparcialidade que é construído em nossos conceitos morais. Já que há mais de um sistema moral compatível com este padrão de imparcialidade, Hume está comprometido com um relativismo limitado, segundo o qual cada um destes sistemas morais são igualmente corretos.

Estes tópicos serão desenvolvidos em ordem linear. A Seção 2 explica a natureza do padrão imparcial em jogo na perspectiva de Hume em relação à correção e uma interpretação de seu sentimentalismo; a seção 3 contém uma interpretação do anti-racionalismo de Hume; a seção 4 discute o desfecho meta-ético de sua interpretação; a seção 5 conclui.

2. A correção de Hume dos sentimentos

2.1 Correção

Hume argumentou que a moralidade é "mais propriamente sentida do que julgada" (T3.2.2.1; SBN 470). Isto é, ele aderiu a uma doutrina que agora é chamada de sentimentalismo e, em alguns aspectos, remonta aos teóricos do senso moral, Hutcheson e Shaftesbury.² Há (ao menos) quatro maneiras distintas de entender esses sentimentos morais.

Sentimentalismo Causal: Juízos morais típicos são formados em resposta a sentimentos morais.

² A versão sentimentalista destes teóricos do senso moral é consideravelmente diferente do sentimentalismo de Hume, claro, especialmente no fato de que Hume era inflexivelmente contrário à visão de Hutcheson de que nosso senso moral não pode ser em termos de nossos sentimentos morais mais familiares.

Sentimentalismo Constitutivo: afirmações morais verdadeiras são verdadeiras em virtude dos sentimentos morais do locutor ou de sua comunidade.

Sentimentalismo epistêmico: os sentimentos morais do locutor ou de sua comunidade *justificam* juízos morais.

Sentimentalismo Semântico: As reivindicações morais típicas *são sobre ou expressam* os sentimentos morais do locutor ou da sua comunidade.

Creio que seja incontroverso que Hume fosse um sentimentalista causal. O filósofo acreditava que, caracteristicamente, nossos sentimentos morais dão origem as nossas crenças morais. Esta dificilmente é uma visão controversa. Mesmo que, em última análise, a moralidade esteja sujeita a discernimento racional, pode-se concordar que os raciocinadores morais comuns normalmente formam crenças morais em resposta aos seus sentimentos.

A maioria concordaria que o sentimentalismo de Hume equivale a algo mais que mero sentimentalismo causal. Há, no entanto, quase nenhum consenso sobre qual das afirmações acima Hume endossaria além do sentimentalismo causal. Embora, eu não acredite que Hume fosse um sentimentalista semântico,³ irei abster-me de discutir questões sobre a visão deste em relação à semântica moral neste texto. Meu foco serão as variedades constitutivas e epistêmicas do sentimentalismo e, abaixo, vou argumentar que Hume, além de ser um sentimentalista causal, era de fato um sentimentalista constitutivo e, em conseqüência, epistêmico, sob uma determinada compreensão desses pontos de vista.

³ Para essa questão, o leitor pode consultar Cohon (2008)

A chave para entender o sentimentalismo de Hume, como veremos, é a sua versão da correção dos sentimentos. Hume reconhecia que os sujeitos morais típicos ligam-se uns aos outros com a expectativa de que seus juízos morais plenos e reflexivos sejam imparciais, ao contrário dos seus sentimentos morais, e de fato, os juízos frequentemente são mais imparciais do que os sentimentos em que os sujeitos baseiam-se.⁴ A seguir, reservo os termos "juízo" e "crença" para os juízos morais completos e reflexivos, em oposição a quaisquer juízos avaliativos que podem ser parte e parcela integrante de nossos sentimentos parciais. Isso pressupõe que, para Hume, há uma distinção entre sentimentos morais, por um lado, e juízos morais propriamente ditos, o que exclui as interpretações não cognitivistas de Hume, segundo as quais nossos julgamentos morais não resultam em crença, e que têm caído em desuso.⁵

Por exemplo, eu me ressinto do ladrão que roubou minha bicicleta mais do que qualquer outro ladrão de bicicleta, mas eu acredito que, todo o mais sendo igual, todos os ladrões de bicicleta merecem a mesma quantidade de ressentimento. Da mesma forma, minha simpatia diz respeito à minha família e aos colegas do meu "círculo estreito", como diria Hume - mais do que estranhos. Ao mesmo tempo, acredito que certos privilégios morais valem para todos os agentes morais igualmente, independentemente do nosso relacionamento e aprovo as mesmas características em meus inimigos como nos membros do meu círculo estreito, mesmo que sejam prejudiciais para mim.

Essa diferença entre os sentimentos morais e juízos morais de todas as coisas consideradas, impõem uma demanda explicativa a sentimentalistas como Hume. Dado que, em qualquer

⁴ Hume fornece uma extensa discussão da parcialidade de nossos sentimentos em T3. 1.1

⁵ Eu concordo com Garret que nossos sentimentos morais e nossos juízos morais considerados deveriam ser entendidos como tipos distintos de avaliações morais. Ver Garret (1997:196). A visão de que Hume era um cognitivista no sentido assumido aqui é agora mainstream. Para uma defesa completa desta visão, ver Cohon (2008)

entendimento do sentimentalismo nossos sentimentos parciais estão intimamente relacionados às nossas crenças morais, o sentimentalista deve explicar como conseguimos alcançar julgamentos morais mais imparciais do que nossos sentimentos.

É por esta razão que Hume apresenta em sua doutrina a correção dos sentimentos morais.⁶ Segundo esta doutrina, nós aprendemos a corrigir a parcialidade dos nossos sentimentos quando formamos crenças morais. O mecanismo psicológico em funcionamento nesta correção, diz Hume, é análogo ao mecanismo pelo qual nós formamos crenças perceptivas estáveis das propriedades físicas dos objetos que percebemos.⁷

No caso perceptivo, a correção para distorções perspectivas e circunstanciais são necessárias: objetos podem parecer pequenos ou grandes para nós, dependendo de sua distância, superfícies elípticas ou circulares dependendo da perspectiva, vermelho ou verde, dependendo da iluminação. No caso moral, os sentimentos morais variam com contiguidade psicológica; aqui, aprendemos a corrigir as maneiras pelas quais nossos sentimentos são influenciados por nossos laços psicológicos em relação aos nossos círculos estreitos.

Uma questão importante diz respeito à natureza do padrão imparcial em jogo na correção de sentimentos: este padrão é uma questão de sentimento? Baseia-se em visões morais independentes? Convenções? Fornecer uma compreensão correta desta norma é uma tarefa crucial e difícil para aqueles que acreditam que Hume defendia o sentimentalismo constitutivo ou epistêmico. Lembre-se que sentimentalismo constitutivo é a visão de que afirmações morais verdadeiras são verdadeiras em virtude de nossos sentimentos morais ou, em outras palavras, que nossos sentimentos morais são aqueles que tornam nossas reivindicações morais verdadeiras.

⁶ Esta doutrina é apresentada em T 3.3.1

⁷ Ver T.3.3.1.16; SBN 582

Mas se aquilo que torna as reivindicações morais verdadeiras são nossos sentimentos, então por que precisamos corrigir nossos sentimentos para então alcançar juízos morais corretos? Lembre-se que o sentimentalismo epistêmico é a visão de que nossos sentimentos morais justificam nossos juízos morais. Se esse for o caso, deve se explicar, da mesma forma, por que, por vezes, deveríamos corrigir distorções em nossa evidência e o que precisamente fornece alguns dos dados proporcionados pelos nossos sentimentos aberrantes.⁸ Enquanto alguém pode, por estas razões, perceber uma tensão entre a versão de Hume da correção dos sentimentos por um lado e sentimentalismo constitutivo e epistêmico por outro, acredito que, de fato, a melhor interpretação da correção de sentimentos favorece uma versão de cada doutrina. A seguir, fornecerei esta interpretação.

2.2 Conceitos Morais

As regras pelas quais corrigimos nossos sentimentos morais estão intimamente ligadas ao nosso uso de linguagem moral e as regras que a governam. Hume enfatiza a importância da linguagem moral (e estética) sempre que volta-se para a correção dos sentimentos e a possibilidade de consenso que tal correção faz surgir. Ele enfatiza que a correção para a tendenciosidade dos sentidos é uma característica comum de todas as línguas, incluindo a linguagem moral.⁹

Correções como essa são comuns para todos os sentidos; na verdade, seria impossível fazer uso da linguagem, ou comunicar nossos sentimentos uns aos outros, se não corrigíssemos a aparência momentânea das coisas, desprezando nossa situação presente. (T.3.3.1.16;SBN 583)

⁸ Cohon (2008: 134 ff) levanta questões similares. Abaixo eu direi mais sobre como minha interpretação se relaciona com a dela.

⁹ Sobre o papel da linguagem moral na construção de consenso, ver Jensen (1977), Garret (1997) e Garret (2015)

Em “Do Padrão de Gosto”, Hume enfatiza, repetidamente, que quando nossos sentimentos diferem, o consenso sobre questões de beleza e deformidade se dá devido às características do nosso discurso. Em particular, Hume alega que os membros de uma comunidade linguística devem concordar em suas aplicações de termos que importam culpa e louvor:

Mas devemos admitir que parte da aparente harmonia na moral pode ser explicada pela própria natureza da linguagem. A palavra virtude, com seu equivalente em toda língua, implica um louvor; vício em culpa: e ninguém, sem a impropriedade mais óbvia e grosseira, poderia reprovar um termo que, na aceitação geral, é entendido em um bom sentido; ou aplicar louvor, onde a expressão requer reprovação

Há certos termos em todas as línguas, que importam a culpa, e outros, o elogio; e todos os homens, que usam a mesma língua, devem concordar em sua aplicação. (EMPL 227-228)

Um termo que é entendido em um bom sentido, para Hume, é aquele que nos compromete por meio de uma exigência linguística¹⁰ a aprovar as coisas a que se refere. Exemplos seriam "bons" ou "bonitos". Dizer que algo é bom ou belo compromete o locutor a certas formas de aprovação. Se no caso ele falhar em aprovar coisas boas ou bonitas o locutor terá cometido um erro lingüístico. Na segunda metade da citação, Hume afirma que é uma pré-condição deste tipo de linguagem que falantes comuns devam concordar nas suas aplicações. Na segunda Investigação, Hume adiciona que conceitos avaliativos que envolvem louvor e culpa estão firmados nos interesses gerais da comunidade:

¹⁰ Garret (2015: Capítulo 5) argumenta que esta é uma das características distintas da função conceitual de termos normativos.

A linguagem geral, portanto, tendo sido formada para uso geral, deve amoldar-se a perspectivas mais gerais e afixar os epítetos de louvor ou censura em conformidade com os sentimentos que brotam dos interesses gerais da comunidade. (EPM 5.42 SBN 228)

A fim de prosseguir, devemos considerar brevemente a explicação de Hume sobre idéias abstratas ou, alternativamente, noções, ou ainda, como poderíamos dizer, conceitos. Como Garrett argumentou em detalhe¹¹, aplicar corretamente tais conceitos e, podemos supor a formação inicial desses conceitos, consiste em quatro etapas.

No caso de conceitos morais (e outros conceitos baseados em sentidos)¹², a fase inicial consiste na ativação repetida de uma resposta, como sentimentos de aprovação e desaprovação no caso moral. Então, na próxima etapa, o indivíduo reconhece as características externas que tendem a dar origem a essas respostas ou, na visão de Hume, as semelhanças nas ideias que dão origem a uma resposta específica. Nesta fase, um raciocinador moral pode ser muito parecido com uma criança que considera o significado de “justo” como sendo mais ou menos sinônimo de “desagradável para mim.”¹³

Os próximos e cruciais passos exigem correção e abstração. Por exemplo, no caso dos juízos perceptivos, chegamos a reconhecer que às vezes a resposta está presente sem o recurso. Alguns objetos parecem vermelhos mesmo não sendo. Outros são vermelhos mesmo que nunca tenham ficado vermelhos para quem os percebe. Correção e abstração são necessárias para gerar o que Don Garrett chamou de o “conjunto de reavivamento” de um conceito: o conjunto de idéias, tanto possíveis quanto reais, a que se aplica, como poderíamos dizer, a sua intenção.

¹¹ Ver Garret (1997: 197 ff.) e para uma explicação detalhada e elaborada, Garret (2015: Capítulos 4-5).

¹² Ver Garret (2015: Capítulo 5) sobre causação e probabilidade

¹³ Cf. Piaget (1965)

Nossa questão, então, diz respeito especificamente ao padrão de correção no caso moral. O que corrigimos, dado que não existe um padrão independente dos nossos sentimentos morais, ou o que classifica alguns dos nossos sentimentos morais como aberrantes e outros como apropriados? Para nos concentrarmos nesta questão, devemos melhor compreender as maneiras pelas quais nossas reações iniciais são distorcidas e precisam de correção. Como Hume enfatizou, uma pré-condição central de nossa capacidade de alcançar o consenso moral é que, em juízos morais, superamos a parcialidade dos sentimentos morais em resposta a maneira pela qual as visões morais são freqüentemente formadas.

Podemos distinguir duas formas pelas quais nossos sentimentos são parciais, ambas discutidas por Hume sob o rótulo de parcialidade. A primeira diz respeito ao escopo estreito de nossos sentimentos morais: eles tendem a se preocupar com o nosso círculo estreito e não generalizar para aqueles que estão fora de tal círculo. Nossas disposições emocionais são limitadas. Não podemos, por exemplo, simpatizar com todos os estranhos ou ficar ressentido com todo malfeitor. Somos ainda mais obviamente incapazes de simpatizar ou nos ressentirmos daqueles que nunca adentraram nossa consciência em absoluto.

O segundo tipo de parcialidade corresponde mais de perto à noção de parcialidade implantada no discurso cotidiano, como quando falamos de um árbitro parcial em um jogo esportivo: nossos sentimentos provavelmente ficarão do nosso lado e de nossos círculos estreitos, assim como um árbitro provavelmente vai ficar do lado da equipe que o pagou, ou, assim como eu provavelmente ressinto-me com mais fervor com o ladrão de bicicletas que roubou minha bicicleta do que com qualquer ladrão de bicicletas. Em geral, eu me ressentirei com os outros por ações que eu, pessoalmente, não me sinto culpado quando submeto os outros aos seus efeitos, e sou mais propenso a simpatizar com a situação do meu círculo estreito do que com a de estranhos.

Hume não formula explicitamente as regras gerais para a correção de preconceito emocional no julgamento moral reflexivo. No entanto, caso ele quisesse dar uma formulação explícita destes, poderia ter nos dito para não atribuímos diferentes direitos, permissões ou obrigações para diferentes agentes e não avaliar seus caracteres diferentemente, a menos que, no fundo¹⁴, nós pudéssemos apontar para uma diferença entre esses agentes ou suas circunstâncias. Ele poderia ter ido além e nos dizer que quando uma diferença moral em direitos, permissões, obrigações, ou caráter é explicada em termos de uma diferença não-moral, a mesma diferença não-moral deve ser correlacionada com a mesma diferença moral ao longo das crenças e afirmações dos raciocinadores (tudo o mais sendo igual). Vamos chamar esses requisitos de *sistematicidade* e *imparcialidade*, respectivamente. Hume poderia facilmente ter introduzido estes como regras gerais de pensamento moral e linguagem explicitamente, apenas como ele explicitamente formulou as regras que devem governar o raciocínio causal. Cada regra é constitutiva do discurso moral e do pensamento, porque isto é uma pré-condição da função coordenadora da linguagem e do pensamento moral.¹⁵ É por isso que Hume diz que os conceitos que implicam o elogio e a culpa estão enraizados nos interesses gerais da comunidade.

Nós podemos explicar o papel que estas restrições desempenham na conversa moral em termos um tanto anacrônicos e permanecer fiel ao espírito de Hume. Cada restrição assegura que, na conversação, nós corrijamos uma de nossas tendenciosidades emocionais. Cada restrição é uma restrição no

¹⁴ O pensamento por de trás desta qualificação é que nós frequentemente explicamos diferenças morais em termos de diferenças morais adicionais, mas que qualquer diferença não moral precisa desencadear em uma diferença não moral.

¹⁵ Um aparte: Hume antecipa aqui o tratamento da superveniência moral defendida por não cognitivistas contemporâneos, como Hare, Gibbard e Blackburn. Veja Gibbard (2003: Capítulo 5), Blackburn (1988b; 1984; 1988a) e Hare (1952: 52). Nestas abordagens contemporâneas, a superveniência é simplesmente um requisito conceitual sobre o pensamento e a linguagem moral, e uma das funções desses requisitos conceituais é promover o consenso e a coordenação.

placar conversacional¹⁶, pois cada uma regula as crenças de que um locutor pode expressar, em conjunto, na conversa. Por exemplo, a crença de Al de que Bob não deve ser promovido não é parcial, individualmente. Em compensação, suas crenças (ou afirmações) de que ele próprio deveria ser promovido e Bob não, quando tomadas em conjunto, seriam parciais caso ele não pudesse explicar por que suas conquistas, mas não as de Bob, justificam a promoção.

Note que a imparcialidade não deve ser confundida com uma demanda por igualdade, para o qual todos devem dar o mesmo peso às preferências de todos os outros ou de acordo com o qual todos contem com o mesmo em todas as circunstâncias¹⁷. Por exemplo, Al não deixa de ser imparcial porque acredita que todos estão permitidos de agir egoisticamente em alguma medida ou que, se houvesse um incêndio na Casa Branca, o presidente deveria ser salvo primeiro. No entanto, ele seria parcial caso pensasse que ele, e a ninguém mais, é permitido agir egoisticamente, no entanto Al se mostrou incapaz de explicar por que, ou se ele pensasse que, caso a Casa Branca estivesse pegando fogo, Barack Obama devesse ser salvo primeiro, sem ser capaz de apontar para alguma propriedade de Obama que explicasse esse privilégio, como por exemplo, ele ser o presidente.

Ao invés de formular a imparcialidade e os requisitos de sistematicidade explicitamente, Hume escolheu uma ênfase diferente quando, ao descrever o ponto de vista geral¹⁸, ele nos deu uma descrição de um juiz cujos julgamentos morais obedeceriam à essas regras precisamente porque suas disposições emocionais

¹⁶ A noção de placar conversacional foi primeira apresentada por Lewis (1979)

¹⁷ O termo parcialidade é frequentemente usado nesta forma na literatura sobre escolha social e versões contratualistas para moralidade como, por exemplo, em Gauthier (1986).

¹⁸ O leitor pode recorrer a Sayre-McCord (1994) para uma discussão aprofundada da psicologia do ponto de vista geral e para Abramson (2001) e as citações nele para discussões sobre o papel da simpatia em particular. Baier (1991) fornece discussões perspicazes de quase todos os aspectos do ponto de vista geral.

seriam favoráveis às suas prescrições. Como é sabido, Hume argumentou que a aprovação moral é, e deve ser, determinada

pelos sentimentos que experimentamos de uma perspectiva intersubjetiva desinteressada. Ele acreditava que, ao assumir este ponto de vista, nós tomamos o que ele chamou de "pesquisa geral"¹⁹. Isso envolve experimentar simpatia de várias perspectivas e não apenas da nossa. Relacionadamente, nossos julgamentos morais são apoiados pelo que Hume chama de “simpatia extensiva”, no *Tratado* e “o princípio da humanidade” na segunda *Investigações*, sendo essa uma capacidade de simpatizar e demonstrar preocupação por qualquer ser humano, o que é facilmente despertado pela imaginação.²⁰

Uma razão para essa diferença de ênfase é uma faceta da psicologia moral de Hume: ele pensava que os raciocinadores morais típicos são de fato capazes de ocupar o ponto de vista geral e que este exercício emocional é o nosso melhor meio de alcançar julgamentos morais imparciais e sistemáticos. Contudo, podemos aceitar a opinião de que sistematicidade e imparcialidade são constitutivas de nossos conceitos morais independentemente da psicologia moral de Hume, e podemos entender o ponto de vista geral da teoria deste, englobando um relato das regras constitutivas do pensamento moral e linguagem, e da psicologia que explica nossa capacidade para estar em conformidade com essas regras. Assim, podemos aceitar a visão de conceitos morais em Hume sem concordar com todas as facetas de sua psicologia moral.

Entretanto, há outra razão filosófica para a ênfase de Hume no ponto de vista geral, que ele mesmo não destaca. A imparcialidade e os constrangimentos de sistematicidade não chegam a garantir o

¹⁹ Ver especialmente T 3.3.3-4

²⁰ Como, por exemplo, nesta passagem do segundo Inquérito: “Virtude, colocada em tal distância, é como uma estrela fixa”... Traga esta virtude mais perto, pelo nosso conhecimento ou conexão com as pessoas, ou mesmo por um considerando eloquente do caso; nossos corações são imediatamente pegos, nossa simpatia avivada, e nossa aprovação fria se converteu nos sentimentos mais calorosos de amizade e respeito” (EPM 5.43; SBN 230).

consenso moral, especialmente para o consenso com aqueles que estão fora dos nossos círculos estreitos. Isto porque a imparcialidade e a sistematicidade não impõem restrições ao que conta como uma diferença não moral relevante. Podemos supor que, dentro de nossos círculos estreitos, sentimentos morais, pelo menos, em certa medida, determinam o que conta como propriedade moralmente relevante: aquelas características que aprovamos e desaprovamos antes da correção iniciar.

Como nossos sentimentos variam com a contiguidade psicológica, não podemos presumir que eles irão determinar de forma semelhante as propriedades moralmente relevantes quando se trata de coordenar com aqueles que estão fora de nossos círculos estreitos. Sistemática e imparcialidade são, por exemplo, perfeitamente compatíveis com a opinião de que aqueles que fazem parte do meu círculo estreito devem gozar de certos privilégios, e que, aqueles que prejudicam os membros do meu círculo estreito são maus, mas não aqueles que prejudicam os membros de outros círculos morais, e assim por diante. No entanto, a razão pela qual realizamos a coordenação com aqueles fora de nossos círculos estreitos tem a ver com a versão de Hume do ponto de vista geral.²¹ Os sentimentos que são distintivos do ponto de vista geral são necessários para tornar certas características, porém não outras, salientes para a generalização.

Um aspecto importante do ponto de vista geral é nossa habilidade de empatizar com outros. É isto que Hume chama de simpatia. A simpatia nos concede o entendimento mútuo das características (não-morais) pelas quais nossos sentimentos morais respondem. Este entendimento mútuo que fornece certas propriedades não-morais salientes quando se trata de formar visões morais imparciais e sistemáticas.

²¹ Meu entendimento do ponto de vista geral deve-se a Mackie (2004: 129 e segs.), Baier (1991: 179 e segs.), Sayre-McCord (1994) e Sayre-McCord (1995). Esses estudiosos rejeitam interpretações objetivistas do ponto de vista geral e cada um fornece uma discussão elucidativa da psicologia especulativa que a discussão de Hume sobre o ponto de vista geral abrange.

Por esta razão, a saliência depende de nossa habilidade de experienciar simpatia²². As propriedades moralmente salientes são, em outras palavras, aquelas para as quais nossos sentimentos morais respondem em nossos respectivos círculos estreitos e estas tornam-se salientes por nossa capacidade de sentir simpatia e nosso conhecimento mútuo dessa habilidade.

Cohon (2008:139ff.) argumentou recentemente que juízos morais imparciais são baseados nos sentimentos que experienciamos enquanto assumimos o ponto de vista geral. Esta interpretação é motivada em parte pela sua tese de que Hume endossou a perspectiva de senso moral, a perspectiva que “nossa consciência básica de vício e virtude é uma apreensão direta via sentimento” (Cohon 2008:103). Embora eu simpatize com muitos aspectos da interpretação de Cohon, não temos que assumir, na verdade, que nosso juízos morais imparciais são apreensões de sentimentos que experienciamos apenas do ponto de vista geral. Pelo contrário: julgamentos morais imparciais procedem de sentimentos morais parciais, exclusivamente. Estritamente falando, o ponto de vista geral é necessário apenas porque a simpatia torna certas propriedades importantes para a generalização entre nós e aqueles fora dos nossos círculos estreitos. Este é um aspecto plausível da visão resultante, pois permite que Hume diga que julgamentos morais imparciais transcendem nossa capacidade de sentimentos pró-sociais em um sentido importante: nossas visões morais envolvem os estranhos longe de nossos círculos estreitos, as futuras gerações e, geralmente, muito mais pessoas pelas quais, provavelmente, poderíamos ter qualquer sentimento.

O raciocínio causal, neste sentido, desempenha uma função na formação de juízos morais comuns, contudo, uma função limitada. Determinar se visões morais de alguém são imparciais e sistemáticas pode exigir juízos inferenciais, especialmente quando,

²² O sentido de saliência no trabalho aqui é a informal que Thomas Schelling explorou pela primeira vez em seu seminal Schelling (1960), embora no contexto de ações salientes em coordenação de jogos.

como é as vezes o caso, nossos sentimentos morais não seguem uma linguagem no processo de correção. Isto é o que Hume tem em mente quando ele diz em T. 3.3.1 que “descobrimos [a razão] não ser se não uma determinação calma e geral das paixões, fundada em uma visão ou reflexão distante.”(T 3.3.1.16) Mas note que, mesmo nestes casos, a razão sozinha é incapaz de discernir os juízos morais em questão. Pois eles ainda são baseados em sentimentos parciais iniciais, assim como nossas crenças perceptivas podem estar baseadas em ilusões perceptivas que nós reconhecemos de alguma forma incorreta.

2.3 Intersubjetividdae sem objetividade

Em minha interpretação acima da doutrina de Hume sobre a correção dos sentimentos, usei livremente palavras como “preconceito” ou “correto”. Esta linguagem espelha as alegações de Hume de que “a experiência logo nos ensina este método de correção dos sentimentos” (T 3.3.1.16 SBN 582) e que “tais correções são comuns no que diz respeito a todos os sentidos” (T 3.3.1.16 SBN 583).

Pode-se pensar que essa retórica exige a noção de um padrão objetivo para o julgamento moral. Para nossos propósitos, podemos dizer que um padrão para julgamento moral é objetivo se for contrafactualmente independente de nossos sentimentos morais, atitudes, crenças e convenções.²³ Talvez a consequência mais chocante da visão de que exista um padrão objetivo para o julgamento moral é que, se nossos sentimentos morais fossem substancialmente diferentes e nossos julgamentos morais diferíssem substancialmente de nossos julgamentos verdadeiros, então nós estaríamos errados.

²³ Isso é freqüentemente chamado de independência mental na meta-ética contemporânea. Muitos pensam em independência mental como a própria marca da objetividade. Veja, por exemplo, Driver(2012).

As definições de virtude e as visões proto-utilitárias de Hume dão origem a pelo menos algumas tentativas de entender Hume em termos objetivistas. Como Garrett²⁴ aponta, Hume define a virtude como "qualidades mentais, úteis ou agradáveis, para a própria pessoa ou para os outros" (EPM 9.1 SBN 268) e, alternativamente, como "qualquer ação ou qualidade mental que comunica ao espectador um sentimento agradável de aprovação" (Anexo 1.10 do EPM SBN 289). Hume também declara sobre a moralidade que "seu único propósito é fazer com que seus devotos e toda a humanidade, em todas as instâncias de sua existência se possível, sejam alegres e felizes". Ele argumenta que "parece ser uma questão de fato, que as circunstâncias de utilidade, em todos os assuntos, são uma fonte de louvor e aprovação" e que a utilidade é "a única fonte dessa alta consideração dada à justiça, fidelidade, honra, lealdade e castidade" (EPM 5.44 SBN 231). Uma maneira de entender essas "definições" da virtude é interpretá-las como um padrão objetivo de correção moral.²⁵

Nesta interpretação, o que torna uma ação certa ou errada ou um caráter bom ou mal, é se eles estão de acordo com um padrão utilitário, independentemente de nossos sentimentos. Se nossos sentimentos fossem radicalmente diferentes, isso de forma alguma ameaçaria sua existência ou minaria sua autoridade; do contrário, julgamentos baseados em sentimentos radicalmente diferentes seriam incorretos se não conseguissem acompanhar o padrão utilitário. Esta compreensão objetivista de Hume é compatível com o sentimentalismo causal e, em certo sentido, com o sentimentalismo

²⁴ Essas definições, observa Garrett, permitem inferir se um ato é virtuoso por meios de raciocínio causal, mas não devem ser entendidos como análises do significado de termos morais ou enunciados; isso, Garrett aponta, seria anacrônico. Veja Garrett (1997: 201). Garrett também não pensa - e isso fica especialmente claro em seu livro recente (veja Garrett 2015) - que o padrão em questão é independente de nossos sentimentos.

²⁵ Como fazem os utilitaristas que acreditam que seguem os passos de Hume. Um famoso exemplo é Bentham (1988: XIV), onde Bentham alega que o próprio Hume demonstrou que "o fundações de todas as virtudes são postas em utilidade." Glossop (1967) argumenta que "as doutrinas essenciais Utilitarismo são declarados com uma clareza e consistência que não podem ser encontradas em nenhum outro escritor do século" (87).

epistêmico, pelo menos na medida em que nossos sentimentos estão de fato em conformidade com o padrão objetivo. Ao mesmo tempo, a força justificatória dos sentimentos morais seria altamente contingente. Claramente, o objetivismo implica na negação do sentimentalismo constitutivo, já que o padrão de correção é contrafactualmente independente de nossos sentimentos morais.

Muitos irão argumentar que o objetivismo não é uma interpretação convincente de Hume. É incompatível com as versões mais interessantes de sentimentalismo pela qual Hume é frequentemente creditado. Além disso, conflita com a afirmação famosa de Hume em T 3.1 de que juízos morais não podem ser derivados da razão. Esta conclusão, assim como o argumento mais famoso a seu favor, está resumida concisamente na seguinte passagem:

Como a moral, portanto, tem uma influência sobre as ações e os afetos, segue-se que não pode ser derivada da razão, porque a razão sozinha, como já provamos, nunca poderia ter tal influência. A moral desperta paixões, e produz ou impede ações. A razão por si só, é inteiramente impotente quanto a esse aspecto. As regras da moral, portanto não são conclusões de nossa razão. (T 3.3.1.10; SBN 457)

Vamos chamar o argumento apresentado aqui de “o argumento de Hume da inercia motivacional da razão” e sua conclusão chamaremos de “o anti-racionalismo de Hume”. Este argumento trata da motivação moral: julgamentos morais vêm junto com disposições emocionais, bem como motivações para agir, todavia nenhuma delas pode ser produto da razão. Portanto, argumenta Hume, os juízos morais não podem ser baseados na razão. Se o objetivismo é verdade, a importância epistemológica do anti-racionalismo de Hume permanece misteriosa. Se existe um padrão objetivo para o julgamento moral correto, parece que o raciocínio causal deveria ter acesso a ele. Na verdade, deveria ser a nossa ferramenta de escolha em discernir e justificar a verdade moral.

Estes podem não ser argumentos decisivos contra uma interpretação objetivista, conquanto, note que podemos evitar estas dificuldades interpretando Hume como um não-objetivista. A interpretação da correção fornecida acima propicia os recursos para fazermos isso. Imparcialidade e sistematicidade não determinam por si se uma visão moral está correta; na verdade, elas apenas restringem pontos de vista morais permissíveis, tendo em conta sentimentos morais como insumos e um conjunto de propriedades salientes não-morais, as quais são determinadas por nossos sentimentos parciais e nossa capacidade para a simpatia. O padrão para correção não é, portanto, independente de nossas disposições emocionais contingentes, nem poderíamos determinar, por meio de inferência, se uma reivindicação moral estaria de acordo com a norma em questão sem tomar nossos sentimentos morais como premissas.

Acontece, então, que a interpretação da correção dos sentimentos de Hume oferecida acima, implica em um tipo de sentimentalismo constitutivo (e, em consequência, epistêmico). Pois, nessa visão, nossos sentimentos morais determinam julgamentos morais corretos em conformidade com a imparcialidade e requisitos de sistematicidade.

Sendo assim, sentimentos morais desempenham um papel justificatório importante, ainda que a relação epistêmica entre os sentimentos morais e julgamentos morais não seja tão simples para fornecer uma imediata inferência de, por exemplo, desaprovação a julgamentos morais negativos, uma vez que julgamentos morais são mediados pelos requisitos discutidos acima e nossa capacidade de simpatia.

Por esse motivo, o padrão de Hume para o julgamento moral correto não é objetivo, contudo, ao assumir que os sentimentos morais de diferentes sujeitos morais em uma comunidade são suficientemente semelhantes, fornece uma base intersubjetiva para

o pensamento moral.²⁶ Esta é uma característica muito atraente da visão de Hume. Como foi observado acima, a função central da linguagem e do pensamento moral, para Hume, é a coordenação social. O que a coordenação social requer é a intersubjetividade, mas não a objetividade. Em outras palavras, o que Hume precisa fornecer é um padrão de correção que se aplique a cada um de nós, dada a nossa composição psicológica compartilhada, convenções e circunstâncias, porém não um padrão que se aplica independentemente dessas questões. Isto, na interpretação dada aqui, é precisamente o que ele fez.

O padrão de correção de Hume não se aplica aos sentimentos sobre os quais nossos julgamentos morais são baseados; isto é, não exige que mudemos estes sentimentos. Em vez disso, impõe restrições às nossas visões morais, dados os sentimentos em que eles são baseados. Essencialmente, isso não quer dizer que nossas visões morais não exigem, de certo modo, que tenhamos em contrapartida certos sentimentos. Na verdade, Hume explicitamente argumenta que certos termos em nossa linguagem nos comprometem com sentimentos como aprovação e culpa quando ele diz que “ninguém, sem a mais óbvia e grosseira impropriedade, poderia declarar reprovação a um termo que, em geral, é entendido em um bom sentido; ou conceder aplausos, onde o idioma exige desaprovação”(EMPL 227-228) e que “toda língua possui um conjunto de palavras que são tomadas no bom sentido, e outras no oposto ...”(EPM1,10; SBN 174). Lembre-se que as palavras que são tomadas no bom sentido, para Hume, são aquelas que, quando usadas em sentenças declarativas, comprometem o falante ou o ouvinte à aprovação. Mas esses sentimentos sociais com os quais nossos julgamentos morais nos comprometem, podemos assumir, são distintos dos sentimentos parciais e não sistemáticos sobre os quais são baseados.

²⁶ Cf. a discussão da intersubjetividade em Cohon (2008: Capítulo 5).

Será útil considerar uma passagem de “Do Padrão do Gosto” e impedir uma possível objeção à visão atual. Nesta passagem, Hume z distinção entre sentimentos saudáveis e deficientes nos quais nossos julgamentos estéticos baseiam-se e, pelo menos, se admitirmos que as opiniões de Hume sobre julgamentos estéticos são amplamente análogas ao seu ponto de vista sobre juízos morais, isso pode parecer contradizer minha interpretação não-objetivista da correção e minha defesa de que, para Hume, os sentimentos sobre os quais nossas opiniões morais baseiam-se não precisam de correção:

Em cada criatura, há um estado sadio e um defeituoso; e o primeiro sozinho se supõe nos proporcionar um verdadeiro padrão de gosto e sentimento. Se, no estado sadio do órgão, houver um todo ou uma uniformidade considerável de sentimento entre os homens, podemos daí derivar uma idéia da beleza perfeita; da mesma maneira que a aparência dos objetos à luz do dia, para o olho de um homem saudável, denomina-se sua verdadeira cor real, ainda que a cor seja apenas um fantasma dos sentidos. (EMPL 233-234)

Hume afirma aqui, com relação ao gosto, que podemos definir um padrão “verdadeiro” de gosto em termos de sentimentos de um sujeito sadio em oposição a um sujeito deficiente. É tentador interpretar “verdadeiro” como objetivo no sentido definido acima. Além disso, é fácil imaginar Hume afirmando que, da mesma forma, podemos definir um padrão de virtude “verdadeiro” em termos das reações de um sujeito sadio. Mas isso não iria contradizer a visão de que não são os sentimentos em si que devem ser corrigidos e, pior, isso não implicaria que existe um padrão para a correção de sentimentos independentes do que sentimos e pensamos?

Acredito que não. Em Hume, a distinção entre disposições emocionais adequadas e deficientes deve ser entendida em termos da função de nossa prática moral: enquanto as disposições emocionais adequadas levam a visões que promovem a função de coordenação da moralidade, as deficientes não o fazem. Se as

emoções são deficientes ou adequadas, são neste aspecto relativas à função da moralidade; no entanto isso não significa que, a menos que nós já queiramos sustentar essa função, somos racionalmente obrigados a desenvolver as disposições emocionais que nos ajudarão a desempenhar tal função ou transmitir os julgamentos que elas sustentam. Isto está, naturalmente, de acordo com a famosa observação de Hume em T 2.3.3 que “não é contrário à razão preferir a destruição do todo mundo a um arranhão em meu dedo”(T 2.3.3.6; SBN 416).

O padrão de correção aqui explicado tem sua base na função social da nossa prática moral. Por esta razão, a correção dos sentimentos é de alguma forma semelhante às virtudes artificiais. Deixe-me explicar. É bem conhecido que Hume concebe as virtudes artificiais em termos de sua função de coordenação social, basicamente: para realizar a coordenação sobre regras de conduta mutuamente benéficas e garantir que ajamos de acordo com tais regras, mesmo quando elas conflitam com nossas paixões, desejos violentos e preferências. Essa compreensão das virtudes artificiais permeia, por exemplo, a discussão de Hume sobre normas de propriedade e dá origem a sua famosa metáfora dos dois homens que estão puxando os remos de um barco para benefício mútuo (T 3.2.2.10; SB 490).

Como vimos, a imparcialidade é crucial para garantir o consenso intersubjetivo em virtudes naturais e artificiais. Então, assim como as virtudes artificiais, as regras da linguagem e do pensamento moral, discutidas acima, são vitais para a função coordenadora.

Nossa prática moral e conformidade são de interesse mútuo. Portanto, nós podemos entender a correção dos sentimentos e as regras que as governam em termos de sua função coordenadora. No entanto, ao contrário das virtudes artificiais, as regras de correção não governam o comportamento social diretamente, mas sim a aplicação correta de conceitos morais de acordo com a função social da moralidade em geral.

Na medida em que a correção se aplica tanto aos juízos relativos às virtudes naturais como às artificiais, conclui-se que, em certa medida, ambas as virtudes estão, em última análise, enraizadas na mesma função social.

Uma questão interessante é saber se a não correção de nossa parcialidade em nossas aplicações de conceitos morais seria irracional.²⁷ Por exemplo, uma pessoa seria irracional se ela tomasse seus sentimentos morais como um valor original e transmitisse julgamentos morais parciais ao seu círculo estreito de uma forma que seria excluída pelo padrão de correção explicado acima?²⁸ Como a racionalidade, para Hume, é uma função dos desejos do agente, isso, claro, depende precisamente da composição motivacional do agente. Agora, como acabamos de ver, a correção é necessária para alcançar o consenso moral e, portanto, para que nós, como um coletivo, decidamos sobre regras de conduta mutuamente benéficas. Fazer parte dos esquemas cooperativos mutuamente benéficos ao qual o uso correto da linguagem moral resulta é indiscutivelmente de nosso interesse e, precisamente por este motivo, para a maioria de nós, nossos desejos tornam o uso correto da linguagem moral racional.

3. Anti -Racionalismo

A interpretação da perspectiva de correção de Hume defendida aqui pode esclarecer seu anti-racionalismo e seu argumento sobre a inércia da razão. Consideremos primeiro a sua afirmação de que as distinções morais não são derivadas da razão. Como esta afirmação deve ser entendida é uma questão de disputa. No que Rachel Cohon chama de “o entendimento tradicional de Hume”, ele argumenta a partir do internalismo de julgamento, a visão de que os julgamentos

²⁷ Agradeço um revisor anônimo por levantar esta questão.

²⁸ Amigos me dizem que “injusto!” Para seus filhos pequenos, é aproximadamente sinônimo de “eu não gosto disso!”

morais necessariamente vêm junto com algum grau de motivação para agir de acordo com eles, e a inércia motivacional da razão contra a visão de que os juízos morais são crenças ou têm condições de verdade. Como consequência, Hume, na interpretação tradicional, é considerado um expressivista antiquado que acredita que os juízos morais não são verdadeiros nem falsos, mas são expressões de sentimentos. Uma grande dificuldade que esta versão enfrenta é parecer estar em desacordo com a correção dos sentimentos pois, se julgamentos morais são meras expressões de emoção, não está claro porque, na formação das crenças morais, devemos corrigir a parcialidade de nossos sentimentos. Outras razões (a meu ver, convincentes) têm sido apresentadas na literatura.²⁹ Se esses argumentos são decisivos ou não, eles fornecem motivos suficientes para interpretar o anti-racionalismo de Hume de uma forma que não impõe a visão tradicional sobre nós.³⁰

Uma compreensão alternativa do anti-racionalismo de Hume que evita a interpretação tradicional torna-se disponível quando levamos em conta sua abordagem funcional à moralidade. Está indicado na seguinte citação da segunda Investigação:

Extingam-se todos os cálidos sentimentos e propensões em favor da virtude, e toda repugnância ou aversão ao vício; tornem-se os homens totalmente indiferentes a essas distinções, e a moralidade não mais será um estudo prático nem terá tendência a regular nossa vida e ações. (EPM 1.8; SBN 172)

Há duas maneiras de entender a afirmação de Hume de que, na ausência de sentimentos, seríamos indiferentes às distinções morais. Uma é que o nosso reconhecimento de distinções morais perderia sua força motivacional, outra, é que seríamos incapazes de reconhecermos as distinções morais. Devemos olhar para isso para entender o argumento de Hume sobre a inércia da razão, pois

²⁹ Ver especialmente Cohon (2008: Capítulo 1).

³⁰ C.f Cohon 2008

é intencionalmente destinado a estabelecer a reivindicação anterior. Na verdade, Hume nega a última afirmação, pelo menos em certa leitura. Há, de fato, um sentido em que apenas a razão – mais especificamente, o raciocínio causal – pode discernir distinções morais:

A única tarefa do raciocínio é discernir em cada um dos casos as circunstâncias que são comuns a essas qualidades; observar as particularidades em que concordam, de um lado, as qualidades estimáveis, e, de outro, as censuráveis; e atingir a partir daí o fundamento da ética, descobrindo os princípios universais dos quais se deriva, em última instância, toda censura ou aprovação. (EPM 1.10; SBN 174)

Fora de contexto, a negação de Hume de que a razão pode discernir a verdade moral e sua ambição de “encontrar esses princípios universais, dos quais toda censura ou aprovação são finalmente derivados” pode parecer intrigante. Pode também parecer intrigante que ele negue que as questões morais são questões de fato, mas ainda assim é óbvio que os princípios morais universais são fatos empíricos que podemos discernir a seguir o método experimental. Mas o que Hume reconhece na citação acima é a possibilidade de verdadeiras generalizações indutivas sobre os juízos morais que os raciocinadores morais concordarão, dadas suas disposições emocionais e seus conhecimentos implícitos das regras do discurso moral e pensamento, e, portanto, ele permite que a razão possa discernir as afirmações morais que seriam aceitáveis para um raciocinador moral ou um grupo de raciocinadores morais.³¹

Esses julgamentos estão corretos em relação às disposições emocionais subjacentes do grupo, contudo, não independentemente dessas disposições.³² Podemos discernir a

³¹ Obrigado a Rachel Cohon pela relevante discussão

³² Em Garrett (2015: Capítulo 4), Garrett faz uma distinção entre definições produtivas e reativas de conceitos morais: podemos definir conceitos morais em termos de nossas reações emocionais (reativas)

intenção de nossos conceitos morais por meio de raciocínio causal que procede de nossas disposições emocionais e nosso conhecimento implícito das regras gerais que regulam a linguagem moral e pensamento. Por exemplo, devido nossa configuração emocional e regras constitutivas do pensamento moral e linguagem vamos, em geral, condenar a crueldade bruta e julgar que crueldade bruta é errada. Alguém familiar com a nossa psicologia, socialização e prática linguística está em posição de acreditar nessa generalização sobre nós e pode, portanto, reconhecer indutivamente que a crueldade bruta está na base comum do "erro". Contudo, o que dá um conceito moral a sua intenção são essas disposições emocionais e regras gerais.

Investigar os "princípios universais" do juízo moral é, de fato, uma das mais importantes preocupações de Hume; é o projeto que ele está envolvido quando ele tenta discernir o que todas as virtudes têm em comum e que, finalmente, o leva às "definições" de virtude e das declarações proto-utilitárias discutidas na Seção 2. Estas, portanto, devem ser entendidas como reivindicações que descrevem, a partir de um alto nível de generalidade e de uma perspectiva que leva em consideração a função coordenadora de nossa prática moral, a extensão de nossos conceitos morais em relação aos nossos sentimentos reais, e não como algo que fornece um padrão objetivo de correção moral independente desses sentimentos. Por conta da abordagem funcional de Hume que ele descarta "as virtudes monásticas"³³ - celibato, jejum, penitência, mortificação, autonegação, humildade, silêncio, solidão - como virtudes. Enquanto alguns de nós podemos estar dispostos a chamá-las de virtudes, estas não são boas para ninguém e, portanto, excluídas da função benéfica da nossa prática moral.

ou, alternativamente, em termos das características que dão origem a estas reações (produtivas). Fazendo uso da distinção útil de Garrett, meu ponto é que, se não fosse pelas reações compartilhadas subjacentes, a possibilidade de dar uma definição produtiva de raciocínio causal seria frustrada.

³³ Ver (E.P.M 9.3) para a discussão de Hume sobre as virtudes monásticas.

Encontrar princípios universais do julgamento moral é, então, uma investigação sobre nossas disposições para julgar, porém, disposições que não estão adequadamente relacionadas à função de nossa prática moral podem ser descartadas como insignificantes quando se trata de determinar a intenção de conceitos morais. Logo, o projeto não é apenas descritivo, todavia, é mais bem compreendido à luz do entendimento de Hume do papel funcional da moralidade e dos conceitos morais. É, por assim dizer, um projeto normativamente carregado nas ciências sociais: discernir os traços de caráter que consideramos virtuosos e explicar como eles sustentam o comportamento cooperativo; e se um traço de caráter for considerado uma virtude sem função, esta virtude é rejeitada como deficiente. Conceitos morais, para Hume, são essencialmente conceitos funcionais.

Qualquer um pode discernir a intenção do vocabulário moral sem incorrer em uma obrigação racional de aplicar o vocabulário a instâncias em consenso. Deixe-me usar uma analogia para ilustrar essa visão: pessoas esclarecidas estavam bastante cientes de como o termo "bruxa" era usado séculos atrás, mas eles não estavam sob nenhuma pressão racional para chamar qualquer pessoa de bruxa (razões prudenciais resultantes de pressão social, não obstante). Tampouco o conhecimento deles sobre o consenso de "bruxa" os motiva a acreditar na feitiçaria ou a matar mulheres inocentes. Da mesma forma, o conhecimento do consenso dos conceitos morais não fornece uma justificativa para nossas crenças morais ou para nossa prática moral.

Em resumo, existem dois argumentos distintos nas passagens muito debatidas sobre a inércia motivacional da razão. Primeiro, embora Hume reconheça que a razão possa produzir crenças corretas sobre as origens das avaliações morais e sobre as bases de nossos conceitos morais, essas crenças não são crenças morais porque não são fundamentadas em sentimentos com força motivacional. A ligação entre julgamento moral e motivação é mais bem explicado em termos da hipótese de que crenças morais são

em grande parte formadas em resposta à sentimentos morais. As crenças morais não são, em geral, produtos de generalização indutiva, mas são, em vez disso, formados em resposta a sentimentos morais, é uma característica da nossa prática moral. Se nossa prática não tivesse esse recurso, não seria uma prática moral, entendida em termos funcionais, dado que não conseguiríamos ser motivados a agir de acordo com nossas crenças morais. Como mencionado acima, deveria levar-se em conta que o raciocínio e a indução abduativos desempenham algum papel, ainda que limitado, na formação de juízos morais, pois essas ferramentas da razão podem ser necessárias para assegurar que crenças morais estejam de acordo com tais regras constitutivas da linguagem moral e pensamento, como imparcialidade e sistematicidade.

Em segundo lugar, o raciocínio causal só pode discernir quais juízos morais estão corretos em relação a um núcleo compartilhado de sentimentos morais, porque Hume endossa um sentimentalismo constitutivo. Em outras palavras, se não fosse pelos sentimentos morais, não haveria nada que a razão pudesse discernir. Isso, creio eu, é o que Hume quer dizer quando ele afirma que "as regras da moralidade ... não são conclusões da razão" (T 3.3.1.10; SBN 457), pois os sentimentos entram na explicação de Hume sobre nossas disposições para aplicar juízos morais e, mesmo que a razão possa discernir correções que nossos sentimentos exigem, o padrão de correção em questão somente aplica-se aos raciocinadores morais que chegam à mesa do debate moral com um conjunto de avaliações morais, emocionais, intuitivas e parciais.

4. Relativismo Restrito

Conceitos morais, para Hume, são conceitos funcionais que tem o propósito de coordenação mutuamente benéfico das regras de conduta. Como a coordenação funciona na visão de Hume? Basicamente, nossos sentimentos morais nos levam a adotar algumas

regras morais que prescrevem um comportamento benéfico para aqueles em nossos círculos estreitos. Assim, um grau de coordenação com as pessoas mais próximas a nós é realizado sem muito esforço reflexivo, simplesmente em virtude das emoções que experimentamos em relação a elas. A coordenação com aqueles que estão fora de nossos círculos estreitos é realizada por meio do cumprimento de outras regras que são constitutivas da linguagem moral, regras que nos obrigam a corrigir a parcialidade e falta de sistematicidade em nossos sentimentos morais. O padrão de julgamento moral correto, como vimos, é constituído em parte por nossos sentimentos morais e em parte por regras gerais: a sistematicidade e os requisitos de imparcialidade.

Uma implicação importante da visão metaética resultante é uma negação de objetividade moral: o que é certo ou errado, bom ou mau, e assim por diante, depende das nossas disposições emocionais e convenções linguísticas. Se ambos fossem diferentes, diferentes julgamentos morais seriam corretos. De qualquer forma, mesmo se mantivermos nossas disposições emocionais fixas, é muito plausível que existam inúmeras maneiras de coordenar sobre conjuntos de regras mutuamente benéficas, de acordo com os requisitos de sistematicidade e imparcialidade. As regras morais sobre as quais tenhamos de fato convergido, isola apenas uma entre muitas maneiras em que poderíamos ter coordenado sobre regras de conduta mutuamente benéficas, dada às disposições emocionais em que nossos julgamentos morais refletidos são baseados. Outra maneira de dizer isso, é que existem diversos conjuntos determinados de visões morais que poderíamos ter adotado e que teriam servido à função coordenadora da nossa moral prática igualmente bem, dadas essas disposições.

Um exemplo ajudará a ilustrar isso. O jogo falcão-pombo é muitas vezes entendido como uma ilustração da necessidade de coordenação social em relação à propriedade. Neste jogo, duas partes competem por um recurso e cada uma tem duas estratégias disponíveis: falcão e pombo. Os falcões estão dispostos a lutar,

enquanto as pombas são concessivas. Quando um falcão encontra um pombo, o pombo vai cedendo e o falcão leva tudo. Quando dois pombos se encontram, eles compartilham o recurso igualmente. Quando dois falcões se encontram, eles irão lutar e cada um incorre em um custo e fazem pior do que dois pombos dispostos a compartilhar. Imagine agora que o jogo representa a competição por território, e imagine aí que os jogadores podem ocupar o papel de proprietário ou intruso. Existem agora estratégias mais complexas, como a "estratégia burguesa", jogar falcão se for proprietário ou pombo se for intruso, e a "estratégia de Robin Hood", jogar pombo se proprietário, falcão como intruso.

Pode-se mostrar que cada uma dessas estratégias é socialmente ótima - uma população cujo todos os membros jogam uma dessas estratégias é melhor do que qualquer outra população. Podemos pensar em uma população jogando a estratégia burguesa como uma população que implementou normas de propriedade bastante próximas das nossas. O ponto importante é que a mesma população poderia ter se coordenado a partir da estratégia de Robin-Hood, sem perdas. Podemos imaginar que cada alternativa é compatível com os sentimentos morais das respectivas populações - a comunidade de Burgueses despreza Robin Hoods e os Robin Hoods desprezam os Burgueses. Além disso, cada norma é obviamente compatível com sistematicidade e imparcialidade: diferenças no comportamento esperado são explicadas em termos das propriedades moralmente relevantes de ser um proprietário ou um intruso.

5. Conclusão

Comecei este trabalho com duas promessas: a primeira, de que entendendo que a ideia de Hume da correção dos sentimentos morais nos julgamentos morais é fundamental para a compreensão do seu sentimentalismo e seu anti-racionalismo; a

segunda, de que tal compreensão renderia uma consequência metaética importante. Essas promessas foram agora cumpridas.

Ao meu entendimento do padrão de imparcialidade envolvido na correção dos sentimentos, a imparcialidade é mais bem entendida em termos de regras gerais que dizem respeito à correta aplicação de conceitos morais: um requerimento para a imparcialidade e um requerimento para a sistematicidade. Estes requisitos, lembrem-se, dizem respeito às nossas crenças morais tomadas em conjunto, não individualmente. Segue-se que os juízos morais ainda são verdadeiros ou falsos em virtude de nossos sentimentos morais, mesmo que estejam sujeitos a outras restrições. Além disso, segue-se que nossos sentimentos morais podem servir como um guia imperfeito da verdade moral. Isso vem a afirmar marcas de sentimentalismo constitutivo e epistêmico, respectivamente.

O argumento de Hume da inércia da razão para o anti-racionalismo consiste, como nós vimos, de duas linhas separadas de pensamento. Primeiro Hume argumenta que, ainda que o raciocínio causal possa, em princípio, discernir juízos morais corretos, um núcleo compartilhado de sentimentos morais, o sentimentalismo causal, é essencialmente verdadeiro de julgamento moral, porque garante que estejamos em grande parte motivados para agir de acordo com nossos julgamentos morais. Em segundo lugar, ele argumenta que o raciocínio causal pode discernir a verdade moral apenas em relação a um núcleo compartilhado de convicções morais, precisamente porque a verdade moral é em parte constituída por esse núcleo. Um importante resultado metaético do sentimentalismo constitutivo de Hume e seu padrão de correção, é uma forma de relativismo moral: não podemos legislar entre diferentes sistemas de crença moral que estejam baseados nos sentimentos compartilhados dentro de uma comunidade e que se encontram em conformidade com o padrão de imparcialidade.

Agradecimentos

Sou grato a Hsue Qu e Dan Waxman, bem como dois revisores anônimos desta revista por fornecer feedback detalhado e muito útil sobre rascunhos anteriores deste texto. Apresentei versões anteriores deste texto duas vezes, uma vez na conferência sobre Sentimento e Razão na Universidade de Buffalo e uma vez em "Kolloquium für philosische Philosophie" de Dominik Perler na Humboldt Universität zu Berlin, e agradeço o feedback inestimável que recebi sobre ambas as ocasiões. Beneficiei-me bastante das discussões com Christian Barth, Sebastian Bender, Rachel Cohon, Don Garrett, Max Hayward, Jennifer Marusic, Dominik Perler, Stephan Schmid, Geoff Sayre-McCord e Kelley Schiffman.

Abreviações

T = David Hume (1978). *A Treatise of Human Nature* (2nd ed.). L. A. Selby-Bigge and P. H. Nidditch (Eds.). Clarendon. Page numbers indicated by "SBN".

EHU = David Hume (1975). *An Enquiry concerning Human Understanding*. In *Enquiries concerning Human Understanding and concerning the Principles of Morals* (3rd ed.). L. A. Selby-Bigge and P. H. Nidditch (Eds.). Clarendon. Page numbers indicated by "SBN".

EPM = David Hume (1975). *An Enquiry concerning the Principles of Morals*. In *Enquiries concerning Human Understanding and concerning the Principles of Morals* (3rd ed.). L. A. Selby-Bigge and P. H. Nidditch (Eds.). Clarendon. Page numbers indicated by "SBN".

EMPL = David, Hume (1874-5). *Of a Standard of Taste*. In T. H. Green and T. H. Grose (Eds.), *The Philosophical Works of David Hume* (Vols. 1-4). Longman, Green.

Referências

ABRAMSON, K. Sympathy and the Project of Hume's Second Enquiry. *Archiv für Geschichte der Philosophie*, v. 83, n. 1, p. 45-80, 2001.

BAIER, A. *A Progress of Sentiments: Reflections on Hume's Treatise*. Harvard: Harvard University Press, 1991.

- BENTHAM, J. **A Fragment on Government**. Cambridge: Cambridge University Press, 1988.
- BLACKBURN, S. Spreading the Word: Groundings. In: **The Philosophy of Language**. Oxford: Oxford University Press, 1984.
- BLACKBURN, S. Supervenience Revisited. In: SAYRE-MCCORD, G. (Ed.). . **Essays on Moral Realism**. Cornell: Cornell University Press, 1988a. p. 59–75.
- BLACKBURN, S. How to Be an Ethical Antirealist. **Midwest Studies in Philosophy**, v. 12, n. 1, p. 361–375, 1988b.
- COHON, R. **Hume's Morality: Feeling and Fabrication**. Oxford: Oxford University Press, 2008.
- DRIVER, J. What the Objective Standard is Good for. In: TIMMONS, M. (Ed.). . **Oxford Studies in Normative Ethics**. Oxford: Oxford University Press, 2012. v. 2p. 28–44.
- GARRETT, D. **Cognition and Commitment in Hume's Philosophy**. Oxford: Oxford University Press, 1997.
- GARRETT, D. **Hume**. Abingdon-on-Thames: Routledge, 2015.
- GAUTHIER, D. **Morals by Agreement**. Oxford: Oxford University Press, 1986.
- GIBBARD, A. **Thinking How to Live**. Harvard: Harvard University Press, 2003.
- GLOSSOP, R. J. The Nature of Hume's Ethics. **Philosophy and Phenomenological Research**, v. 27, n. 4, p. 527–536, 1967.
- HARE, R. M. **The Language of Morals**. Oxford: Clarendon Press, 1952.
- JENSEN, H. Hume on Moral Agreement. **Mind**, v. 86, n. 344, p. 497–513, 1977.
- LEWIS, D. Scorekeeping in a Language Game. **Journal of Philosophical Logic**, v. 8, n. 1, p. 339–359, 1979.
- MACKIE, J. L. **Hume's Moral Theory**. Abingdon-on-Thames: Routledge, 2004.

PIAGET, J. **The Moral Judgement of the Child**. New York: Free Press, 1965.

SAYRE-MCCORD, G. On Why Hume's General Point of View Isn't Ideal—and Shouldn't Be. **Social Philosophy and Policy**, v. 11, n. 1, p. 202–28, 1994.

SAYRE-MCCORD, G. Hume and the Bauhaus Theory of Ethics. **Midwest Studies in Philosophy**, v. 20, n. 1, p. 280–298, 1995.

SCHELLING, T. C. **The Strategy of Conflict**. Harvard: Harvard University Press, 1960.

Capítulo 8

Valor e implicatura

Stephen Finlay¹

Asserções morais caracteristicamente expressam atitudes conativas como, por exemplo, aprovação. A língua pode expressar de várias formas, no entanto, e acomodar essa função expressiva em uma teoria de semântica moral satisfatória tem se provado difícil. A abordagem mais simples a esse problema é o subjetivismo: asserções morais são descrições das atitudes do falante. O subjetivismo implausivelmente faz com que condições de verdade de asserções morais dependam da atitude do falante e tem notória dificuldade em acomodar desacordos morais. Uma melhoria, portanto, é o expressivismo: asserções morais têm uma função semântica de expressar atitudes conativas sem condições de verdade. O expressivismo tem seus próprios problemas – particularmente justificar fortes evidências que asserções morais têm valores de verdade – e as soluções em oferta não tem sido persuasivas para a maior parte dos filósofos. Este trabalho investiga uma visão possivelmente distinta, a teoria da *implicatura*, inspirada pelo trabalho de Paul Grice na filosofia da linguagem.

A noção de implicatura é a de que a informação ou *conteúdo* que é comunicado pelo ato de fala para além do conteúdo que determina a condição de verdade da declaração. O que faz da teoria

¹Stephen Finlay é professor assistente de Filosofia na Universidade do Sul da Califórnia. Tradução: Jade Arbo - Revisão: Luca Igansi.

da implicatura tão promissora é que ela acomoda todas as três seguintes intuições: (1) asserções morais têm valor de verdade; (2) elas caracteristicamente expressam atitudes dos falantes; (3) as atitudes do falante não entram nas condições de verdade de suas declarações morais. A teoria da implicatura tem significância ambígua para o expressivismo, no entanto, por comumente se pensar a implicatura como tendo duas variações: implicatura *convencional*, que é realizada semanticamente (através de convenções de significado socialmente infundido nas próprias palavras), e implicatura *conversacional*, que, por sua vez, é realizada pragmaticamente (através da forma e contexto da declaração – nesse caso, através de regras de etiqueta conversacional)². Teorias de implicatura convencional de conteúdo atitudinal de asserções morais (daqui em diante “teorias de convenção”) têm recentemente sido apresentadas por Stephen J. Barker (2000) e David Copp (2001).

Algumas teorias de convenção são possivelmente compatíveis com o expressivismo, e podem fornecer soluções para muitos de seus problemas: Barker oferece seu relato como uma defesa do expressivismo contra a acusação de Frank Jackson e Phillip Pettit de que o expressivismo não pode ser mantido como uma opção distinta do subjetivismo, e Copp classifica seu posicionamento como “realista-expressivista”³. Se a visão correta é, ao invés disso, uma teoria da implicatura conversacional (doravante “teoria da conversação”), no entanto, o expressivismo é falso: em teorias conversacionais, embora asserções morais

² A etiqueta conversacional é resumida no *princípio cooperativo* de Grice: “faça sua contribuição conversacional da forma requerida, no momento em que ocorre, pelo propósito aceito ou direção da troca conversacional da qual você participa” (1989:26).

³ Classicamente o expressivismo é visto como consistindo de duas alegações centrais, a outra (além de que juízos morais são expressos sem reportarem atitudes) sendo que juízos morais não têm condições de verdade. Essas alegações são separáveis, no entanto, e eu seguirei Copp ao usar “expressivismo” para possibilitar tanto a versão antirrealista quanto a realista.

caracteristicamente expressem a atitude do falante, isso não está entre as suas funções semânticas⁴.

Neste trabalho, eu argumento que teorias da conversação – e uma em particular (a qual eu proponho em Finlay 2004) – têm certa vantagem teórica sobre teorias da convenção e, por extensão, sobre qualquer outra teoria que explique de forma similar conteúdo atitudinal a partir de convenções linguísticas. Na seção 1, examino a teoria *indexical* de Baker (e, como argumentarei, também de Jamie Dreier) do conteúdo semântico comum que determina as condições de verdade de uma declaração moral (doravante “conteúdo asserido”), e na seção 2 observo uma dificuldade técnica para a alegação de Baker de que isso apoia uma teoria da convenção. Eu então argumento (seção 3) que a teoria *indexical* é implausível, e que a teoria *relacional* preferida por Copp e por mim dá conta de forma superior do conteúdo asserido de declarações morais. A seção 4 argumenta que o caso de Copp para uma teoria da convenção é enfraquecida pelo fato de que essa teoria relacional tem os recursos para oferecer conteúdo atitudinal como uma implicatura conversacional, tornando a suposta implicatura conversacional redundante. Além disso (argumento na seção 5), conteúdo atitudinal mostra as características típicas de implicatura conversacional e não de implicatura convencional – particularmente “cancelabilidade”. Concluo meu caso na seção 6, observando que teorias da conversação apelam para o fenômeno incontestado da implicatura conversacional, ao mesmo tempo em que há razões significativas para duvidar da existência de implicaturas baseadas em convenção, e razões particulares para ceticismo quanto às implicaturas alegadas pela teoria da convenção.

⁴ Copp (1995) sugere que uma teoria da conversação pode ser a visão correta mas não oferece essa teoria. Veja também Putnam 1981: 209-10, MacIntyre 1984:13.

1. Teoria indexical de Barker

Qualquer teoria da implicatura de atos de fala morais deve também permitir um conteúdo asserido (ou “explicatura”, na terminologia preferida por Baker)⁵: implicatura é conteúdo parasítico do que é dito juntamente com tanto aspectos pragmáticos quanto semânticos dos atos de fala. Baker, portanto, oferece uma análise de “conteúdo dual”, como se segue:

Se *U* assere a frase “*T* é bom”, então *U* denota uma propriedade *F* através de “bom” e:

- (i) *U* expressa-como-explicatura o conteúdo que *T* é *F*;
- (ii) *U* expressa-como-implicatura o conteúdo que *U* está empenhado na aprovação de coisas-*F*;
- (iii) *U* demonstra que crê no conteúdo em (i) e (ii);
- (iv) *U* demonstra que aprova *T*.

No que se segue, eu me concentro em (i) e (ii), os quais Baker identifica como os componentes do ato locucionário, ao invés de (iii) e (iv), identificados como ilocucionário. Sérios problemas surgem para Baker desse relato do conteúdo asserido.

O que é a propriedade *F*? Barker reconhece que “há complexidades nas dinâmicas gerais da fixação de conteúdo” (2000:277). Um século de tentativas falhas atestam para a dificuldade de identificar uma única propriedade como denotação de “bom”, e assim Barker (como Stevenson, Hare e Edwards) sugerem a identidade de *F* varia de declaração para declaração. “Não há limitação para o que *F* é”, ele diz “além de estar ajustado à perspectiva moral de alguém e ser uma propriedade natural (2000:272).

⁵ Evito essa terminologia no escrúpulo, devido de Kent Bach, de que nem tudo o que o falante assere é explícito.

De que forma *F* é fixado? Não pode ser apenas através do significado convencional de “bom”, pois isso permite praticamente qualquer conteúdo que seja. Isso não é dizer que o significado convencional de “bom” não tem papel algum em fixar sua denotação: o uso do pronome “eu” pode denotar qualquer pessoa, pois a denotação é fixada por um significado convencional que dirige uma audiência a olhar para a identidade do falante. Aqui o significado se combina com o contexto para produzir uma denotação. E quanto a “bom”? Barker dá o seguinte critério:

“bom” – como declarado por *U* – denota *F* de forma que

- (a) *U* está comprometido com a aprovação (moral) de coisas-*F*
- (b) Essa atitude-*F* é partilhada por ou incontroversa para a audiência/intérprete.

Como já foi dito, no entanto, esse critério pode ou subdeterminar ou sobredeterminar a identidade de *F*. O problema da subdeterminação é menos sério: é simplesmente o de que pode haver – de fato, normalmente há – mais do que uma propriedade que satisfaz *F*(a) e *F*(b). A solução é certamente ampliar o *input* de contexto. A propriedade de ser bom é determinada, ao menos em parte, pelo tipo de objeto em consideração (como pode ser visto ao nos atentarmos à observação de Peter Geach do caráter atributivo de “bom”, e ao argumento *ergon* aristotélico). Eu introduzo a noção de uma *situação avaliativa S* incorporando a circunstâncias nas quais a asserção avaliativa é feita. Nós podemos então dizer que “bom” – como declarado por *U* em *S* – denota *F* de forma que:

- (a') *U* está comprometido com a aprovação (moral) de coisas-*F* em *S*.
- (b') Essa atitude-*F* em *S* é partilhada por ou incontroversa para a audiência/intérprete.

Vamos então compreender $F(a)$ como $F(a')$, e $F(b)$ como $F(b')$. Embora ainda possa parecer razoável duvidar que o problema da subdeterminação esteja resolvido, eu me volto agora para o problema da superdeterminação. Os critérios $F(a)$ e $F(b)$ podem produzir resultados compatíveis: um falante pode estar comprometido à aprovação de alguma propriedade sem que essa aprovação seja partilhada por ou incontroversa para a audiência/intérprete, e vice-versa. Se lêssemos $F(a)$ e $F(b)$ como condições juntamente necessárias e suficientes para fixar F , nós teríamos que concluir que quando o critério não for mutuamente satisfeito, “bom” não pode ter uma denotação. No entanto, essa não seria uma teoria plausível de “bom”, e certamente não seria uma teoria amigável ao expressivismo, que toma desacordos morais como a raiz dos desacordos de *atitude*. De qualquer forma, não é o que Barker pretende: ele alega que o resultado é “uma certa instabilidade interpretativa ao invés de uma falha de determinada interpretação” (2000: 277). Mas há um enigma quanto ao status do critério.

Talvez $F(a)$ e $F(b)$ devam ser disjuntivos, critérios individualmente suficientes para fixar F . Por “bom” um falante pode então denotar seja as propriedades as quais ele aprova em S ou aquelas as quais sua audiência aprova em S , como Barker de fato parece sugerir (2000: 277-8). A “instabilidade interpretativa” iria então surgir desses dois usos possíveis de “bom”. Mas a interpretação disjuntiva não pode se sustentar. Primeiramente, está claro que nós devemos escolher entre $F(a)$ e $F(b)$ em casos nas margens do uso significativo de “bom”, enquanto em casos normais, $F(a)$ e $F(b)$ serão juntamente satisfeitos. “Ao privilegiar $F(a)$ para fixar F ... interpretamos [a declaração] como sinalizando a consensualidade da aprovação”, Barker alega, “já que atitudes- F devem ser partilhadas” (2000: 277). Em segundo lugar, se $F(b)$ fornece de forma independente um significado de “bom”, o expressivismo e a teoria da convenção são ambos falsos: é possível que o falante declare que algo é bom sem expressar sua aprovação

sobre o mesmo. A interpretação disjuntiva é, portanto, incompatível com a significância que Barker alega para a teoria. Em terceiro lugar, Barker identifica uma diferença em origem ou status para os dois critérios. $F(a)$ é identificado como semântico (“o caráter da palavra “bom”) enquanto $F(b)$ é identificado como pragmático (vindo do “aspecto pressuposicional da implicatura” [200: 277], sobre o qual eu falarei mais em seguida).

Focando no critério semântico, a teoria da convenção de Barker assinala como o significado convencional de “bom” a *propriedade F de forma que F satisfaça $F(a)$* . Isso é lido como *de re* (2000: 277n); portanto o significado de “bom” é simplesmente F , e o conteúdo asserido de “ T é bom” é simplesmente T é F ⁶. Na teoria de Barker, portanto, “bom” é como um indexical⁷.

É útil apreciar que a proposta de Barker acaba por ser (para os meus objetivos) equivalente à teoria que Jamie Dreier avança com o nome de “relativismo do falante”. Por razões de clareza, entretanto, chamarei essa visão de Dreier/Barker de *teoria indexical da atitude do falante*, ou simplesmente a teoria indexical. Dreier escreve:

Termos morais têm uma semântica de dois níveis do tipo que David Kaplan propõe para indexicais. Indexicais têm como seu significado primário, de acordo com Kaplan, não um conteúdo, mas um caráter, uma regra para determinar o conteúdo de acordo com um contexto [1990:8]

cuja regra, nesse caso, é “uma função das atitudes afetivas do falante no contexto” (1990:9). Portanto “em uma dada ocasião de uso, ‘bom’ tem o conteúdo equivalente a um certo predicado descritivo e naturalista” (1990:18). A preocupação de Dreier, no

⁶ Isso está de acordo com os precedentes expressivistas de Stevenson, Hare e Edwards, e também com a análise de Mackie, onde F é fixado como a propriedade pela qual T ganha a aprovação de U .

⁷ Essa é a significância de Barker atribuir “bom” um “caráter” semântico – termo de David Kaplan para a teoria determinante de conteúdo que constitui o significado convencional de indexicais (ver também Dreier 1990: 8).

entanto, não é com a expressão de conteúdo atitudinal, mas com a condição de “internalismo” no uso apropriado da linguagem moral, o que requer que um falante fazendo uma asserção moral tenha uma atitude conativa correspondente.

Antes que possamos estar confiantes que não estou interpretando incorretamente a proposta de Barker, devo tratar a ponta solta que é F(b) (tenho uma segunda intenção aqui também: F(b) surgirá como significante para a teoria da conversação). É uma pressuposição pragmática, afirma Barker, do uso de *U* de “bom” para denotar que a audiência de *U* partilha ou considera incontroverso a atitude-F: “Atitudes-F são feitas para serem partilhadas”. O que asserimos em conversas normalmente supõe que um conjunto de informações já é compreendida pela nossa audiência, portanto desnecessário de se asserir, e a adequação conversacional de nossos atos de fala geralmente depende da correção dessa suposição: isso é uma “pressuposição pragmática”. Esse fenômeno tem duplo propósito: ele determina condições para o uso aceitável da linguagem, e também tem um importante papel em determinar o conteúdo de implicatura dos nossos atos de fala (pois uma audiência tem a liberdade, pelas máximas conversacionais, de inferir dos nossos atos de fala nossa intenção de comunicar quaisquer condições para as suas performances apropriadas que não estão contextualmente pressupostas ainda). Por essa razão, Barker trata F(a) e F(b) ambos como critério e conteúdo para declarações apropriadas de “T é bom”. Também por essa razão, a restrição do internalismo de Dreier na asserção moral garantirá expressão atitudinal: se asserções morais sinceras requerem a posse de atitudes motivacionais, então as audiências podem inferir a partir de nossas asserções morais que possuímos aquelas atitudes.

Dada a semântica de “bom” na qual chegamos, por que é necessário que “atitudes-F devem ser partilhadas”? Por que um falante com propriedade linguística não pode chamar algo de “bom” por possuir uma propriedade a qual ele, mas não sua

audiência, aprova – como expressivistas mantêm que fazemos? Duas explicações para F(b) vêm a mente. Primeiro, Barker parece estar traçando um paralelo entre sua teoria da implicatura e sua explicação do fenômeno geral de implicatura conversacional, para a qual ele usa o seguinte exemplo:

(1) Até a vovó está bêbada.

O uso da palavra “até” aqui supostamente pressupõe (e contribui com o conteúdo da implicatura) que

1(a) A vovó tem menor inclinação que outros de estar bêbada e que

1(b) a crença que 1(a) é partilhada por ou incontroversa para a audiência.

Enquanto 1(a) é necessário pelo significado convencional de “até”, 1(b) é necessário pelo fato de que o falante negligencia asserir que 1(a). Ao declarar (1) sem asserir que 1(a), *U* pressupõe que a audiência já aceita que 1(a) – p. ex., *U* pressupõe que 1(b) é verdade. 1(b) é, portanto, uma pressuposição pragmática de uma declaração conversacionalmente apropriada de (1) (e é em si comunicado pelo ato de fala).

Existe, na superfície, um paralelo com a teoria da implicatura de Barker. Ao se declarar “T é bom”, é necessário a *U* que F(a): *U* está comprometido com a aprovação de coisas-F em S. Ao declarar “T é bom” sem asserir que F(a), *U* (possivelmente) pressupõe que sua atitude para com coisas-F é incontroversa portanto não precisa ser reportada. F(b) é portanto analogamente uma pressuposição pragmática da declaração de *U*? Não. Existe uma ambiguidade importante na alegação de que a atitude-F de *U* é incontroversa: pode ser tanto que a atitude em si quanto o fato de que a atitude é tida (por *U*) que é incontroverso. A *atitude* em si ser incontroversa quer dizer que ela é partilhada pela audiência e

pelo falante. Ao mesmo tempo em que é claro que Barker usa $F(b)$ como o primeiro (de fato, ele escreve “*partilhado por* ou *incontroverso*”)⁸, apenas o último é pragmaticamente pressuposto pela declaração de U . Enquanto o uso de “até” pressupõe como incontroverso uma certa escala de probabilidade, o uso de “bom” (dado que $F(a)$ fornece seu significado) pressupõe como incontroverso *no máximo*⁹ que U está comprometido com a aprovação de coisas- F ¹⁰. E isso pode ser verdade mesmo onde a atitude- F não é partilhada.

Enquanto a primeira explicação para o critério $F(b)$ falha, um segundo se sugere. Onde $F(a)$ não é asserido pelo falante, seu ato de fala pragmaticamente pressupõe $F(a)$ como incontroverso para a sua audiência. Se a audiência então não tem nenhuma pista independente quanto à propriedade com cuja aprovação U está comprometido em S , a suposição padrão é que é a propriedade com cuja a aprovação todos estão comprometidos em S . $F(b)$ portanto não serve, como Barker parece sugerir, como uma restrição sobre o que um falante pode usar “bom” para denotar, mas como um indicador padrão para uma audiência quanto à denotação pretendida pelo falante. Concluo que Barker superestima a importância de $F(b)$ ao alegar que atitudes- F são feitas para serem partilhadas. Ao reconhecermos $F(a)$ como o único critério semântico para determinar F , podemos designar a propriedade denotada por “bom” como a propriedade F de forma que coisas- F sejam aquilo com cuja aprovação U está comprometido em S .

⁸ Considero óbvio que o “ou” aqui significa clarificação, não disjunção, como quando digo “mexerica ou bergamota” (N. dos T.: “a diaper, or nappy” no original, ambos sinônimos de “fralda”)

⁹ Compare o uso de “eu”: enquanto com frequência pressupõe pragmaticamente a identidade do falante (qual “eu” contribui para as condições de verdade) para ser incontroverso, isso dificilmente é o caso.

¹⁰ Esse contraste é dificultado pela descrição de Barker da escala de probabilidade para “até” como “subjettiva”. Considero isso um erro: uma pessoa proferindo (1) não está pressupondo meramente que ela própria classifica probabilidades de tal forma, mas que essa é a maneira na qual as probabilidades estão de fato.

A visão de Barker, então, combina a teoria indexical da semântica de “bom” com a alegação de isso carrega conteúdo atitudinal como implicatura convencional. Agora quero causar problemas para essa combinação de teorias indexical e de convenção.

2. A teoria indexical respalda a teoria da convenção?

É útil começar observando uma óbvia, mas errônea objeção à teoria indexical. Tendo identificado o significado de “bom” na teoria de Barker como “a propriedade *F* de forma que coisas-*F* sejam aquilo com cuja aprovação *U* está comprometido em *S*”, pode-se pensar que temos aqui uma teoria *subjettivista*, na qual a declaração de *U* incorpora uma descrição das atitudes de *U* em seu conteúdo asserido. O conteúdo atitudinal é então asserido e não conteúdo de implicatura. Essa objeção falha devido à natureza dos indexicais. O significado convencional de um indexical contribui com conteúdo asserido *de re* e não *de dicto*. Existe uma separação, na teoria indexical, entre o significado e o conteúdo de “bom”. No uso de *U* o significado é “a propriedade *F* de forma que *U* está comprometido com a aprovação de coisas-*F* em *S*”, enquanto o seu conteúdo é simplesmente *F*. O conteúdo asserido da declaração de *U*, então, é simplesmente

T possui *F*

a partir do qual nada se segue de não trivial sobre atitudes de *U*. A teoria indexical é então não propriamente uma forma de subjetivismo, e Barker pode parecer justificado em suplementá-la com uma teoria da convenção: pela virtude de seu significado convencional, o uso de “bom” de fato comunicaria a atitude do falante, e não através do conteúdo asserido.

No entanto, há um problema aqui. A função determinante de denotação para qualquer uso de “bom”, na teoria indexical, é

fornecida por $F(a)$, como vimos. Agora veja que $F(a)$ é precisamente o conteúdo que a teoria da convenção de Barker designa como implicatura convencional para a declaração de U . Isso significa que o conteúdo asserido é determinado pela implicatura, o que é bastante estranho. Sugerir que o que U asserir pode às vezes depender de uma implicatura (o que U comunica pela asserção dela) é virar a dependência como normalmente concebida de cabeça para baixo. Contraste uma análise de conteúdo dual de uma palavra como “sádico”, na qual asserir “ A é um sádico” é (a) expressar como conteúdo asserido que A tem a propriedade C , a saber, ter prazer em causar dor a outros, e (b) expressa como implicatura que o falante está comprometido com a desaprovação de coisas- C . Aqui a propriedade em questão, C , é contribuída à implicatura pelo conteúdo asserido. Isso é um problema, pois a noção de implicatura é comumente elucidada como conteúdo a falsidade daquilo que é “compatível com a verdade de uma declaração” (Bach 1999: 331); Na teoria indexical é impossível que “bom” tenha uma denotação sem que $F(a)$ seja verdadeiro. Isso resulta em que é impossível que a declaração de U não pode ser verdadeira sem que $F(a)$ seja verdadeira.

Resta dúvida, portanto, se $F(a)$ é genuinamente implicatura convencional, e portanto também resta dúvida se a teoria indexical realmente respalda a teoria da convenção. Em última instância, no entanto, isso é mera técnica. Primeiramente, a elucidação da implicatura convencional pode ser desafiada: observe que enquanto “ T é bom” não pode ser declarado para alegar uma verdade quando $F(a)$ não obtém, o conteúdo asserido (T possui F) ainda pode ser verdade, já que suas condições de verdade são independentes da verdade de $F(a)$. Copp sugere (em correspondência) que isso nos permite oferecer elucidações concorrentes de implicatura convencional como conteúdo cuja falsidade é compatível com a verdade do conteúdo asserido, o que resgata a reivindicação de $F(a)$ de ser implicatura convencional. Eu não acredito que essa sugestão deva ser aceita: tal “implicatura”

tem uma influência na condição de verdade da declaração, o que entra em conflito com a racionalização básica para o conceito. Mas, em segundo lugar, mesmo que essa proposta seja rejeitada, não há problema aqui para a teoria indexical como uma descrição de conteúdo asserido, ou para a alegação de que nos permite explicar como declarações morais podem expressar convencionalmente conteúdo atitudinal sem asseri-lo. No máximo essa objeção evidencia que nem todo o conteúdo baseado em convenção mas não asserido é conteúdo de implicatura. Irei passar para uma objeção mais substantiva à teoria indexical.

3. Desfragmentando o bem

Ao atribuir conteúdo semântico condicional de verdade a declarações morais, e atribuir propriedades naturais como denotações de “bom”, Barker e Dreier aderem ao expressivismo clássico alegando que não há nada particularmente moral ou normativo naquele conteúdo ou propriedade. A mesma exata proposição poderia ser asserida sem o uso de termos morais ou sequer frases, já que o conteúdo moral não adentra as condições de verdade. Copp assim observa que na visão de Barker, ao realizar uma declaração moral uma pessoa “expressa uma crença empírica comum”, e não há “espaço no seu relato para a existência de propriedades *morais* robustas (2001: 40n). Copp acha isso “bastante implausível”, e eu concordo. Dada a prevalência da dúvida sobre propriedades morais, no entanto, não posso encerrar aqui meu argumento por intuições partilhadas.

A estratégia indexical requer que as propriedades denotadas por “bom” sejam dramaticamente relativas ao contexto e ao falante, um aspecto característico dos indexicais (como “eu” e “aqui”). Enquanto não nego que a denotação de “bom” é relativa ao contexto, de forma que é possível que um falante diga “T é bom” e outro em um diferente contexto diga “T não é bom” sem contradizer o primeiro, a teoria indexical faz com que as condições

de verdade de declarações morais sejam relativas a um grau que não é fiel ao fenômeno – particularmente do desacordo moral.

A fraqueza na teoria indexical pode ser vista quando examinamos o caso onde outro falante *W* concorda com *U* que *T* possui *F* mas não partilha a atitude-*F* de *U*, e, além disso, não acredita que *T* tenha nenhuma das propriedades das quais *W* aprova. Considere essa ilustração não moral: *T* é uma faca, *S* é o contexto de procurar um instrumento para cortar carne efetivamente, *F* é a propriedade de ser serrilhado. Na teoria indexical, *W* deveria considerar que *U*, em asserir que “*T* é bom”, tenha falado a verdade: o “bom” de *U* aqui significa “serrilhado” e *W* acredita que *T* é serrilhado. Isso é contraintuitivo: certamente *W* acredita que seja simplesmente falso que *T* é bom em *S*, não importa quem diga.

Barker irá responder que essa objeção ignora o caráter indexical de “bom”, o que permite duas formas diferentes de desacordo moral ou avaliativo: *baseado em explicatura* (onde *W* partilha da atitude-*F* de *U* mas discorda que *T* possui *F*) e *baseado em implicatura* (onde *W* não partilha da atitude-*F* de *U*). Ele dirá que estou confundindo a negação baseada em implicatura por *W* da declaração de *U* por negação baseada em explicatura: é claro que *W* não pode dizer “Sim, *T* é bom”, já que “bom” em sua boca tem uma denotação diferente, mas isso não quer dizer que ela considere a asserção de *U* falsa.

Considere a analogia com o pronome pessoal. Se *U* asserisse “eu sou ‘*U*’”, o conteúdo asserido seria de que *U* é “*U*” e que o conteúdo de implicatura seria que o falante é “*U*”. *W* não poderia responder “Sim, eu sou ‘*U*’”, mesmo que saiba que *U* falou a verdade, pois “eu” significa algo diferente vindo da sua boca. Até aqui, os casos são paralelos, mas acredito que Barker está errado: a discordância de *W* é baseado em explicatura, e a analogia com o pronome pessoal ajuda a demonstrar isso. Barker acredita que é bastante apropriado para (de fato, no caso moral, moralmente incumbente) *W* protestar *metalinguisticamente* o julgamento de *U*,

pois deve negar a implicatura de *U*. *U* deveria responder “T não é bom. Ser bom não consiste em possuir F”. Observe a falha na analogia: “Eu não sou ‘*U*’. Ser eu não consiste em ser ‘*U*’.” Ao mesmo tempo em que *W* não pode concordar com a declaração de identidade de *U* ecoando-a, também não tem nenhum motivo para protestar sua implicatura.

A falha da analogia será devida, na visão de Barker, à pressuposição seguinte para o uso de “bom”, sem equivalente para o uso de “eu”, de que o falante e a audiência partilham da mesma atitude-F – p. ex., a pressuposição de que F(b) (o equivalente de “eu” seria que o falante e a audiência partilham da mesma identidade). Mas claro, já argumentei que F(b) não pode ser a condição para o uso apropriado de “bom” na teoria indexical. Se eu estou certo quanto a isso, então seria bastante inapropriado para *W* sequer discordar do julgamento de valor de *U*, já que seu conteúdo asserido é simplesmente que T é F (o que *W* aceita), e sua implicatura é simplesmente F(a) – p. ex., que *U* está comprometido com a aprovação de coisas-F em S (o que *W* não tem nenhuma razão para contestar). É claramente absurdo, no entanto, alegar que *W* pode não discordar apropriadamente com a declaração de *U*. Por outro lado, se estou errado em rejeitar F(b) como condição do uso adequado de “bom”, então é um abuso de linguagem para *U* endereçar sua declaração avaliativa para *W* quanto F(b) não é verdadeiro e *W* não partilha sua atitude-F. E isso é inaceitavelmente restritivo para o uso apropriado de “bom”; certamente *U* tem direito de declarar T como bom, mesmo quando sabe que *W* não aprova das propriedades pelas quais *U* aprova T. Expressivistas insistirão nisso.

Para continuar a comprometer a plausibilidade de uma teoria indexical, considere o caso onde, como antes, *W* não partilha da atitude-F de *U*, mas agora *W* agora aprova da propriedade G (digamos, a propriedade de ser afiada), e acredita que T possui G mas não F, enquanto *U* não aprova de G e acredita que T não tem G. *U* e *W* julgariam ambos, por si mesmos, T como “bom”. E ainda

assim cada um consideraria a declaração do outro falsa. Eles não *concordariam* que T é “bom”, embora ambos fariam esse julgamento privadamente. Isso não é uma hipótese incoerente, mas falha significativamente como uma descrição plausível da semântica de “bom” e a dinâmica do discurso moral e normativo. É natural dizer, ao invés disso, que U e W concordam que T é bom mas discordam quanto ao que o torna bom. Parece, portanto, que a teoria indexical erra em confundir a propriedade (ou propriedades) de ser bom com as propriedades que compõe o bom.

Há razão, portanto, para procurar uma teoria que possa identificar uma única propriedade F (ou talvez uma família de propriedades) como denotação de “bom”, distinto das várias propriedades de fazer de forma boa. Embora muitos filósofos ainda considerem que o argumento da questão em aberto G. E. Moore prova a futilidade desse projeto, Copp, eu mesmo e outros mantemos que isso pode ser realizado através de uma *teoria da propriedade padrão-relacional* (ou simplesmente teoria relacional)¹¹. As características genéricas dessa teoria são: (1) as propriedades morais/normativas, incluindo a propriedade denotada por “bom”, são propriedades relacionais; (2) essas relações acontecem entre alguns conjuntos de padrões e objetos, ações, estado das coisas, etc. que tenham a propriedade; (3) eles se consistem nesses objetos etc. de alguma forma, seja atendendo ou não atendendo àqueles padrões; (4) pelo menos algumas pessoas algumas vezes estão motivadas de acordo com esses padrões. Esse modelo básico permite muitas variações: Copp (seguindo Gibbard)

¹¹ Outros incluem Paul Ziff (1960), John Mackie (1977), e, apesar de suas intenções antirrealistas, Allan Gibbard (1990). Dreier falha em distinguir claramente sua teoria indexical da teoria relacional: ele escreve, “O caráter de ‘bom’ [...] é uma função de contextos à *propriedade de ser avaliado positivamente* por um sistema moral do falante no contexto” (1990: 20); “se [o falante] entende o significado de ‘bom’, então sabe que quando usar a palavra está descrevendo [T] como atendendo certos padrões” (1990: 20); “o conteúdo de ‘x é bom’ tem a forma ‘x é aprovado pelo sistema moral M’” (21). Não pode, no entanto, haver dúvidas que Dreier pretende a teoria indexical: é necessário para essa alegação que ao vir a acreditar em uma sentença moral, uma pessoa está acreditando em algo de uma forma nova, ao invés de adquirir uma nova crença (logo, na p. 20, ele identifica a denotação de um uso particular de “bom” com as relevantes “características de fazer de forma boa”).

concebe os padrões relevantes como normas sociais imperativas (onde “atender” consiste em ser permitido por, “não atender” consiste em ser proibido por), enquanto outros (incluindo Ziff, Mackie e eu) preferem identificá-los com fins, metas ou objetivos, como fornecidos por estados motivacionais tais como desejo (de forma que “atender” consiste em avançar ou satisfazer, e “não satisfazer” consiste em obstruir ou falhar em satisfazer). Para o presente propósito, no entanto, essa é uma disputa mutuamente destrutiva. Uma pergunta difícil, a qual não se pode satisfazer aqui, se trata o que em uma descrição relacional torna uma avaliação “moral”. Uma resposta parcial é que um juízo de valor conta como moral apenas se o padrão ao qual ele é relativo conta como moral; digo mais na p. 15. Copp então sugere que dizer que *T* é moralmente errado é dizer que é proibido por um padrão moral relevantemente justificado, onde padrões morais, em sua visão, são distinguidos por sua função social.

Para ilustrar as vantagens da teoria relacional sobre a teoria indexical, considere o que cada uma tem a dizer sobre a discordância de *U* e *W* quanto ao valor de cortar carne das facas. A teoria indexical identifica o quão boa uma faca é, o que é relativo a cada falante, como a propriedade natural pela qual eles aprovam dela – ser serrilhada *versus* ser afiada. A teoria relacional reconhece que existe um padrão comum do qual *U* e *W* partilham com relação à faca – sua capacidade para cortar carne de forma eficaz – e identifica o quão boa ela é para *U* e *W* como a propriedade de avançar essa capacidade. Ser serrilhada e ser afiada então são propriedades naturais as quais *U* e *W* respectivamente acreditam ser propriedades que a tornam boa, e baseados nessa diferença eles de fato, na teoria relacional, genuinamente concordam ou discordam quanto a um assunto em comum: se a faca possui uma propriedade particular, se é boa. A teoria relacional, como a teoria indexical, permite que “bom” tenha denotações diferentes, mas apenas dada uma diferença em padrões básicos ou fins (uma faca boa para cortar carne, afinal de contas,

não é uma faca boa para passar manteiga). No caso moral, portanto, enquanto desacordos morais genuínos entre membros de diferentes sociedades com padrões morais radicalmente diferentes podem ser difíceis de encontrar, eles serão comuns dentro de uma sociedade e entre sociedades com padrões comuns.

4. Conteúdo atitudinal na teoria relacional

A teoria relacional precisa fornecer uma descrição de como uma asserção moral carrega conteúdo atitudinal, já que o significado atribuído a “bom” não envolve diretamente as atitudes do falante. A resposta estará relacionada com a relação do falante com os padrões relevantes. Se um falante está adequadamente motivado de acordo com aqueles padrões (ele “endossa”, “adere a” ou os “internalizou”)¹² então ele irá consequentemente aprovar de coisas na medida em que acredita que elas estejam de acordo com aqueles padrões, e desaprovar de coisas na medida em que acredita que elas não estão de acordo (aprovar de T, nessa visão, não é simplesmente uma reação psicológica bruta a T, mas tem uma explicação subjacente em termos de uma característica relacional percebida em T). Mas como o significado de “bom” assegura essa conexão com a motivação do falante?

A sugestão inicial de Copp (1995) foi que a asserção moral carregava conteúdo atitudinal por implicatura *conversacional*. Quando chega a desenvolver os detalhes (2001), embora ainda reconhecendo a implicatura conversacional, ele coloca ênfase na implicatura convencional. Sua teoria da convenção, no entanto, é feita para ser independente de sua teoria relacional de conteúdo asserido.

Ele escreve,

¹² “Adere” é a terminologia de Copp; “internaliza” é a de Gibbard.

Serviria aos objetivos do discurso moral se uma convenção fosse se desenvolver de forma que uma pessoa asserido uma crença moral básica através do uso de termos morais estaria portanto implicando que ela adere a um padrão moral relevante. Se tal convenção fosse se desenvolver, então a conexão do discurso moral com estados mentais que guiam a ação estaria codificada no sentido dos termos que usamos [2001:34-5].

Irei argumentar agora que isso é um erro, já que (1) na teoria relacional, implicatura convencional é redundante pois o conteúdo atitudinal já é expresso convencionalmente; (2) teorias conversacionais são preferíveis a teorias convencionais, já que se encaixam melhor com os fenômenos (especialmente o aspecto de “cancelabilidade”); (3) há dificuldades com a própria ideia de implicatura convencional, e com essas ditas convenções em particular. Concedo, no entanto, uma vantagem à teoria da convenção de Copp: diferentemente da teoria da conversação defendida aqui, ela não depende de nenhuma teoria específica de conteúdo asserido de juízos de valor, tendo a flexibilidade de funcionar com praticamente qualquer teoria desse tipo, e portanto já ganharia se Copp e eu estivermos ambos enganados em ter visões relacionais desse conteúdo.

Eu diferencio duas formas em que Copp possibilita um conteúdo atitudinal conversacionalmente implicado (2001: 31-3), ambos tendo a ver com o papel social normal de valores morais. Primeiramente, ele escreve que “a razão pela qual a moeda de um código moral pode servir às necessidades da sociedade é que ela envolve uma adesão generalizada ao código dentro da população dessa sociedade”, portanto “o propósito do discurso moral depende [...] da nossa tendência de ter [...] crenças morais que são acompanhadas por adesão a outros padrões correspondentes” (2001: 32). Tipicamente, portanto, as crenças morais das pessoas são acompanhadas pela adesão moral, que é “o estado mental esperado de uma pessoa que expressa uma crença moral básica”; então “a pessoa que faz uma asserção moral conversacionalmente

implica que ela adere a um padrão moral correspondente”. A segunda forma consiste no propósito típico de uma conversa moral: se trata, Copp sugere, de *o que fazer* (o “contexto de decisão”). Oferecidas nesse contexto, asserções morais serão tidas como sendo a contribuição do falante para que se chegue a uma decisão prática, portanto haverá a pressuposição de que o falante de fato adere aos padrões implícitos em sua asserção moral. Ambos os mecanismos soam bastante plausíveis, e irei simplesmente concedê-los isso.

Por que Copp não pensa que a implicatura conversacional, tão fundamentada, é suficiente para explicar o conteúdo atitudinal de asserções morais? A resposta é, eu acho, complexa. Ele sugere primeiramente (2001: 34-5) que dados os objetivos do discurso moral e a natureza de convenções linguísticas, “não seria surpreendente” se o conteúdo atitudinal fosse parte do significado da linguagem moral (sua “coloração”, seguindo Frege). Isso pode ser visto como um argumento de que é apenas de se esperar, moralidade e linguagem sendo o que são, que a teoria da convenção seja verdade. Eu irei criticar essa alegação na seção 6. Em seguida, ele expressa a crença de que “termos morais *de fato* têm essa coloração”. “Seria enganoso, outras coisas mantendo-se iguais”, ele escreve, “que Bob diga a Alice ‘dizer palavrões é moralmente errado’ se ele não adere ao padrão que proíbe dizer palavrões. Se Alice compreende o objetivo do discurso moral, ela assumirá que ele adere a tal padrão” (2001: 35). Embora essa observação seja oferecida aqui para respaldar a teoria da convenção, não se difere de nenhuma forma significativa de sua afirmação anterior para o caso da implicatura conversacional! (De fato, com sua cláusula *ceteris paribus* e seu foco no objetivo do discurso moral, ela é mais adequada para respaldar a implicatura conversacional. É significativo aqui que os argumentos de Copp são direcionados principalmente contra as teorias subjetivista e não descritivista ao invés de teorias da conversação, e por isso minha

crítica o pega despreparado. A alegação importante, portanto, é a que segue:

Eu não acredito que uma pessoa precise ter um entendimento sofisticado da pragmática do discurso moral para que entenda que uma pessoa que asseire uma crença moral básica em termos morais está implicando adesão a um padrão moral correspondente. Me parece que é apenas necessário compreender o que é dito. Se isso está correto, então talvez existam convenções linguísticas que governam o uso de termos morais de forma a serem usados normalmente para expressar adesão a normas. [2001:35]

Uma teoria da conversação não é suficiente, portanto, pois ela forja um elo muito fraco entre asserção moral e expressão de conteúdo atitudinal. Falar de uma “compreensão sofisticada” de pragmática é ambíguo, no entanto. Por um lado apenas um número pequeno de filósofos e linguistas têm uma compreensão sofisticada de pragmática. É certamente verdade que entender o conteúdo atitudinal não requer esse nível de especialidade teórica. Mas por outro lado (o sentido do *know-how*, por assim dizer) todo o falante competente tem uma compreensão sofisticada de pragmática e é exímio em detectar implicaturas conversacionais. Um argumento melhor é talvez o de que teorias da conversação forjam um elo muito *contingente* entre asserções morais e expressão atitudinal. As condições pragmáticas observadas acima não são suficientemente enraizadas para acomodar a proximidade da conexão entre asserção moral e atitude do falante. Muitas pessoas crescem e se tornam adultas em nossa sociedade sem educação moral realizando uma impressão adequada delas. Assuntos morais são frequentemente discutidos em contextos conversacionais que não estão diretamente focados em uma decisão. Ainda assim, Copp sugere, para compreender um falante como expressando adesão a um padrão moral, “é apenas necessário entender o que é dito”: nós não precisamos pressupor a obtenção dessas condições. A implicatura da atitude do falante

deve, portanto, ser uma questão de semântica, não de pragmática – respaldando teorias de convenção ao invés de conversação.

Eu concordo que as formas de implicatura conversacional que Copp delineia são insuficientes para explicar a proximidade de uma conexão entre asserção moral e atitude do falante. Mas como argumentei anteriormente (Finlay 2004), há outro aspecto conversacional de asserções morais bem mais onipresente que Copp ignora e que pode acomodar os fatos aos quais ele recorre. É verdade, como Copp diz, que na maioria dos contextos uma pessoa que faz uma asserção moral comunica deste modo – meramente por sua escolha de palavras – sua adesão ao padrão moral relevante. Mas isso também é algo que teorias da conversação podem explicar. Uma forma pela qual implicaturas conversacionais podem surgir é através da omissão de certas palavras, por parte do falante, de sua declaração, palavras que são componentes necessários para o pensamento completo a ser asserido: falantes estão omitindo constituintes dos seus atos de fala de uma forma significativa.

O que está sendo deixado fora? A teoria relacional de Copp facilita a resposta: o que é omitido é a *especificação do padrão moral relevante*¹³. A teoria relacional permite que diferentes declarações morais possam invocar diferentes padrões morais. Consequentemente o fato de que *U* está julgando *T* relativamente ao padrão *M* é parte importante do sentido completo do que *U* pretende expressar com sua declaração. Esse fenômeno é melhor observado no caso não-moral. Quando *U* adere que a faca é boa, ele está realmente asserindo que a faca é boa *relativa ao padrão de cortar carne de forma eficaz*. Uma omissão como essa é conversacionalmente permitida apenas em um contexto onde pode-se esperar que a audiência pressuponha ou automaticamente complete o conteúdo de

¹³ Os exemplos de Copp têm o qualificador *moralmente* bom, o que em certo grau conta como especificação do padrão relevante. (Veja que juízos morais são mais frequentemente feitos sem sequer esse qualificador, e a implicatura da atitude do falante é até mais forte nesses casos.) Mas em teorias relacionais há mais do que um padrão moral; por isso falar como se houvesse apenas um sugere fortemente (conversacionalmente) que o próprio falante adere ao padrão relevante.

ausente; a omissão de *U* é justificada pelo contexto no qual o fim do corte de carne de forma eficaz é pressuposto.

Audiências estão permitidas, pela “máxima de qualidade” de Grice, de entender como conversacionalmente implicada qualquer informação requerida no contexto para que a declaração do falante faça uma contribuição inteligível para a conversação. Uma declaração avaliativa que está incompleta da forma como descrita acima irá, portanto, conversacionalmente implicar que o falante possui alguma atitude, só para o caso de que ela possuir aquela atitude seja (parte do) o contexto que sua audiência deve pressupor para que possa identificar a relativização-padrão omitida. Há vários tipos de contexto nos quais uma audiência pode identificar o padrão implícito; vários desses são contextos nos quais o próprio falante adere àquele padrão e por isso irá aprovar daquilo que julga que atendê-lo. (Considere analogamente “eu escolho o da esquerda”, declarado em um contexto de jogo dos copos. Eu submeto que a pressuposição padrão é que o falante se refere à esquerda *dele*¹⁴. O princípio geral aqui é que o padrão, quando um ato de fala omite informações importantes de relativização, é que o ato de fala está relativizado às perspectivas do falante.)

No caso não moral, essa pressuposição é facilmente derrotada. Mas no caso moral, tem um embasamento adicional. Uma razão pela qual o padrão relevante pode ser deixado (como se) contextualmente pressuposto é retórico. Ao negligenciar ser explícito sobre a relativização de sua declaração, um falante significa, como Barker observa, que a relativização é incontroversa. Padrões morais são caracterizados em parte por seu status social: são padrões aos quais nossa adesão e conformidade são socialmente demandadas. Essa demanda social ou “expectativa” pode ser expressada ao falar como se a adesão ao padrão relevante fosse partilhada mesmo quando não é. (De fato, para propósitos retóricos não é nem necessário que a audiência possa identificar o

¹⁴ Nesse caso, no entanto, a autorreferência é parte do conteúdo asserido ao invés de da implicatura.

padrão em questão, mas meramente que o falante adere e exige conformidade a ele). Consequentemente, na maioria dos contextos, um falante ao declarar uma sentença moral não relativizada implica conversacionalmente que ele adere àquele padrão (e que espera que todos o façam também) – meramente por sua omissão. Uma analogia não moral pode ser útil: considere um pai que localiza sua filha adolescente que fugiu e diz a ela que “volte para casa”. Que casa é essa? Não “minha casa”, ou “sua casa”, eu acho, mas “nossa casa”. Se a casa do pai ainda é também da filha é uma questão de discórdia entre eles, mas ao falar como se não fosse, o pai conversacionalmente implica que quer que ela more com ele e que demanda que ela aquiesça.

Essa descrição da implicatura conversacional de uma asserção moral compromete a racionalização de Copp para uma teoria da convenção. Na teoria relacional que ambos aceitamos, um aspecto conversacional da asserção moral fornece o mesmo conteúdo atitudinal sem requerer convenções linguísticas adicionais, assim tornando tais convenções supérfluas.

5. Separabilidade e Cancelabilidade de Conteúdo Atitudinal

Teorias relacionais portanto têm os recursos para explicar como a própria escolha de palavras de um falante implica conteúdo atitudinal conversacionalmente. Dada a escolha entre teorias de convenção e essa teoria de conversação relacional, a teoria de conversação é preferível, já que se encaixa melhor aos dados. Copp parece alegar o contrário. “Para testar se o termo “moralmente errado” tem uma coloração do tipo”, ele escreve, “podemos aplicar nossos quatro testes para coloração [...] o teste da verdade de Frege, o teste de separabilidade de Grice, meu teste de cancelabilidade e o teste de mau uso” (2001: 35). Está implicado aqui que se “moralmente bom” passa por todos os quatro testes – e Copp alega que o faz – então deve ter coloração/implicatura conversacional. Mas ele introduz esses testes para distinguir

coloração/implicatura de *conteúdo asserido* (“significado central” [2001: 16-9]), não implicatura conversacional, e não aborda (a não ser de forma oblíqua em uma nota de rodapé) se algum deles serve também para distinguir coloração de implicatura conversacional. É um erro, portanto, tomar um resultado positivo como evidência para uma teoria da convenção ao invés de uma teoria da conversação. Vou agora argumentar que teorias da conversação se saem melhor nesses testes do que teorias da convenção.

Nós podemos dispor de dois desses testes rapidamente. O “teste da verdade” nos diz que uma proposição *p* é implicatura de uma declaração apenas se a declaração fosse verdadeira mesmo que *p* fosse falso. O “teste do mau uso” nos diz que *p* é implicatura de uma declaração apenas se essa declaração fosse ser apropriada mesmo que *p* fosse falso, pois comunicaria que *p*. Esses dois testes juntamente fornecem um teste genérico para implicatura de qualquer variedade e portanto não podem decidir a questão entre eles. Os testes relevantes tratam da separabilidade e cancelabilidade: na visão de Grice esses são aspectos distintivos de implicaturas convencional e conversacional respectivamente.

Conteúdo de implicatura ou *implicata* é separável se o mesmo conteúdo asserido puder ser expresso no mesmo contexto sem a implicatura, meramente substituindo um termo diferente mas coexistente para a palavra carregando a implicatura. Se eles de fato carregam implicatura ou não, pejorativos fornecem uma boa ilustração: o uso de palavras como “bicha” e “crioulo” expressam atitudes de desprezo, mas (possivelmente) exatamente o mesmo conteúdo asserido pode ser expresso sem tal desprezo, usando palavras sem coloração como “homossexual” ou “Afro-Americano”. Como Copp observa (em uma nota de rodapé), Grice vê separabilidade como uma característica distintiva das implicaturas convencionais. Aqui, se em algum lugar, então, pode ser feito um caso para que teorias de convenção sejam preferíveis a teorias de conversação.

Seria o conteúdo atitudinal de asserções morais separável? Pode parecer plausível que possamos achar termos “sem coloração” diferentes mas coexistentes para substituir por termos morais, mas Copp escreve, “me parece que aspas ou ‘aspas irônicas’ podem ser usadas para descolorir termos que normalmente têm coloração (2001: 35). O conteúdo asserido de qualquer declaração moral, em outras palavras, pode ser expresso sem implicatura de atitude através da substituição daquele termo em aspas por qualquer termo moral. Isso parece correto mas também um pouco superficial: esse argumento força o sentido de “termos diferentes mas coexistentes”. Possivelmente um termo em aspas não é um termo diferente mas o mesmo termo usado de uma forma diferente. Qual é então a maneira de usar aspas? Copp sugere que aspas funcionam para separar (e cancelar!) implicaturas. Mas se examinarmos o fenômeno das aspas de forma mais geral, chegamos a um resultado diferente. Se eu falo de os “amigos” de alguém, ou de seus “ganhos” ou “virtudes”, (por exemplo) entre aspas, o que eu estou significando é que eu não os considero genuinamente como amigos, ganhos ou virtudes. Aspas parecem, nesses casos, pelo menos, indicar duas coisas: primeiro, que outros (os quais estou “citando”), referem-se a essas pessoas, objetos, propriedades, etc. dessa forma, e, em segundo lugar, que eu não acredito que essas palavras as denotam apropriadamente. Aspas funcionam primariamente para rejeitar o conteúdo asserido e não o conteúdo de implicatura.

Se essa é também a natureza de uma asserção moral feita entre aspas (o que parece provável), então a solução de Copp falha: “moralmente bom” entre aspas é usado para expressar (1) que certos outros chamariam estas coisas que “moralmente bom”, mas que (2) o falante não considera estes moralmente bons. Deve-se esperar, portanto, asserções morais entre aspas em pelo menos duas ocasiões: primeiro, onde um falante reage às asserções morais de outros os quais considera errados; e segundo, onde um falante é um cético moral radical (como várias pessoas são), que

acredita que não há algo como bondade moral. Copp portanto falhou em mostrar que as implicaturas de atitudes são separáveis¹⁵, e por consequência falhou em demonstrar que há razão para duvidar que sejam conversacionais.¹⁶

Copp não precisa dessa estratégia de aspas, no entrando, se ele abandonar o objetivo de avançar uma teoria convencional que seja autossuficiente, independente de sua teoria relacional. Pois a teoria relacional fornece termos diferentes mas coexistentes para substituir aqueles usados em declarações (elípticas) morais, as quais não carregam implicatura atitudinal: dizer “T é (moralmente) bom” é dizer que T está adequado a algum padrão moral M. Isso falha em respaldar a teoria convencional, no entanto, pois implicaturas conversacionais carregadas por elipses ou a omissão de componentes de conteúdo asserido, diferentemente de implicaturas conversacionais no geral, são de fato separáveis. Pois essas implicaturas surgem da forma que “alguém coloca o que alguém diz” (Bach 1999: 330), não surgem da mesma asserção colocada de forma diferente – p. ex., se palavras diferentes são usadas. A separabilidade de implicaturas atitudinais na teoria relacional de Copp não fornecem portanto nenhuma evidência de que elas são convencionais ao invés de conversacionais.

A implicatura é “cancelável”, no sentido de Grice, se a mesma sentença pode ser proferida, sem impropriedade linguística, sem expressar aquele conteúdo. Uma implicatura é *explicitamente* cancelável se um falante pode (sem abusar do sentido de suas palavras), legitimamente negar que ele quis expressar aquele

¹⁵ Copp discorda em correspondência que isso caracteriza o resultado final do meu argumento, o qual é mais sobre ele estar certo quanto a função separadora das aspas, mas que o teste então falha em distinguir implicaturas convencionais de conversacionais. Eu discordo: na definição técnica de separabilidade dada acima, já que a estratégia das aspas falha em fornecer um *termo diferente mas coexistente*, ela falha em acomodar a separabilidade.

¹⁶ Se bem-sucedida a estratégia das aspas permitiria da mesma forma que implicaturas conversacionais parecessem separáveis. Se eu digo “pelo menos 100 pessoas morreram em um desastre”, eu implico conversacionalmente (pela minha imprecisão) que eu não sei o número exato de mortos. Se eu digo “pelo menos 100’ pessoas morreram em um desastre”, eu não necessariamente implico a mesma ignorância, pois estou pegando emprestado a escolha de palavras de outra pessoa.

conteúdo; é *contextualmente* cancelável se há contextos nos quais a declaração da sentença sequer pareceria sugerir aquele conteúdo. Para mim é bastante claro que o conteúdo atitudinal de asserções morais é cancelável em ambas essas formas: aqui encontramos a figura do amoralista, familiar nessa literatura, que sinceramente faz asserções sobre o valor moral de coisas mas não adere àqueles padrões morais e não expressa aprovação (etc.) através dos seus atos de fala morais. O conteúdo atitudinal pode ser cancelado explicitamente se ele apenas explicar que é um amoralista, ou que tem desprezo ou indiferença para com a moralidade. É contextualmente cancelado se sua audiência já sabe de seu amoralismo (especialmente se os próprios são amoralistas, tornando o amoralismo contextualmente incontestado).

Essa cancelabilidade é um problema sério para o subjetivismo: asserções morais não parecem ser tornadas falsas ou linguisticamente apropriadas pelas atitudes do falante. Não parece um problema muito menor para os teóricos da convenção e expressivistas (incluindo Copp, Barker, Dreier e Gibbard), os quais também unem conteúdo atitudinal a juízos morais diretamente através de convenções linguísticas. A significância de alegar que há uma convenção linguística que falantes proferindo uma sentença moral expressam certas atitudes é, nesse contexto, que proferir uma sentença moral quando não se possui ou não se tem a intenção de expressar tais atitudes é um mau uso da linguagem moral, de certa forma inconsistente com o seu próprio significado. Cancelabilidade é, no entanto, contingente ao contexto conversacional, e é por isso que Grice o vê como aspecto distintivo de implicatura *conversacional*.

O problema é bem conhecido, e várias estratégias foram propostas em resposta. Copp parece oferecer duas respostas distintas (e não completamente compatíveis). Primeiro, ele rejeita a cancelabilidade de implicaturas atitudinais no sentido conversacional de Grice, apelando para uma intuição de que o amoralista *de fato* faz um mau uso de palavras morais. Para

acomodar o uso do amoralista, ele sugere que ele exhibe um tipo mais fraco de cancelabilidade que é um indicador de implicatura conversacional (daí o seu teste de cancelabilidade). Enquanto cancelar implicaturas convencionais é um mau uso de uma palavra, esse mau uso “não seria contraditório, e seria completamente inteligível como uma asserção”, ele escreve, pois, apesar da estranheza linguística, a palavra ainda pode ser usada para esse “significado central” ou como contribuição para o conteúdo asserido (2001: 18.) Isso é insistir que o uso de linguagem moral por parte do amoralista é de fato uma escolha de palavras semanticamente inapropriada, uma violação das regras da linguagem, embora uma violação com um objetivo comunicativo.

O amoralista faz mau uso da linguagem moral? Aqui temos um conflito de intuições: para mim, mas não para alguns outros, me parece claro que podemos proferir juízos morais de forma apropriada sem expressão atitudinal. Ao invés de simplesmente insistir nas minhas intuições em detrimento as quais Copp e outros expressivistas argumentam, eu sugiro que a combinação de teorias relacionais e da conversação possam explicar ambas as intuições. Ela nos permite dar duas descrições diferentes do que queremos dizer ao qualificar um juízo avaliativo como “moral”. Nós podemos querer dizer (i) que é um juízo que diz respeito a um valor de um tipo moral específico, ou (ii) que é um juízo de valor feito de uma maneira de um tipo moral específico (da forma retoricamente exigente descrita anteriormente). Me parece que a noção de juízo moral é ambígua precisamente nesse sentido, e é de fato verdade que o amoralista não faz juízos morais no segundo sentido. Embora isso seja reconhecer alguma verdade na intuição de Copp, isso enfraquece a teoria da convenção que essa intuição visa respaldar: no segundo sentido, ser um “juízo moral” é um aspecto conversacional e não convencional.

A segunda resposta de Copp é restringir as supostas convenções com qualificações. O uso de termos morais por parte de um falante sem conteúdo atitudinal, “outras coisas mantendo-se

as mesmas” (2001: 30, 35, 36). As convenções, então, são contextualmente dependentes: “na maioria dos contextos, o amoralista faria mau uso do termo ‘moralmente errado’” (2001: 36). Ele parece querer permitir, então, que há contextos nos quais o amoralista não faz mau uso do termo. Isso, no entanto, ameaça enfraquecer o caso para a teoria da convenção – dependência contextual é um sinal claro de implicaturas *conversacionais* – e para a alegação de que para entender conteúdo atitudinal nós precisamos entender *o que é dito*. No entanto, não podemos simplesmente contestar que convenções não podem ser dependentes de contexto. Copp sugere que convenções podem ser sensíveis à ocorrência de termos em uma frase, tal como quando termos morais então “embutidos em construções maiores” (2001: 18). Mas isso não é capaz de explicar o amoralista, o qual não é distinguido de outros falantes pela construção sentencial de suas declarações. De fato, o amoralista parece ser capaz de usar termos morais de forma legítima, cancelando implicaturas atitudinais, em todos os contextos nos quais o adepto moral pode (exceto aquela de ser um adepto moral).

É possível que os únicos contextos que Copp tem em mente nos quais o amoralista não faria mau uso de termos morais são contextos nos quais ele os coloca entre aspas: ele sugere que aspas podem funcionar para separar e cancelar implicaturas¹⁷. Como já disse, me parece falso que o amoralista de fato use termos morais apenas desse modo indireto. E eu estiver correto quanto ao fato de que as aspas funcionam primariamente para renegar conteúdo asserido, então é o cético moral (que não acredita em verdades morais) ao invés do amoralista (que é indiferente a elas) que usa a linguagem moral dessa forma.

¹⁷ Observe que aspas não podem ser apropriadamente invocadas para acomodar ambas a separabilidade (que requer o uso de termos *diferentes*) e a cancelabilidade (que requer o uso dos mesmos termos em diferentes contextos).

Dreier confronta o problema de forma diferente. Ele julga a estratégia das aspas implausível, e sugere que “a falha do internalismo é parasítica de um *background* dos seus sucessos” (1990: 11). O amoralista só é possível devido ao caso normal, no qual os falantes realizam declarações morais indexadas ao sistema moral ao qual eles mesmos aderem. A declaração do amoralista é possibilitada pela normalidade social de adesão a um código moral específico: o amoralista pode fazer asserções morais indexadas ao sistema moral corrente em sua sociedade, mesmo quando ele mesmo não adere a este. Isso, no entanto, apresenta uma dificuldade para o “relativismo do falante” de Dreier e a alegação de que juízos morais expressam as atitudes dos falantes: termos morais não precisam estar indexados ao falante, meramente à sociedade do falante.

Dreier resiste o “relativismo de grupos de falantes”, observando a consequência indesejada de que não nos permite expressar desacordos morais com o sistema moral de nossa sociedade. Ao invés disso, ele argumenta que “relativismo do falante” pode acomodar amoralistas, pois “pode permitir que o sistema moral de uma pessoa depende de fatores extrínsecos àquela pessoa” (1990: 21). Ao serem relativos à sociedade do falante, termos morais portanto são relativos ao falante. Mas termos morais *não* seriam necessariamente relativos às atitudes ou adesões morais do falante, o que seria abandonar a teoria indexical *da atitude do falante*. Dreier confessa que não sabe exatamente como a indexação de termos morais funciona (1990: 25), e ele escreve: “Me parece que, dado o contexto de uso, determinamos o conteúdo do termo moral ao buscar o candidato mais adequado para um sistema moral naquele contexto” (1990: 23). Enquanto ele “começou permitindo que os reais estados de motivação do falante determinassem completamente o sistema moral relevante”, a consideração do amoralista o empurra-o em direção a uma abordagem mais contextual. Dreier então reconhece que o problema da cancelabilidade refuta a alegação de que convenções

linguísticas prendem termos morais à atitude do falante, mas não está certo quanto ao que deve ser proposto em seu lugar a não ser que deve ser sensível ao contexto. Parece que a teoria da conversação fornece uma resposta satisfatória.

6. Por que não convenções?

A última questão que vou abordar tem a ver com a sugestão que eu atribuo a Copp de que é apenas de se esperar que convenções prendam termos morais a atitudes do falante. Há dois pontos aqui: o primeiro se trata da “própria ideia” de implicatura convencional em geral; a segunda se trata da forma como a dinâmica do juízo moral pode moldar convenções linguísticas. Meus argumentos nesta seção são consideravelmente mais especulativos e menos decisivos do que considero os argumentos anteriores.

Possivelmente a objeção mais devastadora às teorias da convenção é a alegação de Kent Bach (1999) de que o fenômeno da implicatura convencional não existe: todas as supostas instâncias dela podem ser explicadas como instâncias de outros aspectos da língua. Estou inclinado a concordar, embora eu não tenha nem a habilidade nem o espaço para argumentar o caso aqui. A mera nuvem de suspeita razoável sobre a implicatura convencional, no entanto, é ainda outra razão para preferir a teoria da conversação em detrimento de teorias da convenção, portanto vou brevemente indicar por que nós devemos levar essas preocupações a sério, demonstrando como a própria explicação de Barker da implicatura convencional falha em estabelecer sua existência. Ele escreve:

(1) Até a vovó está bêbada.

Ao declarar (1) o falante U está reportando que a vovó está bêbada e *implica* através de “até” que a) para U há uma escala de probabilidade subjetiva na qual (vovó está bêbada) tem uma probabilidade menor que (N está bêbada) para

qualquer N em uma classe contextualmente determinada;
(b) outros estão bêbados. [2000: 269]

Considere (1) declarado na circunstância de que a vovó é uma alcoólatra, atualmente bêbada, e a classe de comparação relevante consiste inteiramente de pessoas que bebem moderadamente. Se a palavra “até” contribui 1(a) para o conteúdo asserido, então (1) é falso. Mas embora (1) seja inasserível para qualquer um que tenha a escala de probabilidade correta, Barker alega, ela não é falsa. Por quê? Por quê “não se pode dizer”

(2) É falso que até a vovó está bêbada.

Portanto, por ser falso que é falso que até a vovó está bêbada, deve ser verdade que até a vovó está bêbada, embora não seja asserível. E a propriedade de não ser asserível sem falsidade de (1) é evidência para a implicatura convencional.

Esse argumento é muito apressado. A única margem que Barker dá para se pensar que (2) é falso (daí que (1) é verdadeiro) é que (2) é algo que “não se pode dizer”. Isso é meramente dizer que (2) é inasserível – e é uma alegação crucial do argumento de que a propriedade de não ser asserível não indica falsidade! Então, por tudo o que Barker provou, poderia ser (2) e não (1) que é verdade mas não asserível: nessa visão alternativa, (2) tem como conteúdo asserido que não é o caso que

[(1*) vovó está bêbada] e 1(a) e 1(b).

1(a) pode então ser parte do conteúdo asserido de (1), e então o caso de Barker para a implicatura convencional falha. Isso pode parecer um mero truque de mágica. Verdade sem asserabilidade tem sido movida de uma afirmação para a sua negação mas não foi eliminada. Não é suficiente, para fornecer evidências para a implicatura convencional, demonstrar o espaço

entre verdade e asseribilidade? Não – devido à (menos dúbia) variação de implicaturas conversacionais. Enquanto a suposta verdade sem asseribilidade de (1) respalda a implicatura convencional devido ao papel de “até” (o que supostamente contribui para o conteúdo por convenções semânticas que não contribuem com as condições de verdade da declaração), a verdade sem asseribilidade proposta de (2) não respalda a implicatura convencional, já que todo o conteúdo semanticamente determinado está presente e explicado nas condições de verdade. O caráter enganoso de asserir (2) derivaria então da complexidade, e não da escassez, das condições de verdade, o resultado sendo que a audiência não tem pistas verbais para qual parte do conteúdo está sendo negada, e é provável que assumo, incorretamente, que é o conteúdo primário (1*).

Há, portanto, uma interrogação pairando sobre a existência da implicatura convencional. Mas Copp não se sente intimidado por isso: como já observei, ele acha que é apenas de se esperar que tais convenções surjam para termos morais. Eu encerrarei criticando a visão de convenções linguísticas nas quais essa opinião parece estar baseada, e cautelosamente sugerindo que seria surpreendente caso tais convenções existissem. Por que se considera que tais convenções apareçam naturalmente? Eu posso discernir, na descrição de Copp, pistas para dois modelos diferentes de como convenções linguísticas se desenvolvem. Primeiramente ele escreve que “serviria aos objetivos do discurso moral se uma convenção se desenvolvesse [...]” (2001: 34). Isso sugere que um modelo de convenções linguísticas como regras sociais para o bem social, parecida em tipo com as regras da moralidade e da etiqueta. Em segundo lugar, ele escreve (citando Frege), “convenções podem se desenvolver para governar usos de um termo quando este ‘tem sido usado constantemente em casos do mesmo tipo’” (2001: 35). Isso sugere um modelo de convenções linguísticas como normas sociais, no sentido (menos normativo)

do que é *normal* ou costumeiro. Eu não acredito que nenhum dos dois modelos capture a natureza das convenções linguísticas,

No modelo da normalidade, que parece primário na descrição de Copp, a conexão regular entre asserções morais com adesão morais (onipresentes) leva (através de uma forma de associação) esta adesão a ser presa no próprio significado desses termos.¹⁸ É portanto sugerido que termos morais têm coloração na forma como acredita-se que tenham os pejorativos: usos constantes de uma palavra acompanhada por uma atitude de desprezo permeando a palavra em si (como “crioulo”). Duvido que a associação normal por si mesmo determine convenções semânticas, as quais são em parte *limitações* sobre como uma palavra pode ser usada significativa e inteligivelmente. No entanto, a língua tem uma dinâmica natural de facilitar ao invés de restringir o que pode ser expressado, de forma que palavras são ferramentas adequadas por natureza para contextos novos, impensados e estendidos. A aplicabilidade de uma palavra, portanto, não é restringida pela falta de exercício. (A palavra “casa”, por exemplo, não é semanticamente limitada a denotar somente construções que estejam na Terra, mesmo que nós nunca a tenhamos usado para denotar um objeto extraterrestre. O fato de que em algumas comunidades o nome “Satã” nunca é usado, a não ser no sentido negativo não resultou em tal sentido negativo se transformar em uma condição sobre o uso linguisticamente apropriado do nome.) Ao contrário, a aplicabilidade das palavras é restringida apenas pela necessidade de que seja possível discernir mais precisamente entre as diferentes informações as quais os falantes se encontram na necessidade transmitir. Suponhamos analogamente que a palavra “valente” denota uma característica em geral benéfica para a sociedade, por isso seu uso é

¹⁸ Eu não acredito que as declarações amoralistas sejam tão raras quanto esse mecanismo requereria. O verdadeiro amoralista, que sempre está completamente indiferente (no máximo) à moralidade, talvez possa ser raro, mas o ato de fala moral desacompanhado de aprovação é lugar-comum.

normalmente acompanhado por aprovação.¹⁹ Não haveria nenhum motivo no desenvolvimento de uma convenção que limite o uso apropriado de “valente” apenas para aqueles que aprovam da característica. Se alguém quisesse, estranhamente, expressar desaprovação de bravura, seu uso seria pouco comum, mas ainda em acordo com o significado da palavra. Da mesma forma, o fato de que a atribuição de bem moral é normalmente acompanhado de adesão e aprovação moral é certamente razão para se esperar que alguém usando termos morais também adere àqueles padrões morais, mas não é razão para se pensar que o termo estaria semanticamente restrito nessa forma – de fato, não é razão para pensar que não estaria.

Voltando-se para o modelo das regras sociais, deve ser observado que convenções linguísticas são significantemente diferentes de convenções morais. Não é, como no caso da moral, que a não conformidade é vista com maus olhos ou socialmente desencorajada, mas sim que a não conformidade *não é entendida*. Os objetivos sociais que determinam a semântica são objetivos de comunicação, e um uso da linguagem é um mau uso, semanticamente, se é um uso que não é completamente inteligível: um uso que deixa dúvidas dado o que alguém significa pelo uso dessa palavra. O que o falante poderia estar querendo transmitir através dela e por que o falante a escolheu? O contraponto de Copp de que o amoralista faz mau uso dos termos morais deve, então, ser lida como a alegação de que a escolha de palavras do amoralista é confusa para interpretadores competentes: por que essa frase e

¹⁹ Essa descrição de “valente” é controversa. Em uma interpretação alternativa, valentia é uma característica que apenas pode ser percebida de perspectivas afetivas específicas, e sua extensão parece sem forma quando independente desses afetos. No entanto, essa não é uma tese semântica, e não bloqueia aqueles sem o afeto necessário de usar a palavra. Em outra interpretação, a intenção de aprovar é integral à semântica de “valente”. (Ver *Protágoras* 350B de Platão: certas atribuições de coragem são rejeitadas com a justificativa de que “se [tais atos fossem corajosos] coragem seria algo pelo qual sentir-se envergonhado.”) Mas nesse caso, a atitude parcialmente determina a extensão do termo, e por isso sua contribuição ao conteúdo asserido. Assim nenhuma alternativa enfraquece meu argumento de que implicaturas convencionais de atitude são supérfluas.

não outra? (Considere aqui o caso de “até”: se *U* não quer contrastar a tendência da vovó com relação ao álcool de forma favorável com a dos outros, por que motivo ele usa a palavra “até”?) No entanto, como já vimos, Copp alega que a declaração do amoralista é tanto um mau uso da linguagem moral e, ainda assim, “completamente inteligível como uma asserção” (2001: 18). Ele defende essa possibilidade com a observação de que um mau uso de uma palavra por parte de um falante pode ser completamente inteligível se a audiência reconhecer que “ele não sabe ou não consegue lembrar” da palavra apropriada. Claro, estamos supondo que o amoralista seja completamente competente com a língua; a explicação de Copp de sua inteligibilidade é a *não existência* de uma palavra apropriada para o amoralista usar.

Acredito que tais convenções seriam insustentáveis. Supondo que a teoria da convenção de Copp fosse verdade, como um amoralista comunicaria o mesmo conteúdo asserido que nós quando falamos “*T* é bom”? Não há uma forma concisa para que o faça, exceto por também declarar “*T* é bom”. Os intérpretes competentes ficariam confusos com sua escolha de palavras? De forma alguma: dado o que o amoralista quer comunicar, o que ele diz é a forma mais simples e mais natural de dizê-lo. E isso sugere que a nossa suposição é falsa: não pode haver uma convenção linguística que requeira a adesão do falante ao padrão moral relevante, e a teoria da convenção deve ser falsa. As convenções que a teoria propõe teriam o efeito, contrário à função básica das convenções linguísticas, de *limitar* expressões de informação – com efeito, de silenciar amoralistas. Considere novamente os pejorativos, uma descrição de implicatura conversacional a qual eu reconheço não ser sem mérito. Se pejorativos de fato têm coloração de forma convencional, é em parte porque existem na linguagem como *alternativas* a outras palavras com a mesma denotação. Por que um falante chamaria uma pessoa de “bicha” ao invés de homossexual, ou de “crioulo” ao invés de negro ou afro-americano? Essa escolha de terminologia é explicada pela intenção de expressar desprezo com relação a um

grupo. Se esses pejorativos fossem a única forma eficaz que tivéssemos em nossa língua de denotar seus referentes, eles não mais seriam convencionalmente pejorativos. Já que não temos sinônimos eficazes para termos morais, eu então acredito que é no mínimo bastante improvável que eles sejam convencionalmente ligados a uma adesão moral.

Embora tenha achado que a teoria da convenção de conteúdo atitudinal do juízo moral de Copp tem vantagens sobre a teoria da convenção apoiada pela teoria indexical de Barker/Dreier, eu concluo que ela falha da mesma forma, e que uma implicatura conversacional é superior a ambas. Como considere apenas duas versões da teoria da convenção, pode-se pensar que eu ainda não demonstrei que essa teoria da conversação é superior a todas as teorias da convenção, e pode-se também pensar que com o meu foco em implicatura eu não me referi à factibilidade de outros tipos de descrições convencionais. Mas meus argumentos quanto à separabilidade e cancelabilidade e minhas observações de conclusão sobre a natureza das convenções linguísticas devem ser igualmente eficazes contra qualquer teoria de implicatura convencional ou outro tipo de teoria baseada em convenção. Acredito que meu caso fornece maiores razões para rejeitar a alegação básica do expressivismo como falsa: a expressão de atitudes do falante não é uma função semântica de atos de fala morais.²⁰

Referências

- BACH, K. The Myth of Conventional Implicature. **Linguistics and Philosophy**, v. 22, p. 327–66, 1999.
- BARKER, S. J. Is value content a component of conventional implicature? **Analysis**, v. 60, n. 3, p. 268–79, 2000.

²⁰ Sou grato a Kent Bach, Stephen F. Barker, Bridget Copley, David Copp, Gary Ebbs e um revisor anônimo do *Philosopher's Imprint* por seus conselhos.

- COPP, D. **Morality, Normativity, and Society**. New York: Oxford University Press, 1995.
- COPP, D. Realist-Expressivism: A Neglected Option for Moral Realism,. **Social Philosophy and Policy**, v. 18, p. 1-43, 2001.
- DREIER, J. Internalism and Speaker Relativism. **Ethics**, v. 101, p. 6-26, 1990.
- EDWARDS, P. **The Logic of Moral Discourse**. Glencoe, IL: Free Press, 1955.
- FINLAY, S. The Conversational Practicality of Value Judgement,. **The Journal of Ethics**, v. 8, p. 205-223, 2004.
- GIBBARD, A. **Wise Choices, Apt Feelings**. Cambridge, MA: Harvard University Press, 1990.
- GRICE, H. P. **Studies in the Way of Words**. Cambridge, MA: Harvard University Press, 1989.
- HARE, R. M. **The Language of Morals**. Oxford: Oxford University Press, 1952.
- JACKSON, F.; PETTIT, P. A Problem for Expressivism. **Analysis**, v. 58, p. 239-51, 1998.
- MACINTYRE, A. **After Virtue**. 2. ed. Notre Dame, Indiana: Notre Dame University Press, 1984.
- MACKIE, J. L. **Ethics: Inventing Right and Wrong**. Harmondsworth: Penguin Books, 1977.
- PUTNAM, H. **Reason, Truth and History**. Cambridge: Cambridge University Press, 1981.
- STEVENSON, C. C. **Ethics and Language**. New Haven: Yale University Press, 1944.
- ZIFF, P. **Semantic Analysis**. Ithaca, NY: Cornell University Press, 1960.