



**Luan Rafael Marques de Oliveira**

## **LIVRE ARBÍTRIO E SORTE CONSTITUTIVA**

**Uma visão cética do livre arbítrio e da responsabilidade moral**

### **Dissertação de Mestrado**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Filosofia pelo Programa de Pós-Graduação em Filosofia do Departamento de Filosofia da PUC-Rio.

Orientador: Prof. Dr. Ludovic Soutif



**Luan Rafael Marques de Oliveira**

## **LIVRE ARBÍTRIO E SORTE CONSTITUTIVA**

**Uma visão cética do livre arbítrio e da responsabilidade moral**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Filosofia pelo Programa de Pós-Graduação em Filosofia do Departamento de Filosofia da PUC-Rio.

**Prof. Ludovic Soutif**

Presidente

Departamento de Filosofia - PUC-Rio

**Prof. Pedro Merluzzi**

UFABC

**Prof. Rodrigo Guerizoli**

UFRJ

Rio de Janeiro, 30 de setembro de 2022

Todos os direitos reservados. A reprodução, total ou parcial do trabalho é proibida sem a autorização da universidade, do autor e do orientador.

### **Luan Rafael Marques de Oliveira**

Graduou-se em Filosofia na UFRJ (Universidade Federal do Rio de Janeiro) em 2019. Durante a graduação, fez Iniciação Científica pelo PIBIC de 2016-2019.

#### Ficha Catalográfica

Oliveira, Luan Rafael Marques de

Livre arbítrio e sorte constitutiva : uma visão cética do livre arbítrio e da responsabilidade moral / Luan Rafael Marques de Oliveira ; orientador: Ludovic Soutif. – 2022.

75 f. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Filosofia, 2022.

Inclui bibliografia

1. Filosofia – Teses. 2. Livre arbítrio. 3. Sorte moral. 4. Ceticismo. I. Soutif, Ludovic. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Filosofia. III. Título.

CDD: 100

## Agradecimentos

Ao CNPq (número de processo: 131622/2020-0) e à PUC-Rio, pelo auxílio e o apoio institucional concedidos, condições de possibilidade para a realização deste trabalho

Ao meu orientador Professor Ludovic Soutif pelo seu apoio intelectual e emocional no processo de produção desta dissertação, sobretudo na sua reta final.

Aos professores membros da banca examinadora pela sua paciência, acolhimento e *feedback* no processo de avaliação.

Aos meus amigos e colegas do grupo de estudos Saguinus que proveram um tom de diversão ao trabalho intelectual nesses anos difíceis e caóticos de mestrado.

À minha companheira Milena por me doar seus ouvidos, seu acalento, seu olhar para o mundo e tudo de positivo que me falta.

À minha mãe Valdete por sua tolerância e abertura intelectual, sua aceitação das minhas várias mudanças durante meu processo de me encontrar profissionalmente e seu generoso sustento, em todos os sentidos do termo.

Ao meu pai Almir por me ter proporcionado uma despedida positiva, amorosa.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## Resumo

de Oliveira, Luan Rafael Marques; Soutif, Ludovic. **Livre arbítrio e sorte constitutiva: uma visão cética do livre arbítrio e da responsabilidade moral.** Rio de Janeiro, 2022. 75 p. Dissertação de Mestrado - Departamento de Filosofia, Pontifícia Universidade Católica do Rio de Janeiro.

Neste trabalho, defendo a tese de que o livre arbítrio, entendido como o controle necessário para a responsabilidade moral baseada no mérito, não existe, pois é impossível. A tese é um desenvolvimento da visão de Galen Strawson que baseia a impossibilidade da responsabilidade última na impossibilidade da autodeterminação. Aqui, defendo uma abordagem ao problema que conecta os seguintes temas: livre arbítrio, sorte moral e autocriação, mantendo que o fato necessário da sorte constitutiva é o que torna impossível de satisfazer a condição de fonte última do controle necessário para a responsabilidade moral. Minha estratégia argumentativa é mostrar como as tentativas de satisfazer e de rejeitar essa condição falham.

## Palavras-chave:

Livre arbítrio, sorte moral, ceticismo.

## Abstract

de Oliveira, Luan Rafael Marques; Soutif, Ludovic. **Free will and constitutive luck: a skeptical view of free will and moral responsibility.** Rio de Janeiro, 2022. 75 p. Dissertação de Mestrado - Departamento de Filosofia, Pontifícia Universidade Católica do Rio de Janeiro.

In this work, I defend the thesis that free will, understood as the control necessary for merit-based moral responsibility, does not exist, for it is impossible. The thesis is a development of Galen Strawson's view, which bases the impossibility of ultimate responsibility on the impossibility of self-determination. Here, I defend an approach to the problem that connects the following themes: free will, moral luck and self-creation, holding that the necessary fact of constitutive luck is what makes the ultimate sourcehood condition for the control required for moral responsibility impossible to satisfy. My argumentative strategy is to show how attempts both to satisfy and reject this condition fail.

## Keywords:

Free will, moral luck, skepticism.

## Sumário

1. Introdução	8
2. Experiência e prática	11
2.1. Definido “livre arbítrio”	11
2.2. Obstáculos à posição cética	13
2.3. O ônus da prova	15
2.4. A importância da prática	16
2.5. Tipos de responsabilidade moral	18
2.6. Tipos de juízo moral	25
2.7. Agência e deliberação	31
2.8. Atitudes reativas	32
2.9. O sentimento de liberdade	34
3. Três problemas: liberdade, sorte e autocriação	36
3.1. Responsabilidade moral	37
3.2. Livre arbítrio e determinismo	39
3.3. Sorte moral	43
3.4. Autocriação	53
4. Compatibilismo, incompatibilismo e sorte constitutiva	57
4.1. Ação baseada em razões e incompatibilismo	57
4.2. Sorte constitutiva e compatibilismo	60
4.3. Sorte constitutiva e incompatibilismo	63
5. Considerações finais	67
Referências bibliográficas	70

## Introdução

Em 1924, Nathan Leopold e Richard Loeb, dois ricos e brilhantes jovens estudantes judeus norte-americanos influenciados pela filosofia de Nietzsche, tentaram cometer o crime perfeito. Eles sequestraram e assassinaram o menino de 14 anos, Bobby Franks. Foram pegos e, dado que as provas eram incontestáveis, confessaram a culpa. O advogado americano Clarence Darrow, famoso por ter estado na defesa no "julgamento do macaco" de Scopes (um marco na guerra cultural norte-americana do evolucionismo *vs.* criacionismo), também estava na defesa de Leopold e Loeb neste julgamento, e o seguinte é a sua declaração final para que seus clientes fossem poupados da pena de morte:

Vossa excelência, quase fico com vergonha de falar a respeito disto. Mal posso imaginar que estamos no século XX. E, não obstante, há homens que seriamente dizem que, pelo que a Natureza fez, pelo que a vida fez, pelo que o ensinamento fez, Vossa Excelência deveria enforcar estes meninos. (...)

Por que eles mataram o pequeno Bobby Franks? Não por dinheiro, não por despeito; não por ódio. Eles o mataram como matariam uma aranha ou uma mosca; pela experiência. Eles o mataram porque foram *feitos* assim. Porque, em algum lugar nos infinitos processos que contribuem para a formação do menino ou do homem, algo falhou e estes rapazes desafortunados estão sentados aqui odiados, desprezados, degredados, com a comunidade clamando pelo seu sangue. (Em: SOMMERS, 2004.)<sup>1</sup>

Isso chegou a trazer lágrimas aos olhos do juiz. Não admira, tendo sido Darrow um orador fenomenal. Como Jerry Coyne (2013) observa, no julgamento de Scopes, ele "realizou um interrogatório brutal de William Jennings Bryan sobre a veracidade da Bíblia", muito bem dramatizado no filme *O Vento Será Tua Herança* (*Inherit the Wind*) de 1960.

Continuando, Darrow pergunta

---

<sup>1</sup> Todas as citações de fontes originais em inglês neste trabalho foram traduzidas pelo autor da dissertação.



O que este menino tem a ver com isso? Ele não era seu próprio pai; ele não era sua própria mãe, ele não era seus avós. Tudo isso lhe foi entregue. Ele não rodeou a si mesmo de governantas e de riqueza. Ele não fez a si mesmo. E, não obstante, ele é obrigado a pagar. (Em: VIHVELIN, 2013, p. 49.)

Darrow não está negando que seus clientes cometeram o ato criminoso; tampouco que o ato criminoso tem alguma conexão com as suas mentes. Na visão de Darrow (podemos especular pelas suas citações), os réus mataram a vítima pois acreditavam que o ato homicida era um meio para o seu objetivo de experimentar certa sensação ou emoção, e a sua preocupação com a vida ou o bem-estar da vítima não era, para dizer o mínimo, grande o bastante para tornar o ato, da perspectiva deles, uma má ideia. Logo, ele acreditava que o ato se conectava a desejos e crenças na mente dos seus clientes, mas chama atenção para o fato de que eles não criaram as suas próprias mentes (“Ele não fez a si mesmo”). As condições que formaram os indivíduos que eles eram “lhe[s] fo[ram] entregue[s]”; eram externas a eles, estavam fora do seu controle. Darrow acreditava que essas considerações eram de alguma forma relevantes para a questão da responsabilidade moral dos seus clientes (serem “obrigado[s] a pagar”), e assim, para a adequação do veredicto da pena de morte.

Em questões de livre arbítrio e responsabilidade moral, é natural se preocupar que, conforme é angariado mais conhecimento científico das causas do comportamento humano, mais perto “o espectro da exculpação” parece chegar, sorrateiramente (DENNETT, 1984, pp. 156–165). Em algumas ocasiões, é difícil não atribuir certa dissimulação a esse tipo de empreendimento. Por exemplo, esta citação de Steven Pinker ilustra o ponto, com exemplos que fariam bom material para piadas de advogado:

Cientistas cognitivos às vezes são abordados por advogados de defesa criminal na esperança de que um píxel errático numa varredura cerebral possa exonerar o seu cliente (um cenário que é espirituosamente representado no romance *Brain Storm* de Richard Doelling). Quando uma equipe de geneticistas descobriu um gene raro que predispunha os homens de uma família a surtos de violência, um advogado defendendo um réu de um assassinato sem relação com a família argumentou que o seu cliente poderia ter o gene também. Se for o caso, o advogado argumentou, “as ações dele podem não ter sido um produto de total livre arbítrio”. Quando Randy Thornhill e Craig Palmer argumentaram que o estupro é uma consequência de estratégias reprodutivas masculinas, outro advogado contemplou usar a teoria deles para defender suspeitos de estupro. (PINKER, 2002, p. 176.)

No entanto, não vejo razão para crer que Clarence Darrow não era sincero. Este trabalho tenta defender que o advogado tinha certa razão.

---

O livre arbítrio não existe. Embora nossa vida prática e subjetiva dificulte aceitar isso, creio que os argumentos fazem a ideia nos vir bem naturalmente. No capítulo 2, tento lidar com essa resistência que considerações a ver com a prática ("Como vamos lidar com gente se comportando mal?") e a experiência ("Mas eu *sinto* que sou livre, e você também.") levantam contra uma ponderação sóbria dos argumentos. Daí, no capítulo 4, poderemos, na medida do possível, deixar a teoria dominar e ir aonde os argumentos nos levarem.

No capítulo 3, se encontra uma conceptualização do problema do livre arbítrio a partir da qual a argumentação do capítulo 4 se desenvolve. Lá, esclareço em que sentido do termo afirmo que o livre arbítrio não existe, o qual está intimamente vinculado a práticas de responsabilização baseadas em *mérito*. Lá, também deixo claro que a tese da inexistência do livre arbítrio não se baseia em nenhuma tese relacionada a causação determinista – ou indeterminista, aliás. Defendo que o livre arbítrio, de fato, é *impossível*, independentemente de fatos contingentes relativos ao caráter das leis naturais.

Ao meu ver, a impossibilidade do livre arbítrio antes se dá pelo fato de que não podemos assumir o controle sobre nossos traços constitutivos, fatores que fazem de nós quem somos. Em outras palavras, o livre arbítrio é impossível porque não podemos nos *autocriar*, porque necessariamente temos (boa ou má) sorte de sermos quem somos. A partir dessa visão, é crucial conectar o debate relativo ao livre arbítrio e o determinismo com discussões na literatura de sorte moral e problemas que surgem quando tentamos fazer sentido da ideia de autocriação. Com esse alargamento de perspectiva, é minha expectativa que façamos ao menos algum progresso neste perene, fascinante e desconcertante tópico.

## 2

### Experiência e prática

O livre arbítrio não existe. Pode ser difícil de acreditar, mas há um sentido (um sentido *importante*) da expressão “livre arbítrio” em que essa proposição é verdadeira e pode ser justificada como tal. Nesta dissertação, tenho como objetivo defender essa tese, começando por, neste capítulo, fazer um esforço para dissipar essa incredulidade inicial e esclarecer em que sentido é que não temos livre arbítrio. Para isso, teremos que nos demorar um pouco na filosofia moral, o que para alguns poderá parecer um tratamento superficial, mas para outros, uma tangente despropositada. Embora o tema do livre arbítrio não deva ser desatrelado totalmente dela, o foco desta dissertação não é a filosofia moral, mas a metafísica, o que me impede de dar o tratamento pleno às questões éticas. Quanto à questão de por que sequer falar de ética, tenho confiança que a resposta logo ficará clara.

#### 2.1

##### Definindo “livre arbítrio”

O que “livre arbítrio” quer dizer? A resposta a essa questão deve ter duas virtudes: ser esclarecedora e ser inclusiva. Ela deve ser inclusiva porque dar uma explicação sobre o que é o livre arbítrio é exatamente uma das tarefas das diferentes teorias sobre o assunto, e, em consequência, uma definição precisa arriscaria ser tendenciosa contra ou a favor de uma teoria ou outra. Por outro lado, ao tentar evitar isso arriscamos dar uma definição estéril demais. Definir é preciso, no entanto, no mínimo para saber sobre o que os filósofos estão discordando.

Dois caminhos estão disponíveis para esse propósito. O primeiro é o de apelar à **experiência** subjetiva de agência livre: isto é, a como é sentirmos, em primeira pessoa, que somos os autores derradeiros dos nossos pensamentos e ações. O citável ensaísta inglês Samuel Johnson disse sobre o livre arbítrio que

“[t]oda a teoria é contra a liberdade da vontade; toda a experiência, a favor dela” (BOSWELL, 2008, p. 947), e também que “Senhor, (...) *sabemos* que a nossa vontade é livre, e *aqui* a conversa se encerra” (BOSWELL, 2008, p. 411). Dizer que sabemos que temos livre arbítrio pressupõe que é verdade que o temos, de modo que é muito natural interpretar a última frase como uma flagrante petição de princípio da parte do Dr. Johnson, mas há uma interpretação mais caridosa: de que, no mínimo, temos indiscutivelmente a *experiência* de liberdade e que ela é tão irresistível a ponto de tornar a discussão sobre a sua veracidade ou tola, ou irrelevante.<sup>2</sup>

O Dr. Johnson da interpretação caridosa está seguramente certo, não sobre os méritos da questão da veracidade da experiência da liberdade, mas pelo menos sobre a experiência ser real *enquanto experiência*; pois mesmo os que negam a verdade do seu *conteúdo* devem aceitar que existe um *sentimento* generalizado de liberdade, e assim encontramos aqui o primeiro modo satisfatório de entender a expressão “livre arbítrio”.<sup>3</sup> As condições de satisfação do conceito de livre arbítrio seriam herdadas das condições de veracidade da fenomenologia da agência livre (veja HORGAN, 2015), e em consequência, alegar que o livre arbítrio não existe seria alegar que essa experiência é não verídica. É difícil dizer plenamente quais são essas condições, mas um bom começo é pela ideia de que elas consistem em nossas ações serem causadas por nossos pensamentos (desejos, crenças, intenções). Mas isso não é tudo – exige-se que esses pensamentos sejam por sua vez originados por “nós” (mais pessoalmente, por “mim”), enquanto um alguém mental *independente* desses pensamentos particulares (veja STRAWSON, 2010, pp. 82-83, pp. 283-288).

O segundo caminho para uma definição satisfatória de livre arbítrio é apelar à nossa **prática** de responsabilização moral. Julgamo-nos uns aos outros

<sup>2</sup> A. J. Ayer, logo no início de *Freedom and Necessity* (1954) toca nessa interpretação de petição de princípio e, logo após, em algo próximo da interpretação que chamei de mais caridosa.

<sup>3</sup> Sam Harris (2012) escreve bombasticamente: “*A própria ilusão do livre arbítrio é uma ilusão*”, e ele diz coisas como “Não somente não somos tão livres quanto pensamos que somos — não nos sentimos tão livres quanto sentimos que somos”, o que leva a pensar que ele nega o que afirmo sobre a “realidade” da experiência da liberdade. No entanto, o que ele quer dizer com isso não é que não temos uma experiência cujo conteúdo é inverídico, mas que essa experiência não se sustenta diante de uma investigação introspectiva. Como ele coloca: “O nosso senso da nossa liberdade resulta de não prestarmos uma atenção especial a como é sermos nós. No momento em que prestamos atenção, é possível enxergar que o livre arbítrio não se encontra em lugar nenhum, e que a nossa experiência é perfeitamente compatível com essa verdade.” Voltaremos a esse tema sobre a introspecção no final do capítulo.

como culpáveis ou louváveis pelas nossas ações moralmente incorretas ou corretas, e isso nos leva a vermos uns aos outros como alvos apropriados de expressões de certas emoções que os filósofos chamam de *atitudes reativas* (como ressentimento, indignação e gratidão) (STRAWSON, 1962) e como possíveis *merecedores* de culpa ou elogio, punição ou recompensa, pelas nossas infrações ou acatamentos de obrigações morais. Essas práticas são bem difundidas e (aparentemente) irresistíveis na nossa espécie, e como elas envolvem, em vários aspectos, intencionalmente fazer indivíduos sofrerem, sentimos que elas precisam e podem ser justificadas. É de praxe na filosofia moral manter que uma condição para essa justificação é termos certo tipo de *controle* sobre nossas ações; logo, aqui o conceito de livre arbítrio se referiria ao controle sobre nossas ações que é necessário para que essas práticas sejam justificadas, e, em consequência, alegar que o livre arbítrio não existe seria alegar que essas práticas jamais são justificadas.

## 2.2

### Obstáculos à posição cética

Chamar atenção a esses dois pontos, o da experiência e o da prática, é importante por duas razões: (1) eles expõem que o problema em questão não é apenas um debate acadêmico no alto da Torre de Marfim, divorciado de repercussões para a vida comum; (2) eles formam os principais obstáculos para a aceitação do ceticismo quanto ao livre arbítrio, como denominarei a tese de que o livre arbítrio não existe.<sup>4</sup>

---

<sup>4</sup> Falarei mais a fundo sobre a terminologia no próximo capítulo, mas cabe aqui um esclarecimento inicial. Pode causar estranheza chamar de “cética” uma posição que nega a existência de algo, dado que associamos a palavra primeiramente com o ceticismo pirrônico, que simplesmente nos insta a “suspender o juízo”, e não afirma nem nega nada. No entanto, temos razão para empregar o termo além do domínio da epistemologia, empregando-o também no que toca a alegações metafísicas. Pois é bom lembrar que, na metaética, J.L. Mackie (1977) chamou de “ceticismo moral” a sua posição de que não há propriedades morais objetivas. Sim, também a chamou de “teoria do erro”, assim poderíamos, por analogia, chamar as posições que negam a existência do livre arbítrio e/ou da responsabilidade moral de “teoria do erro do livre e/ou responsabilidade moral”. Outro candidato é o chamado “eliminativismo”, seguindo a filosofia da mente (outra opção é “nihilismo”, que não pega muito bem no quesito das Relações Públicas). Mas penso que o termo “ceticismo sobre livre arbítrio e/ou responsabilidade moral” já faz parte do uso dos filósofos da área. Note-se que a *Stanford Encyclopedia of Philosophy* tem um verbete significativamente

Em primeiro lugar, mantém-se que negar uma crença baseada numa experiência tão irresistível é epistemicamente inaceitável, ou no mínimo suspeito. Em segundo lugar, acredita-se que abrir mão da nossa crença de que merecemos ser culpados e louvados pelos nossos atos (i) nos impossibilitaria de lidar apropriadamente com o mau comportamento humano, (ii) tiraria o fundamento das emoções necessárias para a existência de relações interpessoais que são cruciais ao florescimento humano (STRAWSON, 1962) e (iii) tornaria a moralidade em geral uma ficção. A isso, os críticos poderiam acrescentar que a própria experiência de liberdade coloca, além de um problema epistêmico, um problema prático à adoção da posição cética. É plausível pensar (assim o argumento começaria) que não podemos nos livrar da experiência de liberdade, e logo (assim o argumento continuaria) a aceitação simultânea da posição cética nos impossibilitaria de harmonizar as imagens científica e manifesta do homem no mundo – a qual, como Wilfrid Sellars (1991) celebrenemente defendeu, é a tarefa do filósofo. Essa tarefa de harmonização pode ter uma significância existencial, na qual Owen Flanagan parece tocar aqui:

[As duas imagens] importam porque têm um papel crucial na maneira como nos autoconcebemos. A maneira como nos autoconcebemos importa enormemente para a maneira como a vida aparece ou é sentida em primeira pessoa. Assim, a maneira como essas imagens se engrenam ou não se engrenam importa para a maneira como estamos indo, subjetivamente. (FLANAGAN, 2007, p. 5.)

Como Flanagan (2007, p. 6) observa, uma alternativa à harmonização das duas imagens é “apostar tudo em uma imagem em detrimento da outra”. No entanto, se não podemos nos livrar duma experiência considerada ilusória, a adoção simultânea da visão teórica considerada verdadeira implicaria uma experiência instável ou fraturada de nós mesmos e do nosso lugar no mundo. O crítico perguntaria retoricamente: “O ceticismo nos obriga a simplesmente lidar com isso?”

---

intitulado “Skepticism About Moral Responsibility”, da autoria de Gregg Caruso (2018). Uma última coisa: também chamarei a tese de “visão cética”, “posição cética” ou simplesmente “ceticismo”, abreviadamente.

## 2.3

### O ônus da prova

Os problemas experiencial e prático fundamentam a visão de que não basta ao cético oferecer argumentos que, embora não decisivos, fundamentem a sua tese como mais plausível que às de seus rivais – ele tem um ônus de prova bem maior. No entanto, há razões para acreditar que esses problemas podem ser contornados. Sobre o problema da experiência, no tocante à epistemologia, ele só seria um grande empecilho se a fenomenologia tivesse um papel justificador *inderrotável*, que não pode ser sobrepujado por meio de uma rota epistêmica mais indireta – isto é, por meio de argumentos de que as suas condições de veridicidade não são satisfeitas. Mas como Sam Harris observa:

A introspecção não oferece pista alguma de que a nossa experiência do mundo à nossa volta, e de nós mesmos dentro dele, depende de mudanças de voltagem e de interações químicas ocorrendo dentro das nossas cabeças. E não obstante, um século e meio de ciência do cérebro afirma que esse é o caso. (HARRIS, 2010, p. 158.)<sup>5</sup>

Não só a experiência subjetiva pode estar alheia a certos fatos íntimos sobre nós que descobrimos pela teoria; ela pode nos levar a tomar como fato coisas que de fato não o são. Os exemplos mais clássicos e mais interessantes vêm do campo da visão (talvez porque, para primatas, ela pareça o sentido mais confiável): simplesmente pesquise no Google “espiral de Fraser”, “ilusão de Müller-Lyer” e “ilusão do tabuleiro de damas”. A conclusão é: um apelo à experiência não pode excluir categoricamente uma justa apreciação dos argumentos colocados contra a verdade do que ela afirma sobre o mundo.

Pode se discutir se, para contornar a fenomenologia, precisamos prover, além de argumentos sólidos, uma explicação cética plausível sobre as origens causais da experiência da agência livre (isto é, algo que explicasse como ela poderia ter surgido sem ser verídica) – ou se isso seria apenas um bônus teórico da posição cética. No mínimo, tal explicação a tornaria mais *completa*, e para tal ela

<sup>5</sup> Ele também diz: “Eu realmente penso bastante sobre o cérebro humano, mas quase nunca me ocorre que eu realmente tenho um”, numa palestra gravada em vídeo. Disponível em <<https://www.youtube.com/watch?v=ITTxTCz4Ums>>. Acesso em: 16 de nov. de 2021.

teria que passar por várias estações na ferrovia explanatória: as bases cognitivas e neurais da experiência de agência livre, sua formação durante o desenvolvimento, suas funções culturais e/ou evolutivas, etc. Não é meu propósito aqui dar um encaminhamento a essa pesquisa, que de fato já está bem encaminhada (veja CARUSO, 2012), mas os limites dessa dissertação me possibilitam pelo menos apontar que Espinosa estava na direção certa quando disse que “os homens opinam serem livres porquanto são cômicos de suas volições e de seus apetites, e nem por sonhos cogitam das causas que os dispõem a apetecer e querer, pois delas são ignorantes” (ESPINOSA, 2015, p. 111).

Visto que a experiência de liberdade permeia tanto da nossa vida, muitos dos nossos juízos devem pressupor a existência do livre arbítrio, implicando assim que a visão cética é altamente revisionista. Mas, ao contrário do que podemos ser levados a pensar, isso não diz tanto contra o ceticismo, pois, como Neil Levy (2011, p. 6) observa, “na medida em que apenas a teoria interessa, devemos seguir até onde o argumento leva, (...) por mais radical que seja a revisão nas nossas crenças anteriores”. Como ele também observa (LEVY, 2011, p. 7), essa generalidade da pressuposição do livre arbítrio certamente expandirá a gama de considerações relevantes para avaliá-la, dando ao cético mais trabalho para justificar a sua posição, mas isso não é o mesmo que *um aumento nos próprios padrões argumentativos* exigidos.

## 2.4

### A importância da prática

Quanto ao lado prático da história? Poderíamos começar citando Hume, quando disse que:

Em discussões filosóficas, não há método de raciocínio mais comum, mas também mais condenável, que tentar refutar uma hipótese a pretexto de suas consequências perigosas para a religião e a moral. Quando uma opinião nos leva a absurdos, é certamente falsa; mas não é certo que uma opinião seja falsa porque tem consequências perigosas. Argumentos como esses, portanto, deveriam ser evitados, porque não ajudam em nada na descoberta da verdade, servindo



apenas para tornar odiosa a pessoa do adversário. (HUME, 2000, p. 445.)

Óbvio, não estou acusando aqueles que defendem o argumento prático de tentarem destruir a reputação dos outros, apenas transmitindo a ideia de que Derek Parfit tinha alguma razão quando disse que:

Não é o objetivo de uma teoria ser aceita. Se personificarmos teorias, e fingirmos que elas têm objetivos, o objetivo de uma teoria não é ser aceita, mas ser verdadeira, ou ser a melhor teoria. (PARFIT, 1987, p. 24.)

Logo, contanto que foquemos na função que as crenças têm de *rastrear a verdade*, considerações sobre consequências práticas são irrelevantes. No entanto, quando o assunto é *adotar* uma teoria e deixá-la guiar nossas ações, é muito mais plausível que consequências práticas devam ser consideradas. Logo, é apropriada uma palavra sobre o assunto.

Antes de mais nada, vale dizer que não é evidente que as consequências da adoção do ceticismo devam ser, no cômputo geral, mais negativas que positivas, mas pressuponhamos, em prol do argumento, que de fato o sejam. Aqui teríamos um conflito entre a racionalidade teórica (normas que nos dizem o que fazer para termos crenças fidedignas) e a racionalidade prática (normas que nos dizem o que fazer para atingir nossos objetivos, morais ou prudenciais). Pode-se debater sobre se termos crenças fidedignas tem algum valor além da sua utilidade em conseguir-nos o que desejamos ou em satisfazer nossa curiosidade (veja YUDKOWSKY, 2015, p. 15-18), mas caso tenha tal valor, seria certamente uma má situação ficarmos entre as alternativas de ceder a uma ilusão ou ver nossas vidas perderem qualidade em termos de outros bens. A escritora de livros de fantasia P. C. Hodgell disse: “Aquilo que pode ser destruído pela verdade deve ser” (em YUDKOWSKY, 2015, p. 12). Essa atitude estoicamente sóbria tem sua própria beleza, mas parece pressupor que não muito de valor seria destruído – o que é possível, mas não certamente garantido. Portanto, é importante investigar as maneiras como a adoção da visão cética *não* implicaria as consequências (tudo considerado) negativas que muitos creem que implicaria.<sup>6</sup>

<sup>6</sup> No que se segue, eu trato do que a adoção do ceticismo pode nos levar a crer e de como ele pode nos levar a agir *justificadamente*. Assim, não discuto sobre o que podemos ser (ou até provavelmente seremos) levados a crer e fazer, individual e socialmente, devido à nossa confusão sobre o que realmente diz a posição cética ou à nossa ignorância sobre o que é possível, praticamente falando.

## 2.5

### Tipos de responsabilidade moral

Começemos por lembrar o que lemos logo no início deste capítulo: que “há *um sentido* (...) da expressão ‘livre arbítrio’ em que [a] proposição [de que o livre arbítrio não existe] é verdadeira e pode ser justificada como tal” (ênfase acrescida). Dizer que estamos justificados em negar o livre arbítrio, *em um sentido da expressão*, deixa aberta a possibilidade de que haja outros sentidos de “livre arbítrio” em que não estamos tão justificados em negá-lo. Indo direto ao ponto, estou propondo que “livre arbítrio” tenha vários sentidos (ou que se refira a algo que admite vários tipos ou aspectos), e que, em pelo menos um desses sentidos, não temos livre arbítrio, enquanto que, em outros sentidos, o temos – certa, provável ou possivelmente, dependendo das considerações empíricas em jogo.

Como já vimos, os filósofos comumente aceitam que a responsabilidade moral possui uma *condição de controle*, e é natural denominá-la “livre arbítrio”. Proponho também, assim, que “responsabilidade moral” tenha vários sentidos (ou se refira a algo que admite vários tipos ou aspectos – pelo bem da brevidade, deixarei essas qualificações implícitas no que segue), e é um passo natural daí fixar a referência dos diferentes sentidos de “livre arbítrio” como a condição de controle da responsabilidade moral em cada diferente sentido.<sup>7</sup> Assim, negar a satisfação da condição de controle da responsabilidade moral tomada *num certo sentido* não necessariamente implicaria negar a satisfação da condição de controle da responsabilidade moral tomada *num outro sentido*. Nessa visão, poderíamos negar a justificação de certas práticas de responsabilização moral, e ainda assim afirmar ou deixar em aberto a justificação de outras, conforme considerações particulares exigirem.<sup>8</sup>

<sup>7</sup> Correspondentemente, pode haver várias rotas fenomenológicas para a fixação dos variados referentes aqui também. Sobre isso, veja a discussão de (STRAWSON, 2010, 90-94) sobre o “compatibilismo natural” (veja o próximo capítulo para a definição de “compatibilismo”).

<sup>8</sup> Derk Pereboom (2015) faz a seguinte observação sobre o tema dos vários sentidos de responsabilidade moral:

Frequentemente, termos que possuem uma longa história de uso têm uma variedade de sentidos. Ainda que um termo referidor com uma história tal tenha originalmente tido apenas um referente específico, ele é dado com o tempo a ser aplicado a referentes semelhantes, porém

E quais seriam então esses diferentes sentidos de responsabilidade moral? Começamos pelo sentido que mais importa nesta dissertação, aquele em que afirmo que jamais somos moralmente responsáveis porque não temos o requerido controle sobre as nossas ações. Geralmente julgamos que as pessoas *merecem* louvor ou culpa por agirem de modo moralmente correto ou incorreto, e julgamos apropriado fazê-lo *simplesmente* porque elas cometeram o ato correto ou incorreto ao terem plena consciência do que estavam fazendo.<sup>9</sup> Isso é o que Derk Pereboom (2014, p. 2) chamou de **responsabilidade moral de mérito básico**.

Como já tocado acima, sentimos que juízos de culpabilidade e louvabilidade deste tipo nos autorizam a expressar *atitudes reativas*: como o orgulho, a gratidão, a culpa, o ressentimento, a indignação, as quais direcionamos ao que consideramos seus alvos legítimos (que, aliás, também podem ser nós próprios). Também sentimos que eles nos autorizam a recompensar e punir, prover bônus e impor ônus, respectivamente, a benfeitores e a malfeitores – sendo o último, à propósito, algo que julgamos ser um papel legítimo do Estado. Foquemos na punição estatal por um momento aqui, pois ela é um dos aspectos mais moralmente salientes desse tipo de responsabilização. Ela envolve a imposição do que é chamado de um *tratamento duro* sobre o malfeitor: prejudicá-lo ou privá-lo de algum bem. Isso é colocar as coisas abstratamente, e jamais é bom falar apenas em abstrações quando o assunto é o sofrimento humano. Logo, me sinto na obrigação de citar a seguinte observação de Victor Trados neste ponto:

A punição é provavelmente a coisa mais medonha que estados democráticos modernos fazem sistematicamente aos seus cidadãos. Todo estado democrático moderno aprisiona milhares de infratores todo ano, privando-os da sua liberdade, causando-lhes uma grande medida de dano psicológico e às vezes físico. Relacionamentos são destruídos, empregos são perdidos, o risco de um infrator ser machucado por outros infratores é aumentado, e tudo isso a altos custos para o estado. (Em: CARUSO, 2021, pp. 111.)

Para tornar as coisas ainda mais concretas, tome como exemplo a prática do *confinamento solitário*, amplamente utilizada em presídios dos EUA e de outros países, que consiste em isolar detentos em “cubículos sem janelas e à prova

---

distintos, assim adquirindo diferentes sentidos. Os termos “responsabilidade moral” e “culpa” plausivelmente têm esse perfil.

<sup>9</sup> Incluo aqui omissões também.

de som por 23 a 24 horas a cada dia, às vezes por décadas” (CARUSO, 2021, p. 18). Não é preciso ser muito versado nas pesquisas em psicologia positiva para ter noção do sofrimento psicológico que tal privação do contato social causa em membros da nossa espécie, mas é digno de nota como o tal sofrimento pode ter uma fonte diametralmente oposta: na socialização *exagerada* das celas claustrofobicamente superlotadas das prisões brasileiras. Claro, a superlotação é mais um efeito colateral do que um traço de *design* da punição, e em alguns casos a justificativa dada para o confinamento solitário não é exatamente a punição. No entanto, isso nos mostra um outro aspecto preocupante da nossa ideia de que criminosos são meritoriamente culpáveis pelos seus crimes: mesmo quando não nos insta a impor sofrimento e privação, ela parece nos autorizar a não levar em conta o bem-estar de um indivíduo como normalmente fazemos.<sup>10</sup>

A alcunha de *mérito básico* aqui tem como propósito chamar a atenção para o fato de que, nesta concepção de responsabilidade moral, a justificativa para tratar um agente duramente (ou de uma forma correspondentemente positiva, em casos de louvabilidade) não apela a considerações sobre as consequências ulteriores de assim agir (como, por exemplo, desincentivar o mau comportamento futuro). Segundo a responsabilidade moral de mérito básico, o sofrimento, a privação, a ira moralizada impostos a um malfeitor são tomados como um *bem em si mesmo*. Essa concepção de culpa e louvor como mérito independente de considerações *consequencialistas* é capturada vividamente no experimento mental de Kant da sociedade ilha prestes a se dissolver:

Ainda que uma sociedade civil fosse dissolvida pelo consentimento de todos os seus membros (por exemplo, se um povo habitante de uma ilha decidisse se separar e se dispersar pelo mundo todo), o último assassino remanescente na prisão teria primeiro de ser executado, para que cada um lhe tenha feito o que os seus atos merecem e que a culpa de sangue não se atrele ao povo por não ter insistido nessa punição; pois, do contrário, o povo poderá ser considerado como colaborador nessa violação pública da justiça. (KANT, 1996, p. 474.)

A aprovação à pena de morte é demasiado a-contemporânea, e o condicional no final é confuso e difícil de levar à sério se tomado ao pé da letra

<sup>10</sup> Essa atitude faz sentido à luz da seguinte observação de Neil Levy: “Alguém que mantém que agentes podem ser moralmente responsáveis por maus atos ou omissões não está comprometido em manter que esses agentes merecem sofrer; antes, está comprometido em manter que tais agentes não merecem mais a [plena] proteção de um direito ao qual, se fosse de outro modo, eles estariam autorizados: um direito contra ter seus interesses descontados no cálculo consequencialista.” (LEVY, 2011, p. 3.)

(como Dan Dennett (em DENNETT e CARUSO, 2021) observa, estará Kant dizendo que os ilhéus devem matar uma pessoa para salvar a própria reputação, ou para manter suas próprias mãos limpas? E isso não seria um raciocínio consequencialista egoísta?). Mas subtraindo essas estranhezas, Kant aqui nos mostra bem no que consistiria a responsabilidade moral de mérito básico, pois, presumivelmente, punir o assassino não serviria para proteger ninguém, educar ninguém, deter ninguém, tampouco reforçar o respeito pela lei na consciência de ninguém.

Antecipando o argumento principal deste trabalho, é plausível que a justiça demande no mínimo o seguinte: um agente não pode ter sido, em última análise, *sortudo* ou *azarado* por ter agido moralmente correta ou incorretamente, a fim de que ele seja responsabilizado no sentido do termo em questão. Como Pereboom (2014, p. 2) coloca “[u]m agente ser moralmente responsável por uma ação neste sentido consiste em ela ser *dele* de tal modo que ele mereceria ser culpado se entendesse que era errada” (ênfase acrescida), e essa metáfora de *pertencimento* é bastante apta. Parece que uma ação deve pertencer, num sentido particularmente forte, a um agente, para ele ser justa e merecidamente culpável ou louvável pela mesma. Esse sentido forte de pertencimento foi observado por A. J. Ayer quando escreveu que “[d]e uma maneira confusa, atribuímos a nós mesmos e aos outros o que foi por vezes descrito como um poder de *autodeterminação*” (AYER, 2000, p. 96, ênfase acrescida),<sup>11</sup> um poder de criar a si mesmo que, como veremos, é difícil de estabelecer que seja sequer *possível*.

Enquanto, na responsabilidade moral de mérito básico, o tratamento diferenciado (positivo ou negativo) é justificado pelo seu valor intrínseco, numa concepção alternativa de responsabilidade moral ele se justifica exatamente pelo seu valor *instrumental*, e a mesma não exige a autodeterminação do agente como condição de controle, mas a sua simples capacidade de *responder a razões* (FISCHER e RAVIZZA, 1998). Esse sentido de responsabilidade moral é chamado por Pereboom (2014, p. 132) de **answerability**, cuja prática pode muito bem ser descrita, em bom português, como “pedir satisfações”. Uma citação de Hobbes por Daniel Dennett (1984, p. 34) ilustra sucintamente a ideia: “Se as vontades das pessoas são determinadas, perguntam a Hobbes, ‘porque

---

<sup>11</sup> “Originação” é um outro nome que pode ser dado a este poder.

representamos razões a elas?’ Ele responde: ‘porque assim pretendemos fazê-las possuir a vontade que não possuem.’”

Quando uma pessoa comete um ato que julgamos moralmente incorreto, sentimo-nos justificados em exigir-lhe uma justificação, em convidá-la a avaliar as suas ações com base em normas morais. Caso julguemos que a justificativa dada não é adequada, sentimo-nos no direito também de convidá-la a refletir sobre o que essa ação diz sobre as suas intenções e caráter, e além disso a *agir de acordo*: fazendo um pedido de desculpas e comprometendo-se a reformar seu comportamento e disposições no futuro. Claro, refletir sobre si mesmo de uma maneira objetiva produz naturalmente certo desconforto psicológico no *Homo sapiens*. No entanto, na análise de Pereboom (2014, p. 134), tal atribuição se justifica não porque ela, de algum modo, agrega por si só valor ao universo, mas porque a prática que a envolve serve aos bens legítimos da nossa *proteção* contra as consequências nocivas do comportamento imoral, da nossa *reconciliação* com o malfeitor e da *formação moral* do caráter do mesmo.

Como já observei, tal prática de responsabilização teria como condição de controle a responsividade a razões, que é independente da supracitada autodeterminação, e que, logo, ela não é afetada pelo o argumento desta dissertação.<sup>12</sup> Muitos filósofos defendem que a responsividade a razões basta como condição de controle para a responsabilidade de *mérito básico*, mas, como Pereboom bem coloca, essa conexão não é nada clara, e ela parece muito mais adequada à responsabilidade enquanto *answerability*:

É digno de nota que há uma conexão transparente entre a sensibilidade a razões e a noção prospectiva da responsabilidade, e que uma ligação semelhante é faltosa na responsabilidade moral que envolve as atitudes reativas e o mérito básico. A sensibilidade a razões é, por exemplo, evidentemente exigida para que um agente seja submetido à formação moral com base na apresentação de razões, enquanto que o vínculo entre a sensibilidade a razões e a pressuposição do mérito básico ou das atitudes reativas é opaco. (PEREBOOM, 2014, p. 136.)

<sup>12</sup> Meu objetivo aqui é estabelecer que outros sentidos de livre arbítrio e responsabilidade moral não são comprometidos pela negação da autodeterminação e do mérito básico. Há outros argumentos independentes do meu ceticismo para se crer que, por exemplo, de fato não somos tão responsivos a razões como cremos ser. Esses argumentos se baseiam principalmente na pesquisa sobre o *automatismo*, o *inconsciente adaptativo* e o *situacionismo* na psicologia e psicologia social, que levam a crer que as causas das nossas ações não nos são tão transparentes quanto pensamos, e por vezes divergem das razões e justificações que damos para elas. Veja (CARUSO, 2012) para uma boa exposição dessa literatura, e (NAHMIAS, 2007) para uma resposta otimista a ela. Veja a nota seguinte para um comentário semelhante relativo à existência de traços de caráter robustos.

Uma segunda concepção de responsabilidade moral distinta e independente do mérito básico é aquela que envolve a atribuição de traços de caráter positivos ou negativos (virtudes ou vícios) com base nas ações de um agente (ou seja, julgar que estas refletem profundamente quem o agente é). Esta é tipicamente chamada de responsabilidade enquanto **atributibilidade** (MCKENNA e PEREBOOM, 2016, p. 13), e Hume parece chamar atenção à sua prática neste trecho do seu Tratado:

É evidente que, quando elogiamos uma determinada ação, consideramos apenas os motivos que as produziram, e tomamos a ação como signo ou indicador de certos princípios da mente e do caráter. A realização externa não tem nenhum mérito. Temos de olhar para o interior da pessoa para encontrar a qualidade moral. Ora, como não podemos fazê-lo diretamente, fixamos a nossa atenção na ação, como signo externo. Mas a ação é considerada apenas um signo; o objeto último de nosso elogio e aprovação é o motivo que a produziu. (HUME, 2009, p. 517.)

Bem, a tese de tese de Hume de que “[a] realização externa não tem nenhum mérito” pode parecer radical demais, mas não o é a ideia básica de que nos importa atribuir vícios e virtudes aos outros, e que as suas ações são uma base para essa atribuição. Essa é uma prática bem natural e necessária nas nossas vidas, ajudando de um modo particular na nossa conduta para com os outros: por exemplo, não confiamos uma tarefa importante a alguém que julgamos carente em conscienciosidade, mesmo que não o culpemos por ser assim ou lhe exijamos que mude. Mas será que ela nos compromete com a realidade do mérito básico? É plausível que não.

Considere como exemplo um caso em que um homem mata dois jovens inocentes de uma maneira particularmente cruel. Para sermos concretos, digamos que o homem rouba o carro dos jovens, levando-os a uma área afastada; ele atira num deles pelas costas após ter mentido que ninguém se machucaria; ele persegue colina abaixo o outro jovem que tenta fugir, e atira nele quatro vezes; após o homicídio, ele gira a pistola no ar, dá gargalhadas e come o lanche dos dois adolescentes (WATSON, 1987).

Ao ficarmos sabendo do ocorrido, é natural julgarmos o ato um crime horrendo, considerando o assassino viciosamente frio, sentindo uma indignação profunda, culpando-o pelo ato imoral e esperando pela sua merecida punição. No entanto, digamos que logo ficamos sabendo da criação totalmente aquém do

desenvolvimento normal que teve o assassino. Digamos, novamente para sermos concretos, que a sua vida tenha começado após sua mãe ter sido chutada no estômago pelo seu pai, antecipando assim o seu parto; que essa violência conjugal tenha sido causada pelo ciúme do pai, que desconfiava que a mãe tivesse tido o filho com outro homem; que os maus tratos contínuos fizeram a mãe culpar o seu filho pela sua desgraça, cultivando um ódio por ele e despejando a sua ira sobre ele; por exemplo: quando ele buscava seu carinho ela chegava a afastá-lo a chutes.

Sendo informados disso, também é natural considerarmos o homem uma vítima da sua má sorte e ficamos mais pesarosos do que indignados (com o que se passou com o homem assassino, assim como com o jovem assassinado), e começamos a nos questionar sobre se ele realmente merece aquela culpa e punição pelo que fez. Parece que esse questionamento é totalmente compatível com a convicção mantida sobre o que as ações do assassino indicam sobre ele: um caráter cruel, carente de compaixão e apropriada consideração pelos outros. Parece fazer todo sentido perguntar: “Certamente, foi um crime cruel, cometido por alguém cruel, mas podemos mesmo culpá-lo por isso? Será mesmo justo puni-lo por isso?” Assim, a *questão do mérito* é ao menos conceitualmente distinta da *questão da atribuição* (uma questão que, como suposto, nem sequer foi levantada). Poderíamos concluir até que *em todo caso* em que uma pessoa é atribuível por um mau ato ela merece culpa pelo mau ato, mas isso seria uma tese *substancial*, não meramente derivada dos *conceitos* de atribuição e mérito. Mas poderíamos muito bem concluir o contrário: que as extensões dos dois conceitos de fato não são exatamente idênticas.

Assim como com a *answerability* e a sua condição de controle de um agente ser responsivo a razões, há uma conexão transparente entre a atributibilidade e a condição de controle de um agente possuir traços de caráter que produzem ações. Parece que isso é tudo o que é exigido, em termos de controle, para atribuições serem justificadas. Ademais, a satisfação da condição de ter traços de caráter é totalmente independente da satisfação da condição metafisicamente suspeita de uma autodeterminação que escapa à sorte, e, portanto, aquela também não é comprometida pelo meu argumento aqui.<sup>13</sup>

---

<sup>13</sup> Assim como a responsividade a razões e a *answerability*, meu objetivo aqui é estabelecer que traços de caráter e a atributibilidade não são comprometidos pela negação da autodeterminação e do mérito básico. De modo análogo ao que disse na nota anterior, há argumentos contra a existência de traços de caráter robustos, baseados na pesquisa sobre o situacionismo na psicologia



Em resumo, aceitar sentidos variados de responsabilidade moral abre as portas para mantermos a prática de pedir satisfações aos outros e de avaliar o seu caráter, ainda que pensemos ilegítimo julgá-los merecedores de culpa e louvor, punição e recompensa. É plausível que a responsabilidade moral tenha uma condição de controle, mas esta é diferente conforme se varia o sentido daquela. Enquanto, como veremos nos próximos capítulos, é adequado exigir uma autocriação que escape à sorte para merecidamente culparmos e elogiarmos as pessoas, tal coisa não é necessária para avaliarmos o seu caráter com base nas suas ações, e para exigirmos que elas próprias avaliem o seu caráter com base nas suas ações e ajam de acordo com tal reflexão moral. Para as últimas práticas, só precisamos que as pessoas tenham traços de caráter e sejam responsivas a razões.

## 2.6

### Tipos de juízo moral

É claro que todos esses tipos de práticas de responsabilização dependem de juízos morais; não só a prática de culpar e louvar, mas também de pedir satisfações e de julgar o caráter dos outros. Para convocarmos alguém a refletir sobre os seus atos, eles devem ser julgados moralmente *errados*; para justificarmos essa prática, devemos julgá-la como conducente a algum *bem*; para moralmente atribuírmos uma ação aos traços de caráter de alguém, estes devem ser julgados *virtuosos* ou *viciosos*. No entanto, será verdade que juízos normativos sobre ações (de certo e errado), juízos axiológicos sobre resultados (de bom e mau) e juízos areteicos sobre traços de caráter (de virtuoso e vicioso) dependem de juízos de responsabilidade sobre ações (de louvável e culpável)?

À primeira vista, não. E podemos ver isso pela intuição. Consideremos novamente o caso do assassino acima. Podemos considerar um *mau* resultado a

---

social, vindos do desafio situacionista à ética das virtudes (seu *locus classicus* é (DORIS, 2002). Para uma defesa empírica da existência de traços de caráter, veja a discussão sobre o modelo dos cinco fatores da personalidade (ou os “Grandes Cinco”) de (NETTLE, 2009). Para uma discussão de como o modelo se relaciona às virtudes, veja (PARIS, 2017).

morte do jovem, *errado* o homicídio e *vicioso* o caráter do homem, mas ainda assim nos perguntarmos se ele é culpável pelo assassinato. Mas podemos também chegar à resposta por uma rota mais teórica. Sem a intenção de restringir-me ao utilitarismo ético, mas apenas a título de um exemplo simples, num mundo em que ninguém é culpável pelas suas ações, nada impede as pessoas (e outros seres sencientes) de terem seu bem-estar aumentado ou diminuído, que bondade e maldade acompanhem essas mudanças, que a normatividade das ações sejam ao menos em parte uma função de como elas produzem essas mudanças (NORCROSS, 2006) e que, talvez, vícios e virtudes sejam também entendidos em função da tendência de traços de caráter a produzirem essas mudanças (DRIVER, 2001).

No entanto, as dificuldades para a visão cética não são tão facilmente contornadas, pois há uma objeção mais forte à ideia de que ela pode reter a maioria dos nossos juízos morais. É muito plausível a ideia de que, a fim de termos a *obrigação* de agir de modo diferente do que de fato agimos, é necessário que tenhamos o *poder* de agir de modo diferente do que de fato agimos. Esse princípio do “dever implica poder”, juntamente com a ideia de que não “podemos fazer diferentemente” se o determinismo for verdadeiro, acarreta que, se o ceticismo está certo, não temos obrigações morais (ao menos obrigações de fazer coisas que de fato não fazemos).

Esse “poder fazer diferentemente” supracitado é uma habilidade classicamente considerada uma condição necessária para o livre arbítrio: a chamada condição das “possibilidades alternativas”, e tem marcado presença no debate mais celebrenemente no **argumento da consequência**, mais celebrenemente defendido por Peter van Inwagen (1983). O argumento tem como propósito estabelecer a incompatibilidade entre o livre arbítrio e o **determinismo**, e este denomina a seguinte tese. Dois mundos possíveis com o mesmo estado inicial e as mesmas leis naturais deterministas possuirão necessariamente a mesma história subsequente (em outras palavras, eles serão idênticos, de seus respectivos *big bangs* às suas respectivas Mortes Térmicas). Um corolário disso é que mundos com leis deterministas com alguma diferença de estado num dado momento necessariamente possuem ou leis diferentes, ou um passado em cada momento diferente. David Lewis resume as duas ideias aqui:

Um sistema de leis determinista é um sistema tal que, sempre que dois mundos possíveis obedecem às leis perfeitamente, ou eles serão exatamente semelhantes por todo o tempo, ou então eles não serão exatamente semelhantes por nenhuma extensão do tempo. Eles são iguais sempre ou nunca. Eles não divergem, coincidindo perfeitamente nos seus segmentos iniciais, porém não em seguida; tampouco eles convergem. (LEWIS, 1979.)

O argumento da consequência parte dos seguintes pressupostos plausíveis sobre *habilidades* ou *poderes*:

1. Ninguém pode mudar o passado e ninguém pode violar as leis.
2. Considere duas condições: A e B. A é um fato e que A acarreta B também é um fato. Ninguém pode mudar nem A, nem que A acarreta B. Concluimos que B é ou será um fato, e ninguém pode mudar isso. (Por exemplo,: (i) em bilhões de anos o Sol explodirá, e ninguém pode mudar isso; (ii) se em bilhões de anos o Sol explodir, toda a vida na Terra se extinguirá, e ninguém pode mudar isso; (iii) logo, em bilhões de anos a vida na Terra se extinguirá, e ninguém pode mudar isso (baseado em VAN INWAGEN, 1983, p. 98).

E a conclusão é que ninguém pode agir de modo diferente do que age, se o mundo for determinista. Isso é porque, pressupondo o determinismo, as leis e o passado são o que são e elas acarretam o presente e o futuro, e o fato de que ninguém pode mudar esse acarretamento, tampouco as leis ou o passado, nos leva a concluir que ninguém pode ou poderá mudar nada que acontece ou acontecerá no universo. Por exemplo, eu escolho levantar a mão, mas penso que poderia não fazê-lo, mantendo-a abaixada. Digamos que o determinismo é verdadeiro, e que assim é uma contradição que as leis e o passado sejam os mesmos e eu mantenha a mão abaixada. Isso implica que, se eu poderia ter mantido minha mão abaixada, eu poderia ter violado as leis ou mudado o passado, o que é um absurdo. Parece que sou assim levado a crer que na realidade eu *não* poderia ter mantido a minha mão abaixada, ao menos se o determinismo for verdade.

Várias réplicas foram dadas a esse argumento, atacando seja o pressuposto (1), seja o pressuposto (2), acima. Uma delas, que ataca (2), envolve defender uma *análise condicional* de “poder” compatível com o determinismo (e.g., MOORE, 1903). Tal análise mantém que um agente poder agir de outro modo consiste em ser verdade que *caso* ele *quisesse* agir de outro modo, ele *agiria* de outro modo. Se alegações sobre poderes forem condicionais ocultos, poderes serão compatíveis

com o determinismo, porque condicionais são compatíveis com o determinismo. Isso é porque o determinismo só lida com a conexão entre estados do mundo, não com quais estados são dados, incondicionalmente falando.

Outra rota que pode ser tomada para não se aceitar a conclusão da incompatibilidade entre o determinismo e o livre arbítrio por meio do argumento da consequência é negar o princípio das possibilidades alternativas de que ele faz uso. Foram elaborados experimentos mentais em que uma pessoa (chamemo-la “vítima potencial”) não pode fazer diferentemente, porém, mesmo assim, é intuitivamente julgada como livre e responsável pela sua ação. Nos cenários em questão há um chamado “interventor contrafactual” que garantirá (por alguma forma de controle neural) que a vítima potencial aja conforme ele quer, mas apenas *se* houver indícios de que ela não irá assim agir. Alternativamente, o interventor não fará nada se houver indícios de que a vítima, de fato, irá agir conforme ele quer. No fim das contas, a vítima age conforme a vontade do interventor *por conta própria*, tornando desnecessária a intervenção deste. Por não ter de fato intervindo, o interventor não participa da sequência atual de eventos na causação da ação da vítima; o processo causal parece ser do tipo perfeitamente corriqueiro, e é isso o que parece nos dar a intuição de que a vítima é perfeitamente livre e responsável. No entanto, o fato de que o interventor está pronto para intervir, caso a situação exija, nos dá a ideia de que a vítima não pode agir de um modo diferente, tornando assim o cenário um contraexemplo à condição das possibilidades alternativas para o livre arbítrio.

Esse é um dos famosos *casos de Frankfurt* (FRANKFURT, 1969), que também têm gerado bastantes réplicas e repercussões, a principal das quais sendo a mudança de foco das possibilidades alternativas para a *fonte última* como a condição necessária ao livre arbítrio que o determinismo comprometeria. Dessa perspectiva, mesmo que, para um agente ser livre e responsável, não seja necessário que ele possa agir diferentemente, a suas ações não podem ter uma origem última externa à sua própria vontade.

Neste trabalho, seguirei a tradição de focar na condição de fonte última em detrimento da condição das possibilidades alternativas, e de fato não basearei a defesa do ceticismo no determinismo. No entanto, a discussão sobre possibilidades alternativas e o determinismo é importante no que toca ao “poder”

em “dever implica poder”. O cético pode seguir as seguintes alternativas a fim de preservar juízos deônticos, isto é, juízos de obrigação:

1. aceitar o princípio do “dever implica poder”, embora mantendo que ainda temos possibilidades alternativas (mesmo que isso não implique termos o livre arbítrio necessário para sermos culpáveis/louváveis), (i) seja porque o determinismo é de fato falso, (ii) seja porque temos uma análise condicional de “poder” compatível com o determinismo, (iii) seja porque temos uma concepção *epistêmica* de “poder” compatível com o determinismo;
2. aceitando ainda o princípio do “dever implica poder”, segurar as pontas e negar também que tenhamos obrigações morais (ao menos obrigações de fazer o que de fato não fazemos);
3. simplesmente negar o princípio e ficar livre para afirmar deveres morais.

Mantenho que a saída mais promissora é algo como (1, iii), isto é, a solução epistêmica da primeira alternativa, numa leitura de “dever”, e algo como (2), isto é, a solução de negar deveres morais, numa outra leitura de “dever”.

Primeiramente, quais seriam as consequências temidas de se negarem obrigações morais? A mais preocupante, acho eu, é a de que a normatividade das ações seria perdida: isto é, não poderíamos julgar atos como moralmente corretos ou incorretos. Mas será isso verdade? Há duas razões para pensar que não. Primeiro, já vimos que é plausível que acarretar bons resultados fundamente a normatividade das ações. Também, mesmo que não tenhamos um relato teórico da normatividade, juízos intuitivos dão algum apoio à visão da manutenção de juízos normativos sem juízos deônticos. Por exemplo, se uma pessoa insana ou compulsiva mata alguém, parece que podemos julgar a ação errada sem aceitar o juízo de que ela deveria não ter matado, talvez porque, aceitando o “dever implica poder”, julgamos que ela não poderia não ter matado (PEREBOOM, 2014, p. 145). Preservando assim juízos de certo e errado, abrimos o caminho para rejeitar juízos de dever moral sem pagar um tão grande preço.

No entanto, como Pereboom já observou, há usos e usos da linguagem de dever moral. Em um desses usos, quando dizemos coisas como “Você deveria fazer A”, estamos fazendo uma *demanda* de que o ato específico seja cometido. Nesse sentido, é plausível que, se um agente não acatou à demanda, ele deve ter tido o poder de agir diferentemente (isto é, ter acatado a demanda), a fim de que a

demanda seja legítima. Num uso bem diferente da linguagem de “dever”, quando dizemos coisas como “Você deveria fazer A”, estamos fazendo não uma demanda, mas antes uma *recomendação*. Estamos, com efeito, fazendo um juízo axiológico (“Seria bom que você fizesse A”) ao fazer a recomendação. É plausível que, para que essa recomendação seja legítima, o agente deva poder atender à recomendação em algum sentido. Esse sentido, no entanto, não é o metafisicamente robusto, mas o *epistêmico* – é plausível pensar que devemos apenas não estar *certos*, com base nas nossas crenças de fundo, de que ele vá ou não vá atender à recomendação, ainda que acreditemos, pressupondo o determinismo, que ele está determinado ou a atender ou a não atender à recomendação. Também é plausível que devamos acreditar que a recomendação fará uma contribuição causal em formar a motivação do agente a atender à recomendação, o que, à propósito, torna o juízo ainda mais consistente com a ideia do determinismo, e também com uma prática de responsabilização focada na realização de bens futuros (PEREBOOM, 2014, pp. 138-146).

Em suma, um uso legítimo da linguagem de “dever” (um uso que pode ser chamado de *axiológico-recomendatório*) é retido na adoção da visão cética, mantendo ainda uma leitura epistêmica do princípio do “dever implica poder”, independentemente de alegações mais controversas sobre a compatibilidade de obrigações morais, em sentidos mais robustos, com a negação do livre arbítrio em questão.

## 2.7

### Agência e deliberação

O tema das possibilidades alternativas também se conecta com outra objeção geralmente levantada à praticidade da posição cética: a ideia de que ela comprometeria a nossa agência efetiva no mundo, e mais especificamente, a nossa

deliberação. Antes de mais nada (ainda que a visão cética não esteja comprometida com a verdade do determinismo), é bom deixar claro que o determinismo não é o mesmo que o *fatalismo*. Histórias que têm o fatalismo como tema, como a do Rei Édipo, sugerem que, embora possamos percorrer vários caminhos antes de o nosso destino chegar (talvez na tentativa vã de tentar evitá-lo), ele *irá* chegar. Assim, no fatalismo há uma conotação de independência causal entre o passado e o futuro. Nada poderia estar mais distante da tese do determinismo, que implica que os caminhos que tomamos no passado são condições causais (suficientes) para o nosso futuro, embora o caminho específico que tomarmos tenha, por sua vez, condições causais (suficientes) no passado anterior, e assim por diante. Assim, nada no determinismo impede que uma ação que realizamos dependa causalmente, em parte, da nossa intenção de realizá-la, formada pela nossa deliberação, por sua vez baseada nos nossos objetivos e informações (embora, novamente, esses próprios objetivos e informações tenham condições causalmente suficientes além do nosso controle) – de tal modo que, *se* os nossos objetivos ou informações fossem diferentes, a nossa deliberação *seria* diferente, a nossa intenção formada *seria* diferente e, por consequência, a nossa ação *seria* diferente.

No entanto, pode-se aceitar que deliberações são eficazes no mundo, mas questionar se um agente que crê na posição cética possa ser coerente ao deliberar. O ceticismo, novamente, não está obrigado a afirmar o determinismo, mas se considera consistente com ele; assim, vamos pressupor a tese a título de argumentação. A objeção seria a de que, ao cremos no determinismo, cremos que não temos possibilidades alternativas na ação, enquanto que, ao deliberarmos, necessariamente cremos que há possibilidades alternativas abertas a nós (que podemos fazer *A* ou *~A*). Assim, ao deliberarmos enquanto cremos no determinismo, acreditamos que *temos* e que *não temos* possibilidades alternativas, o que é incoerente. Logo, como Van Inwagen coloca, “qualquer um que negue a existência do livre arbítrio deve, inevitavelmente, contradizer a si mesmo com uma monótona regularidade” (VAN INWAGEN, 1983, p. 160).

Será mesmo verdade? Não o seria se o sentido de possibilidades alternativas excluído pelo determinismo fosse diferente daquele necessitado pela deliberação. Caso esse sentido seja o sentido *epistêmico*, a objeção da incoerência perderia a força (PEREBOOM, 2014, pp. 105-115). No entanto, ao passo que é

claro que a circunstância de não termos certeza de que faremos A ou  $\sim A$  é uma condição *necessária* para deliberarmos entre A e  $\sim A$ , é mais duvidoso que seja *suficiente*. Entretanto, o que parece ser exigido a mais não é que acreditemos na indeterminação metafísica, mas que acreditemos que qualquer intenção que formarmos com base na nossa deliberação (seja a de fazer A, seja a de fazer  $\sim A$ ) será eficaz em produzir a respectiva ação (PEREBOOM, 2014, pp. 116-123). E isso, como vimos ao falar rapidamente da análise condicional de “poder”, é totalmente compatível com o determinismo, pois é totalmente compatível com o determinismo que, *se* formarmos a intenção de A, faremos A, e *se* formarmos a intenção de  $\sim A$ , faremos  $\sim A$ .

## 2.8

### Atitudes reativas

Atitudes reativas é a expressão filosófica, cunhada por Peter Strawson (STRAWSON, 1962), para se referir à gama de emoções, como gratidão, ressentimento moral, perdão, amor, cuja expressão é necessária para a participação em nossas relações interpessoais. Estas relações são extremamente importantes para o sentido das nossas vidas, e se perdêssemos as emoções necessárias para participar delas, nossas vidas seriam extremamente diminuídas. A visão cética, pondo em xeque o mérito por culpa e elogio, tornaria as emoções que o pressupõem injustificadas. Uma preocupação é que todas as atitudes reativas dependem do mérito por culpa e elogio, e assim elas todas seriam irracionais dada a aceitação do ceticismo. Ou o cético teria de abrir mão delas, e tomar uma atitude objetiva fria e calculista para com os outros (e consigo mesmo), tendo sua qualidade de vida deteriorada, ou teria de cultivar uma ilusão,



crendo para propósitos práticos em algo que sua racionalidade teórica lhe diz não ser crível.<sup>14</sup>

No entanto, como Derk Pereboom coloca:

Para começar, algumas atitudes reativas que seriam irracionais [dada a crença no ceticismo] não são obviamente necessárias para boas relações interpessoais. Além disso, as atitudes reativas que iríamos querer reter podem não ser comprometidas, ou então poderiam ter aspectos e análogos que não têm falsas pressuposições. Esses aspectos e análogos poderiam muito bem não ser semelhantes à objetividade de atitude de Strawson, e podem muito bem ser suficientes para sustentar boas relações interpessoais. (PEREBOOM, 2001, p. 97.)

Vejamos um exemplo da estratégia de Pereboom aplicada às emoções do ressentimento moral e da indignação. Ressentimento e indignação parecem pressupor que o seu alvo merece ser culpado pelo que fez. Assim, seriam irracionais na aceitação do ceticismo. Mas são essas emoções essenciais a uma relação interpessoal funcional? Uma função que elas possuem é a de comunicar o sofrimento infligido por conta da transgressão cometida. Mas essa função comunicativa da indignação e do ressentimento pode ser exercida simplesmente discutindo esse sofrimento sem essas atitudes. Diante da objeção de que tais discussões, sendo demasiado clínicas, careceriam de profundidade psicológica, Pereboom replica que uma expressão de tristeza ou pesar moral faria bem o trabalho de comunicar o sofrimento com a profundidade emocional necessária sem cometer o erro de culpar o seu alvo (PEREBOOM, 2001, pp. 97- 98).<sup>15</sup>

<sup>14</sup> O céptico teria de cultivar a ilusão não moral de crer que livre arbítrio existe quando não existe. Mas também estaria incorrendo numa ilusão moral, que o faria julgar culpado quem não tem culpa. Essa poderia ser uma ilusão moral *imoral*, porque o faria agir injustamente ao culpar quem não tem culpa. Mas digo apenas que "poderia ser imoral", pois, caso haja, inevitavelmente, concessões a se fazer entre promover a justiça e promover vidas significativas, pode haver casos em que promover vidas significativas leva a melhor sobre promover a justiça.

<sup>15</sup> No que consistiria essa tristeza ou pesar moral? Lembre-se do exemplo do assassino que teve uma criação ruim na seção 2.5. Ao saber do seu crime, podemos reagir com indignação e assim culpá-lo pelo que fez. Mas ao sabermos da sua criação ruim, podemos, por deixar de culpá-lo, gradualmente deixar de estar indignados, porém não ficar emocionalmente frios à situação. Podemos ainda assim ficar tristes, não só pela causa da sua falta de empatia no seu desenvolvimento, mas pelo próprio caráter do homem e a ação horrível que ele cometeu. O tom emocional moral da reação inicial pode ser mantido, mesmo que a própria emoção tenha mudado (veja PEREBOOM, 2001, pp. 97-98, para um exemplo semelhante).

## 2.9

**O sentimento de liberdade**

Termino este capítulo pronunciando-me sobre a segunda coisa que parece militar contra a aceitação da visão cética: o sentimento de liberdade. Repetindo, Samuel Johnson uma vez disse que “[t]oda a teoria é contra a liberdade da vontade; toda a experiência, a favor dela” (BOSWELL, 2008, p. 947). Mas, surpreendentemente, isso pode não ser verdade. Nem *toda* a experiência é a favor do livre arbítrio, embora a maior parte (a parte corriqueira) seja (o que torna o livre arbítrio uma ilusão). Mas Galen Strawson observa que a ideia de que não determinamos nossa vida mental (e assim nossas ações) pode ser vivenciada de forma subjetiva, não meramente teórica. Como ele coloca, “[p]ensamentos simplesmente nos 'ocorrem'; simplesmente 'temos' ideias, elas simplesmente 'nos vêm' (...)" (STRAWSON, 2010, p. 85). Ele cita como ilustração as observações do poeta Rimbaud sobre o seu processo criativo:

É falso dizer: eu penso: deveríamos dizer pensam-me (...) Eu é um outro. Azar da madeira que se descobre violino.

Os românticos (...) provam tão bem que a canção é tão poucas vezes a obra, isto é, o pensamento (...) compreendido do cantor (...) Pois Eu é um outro. Se o cobre desperta como clarim, não é em nada culpa dele. É-me evidente: assisto à eclosão de meu pensamento: contemplo-o, escuto-o (...). (STRAWSON, 2010, p. 85.)

A prática meditativa budista da atenção plena também é um bom modo de experienciar a ideia, o que abre um espaço promissor para a naturalização da espiritualidade (HARRIS, 2014). De fato, Strawson sugere:

É defensável que filósofos que creem que verdadeira responsabilidade é impossível, e que veem a si mesmos como comprometidos com a busca da verdade, devam se submeter a alguma prática de meditação. Pois pode capacitá-los para vir a apreciar a verdade da sua conclusão teórica de um jeito que não podem atingir por nenhum outro meio. (STRAWSON, 2010, p. 103.)

Logo, se há algum problema com o fato de a teoria não se harmonizar com a experiência, há uma possibilidade de ao menos centelhas de harmonização pelo curso das nossas vidas.

### Três problemas: liberdade, sorte e autocriação

Filósofos no debate do livre arbítrio têm notado que podemos ter algum tipo significativo de controle sobre nossas ações (por exemplo, nossos desejos e crenças, e deliberações baseadas neles, podem ter poder causal sobre nossas ações) ainda que o determinismo seja verdadeiro, mas alguns têm notado também como esse tipo de controle é insatisfatório para tornar justificáveis práticas de responsabilização que apelam ao mérito, em contraste com aquelas que focam em boas consequências futuras<sup>16</sup>. O problema é que, quando esses filósofos tentam formular a concepção de liberdade adequada, temos dificuldade de enxergar como o indeterminismo pode ajudar nesse empreendimento. Assim, parece que essa noção de controle realçado que supostamente requer o indeterminismo não pode ser aplicada quer o determinismo seja verdadeiro, quer ele seja falso.

O debate do livre arbítrio está intimamente ligado a questões de responsabilidade moral, mas ele não é o único. Há pelo menos dois mais: aquele relativo ao problema da sorte moral e aquele relativo ao que se pode chamar de problema da autocriação. Nesses debates também surgem problemas sobre como poderia existir algo como o controle necessário para o mérito por culpa e elogio por nossas ações. Dado isso, investigar os problemas surgidos neles pode jogar

---

<sup>16</sup> A atitude de Robert Kane, um libertarista (alguém que crê que há livre arbítrio e que ele é incompatível com o determinismo), é ilustrativa:

*ainda que vivêssemos num mundo determinista*, poderíamos distinguir de modo significativo pessoas que são livres de tais coisas como coibição física, adicção ou neurose, coerção ou opressão política, de pessoas que não são livres dessas coisas, e poderíamos aceitar que essas liberdades seriam dignas de preferência aos seus contrários mesmo num mundo determinista.

O que os incompatibilistas deveriam alegar (e o que eles frequentemente alegaram pela história) é que há *ao menos um* tipo de liberdade que é incompatível com o determinismo, e é um *tipo* *significativo de liberdade que é digna de se querer*. (KANE, 1996, p. 15)

Veja abaixo para a definição de incompatibilismo e o capítulo 2 para uma divergência da ideia de que essa liberdade “é digna de se querer”.

uma luz nos problemas surgidos no debate sobre o livre arbítrio. É isso o que pretendo fazer neste capítulo.<sup>17</sup>

### 3.1

#### Responsabilidade moral

Como vimos no capítulo anterior, quando julgamos que uma pessoa é responsável por uma ação sua, podemos estar fazendo várias coisas. Podemos estar destacando o fato de que a ação reflete quem a pessoa é e não é mero produto da situação em que ela se encontra. Podemos também estar salientando que a ação é, de alguma forma, produto das suas capacidades de raciocínio e que, assim, ela pode prestar contas sobre ela. Julgar as pessoas responsáveis pelas suas ações também pode ter várias significâncias práticas. Se a pessoa pode prestar contas pelas suas ações, faz sentido lhe pedirmos para refletir sobre como elas condizem com os seus valores, se desculpar e se esforçar para mudar no futuro, para não sermos mais prejudicados pelo seu mau comportamento, para ela se tornar uma pessoa melhor ou para nos reconciliarmos com ela (PEREBOOM, 2014). Até independentemente de se ela pode mudar a si mesma no futuro, se o bom ou mau comportamento revela o bom ou mau caráter da pessoa, faz sentido confiar nela com uma tarefa importante, ou tomar cuidado com ela no futuro, por exemplo.

Quando julgamos uma pessoa responsável no sentido em que ela é *merecidamente* culpável ou elogiável por uma ação sua, estamos fazendo algo diferente, no entanto. Estamos mantendo que certas respostas emocionais (como indignação ou gratidão) e punições e recompensas são apropriadas independentemente de bens futuros advindos como aqueles citados no parágrafo anterior. Isso é o que o filósofo Derk Pereboom (2014) denomina de *responsabilidade moral de mérito básico*, e ela, por exemplo, fundamenta a noção

---

<sup>17</sup> Mickelson (2019) é um ótimo recurso com respeito a isso, e este capítulo é fortemente inspirado no seu trabalho.

*retributiva* da justiça criminal, em contraste com aquelas que apelam para a dissuasão, a reabilitação e a incapacitação como justificativa.<sup>18</sup>

Usarei o termo *livre arbítrio* neste capítulo para designar o controle necessário para que seja aplicável a noção de responsabilidade baseada em mérito, em contraste com aquela baseada em consequências. Essa noção de livre arbítrio é ampla o bastante para não excluir de antemão nenhuma teoria substancial do livre arbítrio vinda de variadas partes no debate, mantém os conceitos envolvidos no debate próximos ao senso comum e faz sentido das controvérsias entre as variadas partes que movem o debate (CARUSO e MORRIS, 2017) – afinal, não há muitos filósofos defendendo que as capacidades mentais necessárias para que as pessoas possam mudar de comportamento de modos desejáveis no futuro em resposta à novas informações são incompatíveis com o determinismo, enquanto há vários que defendem apaixonadamente que não podemos merecer recompensas e punições se o determinismo for verdade.

Entender o livre arbítrio dessa forma também mantém em foco o aspecto prático do debate sobre o livre arbítrio. Quando punimos uma pessoa com base na ideia de que ela merece, estamos prejudicando-a intencionalmente sem que algum bem ulterior venha disso. Agir desse modo, presume-se, exige justificativas bastante boas, pois, se elas não forem realmente legítimas, estamos cometendo um erro moral gravíssimo. Esse tipo de prática só é justificado se houver livre arbítrio no sentido explicitado acima. Logo, se não temos suficientemente boas razões para crer que temos livre arbítrio (se, por exemplo, não há respostas suficientemente plausíveis para os desafios dos céticos), isso é razão para descontinuarmos tais práticas, como aquelas que pressupõem a justificativa retributivista para a justiça criminal (CARUSO, 2021, pp. 109-127).

---

<sup>18</sup> Neil Levy tem uma interpretação alternativa desse tipo de responsabilidade que não mantém que o mérito necessariamente torna apropriado prejudicar alguém como um bem intrínseco; ela antes chama a atenção para o fato de que, quando pensamos sobre atribuir bônus e ônus aos indivíduos de uma sociedade, julgamos justo considerar o bem-estar de cada pessoa igualmente. Se um agente merece culpa por um mau ato seu, sentimo-nos autorizados a invalidar esse pressuposto padrão, e assim descontar (levar em menor consideração) o bem-estar do culpado ao agirmos em prol do bem geral. Assim, uma consequência prática de negar esse tipo de responsabilidade moral é, como salienta ele:

que é injustificável descontar os interesses dos malfeitores em cálculos consequencialistas. Se resultar que é, ainda assim, justificável confinar indivíduos, então também teremos o dever para com eles de tentar compensá-los – talvez tornando a sua prisão tão confortável quanto possível – na medida em que o possamos fazer. (LEVY, 2011, p. 4.)

## 3.2

### Livre arbítrio e determinismo

Os filósofos têm se colocado em duas equipes contrárias dependendo da sua resposta à questão: o livre arbítrio é compatível com o determinismo? Os *compatibilistas* respondem que sim; os *incompatibilistas* respondem que não. Mas o que é determinismo?

Imagine dois mundos possíveis. Eles são qualitativamente idênticos em termos de fatos (o estado dos seus objetos e suas relações) num determinado momento. Se eles também são idênticos em termos das leis naturais que descrevem ou governam os padrões dos eventos dentro deles, e essas leis são deterministas, então ambos são também qualitativamente idênticos em *todo* outro momento – ou seja, eles são qualitativamente idênticos como um todo. O corolário é que se imaginamos um terceiro mundo possível cujo estado no presente é minimamente diferente do estado presente desses outros mundos, das duas uma: ou bem suas leis são diferentes, ou então o seu passado é em cada momento diferente.

Imagine uma pessoa humana normal num mundo assim. Chamemo-la de Perséfone<sup>19</sup>. Perséfone, após deliberar sobre levantar a mão ou deixá-la abaixada (ela está votando num júri, imaginemos), escolhe deixá-la abaixada. Dada a história inteira do universo anteriormente à sua escolha (que inclui o estado do seu cérebro anteriormente à escolha, admita-se), juntamente com as leis da natureza, é possível ocorrer apenas um futuro, aquele em que ela mantém a mão abaixada. O corolário é que, se Perséfone tivesse levantado a mão, das duas uma: ou bem as leis naturais teriam de ter sido diferentes, ou então cada momento passado teria de ter sido, em algum aspecto, diferente, desde o estado do seu cérebro em deliberação até o *Big Bang*. Perséfone tem livre arbítrio?

Há duas razões principais (concorrentes) para pensar que o determinismo colocaria em xeque o livre arbítrio (MCKENNA e PEREBOOM, 2016, p. 37-43). A primeira tem a ver com o que chamam de *possibilidades alternativas*: a ideia de

---

<sup>19</sup> O nome *Perséfone* tem a virtude mnemônica de lembrar a palavra *pessoa*, mas também tem o vício histórico de nomear uma deusa grega. Simplesmente esqueça a deusa; Perséfone aqui representa um ser humano normal genérico, que, de modo importante, não pode violar as leis da natureza.

que, para agirmos livremente, temos que ter o poder de agir de modo diferente do que de fato agimos. Perséfone escolheu ficar de mão abaixada, mas (ser humano normal que ela é) provavelmente pensa que poderia tê-la levantado ao invés disso. Por conta de possibilitar apenas um futuro, fixadas as leis e o passado, pensa-se que o determinismo excluiria o nosso poder de “fazer ou agir diferentemente”. A metáfora aqui é a de um jardim em que no meio do caminho encontramos uma bifurcação – podemos ir ou para a direita, ou para a esquerda. A realidade deve ser assim para haver livre arbítrio, dizem alguns incompatibilistas: temos que ter o poder de realizar ações contrárias *dado o passado e as leis*.

A segunda razão para considerar o determinismo como comprometendo o livre arbítrio tem a ver com a ideia de *fonte última*: de que, para agirmos livremente, as nossas ações devem, em última análise, se originar em *nós mesmos*, em vez de outra coisa, externa a nós. Provavelmente (ser humano normal que ela é) há algo que Perséfone experimenta interna e conscientemente como o seu “eu”, e ela pensa que essa é a origem das suas intenções e, assim, suas ações. Se o determinismo é verdadeiro, há fatores externos, fora do nosso controle, que são causalmente suficientes para a ocorrência de nossas intenções (como, por exemplo, a situação em que nos encontramos em conjunto com eventos neuronais aos quais estamos alheios; ou, retrocedendo mais no passado, o estado do nosso genoma em conjunto com as nossas primeiras experiências, e assim por diante). Isso, alguns incompatibilistas defendem, é razão o bastante para concluir que não seríamos a fonte genuína das nossas ações, se o determinismo fosse verdade.

A condição da fonte última tem progressivamente tomado o lugar da condição das possibilidades alternativas como explicação incompatibilista sobre qual condição necessária para o livre arbítrio não pode ser satisfeita num mundo determinista. Isso tem sido o caso não só por causa dos célebres casos de Frankfurt (veja capítulo 2, seção 6)<sup>20</sup>, mas também pelo fato de que argumentos baseados na condição das possibilidades alternativas (como o argumento da consequência) têm infrutiferamente desenterrado discussões antigas sobre análises categóricas (incompatibilistas) e condicionais (compatibilistas) de “poder” que acabam em impasses (veja KANE, 1996, pp. 44-59) (veja o capítulo 2, seção 6). A mudança de foco para a condição da fonte última tem aberto novos caminhos de

<sup>20</sup> Para um caso de Frankfurt que não pressupõe o determinismo, e assim não incorre em petição de princípio contra o incompatibilismo, veja PEREBOOM, 2001, pp. 18-22.



pesquisa, com argumentos que lidam com conceitos como *manipulação*, *sorte* e *autodeterminação* (PEREBOOM, 2001; LEVY, 2011; STRAWSON, 2010) e, como irei argumentar, tem aberto oportunidades para nos valermos, no debate do livre arbítrio, de recursos vindos de outras discussões filosóficas relativas à responsabilidade moral.

No entanto, se ao contemplarmos esses problemas para o compatibilismo somos tentados a aceitar o incompatibilismo, consideremos que ele tem o seu próprio problema. Isso é porque o incompatibilista deve enfrentar o desafio de explicar como o livre arbítrio é compatível com o *indeterminismo*. O indeterminismo diz que dadas as leis e o estado do mundo num dado instante, mais de um futuro é possível (parafraseando Lewis, mundos indeterministas podem divergir). Coloquemos Perséfone nesse mundo indeterminista. Ela, com base nos seus desejos e crenças, e num processo de deliberação baseado neles, escolheu manter a mão abaixada. Mas, segundo incompatibilistas, em um outro mundo possível, em que ela tem os mesmos desejos e as mesmas crenças, e delibera com base neles da mesma forma, ela decide levantar a mão. Como isso pode ser, se ela é responsável pela decisão?

O indeterminismo parece ter introduzido uma desconexão entre a ação do agente e as suas razões<sup>21</sup>. As razões do agente são representações mentais (crenças e desejos no sentido amplo: percepções, conhecimentos, teorias, palpites, hipóteses, heurísticas, ideais, valores, projetos, preferências, inclinações) cujos conteúdos constam como premissas em seus raciocínios práticos e, além disso, caracterizam-no como sendo quem ele é (ALVAREZ, 2016; LEVY, 2011, p. 74). Elas fazem com que uma ação seja realizada por razões e reflita a identidade do seu realizador. Logo, buscando basear a liberdade e a responsabilidade do agente na possibilidade de que a sua ação se desvincule de suas razões e identidade, o incompatibilista parece tornar as ações arbitrárias, caprichosas, aleatórias, irracionais ou inexplicáveis – em outras palavras, *ininteligíveis*<sup>22</sup>.

Os compatibilistas são rápidos para esclarecer que determinismo não é fatalismo: nossos desejos e crenças podem fazer uma diferença no que fazemos,

<sup>21</sup> Há outros lugares em que o indeterminismo pode ser introduzido que não entre as razões e a ação do agente. Uma alternativa é no processo de formação das razões do agente. Exponho as coisas assim, porém, por razões dialéticas. É aí onde os incompatibilistas tendem a situar o determinismo que eles julgam crucial para o livre arbítrio.

<sup>22</sup> Robert Kane chama esse problema para o incompatibilista de *questão da inteligibilidade* (KANE, 1998, pp. 105-107).

mesmo se o determinismo for verdade. Quando não estamos sob o jugo de restrições físicas ou psicológicas, nossa vontade tem eficácia causal sobre nossas ações, e se ela fosse diferente (se nossos desejos, crenças ou deliberação fossem diferentes) uma ação diferente decorreria. Embora essa concepção compatibilista de liberdade implique que, em última análise, esses desejos e crenças têm uma origem determinante externa a nós, ela pelo menos mantém a conexão adequada entre ações e razões. Como é que quebrar essa conexão pode *realçar* essa liberdade que os incompatibilistas consideram insatisfatória para fundamentar a responsabilidade baseada em mérito? Como é que isso pode nos possibilitar satisfazer a condição da fonte última que eles creem que não pode ser satisfeita num mundo determinista? Aqui cabe o questionamento pertinente de Thomas Nagel (1986, pp. 113-14): "Se elas [as possibilidades] estão abertas dado tudo sobre ele [o agente], o que ele tem a ver com o resultado?"

A ideia de que somos a fonte última das nossas ações encontra problemas quando tentamos conceptualizá-la pressupondo o determinismo. Ela parece também encontrar problemas quando tentamos conceptualizá-la pressupondo o indeterminismo. Se esses problemas não puderem ser superados, parece que a condição da fonte última para o livre arbítrio não pode ser satisfeita *independentemente* do que é verdade sobre as leis.

Se a condição da fonte última é necessária para o livre arbítrio, os compatibilistas estão errados, ao crerem que podemos ter livre arbítrio se o determinismo for verdadeiro; mas também estão os incompatibilistas tradicionais, ao crerem que o determinismo é *explicativamente relevante* para a ausência de livre arbítrio em mundos deterministas. O que explica o fato de que não somos as fontes últimas das nossas ações (e assim não temos livre arbítrio) em mundos deterministas deve ser algo verdadeiro em *todo* mundo possível, não algo verdadeiro em alguns mundos possíveis, porém não em outros, como o caráter das leis. Isso também indica que o livre arbítrio deve ser *impossível*, ao invés de meramente impossível com o determinismo. Chame essa posição de que o livre arbítrio é impossível de *impossibilismo*.<sup>23</sup>

<sup>23</sup> Tenho nesta seção, para propósitos expositivos, trabalhado com a taxonomia tradicional do debate do livre arbítrio, cujo terreno é dividido completamente entre os compatibilistas e os incompatibilistas em função da resposta negativa ou positiva à questão da compatibilidade entre livre arbítrio e determinismo. No entanto, a posição impossibilista cria certas complicações para essa classificação tradicional. Um impossibilista é um incompatibilista?

### 3.3

#### Sorte Moral

O problema da sorte moral surge do seguinte fenômeno: somos levados tanto a *rejeitar* quanto a *aceitar* que as pessoas mereçam culpa e crédito por coisas cuja ocorrência é devida ampla ou inteiramente à sorte, isto é, a fatores fora do seu controle. Isso decorre do fato de que aceitamos intuitivamente o que chamam de *princípio do controle*: um agente é responsável por uma coisa somente na medida em que ela está sob o seu controle. Também cremos que as pessoas de fato são, em várias situações, responsáveis por suas ações. No entanto, ao tomarmos consciência da pervasividade do envolvimento da sorte na condição humana, parece que (1) ou bem insistimos na condição do controle, e assim aceitamos que poucos juízos de responsabilidade (ou talvez nenhum deles) se mantêm válidos; (2) ou então mantemos nossos julgamentos corriqueiros, abraçando assim a *sorte moral*, casos em que "um aspecto significativo do que alguém faz depende de fatores além do seu controle, embora continuemos a tratá-lo, com respeito a isso, como um objeto de juízo moral" (NAGEL, 1979)<sup>24</sup>. Diante do fato da sorte generalizada, parece que queremos tanto manter nossos juízos de responsabilidade

---

Depende do que entendemos pelo conceito de compatibilidade em questão. Se ele for entendido como impossibilidade pura e simplesmente, o impossibilismo é um tipo de incompatibilismo, quando *x é meramente impossível com y* quer dizer que necessariamente, se *x* está presente *y* está ausente, sem a implicação de que a presença de *x* é explicativamente relevante para a ausência de *y*. Como ilustra Mickelson (2019) gatinhos fofinhos são impossíveis com círculos quadrados nesse sentido, e segundo o impossibilismo, determinismo e livre arbítrio são impossíveis também nesse sentido. Em todo mundo possível em que há gatinhos fofinhos não há círculos quadrados, mas a presença de gatinhos fofinhos não é explicativamente relevante para a ausência de círculos quadrados. Similarmente, em todo mundo possível determinista não há livre arbítrio, mas o determinismo não é explicativamente relevante para a ausência de livre arbítrio. A impossibilidade de gatinhos fofinhos e círculos quadrados decorre da impossibilidade de círculos quadrados. Similarmente, a impossibilidade do determinismo e do livre arbítrio decorre da impossibilidade do livre arbítrio.

Por outro lado, se o conceito de incompatibilidade implicar relevância explicativa, o impossibilismo não é um tipo de incompatibilismo. Kristin Mickelson (2019) e Kadri Vihvelin (2013) defendem essa interpretação do conceito de incompatibilidade envolvido no debate.

<sup>24</sup> Por "juízo moral" aqui, Nagel certamente quer dizer "juízo de responsabilidade". Ele certamente não quer dizer "juízo axiológico". Por exemplo, ele diz que:

Tal juízo é diferente da avaliação de algo como uma boa ou má coisa, ou estado de coisas. Esta última pode estar presente adicionalmente ao juízo moral, mas quando culpamos uma pessoa pelas suas ações, não estamos meramente dizendo que é ruim que elas tenham acontecido, ou ruim que ela exista; estamos julgando a ela, dizendo que ela é má, o que é diferente de ela ser uma má coisa. (NAGEL, 1979.)

corriqueiros quanto a nossa adesão ao princípio do controle. Daí o paradoxo da sorte moral: tendemos a aceitá-la e negá-la.

A literatura tradicional do debate da sorte moral identifica quatro tipos de sorte relevantes para o problema. Nas palavras clássicas de Thomas Nagel (1979), em primeiro lugar, temos sorte no "tipo de pessoa que você é (...) suas inclinações, capacidades e temperamento". Em segundo lugar, temos sorte no "tipo de problemas e situações que encaramos". Finalmente, temos "sorte em como somos determinados por circunstâncias antecedentes", e ainda "sorte em como nossas ações e projetos resultam". Essas categorias ficaram conhecidas, respectivamente, como *sorte constitutiva*, *sorte circunstancial*, *sorte causal* e *sorte resultante*. Essas quatro categorias não são ontologicamente puras exatamente. Elas combinam fatores relativos a causalidade/leis naturais e fatores relativos a eventos/estados de coisas intrínsecos vs. extrínsecos ao agente. Como Kristin Mickelson sugere, fazemos bem em conceber essas categorias de sorte vindas de Nagel como:

quatro "centros de gravidade" retóricos, ou seja, combinações retoricamente interessantes dos três fatores fora do nosso controle que listei, diferentes casos podendo ser desenvolvidos com base em onde os fatores especificados se obtêm relativamente a uma dada ação. (MICKELSON, 2019.)

Desse modo, Nagel, em seu artigo seminal sobre o assunto, basicamente nos convida a circundar a cena da ação com atenção para esses fatores fora do controle do agente conforme eles ocorrem posterior e anteriormente à ação pela qual, antes de refletirmos, o julgamos moralmente responsável. Sua conclusão desconcertante é a seguinte:

Se não podemos ser responsáveis pelas consequências das nossas ações devidas a fatores além do nosso controle, ou pelos antecedentes das nossas ações que são propriedades do temperamento não sujeito à nossa vontade, ou pelas circunstâncias que impõem as nossas escolhas morais, como então podemos ser responsáveis sequer pelos atos desnudos da própria vontade, se *elas* são o produto de circunstâncias antecedentes fora do controle da vontade?

A área de agência genuína, e, portanto, de juízo moral genuíno, parece encolher, sob esta investigação, até um ponto sem extensão. Tudo parece resultar da influência combinada de fatores, antecedentes e posteriores à ação, que não estão sob o controle do agente. (NAGEL, 1979.)

A seguinte série de casos é bastante útil para entender cada categoria de sorte envolvida neste debate. Quatro mulheres (batizemo-las Ana, Beatriz, Camila

e Daniela<sup>25</sup>) estão numa festa e todas elas sofrem um insulto. Ana quer descarregar a sua raiva na estrada. Ela gira a chave e o motor do seu carro liga. Ela decide dirigir, mesmo que isso seja extremamente irresponsável. Tragicamente, ela atropela um pedestre que estava próximo ao meio-fio.

(1) Versões semelhantes de todos esses eventos ocorrem com Beatriz, exceto o último, pois no seu caso não há pedestre algum próximo ao meio-fio. Isso quer dizer que, apesar de ter dirigido irresponsavelmente, ela não atropela ninguém. No entanto, a única diferença entre Ana e Beatriz é devida a um fator fora do seu controle nas consequências das suas respectivas ações: a presença ou ausência de um pedestre no lugar e no momento crucial. Isso quer dizer que elas são vítimas de (boa ou má) sorte *resultante*. Se culpamos Ana mais que Beatriz, parecemos estar aceitando sorte moral devida aos resultados da ação dos agentes.

(2) No caso de Camila, ela, como Ana e Beatriz, também deseja uma experiência catártica automotiva. No entanto, quando ela gira a chave, o motor não liga. Estipulemos que o fato de o motor ligar ou não ligar está fora do seu controle, assim como está fora do controle de Ana e Beatriz. No entanto, *se* a ocasião permitisse, Camila *iria* decidir dirigir de modo irresponsável, como Ana e Beatriz – aquela tem em comum com estas a *disposição* de assim agir. A diferença crucial aqui entre Camila e as outras, portanto, é devida a um fator fora do seu controle nas situações em que elas se encontram: na presença ou ausência da oportunidade de dirigir perigosamente. Isso quer dizer que elas são vítimas de (boa ou má) sorte *circunstancial*. Se culpamos Ana e Beatriz mais que Camila, parecemos estar aceitando sorte moral devida à situação em que os agentes se encontram.

(3) Finalmente, Daniela se distingue das outras no fato de que ela nem chega a formar o desejo catártico que as outras formam. Isso é porque ela desenvolveu inconscientemente um princípio de conduta de jamais conduzir colericamente, devido ao seu pai ter sido atropelado por um motorista zangado quando ela era criança. Daniela, no entanto, é parecida com Ana, Beatriz e Camila em termos de personalidade o bastante para que seja verdade que, se ela não tivesse esse bloqueio, ela também formaria o desejo de dirigir irada pela estrada diante do insulto. Parece que a diferença crucial aqui entre Daniela e as outras,

---

<sup>25</sup> Estes são uns dos nomes femininos mais comuns no Brasil de acordo com o IBGE. A série de casos se baseia em Hartman, 2017, pp. 7-8.

portanto, se deve a um fator fora do seu controle em uma de suas inclinações, já que a presença ou ausência de um histórico de experiência traumática é algo inteiramente fora do seu controle. Isso quer dizer que elas são vítimas de (boa ou má) sorte *constitutiva*. Se culpamos Ana, Beatriz e Camila mais que Daniela, parecemos estar aceitando sorte moral devida aos traços que caracterizam os agentes.

(4) Podemos ilustrar a sorte causal simplesmente pegando algum desses exemplos e postulando que a ação em questão é produto de uma cadeia causal determinista. A fim de destacar a questão dos antecedentes causais da ação, usemos o exemplo de Beatriz, em que o que está em evidência é a sua decisão de dirigir perigosamente, e não uma consequência dela (atropelar alguém). Se o determinismo for verdade, essa decisão de Beatriz tem uma série de condições antecedentes que determinam a sua ação, dadas as leis. Essas condições retrocedem até um tempo no passado em que ela nem havia nascido, de modo que ela não pode ter controle sobre elas. Dado que ela não tem controle sobre essas condições antecedentes, e dado que elas implicam a sua decisão, então ela não tem controle sobre a sua decisão. Beatriz tem sorte *causal* com respeito a sua decisão.

No entanto, temos que questionar neste ponto: que trabalho independente as causas antecedentes estão fazendo nesse caso para tirar o controle de Beatriz sobre a sua decisão? Um modo de entender a ideia de que Beatriz tem (má) sorte pelos antecedentes causais da sua decisão é contemplando que algo lá no passado distante, sobre o qual ela não tem controle algum, poderia ocorrer diferentemente, acarretando assim que a sua decisão fosse diferente. Mas como o passado só afeta o presente por meio dos passos intermediários, essa diferença no passado distante só seria causalmente relevante dado o seu efeito nas condições logo antes da decisão – mais especificamente, nos fatores relativos à constituição ou à situação de Beatriz<sup>26</sup>. Assim, parece que a sorte causal é supérflua, sendo o seu efeito capturado muito bem ou pela sorte constitutiva, ou pela sorte circunstancial (LEVY, 2011, pp. 39-40; LATUS, 2001).<sup>27</sup>

<sup>26</sup> Assim, condições causais antecedentes fora do controle de Beatriz só poderiam ser relevantes aqui na medida em que elas a tornassem mais similar a Daniela ou a Camila.

<sup>27</sup> Isso é relevante para a discussão anterior sobre o determinismo. O resultado é que a nossa falta de controle sobre as condições determinantes da nossa ação que estão fora do nosso controle só é relevante na medida em que ela implica a nossa falta de controle sobre os nossos traços constitutivos.

Uma abordagem promissora ao problema da sorte moral explica os nossos juízos diferenciais em casos comparativos de sorte como os acima expostos em termos não de uma diferença real no que os agentes merecem, mas da nossa situação *epistêmica* (RICHARDS, 1986). Segundo esta perspectiva, longe de haver envolvimento da sorte na determinação do que cada agente merece, há apenas sorte na disponibilização de *evidências* de culpabilidade (ou louvabilidade) aos observadores. Tomamos o sucesso do agente em agir conforme as suas intenções como o melhor indicador das suas verdadeiras intenções. Tomamos as decisões efetivas do agente como o melhor indicador do caráter do agente. Por acaso, a citação de Hume que reproduzi no capítulo anterior com outro propósito parece ilustrar vívidamente bem essa ideia também:

É evidente que, quando elogiamos uma determinada ação, consideramos apenas os motivos que as produziram, e tomamos a ação como signo ou indicador de certos princípios da mente e do caráter. A realização externa não tem nenhum mérito. Temos de olhar para o interior da pessoa para encontrar a qualidade moral. Ora, como não podemos fazê-lo diretamente, fixamos a nossa atenção na ação, como signo externo. Mas a ação é considerada apenas um signo; o objeto último de nosso elogio e aprovação é o motivo que a produziu. (HUME, 2009, p. 517.)

Assim, esta estratégia epistêmica nos possibilita aceitar um tratamento diferencial em casos comparativos sem com isso nos comprometer com a sorte moral. No entanto, essa estratégia é limitada. Ela pode valer para casos de sorte resultante, e até para casos de sorte circunstancial, mas não para casos de sorte constitutiva, pois neles o que está em jogo é o controle do agente sobre os seus estados e disposições mentais, não sobre os seus supostos “sinais”, em decisões evocadas por situações externas e consequências efetivadas por fatores posteriores.

Com respeito à sorte constitutiva, parecemos bater numa parede. Isso é não só porque a estratégia epistêmica dificilmente é aplicável a ela, mas também porque nos encontramos em dificuldade ao simplesmente tentar fazer sentido dela. Pelo menos, é isso o que alguns na literatura da sorte moral defendem. Essa tese de que a ideia de sorte constitutiva é incoerente pode complementar a composição de uma posição que mantém nossos juízos de responsabilização corriqueiros, ao mesmo tempo em que nega que exista sorte moral. Ela afirmaria os seguintes pontos.

1. A sorte causal é supérflua, no que sua relevância é capturada pela sorte constitutiva e circunstancial.
2. A sorte resultante e a sorte circunstancial só influenciam as evidências para os nossos juízos sobre o mérito do agente, não o seu mérito real propriamente dito.
3. Sorte constitutiva não pode influenciar o mérito do agente porque sorte constitutiva não existe.

Eu nego o ponto 3, e assim afirmo que sorte constitutiva existe. Mas vejamos qual é o argumento para se crer que ela não existe. Esta citação de Nicholas Rescher é um bom começo:

Não se pode dizer, de modo que faça sentido, que uma pessoa tem sorte com respeito a quem ela é, mas somente com respeito ao que lhe acontece. A identidade deve preceder a sorte. Não faz sentido imaginar um precursor desprovido de traços *a priori* que então tem a boa (ou má) sorte de ser equipado com um grupo de traços de caráter em vez de outro. Na teoria da pessoa, assim como na teoria da substância, não há nenhum lugar apropriado para particulares nus que, tendo uma identidade de aspecto indistinto (sem propriedades) *a priori*, pode então ser preenchida com propriedades *a posteriori*. (Em STATMAN, 1993, p. 155.)

Se tenho sorte de ter os traços constitutivos que eu tenho, isso implica que *eu* poderia facilmente ter tido traços diferentes. Mas o que é esse “eu” que não partilha dos meus traços constitutivos? O que ele tem a ver *comigo*? Parece que o conceito de sorte tem um aspecto modal tal que, se um indivíduo tem sorte com respeito a algo (um evento, um estado de coisas), há mundos possíveis em que o indivíduo está presente e esse algo, ausente. É plausível que também esse algo deva ser significante (de valor positivo ou negativo) e sua ocorrência deva estar fora do controle do indivíduo; além disso, esses mundos possíveis devem ser restringidos ao conjunto de mundos próximos (suficientemente semelhantes ao atual) (LEVY, 2011, p. 36). Mas isso não vem ao caso no momento. A questão é que esse aspecto modal parece pôr em xeque a sorte constitutiva, pois, se ele requer que haja mundos possíveis em que o indivíduo está presente, porém aquilo com respeito ao qual ele tem sorte está ausente, então ele não pode ter sorte com respeito a traços que ele possui em todo mundo possível em que ele existe, traços que ele não pode não ter.

Mas temos traços que não podemos não ter, isto é, essenciais? A ideia não é extremamente implausível. Por exemplo, Saul Kripke (1980) defendeu que



nomes são designadores rígidos (eles se referem ao mesmo objeto em todo mundo possível), e que, no caso de uma pessoa, a referência do seu nome é fixada pela sua origem causal num óvulo fertilizado, de modo que em todo mundo possível em que o seu nome se refere a um indivíduo, ele tem o mesmo genoma da pessoa a quem o nome se refere no mundo atual. Isso quer dizer que todos nós essencialmente nos originamos do genoma que possuímos.

A questão crucial é: qual é a relevância disso para os traços constitutivos relevantes no tocante à responsabilidade moral, como traços de personalidade, por exemplo? Isso suscita controvérsias sobre "determinismo genético" e natureza vs. criação, e, claro, esses são assuntos empíricos. Mas vale notar que, no momento, até os psicólogos que mais enfatizam o papel dos genes no desenvolvimento da personalidade deixam claro que a variação genética numa população só explica *metade* da variação em traços de personalidade, conforme entendidos pela teoria dos Grandes Cinco (abertura à experiência, conscienciosidade, extroversão, agradabilidade e neuroticismo), deixando a outra metade para a variação não genética (NETTLE, 2009; PLOMIN, 2018). Isso nos dá alguma razão para duvidar de que o essencialismo kripkeano possa ajudar a eliminar a sorte constitutiva.

No entanto, digo "alguma razão", pois a falsidade do determinismo genético só apoia a ideia de que temos traços constitutivos (neste caso, traços de personalidade) diferentes em alguns mundos possíveis, mundos em que o ambiente de desenvolvimento diverge suficientemente daquele do mundo atual. No entanto, como disse, o aspecto modal do conceito de sorte plausivelmente restringe o escopo dos mundos possíveis em questão não a quaisquer mundos possíveis, mas mundos possíveis *próximos*, e é difícil determinar se esses mundos estão próximos o bastante. Por exemplo, que eu seja exposto a um ambiente que, dado o meu genoma, eu cresço para ser um adulto extrovertido tagarela talvez seja uma possibilidade tão extravagante que o juízo "Luan poderia muito bem ser um extrovertido" resulte falso.

Outra complicação diz respeito a outro tipo de traço constitutivo relevante para a responsabilidade moral: valores. Nossos valores, é plausível pensar, são extremamente influenciados pela nossa cultura. Se a ideia de que tenho sorte de ter os valores que tenho depende da ideia de que eu poderia muito bem ter nascido

em outra cultura, a ideia de que tenho sorte de ter os valores que tenho pode resultar duvidosa, pois, como Neil Levy observa:

Provavelmente é verdade que membros típicos de culturas ocidentais são mais individualistas que membros típicos de culturas orientais. Eles têm sorte constitutiva com respeito a esse traço. Mas o mundo mais próximo em que eu tenha sido criado numa cultura não individualista está muito distante deste mundo no espaço modal. (LEVY, 2011, p. 32.)

Estamos aqui tentando fazer sentido da ideia de que temos sorte de ter os traços constitutivos que temos. A objeção colocada contra ideia foi, basicamente, a de que, se temos sorte por termos a personalidade e os valores que temos, os temos por acaso (no sentido em que poderíamos facilmente não tê-los) e que algumas teses não epistemicamente inaceitáveis sobre referência, modalidade e o papel dos genes e da cultura no desenvolvimento poderiam pôr em xeque a ideia de que temos esses traços constitutivos por acaso, e, assim, a ideia de que temos sorte de tê-los. Creio que é tão intuitiva a ideia de que temos sorte relativamente a traços cujo desenvolvimento em nós é devido a uma forte contribuição genética ou um forte enraizamento cultural que isso é razão para buscar uma concepção de sorte que não implique acaso necessariamente.

Neil Levy propõe uma concepção de sorte que contorna essas complicações supracitadas, tornando assim possível que uma agente tenha sorte de ter os traços constitutivos que tem sem que ele careça deles em muitos mundo possíveis próximos. Lembre-se de que disse acima que a definição de sorte de que temos tratado tem um aspecto relativo a:

- Modalidade: se um agente tem sorte com respeito a algo (um evento ou estado de coisas), há mundos possíveis próximos suficientemente numerosos em que ele está presente e esse algo está ausente.
- Controle: o agente carece de controle sobre a ocorrência do evento ou estado de coisas.
- Significância: o evento ou estado de coisas tem um impacto de valência positiva ou negativa para o agente.

Na concepção de sorte sem acaso de Levy (2011, pp. 32-35), ele mantém o aspecto de controle e significância, mas substitui o aspecto modal por um que considera não mundos possíveis em que o agente individual existe, mas aqueles em que seres humanos como nós existem. Para um agente ter sorte constitutiva com relação a certo traço, esse traço tem que variar suficientemente na experiência humana. Robert Hartman pertinentemente chama essa condição de *condição da instância incomum*, e a exemplifica assim:

Por exemplo, alguém afirma que Leigh tem sorte de ser uma gênica em comparação com os seus semelhantes. Essa afirmação é verdadeira? Leigh carece de controle direto sobre o fato de ser uma gênica, e isso é significativo para ela. Supondo que a maioria dos seus semelhantes não sejam gênios, ser um gênio não é instanciado comumente nesse grupo de referência. Assim, Leigh tem sorte (...) de ser uma gênica. (HARTMAN, 2017, p. 45.)

Não importa a verdade sobre a herdabilidade da inteligência, ou da proximidade de mundos possíveis, Leigh aqui tem sorte por ser uma gênica. Ela tem sorte constitutiva com respeito à sua inteligência.

Há um sentido em que a discussão acima é irrelevante, pois mesmo se a análise correta do conceito comum de sorte implicar que, de fato, não temos sorte de termos os traços constitutivos que temos, isso não elimina por definição o problema da "sorte moral" no que toca à nossa constituição. O problema apenas se manteria de pé com outro nome. Por exemplo, filósofos que trabalham no problema da sorte moral observam que o conceito de fortuna parece ser um parente modalmente estável do conceito de sorte, no que um evento afortunado é significativo e está além do controle do agente, mas não poderia facilmente ter ocorrido de outro modo (HARTMAN, 2017, p. 28). Assim, se o uso comum de "sorte" nos impede de dizer que temos sorte de ter o caráter e os valores que temos, certamente podemos dizer que somos afortunados por tê-los, e isso geraria "o problema da fortuna constitutiva", mesmo para filósofos com altos escrúpulos linguísticos. O aspecto da *falta de controle* presente tanto no conceito de sorte quanto no conceito de fortuna parece suficiente para gerar o problema abordado na literatura da sorte moral (HARTMAN, 2017, pp. 27-29).

Mas estipular que o conceito de sorte no debate da sorte moral deve ser entendido como falta de controle, embora o conceito corriqueiro não se resuma a isso, não introduz complicações e confusões gratuitas? Bem, não se o aspecto da

falta de controle tiver uma relação próxima o bastante com o conceito corriqueiro de sorte. Há várias abordagens na filosofia para entender o conceito de sorte, mas, no mínimo, a falta de controle do agente consta nos casos paradigmáticos do fenômeno. Como Hartman enumera as possibilidades:

Talvez a falta de controle seja uma condição necessária. Ou talvez seja uma "condição necessária aproximada" pois, embora não seja em si uma condição necessária, está intimamente relacionada a uma. Ou talvez a falta de controle seja meramente uma característica central da sorte, pois casos paradigmáticos de sorte estão fora do controle do agente. (HARTMAN, 2017, p. 30.)

Nossos traços constitutivos são pelo menos inicialmente devidos a fatores totalmente além do nosso controle, na célebre dupla dos genes e ambiente inicial (que inclui a transmissão cultural, mas não se limita a ela). Assim, temos sorte quanto a eles, pelo menos no sentido de sorte como falta de controle. Além disso, como vimos, é bem difícil de enxergar como esse tipo de sorte pode ser contornado por uma estratégia epistêmica. Assim, parece que a sorte constitutiva é algo real, e, pelo menos na medida em que sustentamos nossos juízos de responsabilidade corriqueiros, ela deve influenciar o mérito do agente. Isso parece nos obrigar a abrir mão do princípio do controle. Há uma saída, no entanto: afirmar que podemos controlar nossos traços constitutivos e, assim, superar a sorte constitutiva. Na medida em que fazemos isso, somos corretamente responsabilizados por nossas ações.

Dizer que podemos controlar nossa constituição parece ser o mesmo que dizer que podemos nos *autocriar*. E a crença na autocriação pode parecer subjetivamente mais plausível quanto mais culturalmente norte-americano você é.<sup>28</sup> Isso, no entanto, é diferente de dizer que ela é plausível *na realidade*, e na próxima seção veremos o problema que desafia essa aparente plausibilidade.

---

<sup>28</sup> O conceito de *self-made man*, ideia intrinsecamente americana, se refere ao indivíduo que atinge o sucesso apesar de suas origens humildes, devido não à sorte e beneficência externa, mas ao seu trabalho duro e perseverança. Os críticos da ideia destacam a sua contribuição à negligência a questões de justiça social.

### 3.4

#### Autocriação

Há razão para nos preocuparmos que a superação da sorte constitutiva, isto é, a assunção de controle sobre quem somos, é impossível. Esse problema surge na literatura sobre o livre arbítrio, principalmente por meio dos escritos de Galen Strawson. No entanto, como acaba implicando a irrelevância do caráter das leis naturais, é difícil encaixar a problemática na discussão sobre o problema do livre arbítrio e determinismo (e talvez por isso ele tenha sido um tanto negligenciado). Por isso, é apropriado que Kristin Mickelson (2019) dê a ele o nome especial de *problema da autocriação*.

Eis uma versão concisa do que Galen Strawson chama de *argumento básico* para a impossibilidade do livre arbítrio e da responsabilidade moral:

(1) Nada pode ser *causa sui* — nada pode ser a causa de si mesmo. (2) A fim de ser, de modo genuíno, moralmente responsável pelas nossas ações, teríamos de ser *causa sui*, pelo menos em certos aspectos mentais cruciais. (3) Portanto, nada pode ser, de modo genuíno, moralmente responsável. (STRAWSON, 1994.)

Conectando com a discussão sobre a sorte moral, podemos entender a premissa 2 como expressando ideia de que é necessário superar a sorte constitutiva para haver responsabilidade moral. A premissa 1, por sua vez, podemos entender como expressando a ideia de que a sorte constitutiva não pode ser superada, que a sua superação envolveria uma pessoa ser a causa de si mesma. Se cremos irrefletidamente na ideia de autocriação, ninguém melhor que Nietzsche para nos sacudir pela lapela retoricamente a vociferar que a *causa sui*:

é a maior autocontradição já concebida até o momento; é uma espécie de violação ou perversão da lógica. Mas o extravagante orgulho do homem conseguiu se enredar profunda e espantosamente com esta própria asneira. O desejo da "liberdade da vontade" no sentido superlativo metafísico, que ainda impera, infelizmente, nas mentes dos meio-letrados, o desejo de carregar a inteira e derradeira responsabilidade pelas suas próprias ações, e absolver Deus, o mundo, os ancestrais, o acaso e a sociedade, envolve nada menos do que ser precisamente esta *causa sui* e, com uma audácia maior do que a do

Barão de Munchhausen, puxar a si mesmo pelos cabelos existência adentro, dos brejos do nada. (Em STRAWSON, 2010, p. 292.)

Mas precisamos de um argumento também. Strawson o coloca em termos de um regresso (como ele tende a expressá-lo de uma forma bem longa, aqui está o bom resumo de Robert Kane):

1. Você faz o que faz por causa do modo como você é (sua natureza ou caráter).
2. Para ser verdadeiramente responsável pelo que você faz, você deve ser verdadeiramente responsável pelo modo como você é (pela sua natureza ou caráter).
3. Mas, para ser verdadeiramente responsável pelo modo como você é, você deve ter feito algo no passado pelo qual você também foi responsável para tornar a si mesmo, ao menos em parte, do modo como você é.
4. Mas, se você fosse verdadeiramente responsável por fazer algo no passado para tornar a si mesmo o que você é agora, você deve ter sido responsável pelo modo como você era então (pela sua natureza ou caráter) num tempo anterior.
5. Mas, para ter sido responsável pelo modo como você era naquele tempo anterior, você deve ter feito algo pelo qual você foi responsável num tempo ainda mais anterior para tornar a si mesmo do modo como você era naquele tempo mais anterior, e assim por diante. (KANE, 2005, pp. 71-72.)

Tomemos um momento para conectar a discussão sobre autocriação tida até o momento com o que foi discutido sobre sorte moral e livre arbítrio e determinismo. (1) Há certa linha de raciocínio, como o regresso exposto no argumento básico, que indica que a sorte constitutiva é impossível de superar. Se isso for verdade, a sorte constitutiva é, assim, um fato necessário, algo que se obtém em todos os mundos possíveis em que existimos. (2) Vimos que a superação da sorte constitutiva é a única saída para alguém desejoso de manter o seu compromisso tanto com o princípio do controle quanto com os juízos diferenciais em casos comparativos de sorte (nossos juízos de responsabilidade corriqueiros), já que há razões para pensar que a sorte moral não pode ser objeto de uma solução epistêmica nem definicional. (3) Na discussão sobre o problema do livre arbítrio e determinismo, vimos que é difícil ver como a condição da fonte última pode ser satisfeita, seja o determinismo verdadeiro, seja ele falso. Se essas dificuldades não podem ser superadas, isso sugere que o livre arbítrio é impossível, e que o que explica a insatisfação da condição da fonte não é um fato contingente como o caráter (in)determinista das leis naturais vigentes, mas um fato necessário. (4) Mas vimos que há razões que sugerem que a sorte constitutiva

é um fato necessário, de modo que ela parece bem adequada como explicação para a insatisfação da condição da fonte última. Há também uma adequação conceitual entre traços constitutivos, traços a respeito de *nós*, e a propriedade de ser fonte de nossas ações. Somos as fontes genuínas das nossas ações quando elas vêm dos traços que nos caracterizam e somos responsáveis por esses traços.

Podemos resumir, assim, a tese desta dissertação. O livre arbítrio é impossível, independentemente de fatos sobre as leis naturais. A condição da fonte última deve ser satisfeita para haver livre arbítrio, mas ela não pode ser satisfeita. Isso é porque ela exige um ato de autocriação que é impossível, devido ao fato necessário da sorte constitutiva.

Argumentos do regresso são corriqueiros na filosofia (sendo o exemplo mais proeminente encontrado na epistemologia, dentro dos modos céticos do trilema de Agripa), e essa familiaridade com eles da parte dos filósofos pode motivar a resposta que tem sido dada ao argumento: Strawson a descreve como a tendência de "ser rejeitado (...) como errado, ou irrelevante, ou fátuo, ou rápido demais, ou uma expressão de megalomania metafísica" (STRAWSON, 1994). Mas, como Mickelson (2019) sugere, o regresso não é melhor visto como um argumento decisivo para a impossibilidade do livre arbítrio, mas sim como algo que gera um ônus de prova dialético sobre os possibilistas (filósofos que creem que o livre arbítrio é possível), um desafio para eles bolarem modos legítimos de parar o regresso. Até que uma tentativa seja bem sucedida, o impossibilista mantém a vantagem dialética. A tarefa do impossibilista é oferecer critérios claros de modos legítimos de parar o regresso e se manter aberto para avaliar com base nesses critérios as propostas possibilistas em oferta.

Um regresso não é o único modo de expressar a ideia da impossibilidade da autocriação, pois Agnes Callard faz isso por meio de um interessante dilema:

Se, por um lado, o valor que cultivo em mim mesma decorre racionalmente dos valores que já tenho, então não faço criação nenhuma. Pois nesse caso meu "novo" eu já estava contido no meu antigo eu. Se, por outro lado, o novo valor está racionalmente desvinculado dos meus valores anteriores, o eu com que termino pode ser novo, mas não é o produto da minha própria agência. O cultivo racional de valores não é *autocriação*, e o cultivo não racional de valores não é *autocriação*. (Em MICKELSON, 2019.)

Leiamos esse dilema em termos de soluções para o problema do livre arbítrio. Compatibilistas, por um lado, tendem a aceitar que, a fim de ser

responsável, as ações de um agente devem, de algum modo, ser explicadas por aspectos mentais que caracterizam o agente — desejos e crenças, no sentido amplo. Mas eles acabam fazendo pouco caso da preocupação de que o agente não teve papel na produção dessas próprias crenças e desejos. Incompatibilistas, por outro lado, aceitam que há algo de errado na ideia de que um agente pode ser responsável por ações que se devem a desejos e crenças que foram, por assim dizer, entregues a ele, como uma bênção ou uma maldição de(os) Deus(es). Mas daí eles acabam desvinculando as ações do agente das suas crenças e desejos, tornando, sim, a ação não determinada por fatores externos ao agente, mas, perversamente, também não determinada *pelo* agente. Claro que as coisas ganham nuance, tanto para compatibilistas, quando tentam dar conta de explicar como o agente pode assumir a posse do seu caráter, como para incompatibilistas, quando tentam explicar como que a ação indeterminada pode ser baseada em razões. Mas os seus respectivos problemas tendem a voltar: para não caírem no regresso, compatibilistas têm de acabar aceitando a originação externa do caráter do agente, e incompatibilistas têm de aceitar a desconexão entre os traços constitutivos do agente e as suas ações.

Desenvolvamos esse pensamento no próximo capítulo.



## 4

### Compatibilismo, incompatibilismo e autocriação

O seguinte consiste em uma avaliação do argumento básico de Strawson em termos das soluções dadas ao problema do livre arbítrio.

Um modo de resumir o ponto crucial do argumento de Strawson é da seguinte forma: a responsabilidade moral exige que nossas ações estejam conectadas do jeito certo aos nossos traços constitutivos, e que, ao mesmo tempo, não seja o caso que tenhamos sorte com relação a eles. Como não podemos superar a sorte constitutiva, assumindo o controle sobre nossa constituição num ato de autocriação, não podemos ter responsabilidade moral.

#### 4.1

##### Ação baseada em razões e incompatibilismo

Os incompatibilistas poderão questionar qual é essa conexão entre traços constitutivos e ação que é necessária para a responsabilidade moral, desconfiando que se esteja introduzindo causação determinística sorrateiramente. Há um debate filosófico sobre se uma explicação baseada em razões é sequer causal (ALVAREZ, 2016; KANE, 1996, pp. 119-120). Incompatibilistas podem protestar, assim, que a relação entre os estados psicológicos relevantes do agente e a sua ação ou é acausal, ou expressa uma causação de natureza probabilística, ao invés de determinística. Acho que o argumento pode ter sucesso sem tomar partido nessas controvérsias. Acho que basta que se aceite a necessidade de haver uma relação objetiva entre ação e os traços psicológicos que fundamente uma explicação de porque a ação ocorreu ao invés outra em termos dos traços psicológicos do agente – de porque, dados esses fatores mentais, tal ação fez sentido para ele em vez de outra (STRAWSON, 2010, pp. 29-35). A maioria dos incompatibilistas ficará satisfeita que essa relação seja a de causação indeterminística, em que os fatores mentais do agente aumentam a probabilidade da ação, de modo que pressuporei que esse é o caso.

O apelo aos traços psicológicos não nega a existência de uma causa-agente, uma substância com um poder causal indeterminístico irreduzível aos eventos psicológicos que envolvem o agente. Podemos aceitar, em prol do argumento, que essa causa dê o impulso seletivo guiado pelos eventos intrínsecos ao agente na tomada de decisões. Também não estamos pressupondo uma tese de identidade pessoal que exclui uma substância persistente sem propriedades que fundamenta a reidentificação do agente como o mesmo indivíduo através do tempo. Podemos aceitar essa ideia, em prol do argumento, ainda mantendo que eventos psicológicos específicos são necessários para caracterizar a ação como pertencente à pessoa, de modo que a ação faça sentido dado quem ela é (LEVY, 2011, pp. 74-75).

Há dois tipos principais de teorias incompatibilistas do livre arbítrio em voga na literatura relevante: incompatibilismo de causação de eventos e incompatibilismo de causação de agente. O incompatibilismo de causação de eventos (KANE, 1996) mantém que ações e decisões livres devem ser causadas por eventos mentais que envolvem o agente, porém esses eventos ou estados causam a ação ou decisão de modo probabilístico, não determinístico. O incompatibilismo de causação de agente (em KANE, 2011, pp. 309-328), por sua vez, mantém que isso não é suficiente para a liberdade; não explicita o papel do agente na tomada de decisão o bastante. Para eles, na ação livre, além dos eventos intrínsecos ao agente, há uma substância irreduzível a eles, com um poder causal similarmente irreduzível ao poder causal deles, e esse poder causal é indeterminista. Segundo este segundo tipo de incompatibilismo, quando um agente age livremente, não é somente que a ação se deve às suas crenças e desejos que a tornam prováveis, mas também a um poder do agente enquanto substância de causar a ação sem ser determinado a fazer isso por nada externo a ele.

Robert Kane é o principal expoente atual do incompatibilismo de causação de eventos. Sua teoria diz que é necessário que haja decisões produzidas indeterministicamente para haver livre arbítrio, mas não é um requerimento que *toda* decisão livre seja assim. Ele permite que decisões livres e responsáveis sejam determinadas pelo caráter do agente, por exemplo, mas com a condição de que ele tenha feito algo no passado para se responsabilizar por esse caráter, e são essas “ações autoformadoras”, como ele as chama, que não podem ser produzidas deterministicamente (KANE, 1996, pp. 35, 124-125). Assim, será mais apropriado

explorar a teoria de Kane mais à frente, não neste ponto do argumento, pois ela pode estar de total acordo com a ideia de que ação deve se conectar com razões da maneira certa.

Que dizer dos incompatibilistas de causação de agente? É bem tentador para um incompatibilista que busca fazer sentido de como uma causa-agente pode agir por razões tomar a seguinte rota. Comece com a teoria do incompatibilismo de causação de eventos, em que os eventos mentais que envolvem o agente causam as ações de modo probabilístico. Esses fatores mentais intrínsecos ao agente estruturam as razões com que a causa-agente se depara na tomada de decisões. Entra então a causa-agente para complementar a descrição: ela faz o papel de selecionar, com o seu poder causal irreduzível, entre as opções que os fatores mentais intrínsecos ao agente põe diante dela. Essas razões, como Leibniz diria, inclinam, mas não necessitam a decisão; elas deixam espaço para a intervenção do *deus ex machina* da causa-agente.

O problema aqui é que, se a teoria incompatibilista da causação de eventos é insatisfatória no que diz respeito à liberdade, como os incompatibilistas da causação de agente creem, a introdução da causa-agente na figura deve fazer uma diferença relevante à liberdade. Não pode ser que todo o papel relativo a razões seja feito pelos eventos mentais intrínsecos ao agente, pois se for assim, a causa-agente é ociosa. Especificamente, se a causa-agente tem um papel crucial aqui, não pode ser que ela o exercite somente aquiescendo às propensões probabilísticas dos eventos mentais intrínsecos ao agente, simplesmente ratificando as decisões que os traços constitutivos do agente tornam mais racional tomar.

A não ociosidade da causa-agente se mostraria do modo mais vívido no caso em que ela decide *contra* as recomendações dos eventos mentais do agente. Mas, nesse caso, a contribuição da causação de agente é simplesmente introduzir o “poder” de agir fora de caráter, arbitrariamente com relação aos traços constitutivos do agente, o que é difícil de enxergar como contribuindo para a responsabilidade moral.

Outra possibilidade é que a não ociosidade da causa-agente se mostra quando as razões preexistentes do agente não ditam uma opção unicamente racional, e assim o seu papel é o de dar o decisório “impulso final”, sem o qual nenhuma decisão decorreria. Mas nesse caso, a pergunta fica: como então esse

poder de dar o impulso final que expõe a não ociosidade da causa-agente pode ser exercido por razões e não arbitrariamente? Há duas opções. A primeira é admitir que ele não pode ser exercido por razões, o que parece ser uma confissão de derrota da parte do incompatibilista de causação de agente. A segunda opção, por outro lado, é dizer que, longe de escolher sem razão, a causa-agente *cria* razões no processo de seleção. Quando, previamente, nenhuma opção tinha mais peso do que outra, a causa-agente livremente *atribui* mais peso a uma delas, e assim age com base nessa razão autocriada. Como Neil Levy (2011, p. 70) ilustra: “O agente, por assim dizer, põe o dedo em um dos pratos, fazendo assim a balança pender.” Essa é uma opção interessante, mas acho que ela deve ser enquadrada como uma candidata à solução de superação da sorte constitutiva, não como um modo de negar a conexão íntima que razões e ação devem ter, a fim de haver responsabilidade moral.

## 4.2

### Sorte constitutiva e compatibilismo

Aceitando a conexão íntima entre traços constitutivos e ação, a fim de haver responsabilidade moral, o que dizer da necessidade de controle sobre esses os próprios traços? Não importa de onde vieram nossos desejos e crenças? O compatibilista Harry Frankfurt mostra bem o que é "segurar as pontas" diante das duras consequências que ele admite que o seu compatibilismo o faz encarar:

O que precisamos olhar mais essencialmente antes é para certos aspectos da estrutura psíquica que coincide com o comportamento da pessoa (...)

Um manipulador pode ter êxito, através das suas intervenções, em munir a pessoa não meramente de sentimentos e pensamentos particulares, mas também de um novo caráter. Essa pessoa é então moralmente responsável pelas escolhas e conduta às quais ter esse caráter leva. Afinal, somos inevitavelmente moldados e sustentados por circunstâncias sobre as quais não temos controle algum. As causas às quais somos sujeitos também nos modificam radicalmente, sem assim ocasionarem a nossa falta de responsabilidade moral. É irrelevante se essas causas estão operando em virtude de forças naturais que determinam nosso ambiente ou se operam através dos desígnios manipuladores deliberados de outros agentes humanos. (FRANKFURT, 2002, p. 27.)

Frankfurt, essencialmente, mostra aqui uma consequência indigesta de se aceitar a sorte constitutiva: agentes responsáveis podem ser manipulados. Casos em que os traços constitutivos do agente surgem da manipulação de outros agentes são apenas mais uma instância de seus traços constitutivos serem produto de circunstâncias fora do seu controle. Assim, se aceitamos que um agente pode ser responsável por ações que decorrem dos seus traços constitutivos com relação aos quais ele tem (boa ou má) sorte, temos que aceitar que um agente manipulado pode ser moralmente responsável. O problema é que isso é bem difícil de aceitar.

Casos de manipulação de agentes intencionais são apenas um modo vívido de pôr em evidência as circunstâncias causais externas dos estados mentais de um agente. Lembremo-nos da citação de Espinosa do capítulo 2 sobre a sua explicação da ilusão do livre arbítrio: “os homens opinam serem livres porquanto são cômicos de suas volições e de seus apetites, e nem por sonhos cogitam das causas que os dispõem a apetecer e querer, pois delas são ignorantes” (ESPINOSA, 2015, p. 111). Essa tendência pode se manter até quando somos solicitados a imaginar que nossos estados mentais são determinados, não assimilando com plena consciência o que isso envolveria. Um propósito de casos de manipulação, portanto, é contornar essa tendência, e voltar a nossa atenção para os fatores determinantes da vontade do agente.

Porém, alguém pode objetar que o fato de que a circunstância causal externa é um agente intencional tem um papel crucial aqui, de modo que ele em si, e não o fato mais geral da causação externa, explica a nossa intuição de que o agente manipulado não é responsável, bloqueando assim a inferência de que um agente com sorte constitutiva não pode ser livre. Por isso, são úteis casos de “manipulação” em que o que tem o papel originador dos traços constitutivos não é um agente. O seguinte cenário criado por Brian Parks é um candidato:

Enquanto corre no parque, Ana – uma mulher adulta, antes do ocorrido, compassiva e moralmente sensível – é exposta a uma perigosa bactéria comedora de miolos que infecta seu cérebro. A infecção envia todas as suas FPs [“forças psicológicas”, isto é, traços constitutivos mentais] – ela se torna uma sociopata violenta. Certamente, ela não seria moralmente responsável por nenhuma escolha infeliz que ela fizesse em seu estado sociopático. Da mesma forma, ela não seria moralmente responsável por nenhuma

consequência adversa que resultasse dessas escolhas.<sup>29</sup> (PARKS, 2009.)

Uma objeção poderia se desenvolver da seguinte forma: seja admitido que Ana não poderia ser responsável *logo após* de sofrer a infecção, mas se tiver tempo o bastante para pôr em prática suas capacidades de autorreflexão que a possibilitam aceitar ou rejeitar o seu caráter, ela poderá ser responsabilizada pelas suas ações, sim. Isso é semelhante ao que ocorre conosco, pessoas normais não infectadas: recebemos o nosso caráter inicialmente dos genes e do ambiente, sem um pinga da nossa contribuição, mas não somos responsabilizados pelas primeiras decisões que tomamos com base neles. Somente conforme desenvolvemos capacidades de reflexão que nos capacitam para assumir a responsabilidade pelo caráter que recebemos de fontes externas é que podemos ser moralmente responsáveis pelas decisões que decorrem dele.

Essa é uma ideia interessante, mas peca pelo seguinte problema: Ana pode refletir sobre o seu caráter e aceitá-lo ou modificá-lo, de alguma maneira – aceitemos isso. Mas a pergunta é: ela fará essa autoavaliação com base no quê? Parece que, no fim das contas, será com base nos traços constitutivos adquiridos de uma fonte externa (a infecção bacteriana), de modo que Ana tem sorte constitutiva quanto a eles. E, parece plausível pensar, não se pode eliminar sorte constitutiva com mais sorte constitutiva. E o mesmo parece se aplicar a nós, inclusive, que também aceitamos ou modificamos nosso caráter por meio de

---

<sup>29</sup> É relevante, neste contexto, citar este trecho de um artigo de Derk Pereboom (um grande expoente atual do argumento da manipulação) e Gunnar Björnsson:

(...) Björnsson (...) construiu um cenário em que todas as condições compatibilistas padrão são satisfeitas, mas uma causa não agente — uma infecção — vagarosamente torna o agente cada vez mais egoísta sem contornar ou comprometer as suas capacidades enquanto agente. Com base na hipótese de que os sujeitos de pesquisa consideram que a responsabilidade é comprometida quando entendem o agente de uma perspectiva em que a sua deliberação é uma mera variável dependente (...), ele fez a previsão de que, se os sujeitos fossem estimulados a ver o comportamento do agente como dependente da sua causa não agente, isso comprometeria as atribuições de responsabilidade aproximadamente na mesma medida que a introdução de um manipulador intencional. Isso de fato foi o caso: num estudo envolvendo 416 sujeitos, a infecção comprometeu atribuições de livre arbítrio e responsabilidade moral na mesma medida que casos de doutrinação de manipulação intencional. Esse estudo sugere que incompatibilistas podem ser capazes de empregar a mesma estratégia de generalização usada em argumentos de manipulação sem introduzir um controle externo intencional (...). (Em SYTSMA e BUCKWALTER, 2016, pp. 142-157.)

reflexão com base em traços constitutivos que temos (boa ou má) sorte de ter. E a ideia de que sorte constitutiva não pode ser eliminada por sorte constitutiva também parece se manter igualmente de pé aqui também. Neste contexto, é válido expor uma versão do argumento básico de Strawson que ele formula nos termos mais naturais do par hereditariedade e ambiente:

(A) Um indivíduo é do modo como é, inicialmente, como resultado da hereditariedade e das primeiras experiências. (B) Claramente essas são coisas pelas quais o indivíduo não pode ser de modo algum responsável (...). (C) Um indivíduo não pode em estágio futuro algum de sua vida ter a esperança de aceder à responsabilidade última pelo modo como é tentando modificar o modo como já é enquanto um resultado da hereditariedade e da experiência. Pois se pode muito bem tentar mudar a si mesmo, mas (D) tanto o modo particular como alguém é movido a tentar mudar a si mesmo quanto o seu grau de sucesso em sua tentativa de mudança serão determinados pelo modo como esse alguém já é enquanto um resultado da hereditariedade e da experiência. E (E) quaisquer futuras mudanças que alguém pode causar somente após ter causado certas mudanças iniciais serão por conseguinte determinadas, por meio das mudanças iniciais, pela hereditariedade e experiências passadas. (F) Isso pode não ser a história toda, pois pode ser que algumas mudanças no modo como alguém é sejam atribuíveis à influência de fatores aleatórios ou indeterminísticos. Mas (G) é um absurdo supor que fatores aleatórios ou indeterminísticos, pelos quais um indivíduo não é, *ex hypothesi*, de modo algum, responsável, podem em si mesmos contribuir para esse indivíduo ser verdadeiramente ou em última instância responsável pelo modo como ele é. (STRAWSON, 1998.)

### 4.3

#### Sorte constitutiva e incompatibilismo

Claro, incompatibilistas que creem que o livre arbítrio é possível iriam objetar à ideia de Strawson de que é absurdo pensar que a causação indeterminística contribui algo para a responsabilidade moral. Lembre-se de que disse acima que, segundo o incompatibilista de causação de eventos Robert Kane, agentes livres possuem na sua história o que ele chama de "ações autoformadoras", nas quais o agente forma seu caráter. Podemos dizer que nessas situações o agente assume controle sobre seus traços constitutivos, se desfaz da sorte constitutiva, se autocria, em algum sentido. E, Kane mantém, essas ações devem ser causadas de modo não determinístico.

Essas situações de autoformação são situações em que o agente se encontra genuinamente dividido. Exemplos clássicos disso são contextos em que há um conflito interno no agente entre obrigação moral e inclinação ligada ao interesse próprio, ou entre projetos de longo prazo ditados pela prudência e desejos que demandam satisfação imediata. Kane escreve:

Considere uma mulher a caminho de uma importante reunião de vendas que presencia uma agressão ou roubo num beco. Ela deveria parar e chamar ajuda ou prosseguir para evitar perder uma venda crucial para a sua carreira? Ou considere um engenheiro que é um alcoólatra em recuperação tentando reparar seu casamento. Trabalhando tarde da noite com grande estresse, ele fica tentado a tomar uma bebida para ir em frente. A empresária encara um conflito moral; o engenheiro, um conflito prudencial. (KANE, 1996, p. 126.)

Assim, na estrutura psicológica do agente há dois conjuntos de razões; da perspectiva de um, decisão A é a mais acertada e, da perspectiva de outro, decisão B é a mais acertada. Porém, como os dois conjuntos têm peso igual para o agente, nenhuma decisão é mais recomendável considerando a estrutura psicológica do agente como um todo. Essas razões concorrentes geram conflito no agente, o que leva a um esforço de superar esse conflito. Durante esse momento de esforço, nenhum conjunto de razões prevalece, e está indeterminado qual decisão será tomada. Porém, no momento da decisão, e por meio dela, um conjunto de razões prevalece sobre o outro, e a decisão é tomada por essa razão. Neste instante, o peso das razões pende para um lado, e a vontade antes não estabelecida do agente é estabelecida.

Kane (1996, p. 135) acha válido observar que "[o]s agentes (...) não criam as suas razões para escolher ao escolher. Em vez disso, as razões são razões que eles já têm". Isso é importante, pois mantém certa conexão entre o suposto ato de autoformação com os traços constitutivos anteriores do agente, com base nos quais ele toma a decisão autoformadora. Caso contrário, seria difícil fazer sentido da ideia de que a decisão veio *do* agente. Mas então qual contribuição a decisão autoformadora anteriormente indeterminada faz em si para eliminar a sorte constitutiva? Kane (1996, p. 135) diz que "o que [os agentes] fazem ao escolher é fazer um conjunto de razões prevalecer sobre outras ali e então como motivadores da ação. (...) os agentes resolverão a questão sobre qual [opção] eles querem *mais* ao decidir". Assim, os traços constitutivos com relação aos quais o agente tem sorte antes de ocorrer a decisão autoformadora (que supera essa sorte constitutiva)



forneem as razões que, neste momento, têm força igual. O que a decisão autoformadora faz não é produzir novas razões, mas modificar *o peso* das razões preexistentes, tornando assim, apenas no momento da decisão, uma ação mais racional do que outra.

Isso soa como uma sugestão feita na primeira seção deste capítulo na discussão sobre possíveis objeções advindas de incompatibilistas de causação de agente à ideia da necessidade da conexão entre razões e ação para a responsabilidade moral. Lembremos:

A segunda opção, por outro lado, é dizer que, longe de escolher sem razão, a causa-agente cria razões no processo de seleção. Quando, previamente, nenhuma opção tinha mais peso do que outra, a causa-agente livremente atribui mais peso a uma delas, e assim age com base nessa razão autocriada. Como Neil Levy (2011, p. 70) ilustra: “O agente, por assim dizer, põe o dedo em um dos pratos, fazendo assim a balança pender.”

Claro, Kane não concordaria com os termos em que coloquei as coisas, ao descrever uma mudança de peso de razões preexistentes como criação de razões. Mas acho que isso é apenas uma questão de modo de expressão. Posso prosseguir a partir daqui sem falar em criar razões, apenas em atribuir pesos.

Kane também não concordaria com a inserção de uma causa-agente na produção da decisão autoformadora. De acordo com sua teoria, basta que essa decisão seja produzida de forma indeterminística do modo certo, sendo uma causa-agente ou algo supérfluo, ou algo que adiciona mistérios sem necessidade (KANE, 1996, pp. 115-117). Mas acho que discutir (o que aqui construo como) sua solução para superação da sorte constitutiva em conjunto com a teoria da causa-agente não causará dano à sua teoria, pois as objeções que colocarei contra ela será independente da ontologia da causação.

A objeção já deve ser familiar: com base em que a decisão autoformadora, causada indeterministicamente seja por eventos, seja por uma substância, é tomada? O agente, no momento crucial de resolução de conflito volitivo, atribui mais peso, digamos, a opção A do que a opção B; por quê? Uma primeira resposta (a rota incompatibilista de “segurar as pontas” analogamente à resposta compatibilista de Frankfurt acima) é dizer que não há razão alguma. Isso seria o mesmo que o incompatibilista dizer que a decisão está desconectada do

agente enquanto um ser caracterizado, e isso parece ser o mesmo que dizer que a decisão não vem *do* agente.

Uma resposta alternativa é chamar atenção para a conexão da decisão com as razões preexistentes do agente. Antes da decisão, o agente tinha dois conjuntos de razões e estava igualmente envolvido com elas, de modo que, qualquer decisão que tomasse, ele teria razões para agir, e assim a decisão viria dele, estaria a caráter.

Mas essa é só parte da história. Haveria uma explicação baseada em razões para agente ter escolhido *uma ou outra* das opções, sim; no entanto, não há explicação baseada em razões para o agente ter escolhido uma opção *ao invés* da outra. Na medida em que a decisão é explicada por meio de razões, o papel inteiro é feito por traços constitutivos que, em última instância, o agente não teve papel algum em criar (temos que pressupor isso, se a decisão autoformadora for uma que detém o regresso). Isso quer dizer que, nessa medida, a ação é tanto produto de um caráter não autoformado quanto uma decisão produzida deterministicamente por um caráter produzido deterministicamente por fatores externos ao agente, que o incompatibilista não pode aceitar como livre. Por outro lado, na medida em que a decisão supostamente autoformadora adiciona algo ao que havia antes na constituição do agente, modificando o peso das razões anteriores, essa diferença é totalmente arbitrária com relação aos traços que caracterizam o agente. Como então pode-se afirmar que ela vem *dele* num sentido forte que fundamentaria o mérito por culpa e elogio?

O incompatibilista aqui, por mais que tente evitar, acaba caindo no segundo corno do dilema de Callard visto no capítulo anterior:

Se (...) o novo valor está racionalmente desvinculado dos meus valores anteriores, o eu com que termino pode ser novo, mas não é o produto da minha própria agência. (...) o cultivo não racional de valores não é *autocriação*. (Em MICKELSON, 2019.)

## Considerações finais

Nesta dissertação busquei defender os seguintes pontos:

- O livre arbítrio é inexistente, por conta não de teses contingentes sobre a natureza determinista ou indeterminista das leis naturais, mas do fato necessário da sorte constitutiva. Disso decorre que o livre arbítrio não só é inexistente como é impossível.
- A tese do impossibilismo quanto ao livre arbítrio pede um alargamento da discussão sobre o livre arbítrio para considerações relativas ao problema da sorte moral e problemáticas a ver com a ideia de autocriação. Todas essas temáticas se relacionam às perplexidades que encontramos ao fazer sentido do controle requerido para a responsabilidade moral no sentido de mérito básico. Todas elas conspiram para fundamentar a tese de que a condição da fonte última é tanto necessária para o livre arbítrio quanto impossível de ser satisfeita.
- Não é óbvio que a adoção da visão cética implicaria consequências negativas no geral e uma fratura irremediável na nossa experiência subjetiva do mundo. Aliás, também não é claro que a adoção da visão de que há livre arbítrio seria livre de más consequências práticas, pois com ela arriscaríamos infligir direta e intencionalmente sofrimento desnecessário a outros seres sencientes.

Temo que a minha defesa do ceticismo impossibilista quanto ao livre arbítrio tenha sido um tanto esboçada e não tenha feito jus à grandiosidade do tema. No mínimo, ela pede mais investigação. Mas por que é importante investigar a posição que chamo aqui de ceticismo impossibilista? A resposta deve ser dada dividindo a pergunta em duas. Primeiro, por que é importante investigar a posição na medida em que não é apenas cética, mas também impossibilista? A resposta é que ela retiraria do debate a distração da ideia enganosa de que o determinismo especificamente é que é o problema, que nos leva, por conseguinte, à ilusão de que o *indeterminismo* poderia ajudar em alguma coisa, no que toca ao livre arbítrio e à responsabilidade moral.

Segundo, por que é importante investigar a posição no que toca ao aspecto que a distingue de posições libertaristas e compatibilistas, e não de outras posições céticas possibilistas? A resposta é que ela consegue cooptar a comparativa plausibilidade empírica do compatibilismo e a comparativa plausibilidade intuitiva do libertarismo, sem se comprometer com os aspectos de cada posição em que elas pecam.

O compatibilismo tem sido insultado pela história de *subterfúgio desprezível, marabalismo verbal* (KANT, 1996), *atoleiro de evasão* (JAMES, 1956), entre outras coisas. O que isso expressa é que ele tende a ignorar nossas intuições incompatibilistas, mais flagrantemente no que concerne à responsabilidade moral baseada em mérito básico.

O libertarismo, por sua vez, não deixou de receber insultos, o mais famoso tendo sido proferido por P.F. Strawson (1962): *metafísica panícosa*. Como o libertarista naturalista Robert Kane admite, a fim de dar conta da sua concepção de livre arbítrio:

libertaristas tem tradicionalmente invocado fatores como egos numéricos; centros de poder transempíricos; egos cartesianos; “atos volitivos” especiais; “atos de atenção” ou “volições” que não podem em princípio ser determinados; eventos mentais que existem fora da (mas podem também intervir na) ordem natural; formas especiais de causação de agente ou não ocorrente, causação essa que não pode ser explicada em termos de tipos familiares de causação de evento ou ocorrente; a ‘Vontade’ concebida como um tipo de homúnculo não causado dentro do agente; e assim por diante. (KANE, 1996, p. 115.)

Isso expressa o problema do (quase) sobrenaturalismo e do apelo ao mistério que assola o libertarismo.

As posições céticas têm sido menos explicitamente insultadas pela história. Mas isso é melhor interpretado como o julgamento de que o ceticismo nem é digno de tanta atenção, sendo obviamente *incoerente, absurdo* ou *paradoxal*. Mas vimos que as coisas não são bem assim. O cético pode abraçar muito do trabalho valoroso do compatibilista, aliviando assim certas preocupações de implausibilidade, ao mesmo tempo em que aponta para aspectos importantes da nossa experiência de liberdade e responsabilidade de que o compatibilismo não

dá conta. Porém, confrontar esses aspectos pode ter resultados inesperados: as temidas más consequências que se imaginam vir de tal confronto podem ser ilusórias ou exageradas; e boas consequências desapercibidas podem vir à tona (por exemplo, eliminar a justificação de teorias retributivistas da justiça criminal e diminuir o jugo de atitudes emocionais destrutivas, tais como a ira, nas nossas relações interpessoais (PEREBOOM, 2001)). Além disso, o ceticismo abre portas para práticas de investigação introspectiva transformadoras, que prometem contribuir para harmonizar a teoria com a experiência subjetiva e, possivelmente, a prática.

## Referências Bibliográficas

ALVAREZ, Maria. *Reasons for Action: Justification, Motivation, Explanation*. The Stanford Encyclopedia of Philosophy (Winter 2017 Edition), Edward N. Zalta (ed.). Disponível em <<https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/>> Acesso em 15.08.2022.

AYER, A. J. *Freedom and Necessity*, em **Philosophical Essays**, New York: St. Martin's Press, 3–20, 1954.

\_\_\_\_\_. **Hume: A Very Short Introduction**. Oxford University Press, 2000.

BOSWELL, James. **The Life of Samuel Johnson**. New York: Penguin Classics, 2008.

CARUSO, Gregg. **Free Will and Consciousness: A Determinist Account of the Illusion of Free Will**. Plymouth: Lexington Books, 2012.

\_\_\_\_\_. *Skepticism about Moral Responsibility*. The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), Edward N. Zalta (ed.). Disponível em <<https://plato.stanford.edu/archives/spr2018/entries/skepticism-moral-responsibility/>> Acesso em 15.09.2022.

\_\_\_\_\_. **Rejecting Retributivism: Free Will, Punishment, and Criminal Justice**. Cambridge University Press, 2021.

CARUSO, Gregg e MORRIS, Stephen G. *Compatibilism and Retributivist Desert Moral Responsibility: On What is of Central Philosophical and Practical Importance*. *Erkenntnis*, 82(4): 837–55, 2017.

COYNE, Jerry. *Clarence Darrow, Criminality, and Free Will*. Why Evolution Is True, 2 de set. de 2013. Disponível em <<https://whyevolutionistrue.com/2013/09/02/clarence-darrow-criminality-and-free-will/>> . Acesso em 15.09.2022.

DENNETT, Daniel. **Elbow Room: The Varieties of Free Will Worth Wanting**. Cambridge, Mass: MIT Press, 1984.

DENNETT, Daniel e CARUSO, Gregg. **Just Deserts: Debating Free Will**. Polity, 2021.

DORIS, John. **Lack of Character: Personality and Moral Behavior**. Cambridge: Cambridge University Press. 2002.

DRIVER, Julia. **Uneasy Virtue**. Cambridge University Press, 2001.

FISCHER, John Martin e RAVIZZA, Mark. **Responsibility and Control: A Theory of Moral Responsibility**. Cambridge: Cambridge University Press, 1998.

FLANAGAN, Owen. **The Really Hard Problem: Meaning in a Material World**. Cambridge, Mass: The MIT Press, 2007.

FRANKFURT, Harry. *Alternate Possibilities and Moral Responsibility*. Journal of Philosophy, 66: 829–39, 1969.

\_\_\_\_\_. *Reply to John Martin Fischer*, em BUSS e OVERTON (eds.). **Contours of Agency: Essays on Themes from Harry Frankfurt**, pp. 27-31. Cambridge, Mass: MIT Press, 2002.

HARRIS, Sam. **The Moral Landscape: How Science Can Determine Human Values**. Free Press, 2010

\_\_\_\_\_. **Free Will**. New York: Free Press, 2012.

\_\_\_\_\_. **Waking Up: A Guide To Spirituality Without Religion**. New York: Simons & Schuster, 2014.

HARTMAN, Robert J. **In Defense of Moral Luck: Why Luck often Affects Praiseworthiness and Blameworthiness**. Routledge, 2017.

HORGAN, Terrence. *Injecting the Phenomenology of Agency into the Free Will Debate*, em SHOEMAKER, David (ed.). **Oxford Studies in Agency and Responsibility**, 3: 34–61. Oxford: Oxford University Press, 2015.

HUME. David. **Tratado da Natureza Humana**. São Paulo: Editora

Unesp, 2000.

JAMES, William. **The Will to Believe and Other Essays in Popular Philosophy**. New York: Dover. 1956.

KANE, Robert. **The Significance of Free Will**. Oxford: Oxford University Press, 1996.

\_\_\_\_\_. **A Contemporary Introduction to Free Will**. New York: Oxford 48 University Press, 2005.

\_\_\_\_\_, (ed). **The Oxford Handbook of Free Will**. New York: Oxford

University Press, 2011.

KANT, Immanuel. **Practical Philosophy**, tr. Mary Gregor. Cambridge: Cambridge University Press, 1996.

\_\_\_\_\_. **Critique of Practical Reason**, tr. Mary Gregor. Cambridge: Cambridge University Press, 1996.

KRIPKE, Saul. **Naming and Necessity**. Cambridge, Mass.: Harvard University Press, 1980.

LATUS, Andrew. *Moral Luck*. The Internet Encyclopedia of Philosophy, J. Feiser (ed.), disponível em <<https://iep.utm.edu/moralluc/>> Acesso em: 15.09.2022.

LEVY, Neil. **Hard Luck: How Luck Undermines Free Will and Moral Responsibility**. Oxford: Oxford University Press, 2011.

LEWIS, David. *Counterfactual Dependence and Time's Arrow*. Noû 13: 455-476, 1979.

MACKIE, John L. **Ethics: Inventing Right and Wrong**. Harmondsworth, England: Penguin, 1977.

MCKENNA, Michael e PEREBOOM, Derk. **Free Will: A Contemporary Introduction**. New York: Routledge, 2016.



MICKELSON, M. Kristin. *Free Will, Self-Creation, and The Paradox of Moral Luck*. Midwest Studies in Philosophy 43 (1):224:256, 2019.

MOORE, G. E. *Principia Ethica*. New York: Cambridge University Press, 1903.

NAGEL, Thomas. *Moral Luck*, em **Mortal Questions**. New York: Cambridge University Press, 1979.

\_\_\_\_\_. **The View from Nowhere**. New York: Oxford University Press, 1986.

NAHMIAS, Eddy. *Autonomous agency and the threat of social psychology*, em MARRAFFA, M., CARO, M. e FERRETTI, F. (eds.). **Cartographies of the Mind: Philosophy and Psychology in Intersection**. Springer, 2007.

NETTLE, Daniel. **Personality: What Makes You The Way You Are**. Oxford: Oxford University Press, 2007.

NORCROSS, Alastair. *The Scalar Approach to Utilitarianism*, em WEST, Henry (ed.). **The Blackwell Guide to Mill's Utilitarianism**, pp. 217–32. Wiley-Blackwell, 2006.

PARFIT, Derek. **Reasons and Persons**. Oxford: Clarendon Press, 1987.

PARIS, Panos. *Scepticism about Virtue and the Five-Factor Model of Personality*. Utilitas 29 (4):423-452, 2017.

PARKS, Brian D. *Ultimate Moral Responsibility is Impossible: A New Defense of the Basic Argument*. Manuscrito não publicado, 2009. Disponível em: <<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.395.3763&rep=rep1&type=pdf>> Acesso em: 23.09.2022.

PEREBOOM, Derk. **Living Without Free Will**. Cambridge: Cambridge University Press, 2001.

\_\_\_\_\_. **Free Will, Agency, and Meaning in Life**. Oxford: Oxford University Press, 2014.

\_\_\_\_\_. *A Notion of Moral Responsibility Immune to the Threat from Causal Determination*, em CLARK, R., MCKENNA, M. e SMITH, A. M. (eds). **The Nature of Moral Responsibility: New Essays**, pp. 281-296. Oxford University Press, 2015.

PINKER, Steven. **The Blank Slate: The Modern Denial of Human Nature**. New York: Penguin, 2002.

PLOMIN, Robert. **Blueprint: How DNA Makes Us Who We Are**. MIT Press. 2018,

RICHARDS, Norvin. *Luck and Desert*. Mind, 65: 198–209, 1986.

SELLARS, Wilfrid. **Science, Perception, and Reality**. Ridgeview Publishing Company, 1991.

SOMMERS, Tamler. *Giving Up on Ultimate Responsibility*. Naturalism.org, set. de 2004. Disponível em <<https://www.naturalism.org/philosophy/free-will/darrow-and-determinism>>. Acesso em 15.09.2022.

SYTSMA, J. e BUCKWALTER, W. **Companion to Experimental Philosophy**. Blackwell, 2016.

ESPINOSA, Baruch de. *Ética*. Edusp, 2015.

STRAWSON, Galen. *Free Will*. 1998. Em Routledge Encyclopedia of Philosophy, Taylor and Francis, doi:10.4324/9780415249126-V014-1. Disponível em <<https://www.rep.routledge.com/articles/thematic/free-will/v-1>> Acesso em: 15.09.2022.

\_\_\_\_\_. *The Impossibility of Moral Responsibility*. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition, Vol. 75, No. 1/2, Free Will, Determinism, and Moral Responsibility, pp. 5-24, 1994.

\_\_\_\_\_. **Freedom and Belief**. Oxford: Oxford University Press, 2010.

STRAWSON, Peter F. *Freedom and Resentment*. Proceedings of the British Academy, 48: 187–211, 1962.

VAN INWAGEN, Peter. **An Essay on Free Will**. Oxford: Clarendon Press, 1983.

VIHVELIN, Kadri. **Causes, Laws, and Free Will: Why Determinism Doesn't Matter**. Oxford University Press, 2013.

WATSON, Gary. *Responsibility and the Limits of Evil: Variations on a Strawsonian Theme*, em FISCHER e RAVIZZA (eds.). **Perspectives on Moral Responsibility**, pp. 119-148. Cornell University Press, 1987.

YUDKOWSKY, Eliezer. **Rationality: From AI to Zombies**. Machine Intelligence Institute, 2015.