



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

Letícia Palazzo Rodrigues

Liberdade, determinismo e possibilidades alternativas de ação

FLORIANÓPOLIS

2024

Letícia Palazzo Rodrigues

Liberdade, determinismo e possibilidades alternativas de ação

Dissertação submetida ao Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Mestra em Filosofia.

Orientador(a): Prof. Dr. Alexandre Meyer Luz

FLORIANÓPOLIS

2024

Rodrigues, Leticia Palazzo

Liberdade, determinismo e possibilidades alternativas de ação
Leticia Palazzo Rodrigues ; orientador, Alexandre Meyer Luz, 2024.

125 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Filosofia e Ciências Humanas, Programa de Pós-Graduação em Filosofia, Florianópolis, 2024.

Inclui referências.

1. Filosofia. 2. Metafísica. 3. Filosofia da Ação. I. Luz, Alexandre Meyer. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Filosofia. III. Título.

Letícia Palazzo Rodrigues

Liberdade, determinismo e possibilidades alternativas de ação

O presente trabalho em nível de Mestrado foi avaliado e aprovado, em 11 de dezembro de 2023, pela banca examinadora composta pelos seguintes membros:

Prof. Dr. Denis Coitinho Silveira
UNISINOS

Prof. Dr. Pedro Merluzzi
PUC-Rio

Certificamos que esta é a versão original e final do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestra em Filosofia.

Coordenação do Programa de Pós-Graduação

Prof. Dr. Alexandre Meyer Luz
Orientador

Florianópolis, 2024.

AGRADECIMENTOS

Um número considerável de filósofas e filósofos generosamente leu trechos, fez comentários e sugeriu modificações importantes para esta dissertação. Agradeço em primeiro lugar ao meu orientador, o professor Alexandre Meyer Luz, e à banca, os professores Denis Coitinho e Pedro Merlussi. Seus vários incentivos, observações, objeções e comentários estratégicos foram fundamentais para que este trabalho adquirisse a forma que agora possui.

Dois grupos de pesquisadores detiveram-se com muita atenção em capítulos deste trabalho: o grupo de estudos em livre-arbítrio e responsabilidade moral Ignorantes Culpáveis e o Grupo de Escrita para Mulheres na Filosofia (GEMF). Em especial, agradeço à professora Beatriz Sorrentino Marques, que compõe ambos os grupos, e a Alex Bispo e Gabriel Maruchi – pessoas com quem tenho discutido minha pesquisa há alguns anos de maneira muito frutífera. Os outros integrantes (atuais ou anteriores) desses grupos são Thiagus Batista, Marcelo Fischborn, Vinícius Bianchi, Eduarda Calado, Nara Figueiredo e Raquel Krempel.

Deixo um agradecimento ao professor Jerzy Brzozowski e ao seu grupo de estudos em Metafísica (GELMA) pelas discussões importantes ao longo de minha pesquisa de mestrado. Meu trabalho também se beneficiou dos comentários dos públicos do XIX Encontro Nacional de Filosofia da ANPOF, do IV Colóquio Nacional e III Colóquio Internacional de Pesquisa em Filosofia da Universidade Federal de Santa Catarina e do III Colóquio de Metafísica Analítica.

Pelo trabalho diligente e auxílio aos estudantes, agradeço à secretaria do Programa de Pós-Graduação em Filosofia da UFSC, à coordenação e ao corpo docente.

Por fim, agradeço ao meu companheiro Silvio Kavetski pela leitura de trechos desta dissertação e por suas sugestões sempre úteis, além de pela parceria sensível, acolhedora e que inspira o melhor de mim.

“Em particular, atribuições de responsabilidade moral precisam ser *justas*. A justiça envolve, de modo central, que se trate pessoas diferentemente apenas se há diferenças relevantes entre elas. É injusto dar a alguém uma parcela maior de um recurso sobre o qual sua demanda não é maior do que a de outros somente em razão de sua raça, sexo ou altura. Essas são diferenças moralmente irrelevantes entre as pessoas. Apenas o que está (de certa maneira e em certo grau) sob seu controle é moralmente relevante, mas o que é questão de sorte para elas está, por definição, *fora* do controle. Culpar uma pessoa por aquilo que é questão de sorte para ela é moralmente semelhante a culpá-la pela cor de sua pele; trata-se de uma ofensa à justiça.”

Neil Levy, *Hard Luck*

RESUMO

Aparentemente, a liberdade é uma condição necessária para que uma agente tenha responsabilidade moral pela maneira como agiu e seja passível de reações como censura, punição, elogio e recompensa. Não está claro, no entanto, qual é a natureza dessa liberdade. Na literatura filosófica contemporânea, grandes avanços foram feitos nessa investigação a partir do pano de fundo teórico do Problema da Compatibilidade entre liberdade e determinismo, mas trata-se de um debate ainda em aberto. Essa discussão é o alvo geral deste trabalho. O alvo mais específico diz respeito ao conceito de liberdade de agir diferentemente, que pressupõe que, em algum sentido, a agente tinha acesso a possibilidades alternativas de ação. É plausível supor que esse acesso nunca ocorreria caso o determinismo fosse verdadeiro. Exploro então a seguinte questão: o acesso a possibilidades alternativas no momento relevante anterior a uma ação é de fato necessário para a responsabilidade moral da agente? A pergunta é importante porque, embora uma resposta afirmativa seja mais intuitiva, o influente argumento dos Casos Estilo-Frankfurt endossa uma resposta negativa. O sucesso dessa contribuição, porém, divide posições entre filósofos(as) há décadas. Tendo em vista os problemas expostos, esta dissertação possui três objetivos. O primeiro é apresentar as noções de liberdade e de responsabilidade moral, bem como o vínculo entre elas, tais como trabalhadas na Metafísica do Livre-Arbitrio e Filosofia da Ação contemporâneas. O segundo é apresentar o Problema da Compatibilidade e as principais posições em disputa (o compatibilismo e o incompatibilismo), oferecendo um panorama de argumentos influentes e alguns comentários sobre a discussão. Por fim, o terceiro objetivo é oferecer uma contribuição original para o debate em torno dos Casos Estilo-Frankfurt, defendendo uma formulação do tipo Caso Zona-Tampão Modificado que acredito ser bem-sucedida frente a objeções recentes. A posição sustentada ao final deste trabalho é, portanto, a de que possibilidades alternativas de ação não são necessárias para a responsabilidade moral.

Palavras-chave: Livre-arbitrio, Culpa, Problema da Compatibilidade, Compatibilismo, Casos Estilo-Frankfurt.

ABSTRACT

It seems that free will is a necessary condition for an agent to have moral responsibility for the way she acted, and to be subject to reactions such as blame, punishment, praise and reward. It is not clear, however, what is the nature of free will. In the contemporary philosophical literature, great advances have been made in this investigation from the theoretical framework of the Compatibility Problem between freedom and determinism, but this is still an open debate. Such discussion is the general aim of this work. The more specific target concerns the concept of freedom to act otherwise, which presupposes that, in some sense, the agent had access to alternative possibilities of action. It is plausible to assume that this access would never occur if determinism were true. I then explore the following question: is the access to alternative possibilities at the relevant moment prior to an action actually necessary for the agent's moral responsibility? This question is important because, although an affirmative answer is more intuitive, the influential Frankfurt-style Cases Argument endorses a negative answer. The success of this contribution, however, has divided positions among philosophers for decades. In view of the problems exposed, this dissertation has three objectives. The first is to present the notions of free will and moral responsibility, as well as the link between them, as worked on in contemporary Metaphysics of Free Will and Philosophy of Action. The second is to present the Compatibility Problem and the main positions in dispute (compatibilism and incompatibilism), offering an overview of influential arguments and some comments on the discussion. Finally, the third objective is to offer an original contribution to the debate around Frankfurt-Style Cases, defending a Modified Buffer Zone Case-type formulation that I believe is successful in the face of recent objections. The position supported at the end of this work is, therefore, that alternative possibilities of action are not necessary for moral responsibility.

Keywords: Free will, Blame, Compatibility Problem, Compatibilism, Frankfurt-style Cases.

SUMÁRIO

1 INTRODUÇÃO	10
2 CAPÍTULO 1: O VÍNCULO ENTRE LIBERDADE E RESPONSABILIDADE MORAL	15
1.1 INTRODUÇÃO	15
1.2 CONDIÇÕES E VARIEDADES DA RESPONSABILIDADE MORAL	18
1.2.1 Níveis de avaliação e competência normativa	18
1.2.2 Culpa, mérito moral e merecimento básico de responsabilização	24
1.2.3 Ações intencionais e ações livres	29
1.3 O PROBLEMA DA COMPATIBILIDADE: LIBERDADE E DETERMINISMO	32
1.3.1 A tese do determinismo causal	33
1.3.2 Duas ameaças à liberdade	36
1.3.3 Argumento da Consequência: um poderoso argumento incompatibilista	41
1.3.4 Algumas estratégias para o compatibilista	45
3 CAPÍTULO 2: LIBERTISTAS E COMPATIBILISTAS ENTREGUES À PRÓPRIA SORTE	50
2.1 INTRODUÇÃO: O DILEMA DO DETERMINISMO	50
2.2 COMPATIBILISMOS BASEADOS EM CONTROLE	55
2.2.1 As Visões do Eu Profundo	59
2.2.2 O Semicompatibilismo de Fischer e Ravizza	63
2.2.3 Algumas críticas	70
2.3 UMA TEORIA LIBERTISTA CONTEMPORÂNEA	75
2.3.1 Kane e o controle voluntário plural	77
2.3.2 Sorte entre mundos e aumento de controle	84
4 CAPÍTULO 3: CASOS ESTILO-FRANKFURT E A IRRELEVÂNCIA DE POSSIBILIDADES ALTERNATIVAS DE AÇÃO PARA A RESPONSABILIDADE MORAL	88
3.1 INTRODUÇÃO: OS CASOS ESTILO-FRANKFURT	88
3.2 A DEFESA DO DILEMA E ALGUMAS RÉPLICAS	91
3.2.1 Casos Estilo-Frankfurt deterministas	92
3.2.2 Os casos Zona-Tampão	96
3.2.3 A Defesa Timing	99
3.3 UMA DEFESA DOS CASOS ESTILO-FRANKFURT	102
3.3.1 Outro desafio a “Corte de Imposto” e uma revisão de Robustez	103
3.3.2 Um caso Zona-Tampão Modificado	107
3.3.3 A irrelevância de possibilidades alternativas de ação para a responsabilidade moral	111
5 CONCLUSÃO	118
REFERÊNCIAS	121

1 INTRODUÇÃO

No filme *Dançando no Escuro* (2000)¹, a protagonista Selma Jezkova (interpretada pela cantora e compositora Björk) é uma imigrante tcheca que trabalha em uma fábrica nos Estados Unidos. Ela cuida sozinha de seu filho Gene e sofre de uma doença degenerativa que está prestes a deixá-la completamente cega. Selma mantém sua doença em segredo. Com muita dificuldade (tanto devido à visão prejudicada quanto ao baixo salário que recebe mesmo trabalhando em turnos extras), ela consegue juntar quase todo o dinheiro necessário para pagar a cirurgia que é a única maneira de impedir que Gene também comece a perder a visão. Para que não seja tarde demais, a cirurgia deverá ser feita dali a alguns meses, no aniversário de treze anos do garoto. Selma e Gene vivem em um trailer estacionado na propriedade de Bill e Linda, um casal com quem desenvolvem uma amizade. Bill trabalha como policial e leva o garoto à escola todos os dias.

Certa noite, Selma e Bill trocam segredos enquanto conversam sozinhos no trailer. Ele revela que teme perder a esposa caso ela descubra que a situação financeira do casal chegou a um estado preocupante, sobretudo por causa dos gastos de Linda. Bill recebera uma herança, mas esse dinheiro já havia acabado e os dois corriam o risco de perder a própria casa. Selma incentiva que ele converse com a esposa e revela também o seu segredo, sobre a doença e sobre os esforços para a operação de Gene. Pouco tempo depois, após outra conversa com Selma, Bill finge ter saído do trailer e dá-se conta de que ela não consegue vê-lo parado ao lado da porta e observando-a. Ele descobre que Selma guarda todas as economias em uma lata de bombons, dentro de um pequeno armário, e decide roubar o dinheiro.

Quando percebe que foi roubada, Selma vai à casa de Bill para pedir suas economias de volta. Ao encontrá-lo guardando as cédulas no cofre em seu escritório, Selma coloca o dinheiro na bolsa, mas é impedida de deixar a casa. Bill a ameaça com uma arma e puxa a bolsa para si. Durante a briga, a arma acaba disparando e ele é baleado. Bill cai no chão, mas continua segurando a bolsa com firmeza e Selma não tem força para pegá-la. Nesse momento, Bill diz que ela terá que matá-lo se quiser recuperar suas economias e que prefere estar morto do que perder o amor de Linda. Selma está com a arma na mão e, desesperada, depois de várias tentativas inúteis de puxar de volta a bolsa com o dinheiro que ela simplesmente não teria condições de juntar outra vez, ela finalmente atira. Dois tiros não são suficientes para fazê-lo soltar o dinheiro. Selma pega então o cofre vazio em cima da mesa e o golpeia na cabeça de Bill até sua morte. No fim do filme, ela é julgada pelo assassinato e não diz nada

¹ Roteiro e direção de Lars von Trier.

em seu favor, já que havia prometido ao falecido amigo não revelar seu segredo a ninguém. Consequentemente, ela é condenada à morte por enforcamento.

A história de Selma é uma tragédia por uma série de razões. Ela é uma mulher de caráter irrepreensível, sonhadora e leal, que não teria feito o que fez se as circunstâncias fossem apenas um pouco diferentes. O maior objetivo de sua vida era evitar que Gene sofresse com a cegueira que ela agora enfrentava – um objetivo nobre que ela buscava realizar de maneira honesta e diligente. Selma é então colocada em uma situação extremamente injusta por alguém em quem confiava. Depois de ameaçada e esgotada, ela não vê alternativa exceto atender àquele que era o pedido do próprio Bill. Quando compreendemos tudo o que estava em jogo nas circunstâncias da ação de Selma, nossa reação a essa personagem tem duas faces: se por um lado reconhecemos que ela cometeu um assassinato brutal, por outro, atenuamos significativamente nossa atribuição de culpa², por causa daquilo que Bill fez a ela em primeiro lugar. Poderíamos inclusive dizer a Selma que ela “apenas fez o que tinha que fazer”, já que se encontrava sob uma espécie de coação.

Casos reais ou ficcionais podem ser um instrumento muito frutífero para a investigação filosófica. As características de uma história como essa podem instigar perguntas, motivar teorias e, na outra ponta, serem cruciais para testar a adequação dessas mesmas teorias. Além disso, filósofas e filósofos podem construir seus próprios casos ficcionais e estipular condições extremas para que suas teorias, ou as adversárias, sejam levadas ao limite. No espírito de uma investigação iluminada por casos, reflitamos mais um pouco a respeito da culpa de Selma pelo assassinato de Bill. Observe que, nesse contexto, nós julgamos que a agente não agiu *livremente* (ao menos não de modo pleno), embora tenha agido de maneira intencional. As circunstâncias de ação eram tais que privaram Selma da *liberdade que parece necessária para que ela seja completamente culpada* por aquilo que fez.

Imagine agora que uma outra pessoa comete um assassinato, mas que ela habita um *mundo determinista* – em linhas gerais, um mundo em que as leis da natureza fazem com que, para o estado de coisas no universo em qualquer instante, haja um único futuro possível. O assassinato teria sido então determinado pelo passado e pelas leis da natureza. A comparação entre esses dois casos pode nos levar a algumas questões importantes: será que a reação que

² Ao longo deste trabalho, utilizo os termos “culpa” e “culpado(a)” como equivalentes a, respectivamente, *blameworthiness* e *blameworthy*. Esses termos, como veremos, referem-se à responsabilidade moral de uma agente por uma ação, omissão, consequência etc. que constitui uma falha moral. A responsabilidade difere do ato de responsabilizar/culpar/censurar: ela é ontologicamente mais básica. Uma agente só é moralmente responsável por uma ação x se possui certa *relação* com essa ação ou omissão - uma relação que não se reduz ao fato de a agente ter realizado x.

sustentamos perante a ação de Selma seria justificadamente sustentada também perante esse outro agente, pelo simples fato de ele viver em um mundo determinista? Ou, em outros termos, será que esse outro agente também não é plenamente culpado pelo modo como agiu? O que a resposta a essas perguntas nos diz sobre a *natureza* da liberdade que parece estar associada à culpa por uma ação? E qual seria, afinal, essa natureza?

Há pelo menos uma boa razão para nos preocuparmos com essas questões e seus desdobramentos. Como afirma Neil Levy (citado na epígrafe deste trabalho), atribuições de *responsabilidade moral* (por exemplo atribuições de culpa) precisam ser justas. Não parece justo tratar diferentemente duas pessoas sem que a diferença entre elas seja moralmente relevante. Nesse sentido, ao menos alguns elementos que privam uma agente do controle ou da liberdade em relação a uma ação pela qual essa agente é punida (como no caso de Selma) assemelham-se a contextos de punição ou dano devido a raça/sexo – diferenças que não oferecem um fundamento para esse tratamento. Ao final de *Dançando no Escuro*, sentimos que uma grande injustiça foi cometida contra Selma. A investigação da natureza da liberdade necessária à responsabilidade moral e dos fatores que representam uma ameaça a essa liberdade pode tornar mais claras as circunstâncias em que estaríamos cometendo uma injustiça desse tipo contra uma agente.

Na literatura filosófica contemporânea, a natureza da liberdade associada à responsabilidade moral por uma ação tem sido investigada a partir do pano de fundo do *Problema da Compatibilidade*: liberdade e determinismo são compatíveis? Enquanto alguns autores pensam que as condições nas quais é apropriado atribuir por exemplo culpa a uma agente podem se dar mesmo que o determinismo seja verdadeiro, outros argumentam que o determinismo precisa ser falso (isto é, o *indeterminismo* precisa ser verdadeiro), ou ninguém é de fato culpado pela maneira como age. Acredito que a tese do determinismo, mesmo se for descoberta falsa, oferece um modelo teórico em alguns sentidos útil para que a natureza da liberdade (e da própria responsabilidade) seja descrita de maneira mais acurada.

Essa investigação, tal como enquadrada pelo Problema da Compatibilidade, compõe um debate ainda em aberto. Ela será o alvo fundamental deste trabalho. Explorarei também uma questão mais específica: *o acesso a possibilidades alternativas no momento relevante anterior a uma ação é necessário para que uma agente seja responsável por realizar essa ação?* Essa pergunta é importante porque, intuitivamente, a liberdade necessária para a responsabilidade moral envolve que a agente tenha a capacidade de agir de ao menos dois modos diferentes. Pensemos outra vez no caso de Selma: se assumirmos que ela, embora tenha agido intencionalmente, não agiu livremente porque foi coagida e apenas “fez o que

tinha que fazer”, um elemento que parece estar por trás dessa explicação é que Selma não poderia ter agido diferentemente. Em certo sentido, é claro que ela poderia: a qualquer momento do início do conflito com Bill, Selma poderia ter desistido de recuperar seu dinheiro e ido embora. Mas observe que, nesse exemplo, a noção de “poder agir de outro modo” que importa diz respeito a alternativas *razoáveis* para o contexto dessa agente. Quando familiarizados com a história de Selma, compreendemos de imediato que não era uma opção simplesmente desistir da única maneira pela qual ela poderia salvar o filho de uma grave doença degenerativa.

O princípio de que uma agente precisa ter podido agir de outro modo (em uma oportunidade minimamente razoável) para que seja responsável pela ação que realizou foi, porém, objetado pelo influente argumento dos *Casos Estilo-Frankfurt*. É nesse ponto que a questão exposta acima deixa de ter uma resposta afirmativa trivial e ganha importância: se há um contraexemplo bem-sucedido a uma ideia tão intuitiva, como pretendem ser os casos elaborados por Harry Frankfurt (1969), então *talvez seja preciso repensar muito daquilo que associamos à noção de liberdade e à responsabilização justa de pessoas*. A discussão sobre possibilidades alternativas também tem uma óbvia relevância no contexto do Problema da Compatibilidade: se as leis da natureza deterministas fazem com que haja um único futuro possível a qualquer instante e se o determinismo é verdadeiro, então parece não haver possibilidades alternativas a qualquer evento ou cenário que se efetiva – incluindo, é claro, as ações humanas. Mas isso não seria um obstáculo para a existência de responsabilidade moral.

No terceiro capítulo deste trabalho, endossarei a posição de Frankfurt e defenderei que uma formulação mais complexa de um Caso Estilo-Frankfurt é bem-sucedida em objetar o que se convencionou chamar de *Princípio das Possibilidades Alternativas*. Esse capítulo contará então com uma contribuição original para o debate e será mais específico e denso do que os anteriores.

Antes de chegar a essa questão, porém, apresentarei os principais conceitos, argumentos e posições presentes no debate contemporâneo mais amplo. A discussão filosófica sobre liberdade e responsabilidade moral é vasta e tem sido especialmente frutífera e ramificada desde a década de 1960. Por essa razão, este trabalho tenta apresentá-la a partir de uma certa linha de raciocínio, priorizando um pouco menos a amplitude de posições e argumentos abordados do que algum nível de aprofundamento naqueles aspectos do debate que considero mais importantes e nos caminhos que me parecem mais promissores. Além disso, busquei organizar, em especial no primeiro capítulo, alguns pressupostos da discussão do Problema da Compatibilidade que poucas vezes são tornados explícitos pelas filósofas e

filósofos que o lideram. A interpretação que apresento desses pressupostos talvez leve a algum tipo de compromisso teórico, mas tentei ser tão neutra quanto possível ao lidar com essa lacuna.

Este trabalho possui então três objetivos gerais: (1) apresentar as noções de liberdade e de responsabilidade moral, bem como o vínculo entre elas, tais como trabalhadas na Metafísica do Livre-Arbitrio e Filosofia da Ação contemporâneas; (2) apresentar o Problema da Compatibilidade e as principais posições em disputa (o compatibilismo e o incompatibilismo), oferecendo um panorama de argumentos influentes e alguns comentários sobre a discussão; e (3) defender um Caso Estilo-Frankfurt do tipo *Zona-Tampão Modificado* que chamei de *Decepção*, tendo em vista objeções recentes às formulações mais sofisticadas dos Casos. O objetivo (1) será realizado no primeiro capítulo, *O vínculo entre liberdade e responsabilidade moral*. O objetivo (2) está distribuído entre o primeiro e o segundo capítulo, de título *Compatibilistas e libertistas entregues à própria sorte*. Já o terceiro capítulo, *Casos Estilo-Frankfurt e a irrelevância de possibilidades alternativas de ação para a responsabilidade moral*, contará com uma tentativa original de contribuir para o debate por meio da realização do objetivo (3).

2 CAPÍTULO 1: O VÍNCULO ENTRE LIBERDADE E RESPONSABILIDADE MORAL

1.1 INTRODUÇÃO

Em novembro de 2022, no trecho da BR-364 entre duas cidades do estado de Mato Grosso, um pai tentava levar o filho para a realização de uma cirurgia oftalmológica sem a qual o menino corria o risco de perder um dos globos oculares. A cirurgia havia sido marcada com alguma dificuldade e precisava ser feita com urgência. Um ponto da BR, porém, estava interditado por apoiadores do ex-presidente Jair Bolsonaro que se manifestavam contra a recente derrota do candidato nas eleições. Chegando ao bloqueio, já à noite, o pai discute com os apoiadores e pede passagem. Mas um deles responde que não se importa caso o menino fique cego. Pai e filho são então ameaçados com um facão e impedidos de prosseguir com a viagem³.

Imagine que o pior tivesse acontecido e que, devido ao bloqueio que o impediu de chegar à cirurgia, o menino de fato perdesse um dos olhos. Suponha também que o grupo de inconformados fosse liderado por um homem, Jorge, e que esse homem tomasse todas as decisões importantes relativas a, por exemplo, deixar ou não alguém passar pelo bloqueio. Nessa versão semificcional da história, é Jorge quem proíbe o pai de seguir adiante, ameaçando-o com o facão e dizendo que não se importa caso o menino fique cego. A ação relevante aqui será descrita como “impedir o menino de chegar à cirurgia”. Considerando que é razoável supor que Jorge estava ciente do que fazia e da consequência provável de sua ação (o menino ficando cego de um dos olhos), tinha controle sobre aquilo que dizia e fazia e, no momento da decisão, não estava sujeito a coerção/manipulação⁴ etc., também é razoável supor que esse agente seja moralmente responsável por aquilo que fez, assim como pela consequência terrível de sua ação.

Neste trabalho, a preocupação mais fundamental será *a natureza da liberdade que parece estar associada à responsabilidade moral de uma agente por uma ação, bem como os*

³ Após o ocorrido, pai e filho conseguiram remarcar a cirurgia. Cf. reportagem de Vinícius Prates, “Pai se revolta com bloqueio de bolsonaristas que impedia cirurgia do filho”, no *Correio Braziliense* (2022).

⁴ Provavelmente, uma pessoa como Jorge que viva o contexto político brasileiro e que tenha sido cooptada pelo bolsonarismo está sim sujeita a algum grau de manipulação. A compreensão da extensão e das consequências dessa forma de manipulação é um assunto que merece investigação por si próprio, tanto da parte de filósofos(as) quanto de cientistas sociais, psicólogas(os) etc. A suposição importante aqui, porém, é que essa manipulação “sistemática”, digamos, não parece privar Jorge do controle relevante sobre a ação de impedir o menino de chegar à cirurgia, de modo que seria adequado responsabilizá-lo por essa ação em particular.

fatores que talvez representem uma ameaça à existência dessa liberdade. Jorge não apenas impediu o menino de chegar à cirurgia como também parece tê-lo feito *livremente*. Esse elemento é crucial para que seja apropriado responsabilizá-lo: tanto os nossos (prováveis) sentimentos de indignação e raiva quanto o juízo de que Jorge merece ser punido por ter feito o que fez estão ancorados nessa suposição. Como veremos, no entanto, não é tarefa muito fácil dizer o que é a liberdade no sentido aqui em questão. A dificuldade é revelada na medida em que avaliamos fatores que estão (ou poderiam estar) presentes nas sequências causais levando às nossas ações e que parecem incompatíveis com a liberdade.

O próprio termo “responsabilidade moral” possui diferentes usos no vocabulário cotidiano; e pode ser que haja variedades de responsabilidade, de modo que é preciso especificar aquilo de que estamos falando. Deixando de lado a questão sobre a legitimidade política/moral ou legal da forma de protesto escolhida pelos apoiadores de Bolsonaro, não é irrazoável fazer, para nosso caso semificcional, a seguinte suposição: na posição de líder, era responsabilidade de Jorge que o bloqueio, ao mesmo tempo em que cumprisse sua função de protesto, causasse apenas o dano “necessário”, ou “tolerável”, derivado da interdição da BR-364. Particularmente, podemos dizer que ter permitido que pai e filho passassem pelo bloqueio não apenas não afetaria a efetividade do protesto, como também era um dever de Jorge, dada a urgência da cirurgia e a consequência trágica de não a realizar. Jorge, então, era responsável no sentido de que tinha uma *obrigação moral* a cumprir: em casos de dano grave a alguém, o bloqueio é que deveria ser momentaneamente interrompido, em vez de priorizado às custas da pessoa sofrendo o dano. Tendo contrariado sua obrigação moral, ou falhado em cumprir o seu dever, Jorge realizou uma ação moralmente errada.

Mas não é esse o sentido de “responsabilidade moral” que nos interessa aqui. O outro sentido em que Jorge é responsável é *retrospectivo*, isto é, diz respeito a algo que já aconteceu: a ação e sua consequência. Além disso, ele diz respeito a uma *relação* entre Jorge e sua ação que nos “diz algo” sobre a agência moral desse indivíduo. Ele é *culpado* por ter impedido o menino de chegar à cirurgia, resultando na perda de um de seus globos oculares. Observe que essa não é uma relação trivial. Em um cenário diferente, talvez Jorge tivesse causado o impedimento da passagem do pai e do filho – sendo então “causalmente responsável” – sem, no entanto, tê-lo feito intencionalmente, ou ciente do que fazia. Mas não é isso o que ocorre; na verdade, Jorge chega a dizer que não se importa com o que acontecerá ao menino. A responsabilidade moral é, pois, algo mais exigente do que a mera responsabilidade causal. O gato que sobe em minha mesa e derruba o meu estojo de canetas no chão é causalmente, não moralmente, responsável pela queda do estojo. Essa ação não diz

(embora eu possa suspeitar das más intenções do gato) nada de substancial sobre sua agência moral, já que nem mesmo pensamos que um gato seja o tipo de indivíduo ao qual atribuímos agência moral⁵. Esse exemplo mostra também que ser causalmente responsável não é suficiente para ser moralmente responsável por uma ação ou consequência de ação.

Os exemplos discutidos acima levantam algumas questões a respeito de *responsabilidade moral*.⁶ Primeiramente, cabe esclarecer o papel que o termo “moral” desempenha aqui. Em segundo lugar, nem todos os indivíduos são agentes morais, no sentido de serem passíveis de avaliação moral e, em especial, de poderem exhibir com suas ações (ou omissões, ou consequências dessas coisas⁷) o tipo de relação “ser moralmente responsável”. Mesmo entre agentes morais, inclusive, haverá instâncias de ações que não satisfazem todas as exigências para a responsabilidade moral. Podemos então questionar que características distinguem cada um desses casos. Uma outra questão é como exatamente a liberdade e a responsabilidade moral estão relacionadas; e sob que condições essa liberdade é possível. Por fim, podemos ainda nos perguntar sob que condições essa liberdade (que tão imediatamente julgamos existir, como no caso de Jorge) *não* seria possível.

O objetivo deste capítulo é apresentar um tratamento inicial a todas essas questões, com base nas distinções, conceitos e argumentos da literatura filosófica contemporânea sobre liberdade e responsabilidade moral. Na seção 1.2, farei uma apresentação ampla das condições necessárias e variedades da responsabilidade moral, considerando algumas das circunstâncias em que atenuamos ou suspendemos nossos juízos de responsabilidade e as razões que sustentam essas reações. Na seção 1.3, tratarei mais diretamente da noção de liberdade e de como a sua natureza tem sido investigada sob o recorte do *Problema da Compatibilidade* (com o determinismo causal). Por fim, apresentarei uma das contribuições centrais para esse debate: o Argumento da Consequência. Os pontos introduzidos neste capítulo servirão como subsídio teórico para as discussões dos capítulos seguintes.

⁵ O gato é, porém, um sujeito moral, no sentido de ser digno de considerações morais: há uma série de ações contrárias aos interesses de um gato cuja realização, por parte de uma pessoa, é moralmente errada. Essa distinção será apresentada a seguir.

⁶ É claro que o sentido de responsabilidade que interessa a este trabalho também difere do uso desse termo como um sinônimo de prudência, ou de cumprimento das obrigações.

⁷ Daqui em diante, farei referência quase sempre apenas a ações, por uma questão de simplicidade. É importante notar, porém, que os objetos *diretos* da responsabilidade moral no domínio dos comportamentos (isto é, excluindo a responsabilidade que talvez tenhamos por nossos sentimentos, traços de caráter moralmente relevantes etc.) são apenas ações e omissões. Somos responsáveis por algumas consequências de nossas ações e omissões de maneira *indireta*, (grosso modo) em virtude da responsabilidade por alguma ação ou omissão anterior e de se essas eram consequências razoavelmente previsíveis. Nesse trabalho, não tratarei das especificidades da responsabilidade moral por consequências. Para um tratamento dessa questão, cf. Fischer e Ravizza, 1998.

1.2 CONDIÇÕES E VARIEDADES DA RESPONSABILIDADE MORAL

Nesta seção, a estratégia principal será a consideração de uma série de casos, por meio dos quais delimitarei tanto o assunto quanto os pontos de partida teóricos da discussão abordada nesse trabalho. Na primeira seção, o objetivo será distinguir os dois níveis de avaliação (da ação e da agente) presentes em qualquer atribuição de responsabilidade moral, considerando em seguida duas classes de casos úteis à investigação sobre suas condições necessárias. Uma primeira condição será introduzida: a competência normativa. Na segunda subseção, prosseguirei com a delimitação de assunto ao apresentar os tipos paradigmáticos (culpa e mérito moral) e as variedades de responsabilidade que talvez estejam presentes em nossas práticas, restringindo o objeto deste trabalho (caso o pluralismo seja verdadeiro) à responsabilidade moral no sentido associado a merecimento básico de responsabilização. Na terceira subseção, por fim, farei uma distinção entre ações intencionais e ações livres, enfatizando que na segunda condição aparentemente necessária para a responsabilidade moral – a ação deve ser intencional – está inclusa uma condição epistêmica. O ponto a ser atingido ao final desta seção é o seguinte: apesar de a realização intencional e normativamente competente de uma ação ser necessária para que uma agente aja livremente e seja responsável por essa ação, essa não é ainda uma condição suficiente.

1.2.1 Níveis de avaliação e competência normativa

Retomemos o exemplo de Jorge. Ao distinguir a noção de responsabilidade moral que está aqui em questão da noção de responsabilidade como obrigação, aponte o fato de que Jorge realiza uma ação moralmente errada. A afirmação de que esse agente é responsável por uma ação errada *equivale* à afirmação de que ele é culpado⁸. Observe que estão aqui em jogo dois níveis de avaliação distintos. Quando dizemos que foi moralmente errado impedir o menino de chegar à cirurgia, estamos nos referindo à *ação* – para qualquer outra pessoa que ocupasse o lugar de Jorge, a ação ainda seria errada. Nesse nível, o que importa é a *relevância* e a *valência moral* (ou qualidade moral) da ação; isto é, se ela é ou não significativa de um ponto de vista moral (por exemplo, se afeta outras pessoas de uma maneira que produza dano/malefício, ou que produza benefício), e se ela tem valência negativa (é errada) ou

⁸ Há autores que defendem uma “visão subjetiva” acerca da culpa, segundo a qual uma agente não precisa agir de maneira objetivamente errada para ser culpada – bastaria a “condição subjetiva de que a agente *considera* sua conduta como objetivamente (...) errada” (MCKENNA, 2012, p. 14). Nesse trabalho, porém, assumirei a posição majoritária de que a ação *ser* moralmente errada é uma condição necessária para culpa.

positiva (é correta, ou é louvável). Quanto à valência, há também a possibilidade de que uma ação seja moralmente neutra, no sentido de que ela possui relevância moral, mas é meramente permissível, ou apenas satisfaz uma obrigação moral sem atingir o estatuto de ação louvável (ou de boa ação para além do que era razoável esperar da agente naquele contexto). Há um debate sobre se a avaliação de ações em termos de certo e errado é a única forma de avaliação pertinente, ou se deveríamos adotar uma visão mais inclusiva, também em termos de bom e mau, virtude e vício etc. (MCKENNA, 2012). Nesse trabalho, porém, estarei focada em casos nos quais “as valências morais pertinentes são enquadradas em termos de certo e errado, sendo que certo e errado podem ser entendidos em termos das noções deontológicas de obrigação, dever, permissibilidade e sobreterrogação”⁹ (MCKENNA, 2012, p. 15).

Quando identificamos, por outro lado, que Jorge exhibe com sua ação errada a relação “ser moralmente responsável”, ou “ser culpado”, a avaliação é acerca do *agente*. Trata-se, portanto, de uma avaliação a ser feita separadamente da primeira. Se houvesse algum modo razoável de Jorge argumentar que, dado o contexto e a importância do protesto, sua ação na verdade não fora errada (apenas suponha, por mais implausível que isso pareça), então ele teria uma *justificativa* para agir como agiu e não seria culpado, mas somente em virtude de a ação não ser errada¹⁰. Esse não é, no entanto, o único tipo de apelo ou de contestação de responsabilidade moral disponível a um agente (ou a quem queira defendê-lo). Existem espécies de considerações diretamente acerca da agência de uma pessoa e que nos fazem modificar, atenuar ou até mesmo suspender por completo o nosso juízo de que ela é moralmente responsável por uma ação (STRAWSON, [1962] 2008). Essas considerações também podem ser capazes de alterar os nossos sentimentos associados a esse juízo – como indignação, raiva, rancor, ressentimento etc.

Estipulemos algumas versões diferentes da história de Jorge, com o objetivo de examinar que tipos de considerações emergem sobre sua responsabilidade por impedir o menino de chegar à cirurgia¹¹. Primeiramente, suponha que Jorge estava sujeito a condições bastante comprometedoras de sua agência, de uma maneira global. Uma possibilidade é a de que ele possuísse uma grave doença mental, ou fosse neurodivergente de uma forma muito relevante, de modo que não diríamos que ele possui suficiente controle sobre suas ações.

⁹ Todas as citações diretas deste trabalho são traduções livres minhas.

¹⁰ Cf. Brink e Nelkin (2013): “Justificativas (...) negam que a ação tenha sido errada, insistindo que o comportamento que de outra maneira teria sido errado não o era naquelas circunstâncias” (p. 290).

¹¹ O movimento a ser feito a seguir é uma reconstrução das duas classes de apelos ou contestações de responsabilidade moral plena identificadas no conhecido artigo *Freedom and Resentment*, de P. F. Strawson ([1962] 2008). Esse texto é o fundador das abordagens da responsabilidade moral na filosofia contemporânea. Cf. tradução para o português em Strawson, 2015.

Outra possibilidade é a de que Jorge tivesse anteriormente sido vítima de uma lavagem cerebral tão violenta e profunda, ou de algum dano neurológico tão trágico¹², que já nem pudéssemos dizer que esse agente era “ele mesmo”. Por fim, imagine uma versão um pouco diferente da história. Suponha que, por algum descuido (e devido à péssima ideia de um dos apoiadores inconformados de levar o próprio filho para fazer parte do bloqueio na BR), uma criança de 11 anos de idade é quem acaba tomando a decisão de impedir o menino de chegar à cirurgia. Observe que, em todos esses casos, nós não julgaríamos que o agente é um alvo apropriado de responsabilização plena e de punição – mas de tratamento, ou proteção, ou reclusão¹³.

P. F. Strawson ([1962] 2008)¹⁴, ao dirigir a atenção para casos como esses, notou o seguinte:

Eles não nos convidam a encarar a ação do agente de uma maneira consistente com a plena retenção das atitudes interpessoais habituais e meramente inconsistente com uma ação particular. Eles nos convidam a ver o agente *ele próprio* sob uma luz diferente daquela em que normalmente veríamos alguém que agiu como ele agiu. (p. 9, grifo meu)

Considere agora a seguinte “taxonomia” de indivíduos de uma comunidade moral, descrita por Michael McKenna (2012):

A classe de agentes morais está contida dentro e é mais restrita do que a classe mais ampla das pessoas, e a classe de agentes moralmente responsáveis está contida dentro e é mais restrita do que a classe mais ampla de agentes morais. Um agente moral é uma pessoa capaz de ação que pode ser moralmente avaliada como boa ou má, certa e errada, virtuosa ou viciosa. Algumas pessoas não são capazes de ação que possa ser moralmente avaliada. (...) É claro, todas essas pessoas, as que são e as que não são agentes morais, são sujeitos adequados de consideração moral. Elas são sujeitos morais. Toda pessoa enquanto pessoa é um sujeito moral. Portanto, elas possuem direitos distintos das pessoas, como o direito à vida; e outros agentes morais certamente têm deveres e obrigações correlatos em relação a elas. Observe, além disso, que a própria classe dos sujeitos morais é mais ampla do que a classe das pessoas. Vários animais além das pessoas são

¹² Casos como o de Phineas Gage, o mais famoso das neurociências (cf. DAMASIO, 2012), exemplificam bem esse ponto. Podemos dizer que, depois do acidente, Gage deixou de ser parte do grupo dos agentes morais, embora seja ainda uma pessoa e, como tal, um sujeito digno de receber consideração moral. Uma outra maneira de colocar o mesmo ponto é que Gage perdeu sua competência normativa. Apresentarei essas noções a seguir.

¹³ Essas atitudes são aquilo que P. F. Strawson chamou de *atitudes objetivas*, em contraste com as *atitudes reativas* de envolvimento ou participação em uma relação interpessoal típica – como as já mencionadas indignação, raiva, rancor e ressentimento, mas também como gratidão, perdão, entre outras. Localizarei o papel das atitudes reativas, inclusive as de valência positiva, na subseção 1.2.2.

¹⁴ A posição de Strawson no debate sobre liberdade e responsabilidade moral será explorada (de modo breve) no próximo capítulo.

sujeitos morais e, portanto, candidatos à consideração moral. (MCKENNA, 2012, p. 11)¹⁵

Que conclusão podemos extrair acerca desses casos? Parece que o grupo de considerações que acabo de elencar – possuir uma doença mental grave, estar sujeito à lavagem cerebral extrema, ser uma criança etc.¹⁶ – forma uma classe de indivíduos que, embora sejam pessoas (e, portanto, sujeitos morais, cujos interesses devem ser levados em conta), não contam como *agentes morais*. Não se trata de dizer que esses indivíduos são isentos de culpa meramente por impedir a criança de chegar à cirurgia; o ponto é que, de maneira global, eles não são capazes (ou são significativamente *menos* capazes) de realizar ações que possam ser moralmente avaliadas. Como consequência, essas pessoas não são moralmente responsáveis por suas ações em geral – ao menos temporariamente, como no caso das crianças (cuja agência moral está em desenvolvimento). Mas resta ainda fazer a seguinte questão: o que é que esses casos teriam em comum?

Uma boa hipótese é que essas agentes não possuem *competência normativa*¹⁷. Considere uma criança muito pequena. É inegável que crianças pequenas ajam intencionalmente, com base naquilo que desejam e da forma que pensam ser a adequada para a satisfação desses desejos. Se elas deliberam a respeito de uma ação, no entanto, e se em função disso exercem algum grau de autocontrole sobre seus desejos mais fortes, trata-se “primariamente de raciocínio instrumental, a serviço de objetivos fixados” (BRINK & NELKIN, 2013, p. 292). Agentes morais, por outro lado, precisam “não simplesmente agir baseados em seus desejos mais fortes, mas ser capazes de dar um passo atrás a respeito desses desejos, avaliá-los, e agir por boas razões. Isso exige que (...) possam reconhecer e responder a razões para agir” (BRINK & NELKIN, 2013, p. 292). O fato de uma ação ser intencional, pois, não é suficiente para que a agente seja moralmente responsável.

¹⁵ Observe que McKenna distingue o grupo dos agentes morais de um grupo mais restrito, o de agentes moralmente responsáveis. Essa distinção se justifica pela presença, na literatura sobre liberdade e responsabilidade moral, de filósofos(os) que sustentam um ceticismo sobre a existência da responsabilidade na variedade que está aqui em jogo (*accountability*, que será mais bem esclarecida a seguir), seja em absoluto, seja se o determinismo causal for uma tese verdadeira sobre o mundo. Essas posições serão discutidas ao longo do trabalho. O ponto aqui é somente que da inexistência de agentes moralmente responsáveis não se segue a inexistência de agentes morais.

¹⁶ Filósofos(os) tipicamente incluem nesse grupo alguns psicopatas. Cf. por exemplo Fischer e Ravizza, 1998.

¹⁷ A ideia de competência normativa tem sido amplamente desenvolvida pelos defensores do compatibilismo no Problema da Compatibilidade (que será apresentado na seção 1.3). Cf., por exemplo, Wolf, 1990; Fischer e Ravizza, 1998; Nelkin, 2011; e McKenna, 2012. Suas exigências serão melhor desenvolvidas no capítulo 2 desse trabalho. Que a competência normativa seja necessária para a liberdade e a responsabilidade moral, porém, é um ponto pacífico entre os adversários no Problema da Compatibilidade: cf., por exemplo, Kane, 2016; e Franklin, 2017 (ambos defensores do incompatibilismo).

Christopher Franklin (2017) oferece a seguinte caracterização do que é, para uma agente, ser normativamente competente:

Competência normativa é uma capacidade complexa que consiste em um conjunto de capacidades volitivas e cognitivas. Quanto a capacidades cognitivas, a competência normativa exige uma apreensão dos conceitos morais em jogo na responsabilidade moral, tais como os de ações erradas, obrigação e culpa, assim como a capacidade de trazê-los para operar em uma ampla gama de situações nas quais eles se aplicam – isto é, ser capaz de antecipar demandas e ajustar o próprio comportamento à luz de outros princípios morais. Quanto a capacidades volitivas, a competência normativa exige as capacidades de controlar e regular o próprio comportamento à luz dessas razões. Em outras palavras, a agente precisa possuir *arbitrio*¹⁸: a capacidade de dar um passo atrás e avaliar criticamente suas motivações à luz de razões morais, de fazer escolhas com base em suas avaliações, e de traduzir essas escolhas em ação. (p. 40)

Nas circunstâncias listadas há pouco, os agentes plausivelmente *falham* em atender às exigências da competência normativa (ao contrário de Jorge na história inicial). Nessa classe de indivíduos, as exceções, a atipicidade ou a imaturidade psicológica/moral podem fazer com que o agente por exemplo não reconheça a urgência da cirurgia oftalmológica como uma razão suficiente para deixar que pai e filho passem pelo bloqueio na BR. A falta de competência normativa, no entanto, não é a única fonte de considerações acerca da agência de uma pessoa que podem nos fazer suspender, por completo ou parcialmente, nossos juízos de responsabilidade moral e os sentimentos a eles atrelados.

Em segundo lugar, considere a figura de Jorge à luz de outras duas versões do caso. Suponha que ele não tivesse recebido do pai a informação mínima relevante sobre o que estaria em jogo caso os dois fossem impedidos de passar pelo bloqueio. Ou então imagine que Jorge na verdade não tomou sua decisão de maneira usual: ele fora diretamente coagido por um organizador dos bloqueios de estradas a nível nacional. Essa pessoa ameaçara atentar contra a segurança física da família de Jorge caso ele abrisse quaisquer exceções à interdição na BR-364 (suponha também que esse organizador estivesse presente no momento da ação, para se certificar de que sua coação fosse efetiva)¹⁹.

Quanto a essa classe de casos, Strawson observa que

¹⁸ N. T.: no original, *will*. A discussão aqui ainda não diz respeito, porém, ao *livre-arbitrio*. Isso ficará claro mais adiante no texto, na subseção 1.2.3 e na seção 1.3.

¹⁹ Outras considerações dessa mesma classe poderiam ser estresse/fadiga extremos, ou dependência química severa (desde que a agente não seja responsável por ter se colocado nessas circunstâncias de ação). Em geral, parece que ações “compulsivas” – motivadas por impulsos e desejos da agente tão difíceis de resistir que pensaríamos que as ações estavam fora de seu controle – são ações pelas quais há menor grau, ou nenhum grau, de responsabilidade moral. Explorarei um exemplo desse tipo mais adiante, na subseção 1.2.3.

Nenhum deles nos convida a suspender perante o agente, no momento da ação ou em geral, nossas atitudes reativas ordinárias. Eles não nos convidam a enxergar o *agente* como alguém a respeito de quem essas atitudes são de algum modo inapropriadas. Eles nos convidam a enxergar o *dano* como algo a respeito do qual uma dessas atitudes em particular é inapropriada. (...) Eles nos convidam a ver o *dano* como algo pelo qual ele [o agente] não é plenamente, ou não é em absoluto, responsável. ([1962] 2008, p. 8)

As duas possibilidades acima levantadas diferem relevantemente da classe anterior²⁰. Em primeiro lugar, a competência normativa de Jorge – e conseqüentemente seu estatuto enquanto agente moral – não estava comprometida. Além disso, podemos dizer que há alguma ação que esse agente realiza intencionalmente: na primeira versão, uma ação cuja descrição não incluirá fatores importantes do risco que o menino corria ao não passar pelo bloqueio (porque Jorge não estava ciente desses fatores); na segunda versão, uma ação de cuja descrição relevante fará parte a circunstância de coação. Não é o caso, porém, que Jorge livremente impeça o menino de chegar à cirurgia, e livremente provoque a perda de seu globo ocular. Nossa tendência seria a de fazer considerações como “Ele não sabia o que fazia” e “Ele foi obrigado”, ou “Ele não tinha alternativa”. À luz dessas considerações, nós avaliariamos que Jorge não é culpado, ou que não é plenamente culpado. Isso significa que, embora seja necessária, a competência normativa não é suficiente para que uma agente seja moralmente responsável por uma ação.

As razões pelas quais *nesses* casos não há responsabilidade moral plena são mais complexas do que na classe anterior. Observe que a espécie de capacidade exigida para competência normativa é bastante geral, de modo que nem estamos ainda exigindo que a agente tenha a oportunidade de exercê-la naquele contexto de ação em particular. Para explicar aquilo que ameniza ou suspende a responsabilidade moral da agente na primeira classe de considerações, não há necessidade de apelar para a noção de liberdade (FRANKLIN, 2017). Os fatores que, como veremos, parecem ameaçar a liberdade de uma agente (ou de todos os agentes morais) não representam razões para duvidar de que pessoas ordinariamente possuem competência normativa nem desempenham papel explicativo nos casos de ausência dessa competência. Por outro lado, as considerações da segunda classe dizem respeito a algo na agência do indivíduo que não ocorreu da maneira necessária para a existência de responsabilidade *naquele* contexto. Esses casos desempenham um papel importante na investigação sobre as condições para a responsabilidade moral: por meio de exemplos de ausência, podemos nos perguntar o que há de especial nos exemplos de presença

²⁰ Essas duas classes de apelos ou contestações de responsabilidade moral plena ficaram consolidadas na literatura contemporânea respectivamente pelos termos *exemptions* e *excuses*. Cf. Brink e Nelkin, 2013.

de responsabilidade (assumindo a competência normativa da agente). Ao longo do trabalho, essa estratégia será mais bem explorada, começando pela seção 1.2.3 abaixo. Na próxima subseção, porém, continuarei especificando o alvo da discussão por meio de dois esclarecimentos: o papel do termo “moral” na expressão “responsabilidade moral” e a variedade de responsabilidade aqui em questão (considerando que alguns autores defendem a existência de variedades).

1.2.2 Culpa, mérito moral e merecimento básico de responsabilização

Na subseção anterior, ao distinguir os dois níveis de avaliação (da ação, do agente) em jogo quando uma pessoa por exemplo possui culpa, mencionei que estamos atentos tanto à relevância quanto à valência moral dessa ação. Há pelo menos uma boa razão para atrelar a investigação sobre liberdade à responsabilidade *moral*. Parece que os momentos em que mais nos importamos com a existência e a natureza da liberdade são aqueles que envolvem moralidade em sentido amplo – tanto as ações que afetam o bem-estar e os interesses de outras pessoas quanto as ações afetam o *nosso* bem-estar e interesses, na medida em que entendamos que existem deveres em relação a nós próprios. Também nos importamos com a liberdade em sentido político; mas a noção de liberdade nesse âmbito de discussão muitas vezes difere daquela que é alvo deste trabalho²¹. Assim, tratarei exclusivamente de contextos moralmente relevantes e particulares, em relação a um indivíduo. Nesta subseção, veremos que uma agente pode ser responsável por ações de três valências: positiva, negativa e neutra. Apontarei também que a responsabilidade moral no sentido aqui em questão envolve que a agente mereça ser responsabilizada, e de maneiras específicas.

Consideremos mais cuidadosamente aquilo que P. F. Strawson chamou de *attitudes reativas*: “atitudes e reações não-distanciadas de pessoas diretamente envolvidas em relações umas com as outras; (...) atitudes e reações de partes ofendidas e de beneficiários; de coisas como gratidão, ressentimento, perdão, amor e sentimentos feridos” ([1962] 2008, p. 5). Strawson chama a atenção para o fato de que esse grupo de sentimentos é parte central das nossas *práticas de responsabilização* – e, segundo o autor, são a própria fundação da responsabilidade moral (MCKENNA & PEREBOOM, 2016). Dirigir uma atitude reativa a outra pessoa (ou sustentar uma atitude “autorreativa”, como sentir-se culpado e arrependido) é responsabilizá-la. Na subseção anterior, apontei que afirmação de que Jorge é responsável por

²¹ Ela pode dizer respeito, por exemplo, a poderes ou limitações na atividade do Estado pensadas para que os indivíduos tenham a capacidade de agir como desejarem, sem barreiras ou obstáculos, ou então em busca de autorrealização (CARTER, 2022). O alvo aqui é deslocado para a disponibilidade de ações em princípio possíveis para as agentes.

uma ação moralmente errada *equivale* à afirmação de que ele é culpado; e mencionei algumas atitudes reativas que provavelmente dirigiríamos a esse agente: indignação, raiva e, no caso do pai do menino, rancor e ressentimento. Observe, porém, que as atitudes associadas a uma valência moral negativa da ação não são as únicas que compõem esse conjunto.

Strawson faz menção à gratidão e ao “tipo de amor que podemos dizer que às vezes dois adultos sentem reciprocamente, um pelo outro” ([1962] 2008, p. 10). Mas deixemos o segundo de lado²². Sobre a gratidão, é razoável dizer que, quando nos sentimos gratos a uma pessoa, estamos responsabilizando-a: um filho grato ao pai pelo modo como foi por ele criado é um reconhecimento de que um grupo de ações de valência moral positiva (que expressam cuidado atencioso e respeitoso, carinho, cumplicidade etc.) “diz algo” acerca da agência do pai. Acerca de qualquer uma dessas ações, diremos então que o pai possui *mérito moral*²³; e a afirmação de que uma agente possui (ou é digna de) mérito moral *equivale* à afirmação de que ela é moralmente responsável por uma ação moralmente louvável, ou por uma boa ação para além do que era razoável esperar dessa agente no contexto.

Culpa e mérito moral são os tipos paradigmáticos de responsabilidade moral. Como já assinalai, porém, uma ação talvez possa ser moralmente relevante e exibir uma valência neutra. Aqui também haveria responsabilidade moral, como sugere a seguinte estrutura de exemplo proposta por McKenna (2012). Imagine que João está fazendo sua caminhada diária e se depara com uma pessoa precisando de ajuda. Ele pode ajudá-la; suponha que isso demandaria, no entanto, uma ação que vai muito além do que é razoável esperar dele. Por outro lado, João também poderia facilmente tirar vantagem dessa pessoa e conseguir um dinheiro extra. No fim das contas, ele decide apenas continuar sua caminhada, sem fazer nada além daquilo que devia: evitar tirar vantagem da pessoa precisando de ajuda. Nesse caso, a ação é moralmente significativa porque João “exercitou suas capacidades racionais ao agir dentro das fronteiras do que a moralidade exige” (p. 17). Não se trata, porém, de uma ação errada, nem de uma ação louvável. Mas isso não implica que não se trate de uma ação pela qual o agente é moralmente responsável.

²² Diferentemente do que afirma Strawson, o amor não parece uma atitude reativa similar à gratidão, ou à forma talvez paradigmática de gratidão. Pereboom (2007) faz o seguinte comentário: “As razões pelas quais amamos os outros são certamente variadas e complexas. (...) Mesmo se há algum aspecto do amor que concebemos como uma resposta merecida a qualidades moralmente admiráveis, é improvável que ele fosse ameaçado caso acreditássemos que essas qualidades não são produzidas ou sustentadas por decisões livres.” (p. 121)

²³ Optei por essa expressão para traduzir os termos *praise* e *praiseworthiness*. No caso do adjetivo *praiseworthy*, utilizo “louvável” para ações de valência moral positiva e “digno de mérito”, ou “digno de elogio”, para agentes moralmente responsáveis por essas ações.

Na história da discussão filosófica sobre responsabilidade moral, alguns autores propuseram que nossas práticas de responsabilização se justificam *exclusivamente* pelas consequências benéficas do engajamento nessas práticas: encorajar o comportamento moralmente desejável e desencorajar o comportamento moralmente condenável (TALBERT, 2022). Essa visão contrasta com a que estive supondo até agora. Segundo Moritz Schlick ([1930] 1966), por exemplo, a ideia de que a punição é uma retaliação natural da agente pela má ação realizada no passado não deveria mais ser defendida; o propósito de culpar/punir, ou de elogiar/recompensar, seria alterar a incidência, no futuro, de certos tipos de ações. Esse tipo de abordagem grosseiramente “consequencialista”²⁴ e prospectiva, porém, não apenas deixa espaço para que (em casos excepcionais) tratemos como culpadas pessoas inocentes, como também oferece uma descrição implausível das nossas atitudes reativas como meras ferramentas sociais de mudança de comportamento (STRAWSON, [1962] 2008; MCKENNA & PEREBOOM, 2016).

Diferentemente do que é sugerido por essa abordagem (e talvez por abordagens que justificam as práticas de responsabilização por meio de considerações sobre a existência de um contrato social, que legitima por exemplo a punição²⁵), parece que nossas atitudes reativas estão intimamente relacionadas à noção de *merecimento básico* (PEREBOOM, 2015). Isso significa que elas pressupõem que os agentes *merecem* que elas lhes sejam dirigidas; e esse merecimento é básico no sentido de não derivar de quaisquer considerações consequencialistas (ou contratualistas). Na forma básica do merecimento,

(...) alguém que tenha agido mal por más razões merece ser culpado e talvez punido *apenas* porque agiu mal por essas razões, e alguém que tenha realizado uma ação moralmente exemplar por boas razões merece crédito, elogio, e talvez recompensa *apenas* porque realizou essa ação por essas razões (...). (PEREBOOM, 2015, p. 281, grifo meu)

Cabe agora questionar se não haveria alguma variedade de responsabilidade moral que, ao contrário da descrição acima, não envolve que a agente mereça ser alvo de atitudes reativas, ou de punição/recompensa²⁶. Segundo Sofia Jeppsson (2016), há uma visão muito

²⁴ “Grosseiramente consequencialista”, nesse contexto, significa uma abordagem da responsabilidade moral que atribui culpa/mérito moral etc. exclusivamente em função dos resultados dessas atribuições. Não estou me referindo a uma teoria sobre aquilo a que me referi (na subseção anterior) como o primeiro nível de avaliação – o das ações. Estamos no nível de avaliação da agente.

²⁵ Considerações exclusivamente consequencialistas e contratualistas acerca do que é a responsabilidade moral (e do que justifica as práticas de responsabilização) têm sido recuperadas por autores contemporâneos com abordagens mais sofisticadas (cf., por exemplo, Vargas, 2013). Essa discussão, porém, não será contemplada nesse trabalho.

²⁶ Nesse trabalho, minha posição será simplesmente a de deixar em aberto a possibilidade do pluralismo. Para uma discussão sobre variedades de responsabilidade moral, ver Jeppsson, 2022.

difundida de que podemos discutir ao menos duas variedades de responsabilidade moral²⁷, sendo a primeira mais exigente em termos de liberdade/controle do que a segunda: (i) a responsabilidade moral associada a merecimento e (ii) a responsabilidade moral enquanto *atributabilidade*²⁸. O segundo tipo estaria interessado simplesmente em “se a agente agiu por boas razões e/ou correspondeu a certas demandas morais justificadas” (p. 682). As variedades (i) e (ii) podem ser consideradas tipos genuinamente diferentes de responsabilidade moral se autorizam tratamentos distintos da agente e/ou se oferecem justificações distintas desse tratamento (JEPPSSON, 2022). Embora esse seja um debate em aberto, é razoável dizer ao menos a princípio que essa condição é satisfeita.

Começando por (ii), uma agente é moralmente responsável no sentido de atributabilidade se faz sentido perguntá-la “por que você fez isso?” – isto é, demandar explicações – e se essa agente é alvo apropriado de crítica ou elogio moral (JEPPSSON, 2016). Parte do objetivo de fazer essa questão seria investigar se a agente possui, e reconhece, por exemplo uma disposição para agir mal naquele gênero de contexto; e se as razões oferecidas em resposta confirmam que há essa má disposição, então se torna apropriado exigir da agente um esforço para modificá-la ou eliminá-la (PEREBOOM, 2015). Se chegássemos à conclusão de que uma determinada pessoa não agiu livremente (em algum sentido relevante) ao realizar a ação moralmente errada *x*, e por isso não seria apropriado sentir rancor, ressentimento etc. em relação a ela, ainda assim poderíamos qualificá-la (em geral ou quanto àquela ação) como cruel, ou egoísta, ou insensível, e talvez sustentar sentimentos de tristeza e desapontamento – atitudes reativas mais “brandas” e menos duramente direcionadas à agente²⁹. Embora o merecimento básico não esteja envolvido na atributabilidade, as práticas

²⁷ A própria Jeppsson, porém, não endossa essa visão.

²⁸ O primeiro tipo corresponde ao que é chamado na literatura de *accountability*, enquanto o segundo tipo fará referência ao que alguns autores chamam de *answerability* (cf. Smith, 2012; Pereboom, 2015; e Jeppsson, 2016). Há ainda um terceiro termo, *attributability*, proposto no conhecido artigo de Gary Watson, *Two Faces of Responsibility* (1996). Embora a palavra “atributabilidade” seja uma tradução literal do terceiro termo, não do segundo, minha opção de tradução justifica-se pelo fato de ser muito duvidoso que haja referentes distintos para esses dois termos (cf. Jeppsson, 2022). Além disso, não há termos sinônimos de responsabilidade moral em língua portuguesa do modo como há em língua inglesa; e a terminologia em português para fazer referência às variedades de responsabilidade ainda não está fixada.

²⁹ Pereboom (2007, 2014) sustenta que o conjunto de atitudes reativas delineado por P. F. Strawson não é homogêneo: apenas algumas delas, como o ressentimento e a gratidão (ao menos em uma de suas formas), pressuporiam merecimento por parte da agente. Uma ideia similar, embora do âmbito de uma outra discussão, parece estar presente em Fricker (2007): a autora afirma que agentes que inadequadamente conferem menos credibilidade à fala de uma pessoa devido a preconceitos identitários (como o racismo ou o sexismo) podem não ser *culpados* devido a certos contextos históricos em que se encontravam, mas ainda assim são alvos apropriados do nosso *desapontamento moral* (uma forma mais branda de ressentimento moral). Nesta subseção, estou supondo a

de crítica e elogio moral estariam justificadas porque contribuem com três *desiderata* morais: “proteção de vítimas em potencial, reconciliação em relacionamentos pessoais e com a comunidade moral, e formação do caráter moral”³⁰ (PEREBOOM, 2015, p. 285).

Observe que a justificação de práticas de responsabilização no sentido de atributabilidade é amplamente prospectiva: proteção, reconciliação e formação moral *futuras*. Mas como é que a noção de responsabilidade aqui em jogo poderia descrever o papel mesmo de atitudes reativas mais brandas sem carregar as mesmas fragilidades da noção consequencialista mencionada há pouco? Jeppsson (2016) evidencia que, nesse contexto, o melhoramento moral não é concebido como mera alteração da frequência de certas ações, mas como a compreensão mais sofisticada de razões morais e a mudança de comportamento à luz dessas razões. Além disso, as considerações envolvidas na noção de atributabilidade não são *exclusivamente* prospectivas: é preciso que a agente, dotada (no mínimo) de competência normativa, de fato tenha realizado uma ação com a valência moral correspondente à atitude reativa e ao qualificador dos quais esteja sendo o alvo.

Neste trabalho, a variedade de responsabilidade moral em questão será (i) – o sentido associado a merecimento básico de responsabilização. Nessa acepção, *a agente merece responsabilização simplesmente pelo fato de ter agido livremente*. No início deste capítulo, apontei que considerávamos Jorge culpado por ter impedido o menino de chegar à cirurgia e pela consequência dessa ação (o menino perde um dos globos oculares) porque era razoável supor que ele estava ciente do que fazia e da consequência provável, tinha controle sobre o que fazia e não estava sujeito a coerção/manipulação etc. Tradicionalmente, filósofos(as) dividiram as condições necessárias e suficientes para a responsabilidade moral em uma *condição epistêmica* (aquilo de que a agente precisa estar ciente sobre a ação que realiza) e uma *condição metafísica* (a condição de controle, ou de liberdade) (FISCHER & RAVIZZA, 1998; RUDY-HILLER, 2019). Essa divisão em duas condições separadas, porém, parece não posicionar adequadamente uma outra expressão que frequentemente utilizamos ao nos referir a agentes responsáveis: “agir intencionalmente”. Na próxima subseção, apontarei que a condição epistêmica mais fundamental para a responsabilidade moral está inclusa no que é agir intencionalmente; mas que agir intencionalmente, com competência normativa, não é suficiente (embora necessário) para agir livremente.

possibilidade de subespécies de atitudes reativas – possibilidade que sustenta a hipótese do pluralismo sobre a responsabilidade moral.

³⁰ E, no caso de ações de valência moral positiva, “reconhecimento de sucessos e realizações em uma relação” (PEREBOOM, 2015, p. 286).

1.2.3 Ações intencionais e ações livres

Considere o exemplo a seguir³¹. Wilma sofre de uma condição chamada agorafobia, caracterizada pelo medo de se colocar em situações das quais não seja fácil escapar (como locais fechados, ou locais abertos desconhecidos e/ou com muitas pessoas) e cuja manifestação mais intensa leva o indivíduo a ficar fechado em casa – o que é precisamente o caso dessa agente. O filho de Wilma se casará dali a algumas semanas e, é claro, convidou a mãe para a cerimônia. Mas Wilma sente um medo tão forte de sair de casa que, se o casamento fosse no dia de hoje, ela certamente perderia a cerimônia. Ela então empreende várias tentativas de aliviar seu medo nas semanas que ainda lhe restam, como terapia cognitivo-comportamental e exposições controladas às situações que lhe causam ansiedade. Wilma faz todo o esforço que poderíamos esperar de uma mãe que compreende a importância que sua presença no casamento possui para seu filho. No dia da cerimônia, porém, ela é tomada pelo medo e decide ficar em casa assistindo televisão.

Observemos esse caso com atenção. Wilma é, é claro, causalmente responsável pela ação de ficar em casa assistindo televisão. Também claramente, a ação possui valência moral negativa – faltar ao casamento do próprio filho para assistir televisão é moralmente condenável. Por tudo o que sabemos sobre Wilma, ela é uma agente moral (isto é, possui competência normativa). Seu caso contrasta, pois, com a primeira classe de casos explorada na subseção 1.2.1. Em relação às versões que explorei na segunda classe de casos, a história de Wilma também possui algumas diferenças, embora pertença à mesma classe. O propósito daquelas versões era mostrar que, apesar de a competência normativa ser necessária para a liberdade e a responsabilidade moral, não é ainda suficiente. As considerações evocadas para atenuar ou suspender a atribuição de culpa eram que a ação envolvia coação, ou que o agente não tinha acesso às informações relevantes do contexto. Não é o caso que Wilma tenha sido coagida. Já quanto a estar ciente das informações relevantes do contexto, a diferença entre a história de Wilma e a versão da história de Jorge em que não lhe são contadas todas as informações relevantes é que a ação desse segundo agente *não era intencional*.

Estipulemos que ações intencionais são aquelas causadas, da maneira apropriada, por desejos ou intenções da agente. Ao menos aparentemente, intenções estão relacionadas a desejos e a crenças, como a de que o desejo provavelmente será realizado. Ao realizar uma ação intencional, a agente *está ciente do que faz, sob uma descrição apropriada*: não faz sentido dizer que um garçom intencionalmente serviu comida envenenada a uma cliente se ele

³¹ Adaptado de Mele, 1996. Agradeço a Beatriz Sorrentino Marques pela sugestão do exemplo para essa subseção.

não sabia que servia comida envenenada (mas simplesmente que servia comida). Já no caso de Wilma, podemos caracterizar a ação de ficar em casa assistindo televisão como intencional: se ela tenta ir ao casamento e é tomada pelo medo, é duvidoso que no dia da cerimônia ela de fato tenha a intenção de ir ao casamento; além disso, ela está ciente de todas as informações relevantes sobre sua decisão e age, é claro, motivada pelo desejo de ficar em casa.

Assim como os outros casos da segunda classe, porém, a agente não possui responsabilidade moral – ela não é culpada. Parece que nós ou atenuamos muito ou suspendemos por completo a responsabilidade de Wilma, mesmo no sentido de atributabilidade especificado na subseção anterior: aquilo que sabemos sobre o transtorno da agorafobia e sobre sua tentativa anterior nos impede de atribuir a essa agente qualquer disposição de indiferença aos sentimentos do filho, por exemplo. Wilma age “compulsivamente”, motivada por um impulso talvez incontrolável derivado de sua condição psicológica. Por essa razão, também não pensamos que Wilma *mereça* ser alvo de ressentimento, desapontamento ou reprovação. Nossos sentimentos perante essa agente serão, no máximo, os de pena e lamento pelo que ocorreu.

A análise do caso de Wilma nos dá razões para diferenciar ações intencionais de ações livres³². Considerando que essa agente não é culpada por faltar o casamento do filho para assistir televisão, podemos concluir que agir livremente é algo mais exigente do que agir intencionalmente (com competência normativa), embora a segunda seja condição necessária da primeira³³. Essa distinção é mais ou menos análoga àquela proposta por Harry Frankfurt (1971) entre liberdade de ação e liberdade da vontade/do arbítrio; em que a primeira significa a capacidade de converter os próprios desejos ou intenções em ações sem impedimento (e

³² Considerando o debate entre compatibilistas e incompatibilistas que apresentarei na seção a seguir e no capítulo 2, uma preocupação aqui é a de que a distinção entre ações intencionais e ações livres seja uma petição de princípio para o incompatibilista justamente pelo fato de que o compatibilista pode não ter muita facilidade em justificar essa distinção. O ponto desta subseção, porém, é que ao menos uma distinção conceitual é possível, mesmo que alguma caracterização de “ação intencional” seja demonstrada equivalente a “ação livre” (da maneira necessária para responsabilidade moral). Agradeço a Beatriz Sorrentino Marques por esse comentário.

³³ Frequentemente, julgamos que pessoas são culpadas por ações ou omissões *negligentes*, involuntariamente realizadas. Nesses casos, pode ser que a agente não estivesse ciente do que fazia sob a descrição apropriada, e por isso a ação não seria intencional. Uma boa hipótese aqui é, porém, a de que a culpa da agente – se de fato há culpa – não é diretamente pela ação/omissão negligente. Trata-se de culpa indireta, derivada de um momento anterior em que a agente intencionalmente deixa de prevenir as condições para a negligência posterior. Por essa razão, seria justificado dizer que agir intencionalmente é condição necessária para agir livremente e ser moralmente responsável pela ação. Para um tratamento de casos de omissões involuntárias que aponta nessa direção, ver Nelkin e Rickless, 2017.

então agir intencionalmente) e a segunda significa que os desejos que se tem a capacidade de converter em ações estão sob o controle do sujeito do modo necessário para a responsabilidade moral – o que associarei, ao longo deste trabalho, a agir livremente. Seguindo McKenna e Pereboom (2016), entenderei a liberdade como a “capacidade única de pessoas de exercer *o mais forte sentido de controle* sobre suas ações necessário para responsabilidade moral”³⁴ (p. 6, grifo meu). Mas resta dizer o que isso significa – essa é a tarefa difícil.

Para que uma agente seja moralmente responsável por uma ação, ela precisa estar ciente³⁵ de certas coisas a respeito das circunstâncias daquela ação. Essa observação aponta para a existência, como mencionei no fim da subseção anterior, de uma condição distintamente epistêmica da responsabilidade moral. Uma das questões cruciais em torno das quais o debate contemporâneo sobre a condição epistêmica se concentra é no *conteúdo* dessa condição, para o qual há ao menos três candidatos plausíveis: estar ciente da ação, da valência moral, e de alternativas (RUDY-HILLER, 2019)³⁶. Começando por estar ciente da ação, considere mais um exemplo, de John Fischer e Mark Ravizza (1998). Um homem tira o carro da garagem sem se dar conta de que um gatinho cochilava próximo à roda traseira. Ele não tinha razão para supor que ali haveria um gatinho e acaba causando a morte do animal. Diríamos então que esse agente não é culpado pelo atropelamento; e o que justifica a afirmação de que ele nem mesmo agiu intencionalmente (e, portanto, não agiu livremente) é que ele não estava ciente do que fazia sob a descrição apropriada – “atropelar o gatinho”. Também não é o caso que o homem *deveria* estar ciente. A avaliação desse exemplo “captura a intuição de que um agente é responsável apenas se ele conhece os fatos particulares em torno de sua ação e age com o tipo apropriado de crenças e intenções” (FISCHER & RAVIZZA, 1998, p. 13).

O segundo candidato, estar ciente da valência moral da ação, envolve ou que a agente tenha a crença de que a ação é errada/moralmente exemplar (ou neutra, apesar de moralmente relevante), ou que a agente esteja ciente das *características* que compõem a valência moral da ação³⁷. Já o terceiro, estar ciente de alternativas, envolve que a agente acredite que

³⁴ Isso é consistente com a existência de ações livres em contextos sem relevância moral, desde que elas satisfaçam as exigências tais que, *se* houvesse relevância moral, a agente seria moralmente responsável (MCKENNA & PEREBOOM, 2016).

³⁵ É um objeto de discussão à parte o que “estar ciente” significa: conhecimento, ou crença verdadeira, ou crença justificada (RUDY-HILLER, 2019). Não tratarei desse ponto nesse trabalho.

³⁶ Há ainda um quarto candidato mencionado no verbete: estar ciente de consequências. Porém, dado que essa exigência se aplica à responsabilidade moral por consequências (não diretamente por ações), não a incluí nessa lista.

³⁷ Cf. Zimmerman, 1997; Levy, 2011; e Talbert, 2013.

“alternativas [de ação] estavam disponíveis (...), e entenda o significado dessas alternativas” (LEVY, 2011, p. 111). A discussão da condição epistêmica da responsabilidade moral está fora do escopo deste trabalho. Parece-me, porém, que a satisfação da segunda e da terceira exigência depende da satisfação da primeira: a agente precisa estar ciente do que faz para estar ciente da valência moral e do significado de alternativas de ação que talvez estejam disponíveis. Além disso, tanto a segunda quanto a terceira exigência são bastante controversas³⁸; e pode ser que “estar ciente das características que compõem a valência moral da ação” signifique simplesmente a posse de certas crenças *factualis* – de modo que essa condição estaria incluída em “estar ciente da ação”. Parece então que “estar ciente da ação” é, se não a única, a condição epistêmica mais fundamental da responsabilidade moral – que, como apontei, está inclusa no que é “agir livremente”. Se essa é a única condição epistêmica, aliás, não há condição epistêmica para a responsabilidade moral que não seja também uma condição de controle/de liberdade³⁹. Feitas as distinções e considerações anteriores, passemos agora àquela que é a preocupação central desse trabalho: a condição de liberdade da responsabilidade moral.

1.3 O PROBLEMA DA COMPATIBILIDADE: LIBERDADE E DETERMINISMO

Considere um sentido “mundano” de liberdade. Aquela liberdade bastante trivial que sentimos que temos quando, no fim de um dia cansativo, escolhemos o que fazer do tempo restante antes de dormir. Ou a liberdade que sentimos que temos durante uma tomada de decisão importante, em que é preciso considerar perdas e ganhos de cada opção ao nosso dispor. Ou o controle sobre a própria ação de um amigo que mente apenas para evitar problemas conosco. Também mundanamente, às vezes refletimos sobre como os valores e critérios a partir dos quais deliberamos e decidimos são resultado de uma cultura específica, ou de costumes arbitrários que acabamos por aprender. Nas relações com outras pessoas, consideramos a todo tempo o porquê de alguém ter agido como agiu, ou as nossas próprias razões para agir. Nesses contextos, podemos desculpar alguém que nos tratou grosseiramente, com a descoberta de que essa pessoa sofrera uma grande injustiça logo antes de nos encontrar; ou podemos deixar de sentir gratidão pela boa ação de alguém se soubermos que ela foi motivada por segundas intenções. Se ampliamos o alcance e a profundidade dessas reflexões, descobrimos a quantidade de acasos e de circunstâncias fora de nosso controle que nos fazem

³⁸ Cf., por exemplo, Talbert, 2013; e Sartorio, 2017.

³⁹ Como sugerido em Mele, 2011.

agir do modo como agimos todos os dias. É nesse plano, o dos acontecimentos mais ou menos triviais de vidas humanas, que surge o problema da liberdade.

Em uma primeira aproximação, o termo “liberdade” será entendido aqui do seguinte modo: o controle que pessoas parecem (ao menos às vezes) possuir sobre suas ações e omissões e que as torna moralmente responsáveis por essas ações e omissões, além de por algumas consequências dessas coisas. A liberdade seria, então, a condição de *controle* da responsabilidade moral, ou sua condição metafísica – em contraste com as condições epistêmicas que explorei brevemente na seção anterior. Como ficará claro ao longo deste trabalho, há diferentes modos de pensar sobre a natureza da liberdade necessária para responsabilidade, e há “sérios desacordos sobre o que constituiria uma teoria adequada da liberdade” (FISCHER, KANE, PEREBOOM & VARGAS, 2007, p. 1). Mas o que motiva esses desacordos? Derk Pereboom (2022) aponta que uma fonte tradicional de preocupação em relação à liberdade humana é teológica: talvez tudo o que ocorre no mundo, incluindo nossas ações, seja resultado de uma vontade divina. Outra fonte é não-teológica e diz respeito a uma tese que poderia ser verdadeira sobre as leis da natureza⁴⁰: o *determinismo causal*. Nas subseções a seguir, explicarei o que é o determinismo causal, as razões pelas quais ele parece uma ameaça à existência de liberdade e o modo como o debate em torno dessa tese leva a desacordos sobre *qual é a natureza da liberdade necessária para responsabilidade* – aquilo que estou chamando, genericamente, de *problema da liberdade*.

1.3.1 A tese do determinismo causal

No “The Free Will Lexicon”⁴¹, o determinismo é apresentado da seguinte forma:

Determinismo. A tese de que uma descrição completa das leis da natureza, junto a uma descrição completa do estado do universo inteiro em qualquer ponto do tempo, implica logicamente uma descrição completa do universo inteiro em qualquer outro ponto do tempo. O determinismo (que é às vezes referido como “determinismo causal”) é uma tese mais forte do que a tese de que todo evento tem uma causa ou um conjunto de causas. Determinismo é uma tese sobre a natureza das leis causais mais fundamentais (ou “leis da natureza”) – a saber, que as leis descrevem padrões dinâmicos que sempre implicam *outputs* únicos para cada *input* completo. Isso contrasta com visões segundo as quais as leis fundamentais são meramente estatísticas ou probabilísticas: essas visões sustentam que para cada *input* completo, há uma gama de possíveis *outputs* ponderada por probabilidades objetivas ou

⁴⁰ Nesse trabalho, não oferecerei nenhuma definição de “leis da natureza”. Adotarei um entendimento mais ou menos intuitivo do que elas sejam – uma expressão de regularidades factuais que “aparentemente envolvem necessidades, limites, restrições, ou condições sob as quais o mundo natural se desdobra” (MCKENNA & PEREBOOM, 2016).

⁴¹ O *The Free Will Lexicon* é um apêndice do livro *Surrounding Free Will* (2015) que oferece definições introdutórias de termos-chave da pesquisa contemporânea sobre liberdade.

tendências. Se o determinismo é falso, uma ação pode ser causada, mas não será deterministicamente causada. (MELE, 2015, p. 322)

Supondo descrições completas do estado do universo em um momento *x* e das leis da natureza que operam nesse universo, e dada a tese do determinismo, seria possível (por uma inferência dedutiva) extrair uma descrição completa do estado do universo em um momento *y*. Aqui é necessário distinguir cuidadosamente o registro *metafísico* do registro *epistêmico*. O determinismo é uma tese sobre como as leis da natureza *são* – isto é, pertence ao registro metafísico da discussão sobre essas leis.⁴² A referência a descrições completas na definição acima não implica que seres humanos precisem conhecer essas descrições, nem que sejam capazes de articular uma teoria completa das leis da natureza⁴³. Basta que aceitemos que, em princípio, tudo o que está ocorrendo no universo em um ponto específico do tempo e as leis que expressam as *regularidades* dos processos da natureza são passíveis de serem descritos (ainda que com descrições inimaginavelmente longas e complexas). Se, supondo essas descrições, o resultado for uma inferência dedutiva, então as leis da natureza são deterministas.

Uma outra maneira de formular a tese do determinismo é a seguinte: as regularidades da natureza são de um modo tal que, a qualquer momento, apenas um futuro é fisicamente possível⁴⁴. Há um sentido em que existem muitos futuros possíveis para o momento presente. Trata-se do sentido de *possibilidade lógica*: é ao menos concebível para mim que aconteça um eclipse total do Sol nessa manhã, enquanto escrevo. Posso dizer isso sem contradizer a mim

⁴² Alguns autores, como Beebe e Mele (2002), afirmam que o determinismo não deve ser entendido metafisicamente. Essa é, porém, uma posição minoritária.

⁴³ Previsibilidade e determinismo estão misturados na famosa citação de Laplace (2009): “Nós podemos considerar o estado presente do universo como um efeito de seu passado e uma causa de seu futuro. Um intelecto que em certo momento pudesse estar ciente de todas as forças que colocam a natureza em movimento, e de todas as posições de todos os itens dos quais a natureza é composta, se esse intelecto fosse também vasto o suficiente para submeter todos esses dados a análise, comportaria em uma única fórmula os movimentos dos maiores corpos do universo, assim como os do menor átomo; para tal intelecto, nada seria incerto e o futuro, assim como o passado, estaria presente perante seus olhos.”

⁴⁴ A tese do determinismo, no entanto, difere da tese do *fatalismo*: “Fatalismo é a tese de que todos os eventos (ou em algumas versões, pelo menos alguns eventos) estão destinados a ocorrer, não importa o que façamos. A fonte da garantia de que esses eventos acontecerão está localizada na vontade dos deuses, ou em sua presciência divina, ou em algum aspecto teleológico intrínseco do universo, e não no desenrolar dos eventos sob a influência de leis naturais ou de causa e efeito. relações. O fatalismo é, portanto, claramente separável do determinismo, pelo menos na medida em que se pode separar as forças místicas e as vontades e presciências dos deuses (sobre assuntos específicos) da noção de lei natural/causal. (...) de modo geral, podemos imaginar que certas coisas estão fadadas a acontecer, sem que isso seja resultado apenas de leis naturais determinísticas; e podemos imaginar o mundo sendo regido por leis deterministas, sem que nada esteja fadado a ocorrer (talvez porque não existam deuses, nem forças místicas/teleológicas merecedoras do título *destino*, e em particular nenhuma determinação intencional das “condições iniciais” do mundo).” (HOEFER, 2022)

mesma, e um ouvinte pode imaginar como o eclipse seria. Mas o sentido relevante de “possível” na tese do determinismo é *físico*, não lógico. Há um sentido óbvio em que esse futuro eclipse é fisicamente impossível, não porque a ocorrência de eclipses solares contrarie as leis da natureza, mas porque seria preciso que o passado fosse diferente do que é para que houvesse um eclipse essa manhã. Feita essa distinção entre possibilidades lógicas e possibilidades físicas, podemos compreender que o determinismo é a tese de que as regularidades da natureza são bastante restritivas. Isso é o que a caracterização em termos de *inputs* e *outputs* em uma função pretende capturar – um conjunto qualquer de dados de entrada tem exatamente *um* conjunto de dados de saída. Considerando que nossas ações envolvem movimentos corporais e processos neurais correspondentes, que são partes do mundo físico, a consequência da verdade do determinismo não seria trivial: nossas ações, a qualquer momento, seriam as únicas ações possíveis.

Observe que a tese do determinismo é diferente da tese da causação universal – isto é, a tese de que todo evento tem uma causa⁴⁵. Um mundo em que apenas um par de eventos tivesse uma relação não-determinista de causalidade – ou seja, um mundo em que as leis da natureza são tais que, dado o evento que é a causa nesse par, havia mais de um evento possível como seu efeito – já poderia ser chamado de *indeterminista*⁴⁶ e permaneceria compatível com a existência de causação universal. Isso é importante porque pode ser o caso, em um mundo indeterminista, que parte dos eventos cuja relação causal não é determinista diga respeito a decisões e ações humanas. Ao menos em princípio, não parece que essas coisas sejam não-causadas, mesmo que suas causas apenas restrinjam uma gama de efeitos prováveis em vez de determiná-las. Como veremos ainda nesta seção, a exigência de causalidade indeterminista em decisões e ações humanas é um ingrediente fundamental do entendimento da liberdade de uma parcela dos incompatibilistas (os libertistas⁴⁷).

Foi dito anteriormente que o determinismo é uma tese que *poderia ser verdadeira* sobre as leis da natureza – embora a maioria dos físicos e filósofos pensem atualmente que ela é falsa (FISCHER, KANE, PEREBOOM & VARGAS, 2007). O que justifica a afirmação dessa possibilidade? Em *Four Views on Free Will* (2007), os autores apontam algumas razões. A primeira é que é notoriamente difícil prever como a ciência se desenvolverá, e pode ser que futuros físicos formulem bons argumentos para uma interpretação determinista das leis da

⁴⁵ Para uma argumentação alternativa por essa mesma conclusão, cf. van Inwagen, 1983.

⁴⁶ “Indeterminista” é simplesmente a negação de “determinista”.

⁴⁷ Optei por traduzir os termos *libertarian* e *libertarianism* por “libertista” e “libertismo” para distinguir essa posição dentro do Problema da Compatibilidade das posições libertárias em filosofia política, já que elas nada têm em comum além do nome.

mecânica quântica⁴⁸. A segunda razão é que, mesmo se o universo não for inteiramente determinista, pode ser o caso que o determinismo seja verdade localmente, “ou como uma questão de como o espaço-tempo local é construído, ou como uma questão de como a Física para objetos não-quânticos opera” (p. 2). Para a maioria dos filósofos que trabalham com liberdade e responsabilidade moral, a verdade do determinismo deve ser avaliada empiricamente pelos físicos (CARROL & MARKOSIAN, 2010). A mera possibilidade (não-trivial) de que essa tese seja verdadeira, porém, já é encarada como suficiente para tornar relevante a pesquisa contemporânea sobre liberdade a partir do *Problema da Compatibilidade*. Explicarei esse problema a seguir.

1.3.2 Duas ameaças à liberdade

A possibilidade de que o determinismo seja o caso oferece dois tipos de desafios à existência da liberdade humana necessária para responsabilidade moral (MCKENNA & PEREBOOM, 2016). É importante distingui-los, já que cada desafio opera a partir de um aspecto diferente do que parece ser a nossa noção mais recorrente da liberdade, no sentido “mundano” que apresentei. Para introduzir o primeiro desses desafios, que diz respeito à *liberdade de agir diferentemente*, considere a definição do “The Free Will Lexicon” para “escolha” (que tratarei como sinônimo de “decisão”):

Escolha. Um ato de selecionar ou estabelecer um curso de ação. Escolher é diferente de qualquer deliberação que precede a escolha. Agentes escolhem entre opções, ou aparentes opções. Muitas escolhas são respostas à incerteza quanto ao que fazer. Ações intencionais rotineiras (como destrancar a porta quando você chega em casa) tipicamente não são precedidas da escolha de realizá-las em parte porque os agentes que as realizam não estão em dúvida sobre o que fazer. Mas talvez possamos dizer que agentes fazem escolhas em algumas situações nas quais estão certos do que fazer. Por exemplo, se oferecemos a uma pessoa faminta ou sua refeição favorita ou uma banana para o jantar, podemos dizer que essa pessoa faz uma escolha pela primeira opção. (MELE, 2015, p. 320-321)

Imagine que, no segundo turno das eleições presidenciais brasileiras, um eleitor vota no candidato B em vez de votar no candidato A. Esse eleitor faz uma escolha (terrível) entre duas opções disponíveis, dado que o candidato B é um ex-presidente que deliberadamente se omitiu em providenciar vacinas para a população no momento adequado durante uma pandemia global – informação da qual o eleitor estava, ou deveria estar, ciente. Suponha que ele decidiu seu candidato pouco tempo antes do dia da votação, e agiu conforme essa decisão

⁴⁸ Segundo Balaguer (2009), não há ainda bons argumentos nem por uma interpretação indeterminista, nem por uma interpretação determinista das leis atualmente probabilísticas da mecânica quântica (cf. p. 6-7).

sem nenhum tipo de manipulação ou coerção que lhe privasse de controle. Uma maneira muito imediata de pensar sobre a liberdade desse agente em relação ao voto em B é a seguinte: ele é moralmente responsável pelo seu mau voto (ou ele é culpado pelo voto) devido ao fato de que, em um momento relevante anterior a essa ação, havia uma alternativa que poderia ter sido selecionada; esse agente era capaz, portanto, de agir diferentemente. Considere então a seguinte caracterização:

Liberdade de agir diferentemente. “é a capacidade de uma agente, em um dado momento, de agir ou de abster-se: isto é, se uma agente age livremente, então ela poderia, naquele momento, ter se abster de agir como agiu.” (PEREBOOM, 2022, p. 1)

Para que alguém possua a liberdade de agir diferentemente, nem sempre é necessário que a alternativa disponível à ação efetiva da agente seja uma outra ação: a opção de simplesmente se abster já seria suficiente para a existência desse tipo de liberdade. Essa noção dá origem a um princípio bastante intuitivo a respeito da conexão entre liberdade e responsabilidade moral, o *Princípio das Possibilidades Alternativas* (PPA):

PPA. “uma pessoa é moralmente responsável pelo que faz apenas se podia ter agido diferentemente.” (FRANKFURT, 1969)

Dispondo desses recursos conceituais, é fácil visualizar como o determinismo oferece uma ameaça à liberdade caso seja verdadeiro: ele faz com que o eleitor nunca tenha tido uma alternativa genuinamente disponível de ação, já que seu voto em B era a única ação fisicamente possível dados o passado e as leis da natureza. A tensão gerada é a seguinte:

1. O determinismo é verdadeiro. [Suposição]
 2. Ao menos algumas pessoas [aparentemente] são moralmente responsáveis pelo que fazem.
 3. Responsabilidade moral exige liberdade.
 4. Liberdade exige a capacidade de agir diferentemente.
 5. A capacidade de agir diferentemente é incompatível com o determinismo. (MCKENNA & PEREBOOM, 2016, p. 41)
- [Logo, ninguém é moralmente responsável pelo que faz.]

Por que há aqui uma tensão? Dois pontos são centrais para o entendimento adequado do problema. O primeiro é que, se o determinismo é verdadeiro, esse já é o caso independentemente de nossas crenças a respeito dessa tese. Embora essa seja uma observação trivial, é importante ter em mente que a descoberta da verdade do determinismo seria uma

mudança meramente epistêmica, não metafísica, a respeito dos eventos no mundo e das nossas ações como parte desses eventos. Já o segundo ponto está expresso em 2. Em comunidades humanas, compartilhamos a crença (e conjuntos complexos de práticas relacionadas) de que pessoas ordinariamente são responsáveis pelo que fazem. Aplicamos punições, expressamos atitudes reativas e julgamos que ações nos dizem algo a respeito de seus autores porque estamos comprometidos com a ideia de que agentes merecem esses juízos e reações. Se a liberdade que fundamenta o merecimento básico de responsabilização é incompatível com o determinismo (supondo que ele seja uma tese verdadeira), como preservar a crença de que ao menos algumas pessoas (ou a maior parte delas) são moralmente responsáveis por suas ações?

Considere agora a ação de votar no candidato B sob uma outra perspectiva. Se concordarmos que o eleitor é moralmente responsável pelo modo como agiu, parece que há um sentido forte em que consideramos que esse agente é a “fonte”, ou a origem, da própria ação. Não se trata meramente de dizer que ele causou o registro do voto em B na urna eletrônica, por meio de certos movimentos corporais de digitação de um número e do aperto de um botão de confirmação do voto. Pensamos que o agente contribui substantivamente com a ocorrência da ação, iniciando-a de um modo direcionado que nos diz algo sobre a própria agência do indivíduo devido à sequência de elementos (crenças, desejos, deliberação, escolha, movimentos corporais etc.) *efetivamente conduzindo a essa ação*. Farei referência a essa segunda noção de liberdade, que não diz respeito à posse de escolhas alternativas, como a *liberdade da fonte genuína*:

Liberdade da fonte genuína. Uma agente é fonte genuína de sua ação apenas “se ela voluntariamente contribui com alguma condição C substantiva e necessária para as condições que de fato conduzem à sua ação, e não há condições suficientes para C que ocorram independentemente de sua agência.” (MCKENNA & PEREBOOM, 2016, p. 150)

A verdade do determinismo ofereceria uma ameaça a essa noção de liberdade da seguinte maneira: se as descrições de todas as ações humanas podem ser dedutivamente inferidas de descrições sobre o passado e as leis da natureza, então parece não ser o caso que o *eleitor* contribua substantivamente com uma condição C necessária à ocorrência de seu voto em B, já que as leis deterministas seriam tais que o *input* de um momento qualquer no

passado ofereceria condições suficientes para o futuro voto desse eleitor no candidato B. A tensão gerada dessa vez é a seguinte:

1. O determinismo é verdadeiro. [Suposição]
2. Ao menos algumas pessoas [aparentemente] são moralmente responsáveis pelo que fazem.
3. Responsabilidade moral exige liberdade.
4. Liberdade exige a capacidade da agente de ser a fonte genuína⁴⁹ de suas ações.
5. A capacidade de ser a fonte genuína de ações é incompatível com o determinismo. (MCKENNA & PEREBOOM, 2016, p. 42).
[Logo, ninguém é moralmente responsável pelo que faz.]

O *Problema da Compatibilidade* pode ser sintetizado na seguinte pergunta: *a liberdade necessária para a responsabilidade moral é compatível com o determinismo?* Esse problema soma-se às questões que ocupam a atenção de filósofos contemporâneos. Apesar de se tratar de um debate filosófico clássico⁵⁰, o Problema da Compatibilidade ganhou os contornos contemporâneos que acabo de apresentar a partir da década de 1960 (COATES & MCKENNA, 2018)⁵¹. Dados os dois desafios que o determinismo parece colocar, tanto para a liberdade de agir diferentemente quanto para a liberdade de ser a fonte genuína das próprias ações, a resposta mais imediata à pergunta pela compatibilidade é “Não”. Essa resposta constitui a posição *incompatibilista*:

Incompatibilismo. “nossa posse de liberdade não é compatível com o determinismo causal, com todas as nossas ações sendo causalmente determinadas por fatores fora de nosso controle.” (PEREBOOM, 2022, p. 2)

Segundo parte dos incompatibilistas, os *libertistas*, a natureza da liberdade humana envolve a presença de indeterminismo em ao menos algum processo causal levando a uma decisão e ação. Sem essa indeterminação, não seria adequado, ou justo, responsabilizar o eleitor por seu voto condenável no candidato B. Mas é crucial observar que os dois desafios expostos não constituem uma defesa do incompatibilismo. Há ainda a posição compatibilista – que é, inclusive, a posição defendida pela maioria dos filósofos trabalhando nessa discussão atualmente⁵²:

⁴⁹ N. T.: no original, *initiating source*.

⁵⁰ Cf. Hobbes, 1839; Hume, 2009; Kant, 2016; Locke, 1999.

⁵¹ Os principais responsáveis por esses novos contornos do problema foram P. F. Strawson (1962), H. Frankfurt (1969) e o argumento da consequência (cf. Ginet, 1966; e van Inwagen, 1983). Tratarei dessas contribuições ao longo do trabalho.

⁵² Segundo a “2020 PhilPapers Survey” (BOURGET & CHALMERS, 2023), cerca de 60% dos filósofos profissionais contemporâneos de departamentos que publicam majoritariamente em língua

Compatibilismo. “nossa posse de liberdade é compatível com o determinismo causal, com todas as nossas ações sendo causalmente determinadas por fatores fora de nosso controle.” (PEREBOOM, 2022, p. 2)

Mas o que justifica essa posição? Uma estratégia compatibilista amplamente empregada é argumentar que a natureza da liberdade necessária à responsabilidade é de um outro tipo: bastaria a capacidade da agente de agir “racionalmente, isto é, de um modo que é responsivo a razões [para agir]” (PEREBOOM, 2022, p. 2). Considerando que seres humanos, em parte significativa do tempo, são capazes de detectar, compreender e agir segundo razões (por exemplo, abster-se de realizar uma ação x em virtude de essa ação parecer moralmente errada, ou agir do modo y porque y desempenharia um papel instrumental em relação a um objetivo que a agente deseja atingir, ou ainda agir do modo z porque parece a coisa mais prudente a se fazer), pode ser que essa propriedade seja suficiente para indicar o que significa, para uma agente, ser a fonte genuína da própria ação, do modo apropriado para fundamentar responsabilidade moral. Apresentarei essa estratégia no capítulo seguinte. Um outro ponto importante é que, como veremos no próximo capítulo, é um desafio por si só explicar como o indeterminismo contribui com a liberdade de uma agente.

Nesta subseção, apresentei o Problema da Compatibilidade tal como é geralmente compreendido na filosofia contemporânea. Trata-se de um debate ainda em aberto (apesar da mencionada vantagem numérica para o compatibilista). Ao longo da exposição, ofereci caracterizações de conceitos que desempenham papel central nessa discussão: determinismo, escolha, liberdade de agir diferentemente, princípio das possibilidades alternativas, liberdade da fonte genuína, incompatibilismo e compatibilismo. A questão controversa aqui, que motiva o desacordo entre compatibilistas e incompatibilistas (assim como os desacordos entre filósofos no interior de cada uma dessas posições), é como entender a natureza daquela liberdade em sentido bastante mundano, em relação a ações e omissões particulares (e algumas de suas consequências). Em outras palavras, argumentos substanciais para o compatibilismo ou para o incompatibilismo dependem de um entendimento da condição de controle da responsabilidade moral que não está dado: ele é o alvo principal da investigação. Para finalizar esta seção, tratarei a seguir de uma das contribuições mais importantes já feitas

inglesa aceitam ou tendem ao compatibilismo. A menção a essa pesquisa, no entanto, consta aqui como mera curiosidade. Não é minha intenção defender que haja tendência de convergência à verdade em filosofia.

para a investigação do Problema da Compatibilidade, o Argumento da Consequência, e de algumas estratégias compatibilistas que pretendem objetá-lo.

1.3.3 Argumento da Consequência: um poderoso argumento incompatibilista

O Argumento da Consequência é uma poderosa ferramenta em favor do incompatibilismo⁵³. Ele foi proposto por Carl Ginet (1966), desenvolvido sobretudo por Peter van Inwagen (1983) e amplamente debatido nas décadas subsequentes como “uma das três mais significativas influências nos debates contemporâneos sobre liberdade” (MCKENNA & PEREBOOM, 2016, p. 72). As outras duas contribuições são a posição compatibilista de P. F. Strawson e os Casos Estilo-Frankfurt, que serão melhor exploradas posteriormente nesse trabalho. A seguinte apresentação informal do Argumento da Consequência é frequentemente citada na literatura:

Se o determinismo é verdadeiro, então nossas ações são as consequências das leis da natureza e dos eventos no passado remoto. Mas não depende de nós o que ocorreu antes de nosso nascimento, nem depende de nós quais são as leis da natureza. Portanto, as consequências dessas coisas (incluindo nossas ações presentes) não dependem de nós. (VAN INWAGEN, 1983, p. 56)

O Argumento da Consequência diz respeito à questão da compatibilidade entre o determinismo e a liberdade de agir diferentemente e emprega algumas ideias bastante intuitivas. Primeiramente, nossas crenças a respeito dos eventos do passado e dos eventos do futuro são assimétricas: pensamos que, enquanto o passado está “fixado” (e, portanto, não teríamos qualquer escolha sobre ele), o futuro está em aberto. Do mesmo modo, pensamos que as leis da natureza estão fixadas e “impõem” (ou “governam”) os acontecimentos; assim, elas não estariam sob nosso poder. Uma forma mais precisa de colocar essas ideias é a seguinte: não há nada que sejamos capazes de fazer que afetaria causalmente o passado e as leis da natureza. Se o determinismo é verdadeiro, parece então que não há nada que sejamos capazes de fazer que afetaria causalmente as ações que efetivamente realizamos, já que elas compõem as únicas consequências possíveis do passado e das leis.

Em *An Essay on Free Will* (1983), van Inwagen apresenta três formulações para o Argumento da Consequência. Consideremos a terceira formulação, a de um argumento modal (p. 93-104)⁵⁴. Esse argumento será formalizado em uma lógica que contém o operador **N**, em

⁵³ Outra ferramenta é o Argumento Direto (cf. van Inwagen, 1983), que estabelece que a responsabilidade moral é incompatível com o determinismo. Esse argumento não será explorado ao longo do trabalho, com exceção de uma rápida menção no capítulo 3.

⁵⁴ Esse argumento fará uso de dois operadores modais, isto é, o tipo de operador que “vincula-se a sentenças que têm (...) valor de verdade, para formar sentenças que têm valor de verdade; e o valor de

que “ Np ” especifica a forma lógica de “ p e ninguém tem, ou já teve, qualquer escolha a respeito de se p ”. Nessa lógica, as duas seguintes regras de inferência serão adotadas:

Alpha: De $\Box p$, podemos inferir que Np .

Beta: De Np e $N(p \supset q)$, podemos inferir que Nq .

Estipularemos que: “ L ” abrevia uma sentença que é a conjunção de todas as leis da natureza; “ M ” é uma sentença sobre o estado total de coisas no mundo em algum instante do passado remoto anterior ao nascimento de quaisquer agentes humanos; “ \Box ” é um operador para “é logicamente necessário que”⁵⁵; “ \supset ” é o operador para a condicional material (se p , então q); e “ P ” é uma letra a ser substituída por qualquer sentença verdadeira. A regra Alpha expressa que, se p é logicamente necessário, então ninguém tem, ou já teve, escolha sobre se p . Já a regra Beta expressa que do fato de que ninguém tem, ou já teve, escolha sobre se p , e que ninguém tem, ou já teve, escolha sobre se p implica materialmente q , podemos inferir que ninguém tem, ou já teve, escolha sobre se q (VIHVELIN, 2022). Temos então o seguinte argumento:

1. $\Box((M \& L) \supset P)$ definição de determinismo
2. $\Box(M \supset (L \supset P))$ de 1 (exportação⁵⁶)
3. $N(M \supset (L \supset P))$ de 2, pela regra Alpha
4. NM premissa de fixidez do passado
5. $N(L \supset P)$ de 3 e 4, pela regra Beta
6. NL premissa de fixidez das leis da natureza
7. NP de 5 e 6, pela regra Beta

Enquanto as premissas 1 e 2 seguem-se do determinismo (por definição, dado o determinismo, é logicamente necessário que a conjunção entre uma sentença sobre o estado total de coisas no mundo no passado remoto e uma sentença contendo todas as leis da natureza implique materialmente uma sentença verdadeira qualquer), a premissa 3 se segue da regra Alpha, que ao menos em princípio parece incontroversa (VIHVELIN, 2022). Também parecem incontroversas as premissas 4 e 6 – voltarei a esse ponto mais adiante. Por fim, a regra Beta é bastante plausível: ela expressa a forma lógica de uma inferência como i. ninguém tem ou já teve escolha sobre se o Sol explode no ano 2000; ii. ninguém tem ou já

verdade de uma sentença formada desse modo não é em todo caso uma função do valor de verdade da sentença à qual o operador está vinculado” (VAN INWAGEN, 1983, p. 93).

⁵⁵ Uma sentença é logicamente necessária quando sabemos, com base apenas em suas condições de verdade, que ela é verdadeira e que não poderia ser falsa (MURCHO, 2019).

⁵⁶ Exportação é uma regra derivada para deduções que permite a passagem de “ $(A \& B) \supset C$ ” para “ $A \supset (B \supset C)$ ”.

teve escolha sobre se, se o Sol explode no ano 2000, então a vida na Terra acaba no ano 2000; iii. logo, ninguém tem ou já teve escolha sobre se a vida na Terra acaba no ano 2000 (VAN INWAGEN, 1983). Suponha que P é uma sentença sobre a minha ação de tomar um gole de água nesse exato momento. Se o Argumento da Consequência é cogente (isto é, válido, com premissas verdadeiras e mais plausíveis do que a conclusão etc.), e se o determinismo é verdadeiro, segue-se que eu tomo um gole de água nesse exato momento e ninguém tem, ou já teve, qualquer escolha a respeito de se eu tomo um gole de água nesse exato momento. Eu não tinha, portanto, a liberdade de agir diferentemente do que agi – eu não era livre, por exemplo, para continuar digitando esse parágrafo em vez de interromper a digitação para tomar um gole de água.

O ponto crucial dessa formulação do Argumento da Consequência é a regra Beta. Como observam McKenna e Pereboom (2016), essa regra desempenha o papel de ser um princípio teoricamente neutro em relação à disputa entre compatibilistas e incompatibilistas e que, no entanto, ajuda a decidir essa disputa, autorizando a transição (supondo a verdade do determinismo) entre nossa falta de escolha a respeito dos eventos do passado e das leis da natureza e nossa falta de escolha a respeito de nossas próprias ações. Assim, por meio de um princípio com o qual o compatibilista poderia concordar, o incompatibilista obtém uma fonte de justificação para sua conclusão. Mas, “se fosse o caso que a regra [Beta] (...) não fosse afinal um padrão válido para raciocínios de maneira muito geral (...), isso traria à tona a preocupação de que uma aplicação da regra Beta em si mesma quando usada no Argumento da Consequência não seja inocente de viés teórico” (MCKENNA & PEREBOOM, 2016, p. 88). De fato, o próprio van Inwagen concedeu (2000) que a regra Beta foi descoberta inválida por Thomas McKay e David Johnson (1996). McKay e Johnson demonstraram que a regra Beta está comprometida com a regra da aglomeração para o operador N e, sendo a aglomeração inválida (como veremos a seguir), Beta seria também inválida.

Inicialmente, considere a sentença “ $(P \supset (Q \supset (P \& Q)))$ ”. Essa sentença é tal que, independentemente dos valores de verdade de P e de Q, será sempre verdadeira. Quando uma sentença complexa “na linguagem da lógica de primeira ordem é verdadeira em qualquer interpretação dos valores de verdade de seus constituintes, ela é uma verdade necessária” (MCKENNA & PEREBOOM, 2016, p. 89). Podemos então inferir que “ $\Box(P \supset (Q \supset (P \& Q)))$ ”. Suponha agora a verdade de duas sentenças P e Q sobre estados de coisas a respeito dos quais ninguém tem ou já teve escolha. Temos então o seguinte (MCKAY, JOHNSON, 1996, p. 115):

1. NP premissa

2. $\mathbf{N}Q$ premissa
3. $\Box(P \supset (Q \supset (P \& Q)))$ necessidade de uma verdade lógica
4. $\mathbf{N}(P \supset (Q \supset (P \& Q)))$ de 3, pela regra Alpha
5. $\mathbf{N}(Q \supset (P \& Q))$ de 1 e 4, pela regra Beta
6. $\mathbf{N}(P \& Q)$ de 2 e 5, pela regra Beta

Desse modo, inferimos “ $\mathbf{N}(P \& Q)$ ” de “ \mathbf{NP} ” e “ \mathbf{NQ} ” – aglomeração – por meio da regra Beta. McKay e Johnson, porém, oferecem um contraexemplo à aglomeração. Imagine que Clarice tem uma moeda no bolso, que poderia ser lançada e então teríamos, é claro, cara ou coroa. Suponha agora que “ P ” abrevia “a moeda não vira cara”, “ Q ” abrevia “a moeda não vira coroa” e “ $P \& Q$ ” abrevia “a moeda não vira cara e a moeda não vira coroa”. Essas três sentenças são verdadeiras. Considerando que Clarice não tem nenhum tipo de habilidade especial para controlar o lado em que viram moedas lançadas no ar, as sentenças “ \mathbf{NP} ” e “ \mathbf{NQ} ” também são verdadeiras. Mas “ $\mathbf{N}(P \& Q)$ ” é falsa: se Clarice tivesse decidido tirar a moeda do bolso e lançá-la no ar, ela teria tido, sim, escolha sobre se a moeda não virou cara ou a moeda não virou coroa.

Independentemente de a objeção de McKay e Johnson ser bem-sucedida⁵⁷, van Inwagen propôs que o Argumento da Consequência fosse reformulado substituindo o operador \mathbf{N} por \mathbf{N}^* , em que “ \mathbf{N}^*p ” abrevia “ p e ninguém pode, ou já pôde, fazer algo tal que, se fosse feito, p seria falso” – para esse operador, aglomeração seria válida (VIHVELIN, 2022). Além disso, vários incompatibilistas sugeriram princípios alternativos à regra Beta e que poderiam desempenhar o mesmo papel (MCKENNA & PEREBOOM, 2016). Suponhamos então que existe uma formulação modal válida do Argumento da Consequência. Mas resta ainda a seguinte questão: o argumento é sólido (isto é, suas premissas são verdadeiras)? Também podemos nos perguntar se, sendo sólido, o argumento desempenha o papel decisivo que pretende desempenhar em favor do incompatibilismo no Problema da Compatibilidade. Para finalizar esta seção, explorarei brevemente as principais objeções quanto à solidez do Argumento da Consequência, que dizem respeito às premissas 4 e 6 – a fixidez do passado e das leis da natureza. Já a questão sobre o papel desse argumento na investigação do Problema da Compatibilidade será explorada no capítulo seguinte.

⁵⁷ Para uma discussão em português das principais objeções a Beta, cf. Merluzzi, 2013.

1.3.4 Algumas estratégias para o compatibilista

Em “Are We Free to Break the Laws?” (1981), David Lewis distingue uma tese forte e uma tese fraca sobre agentes terem escolha quanto às leis da natureza, supondo a verdade do determinismo. Segundo Lewis, enquanto a tese forte – “Eu sou capaz de quebrar uma lei da natureza” – é totalmente implausível, a tese fraca – “Eu sou capaz de fazer alguma coisa tal que, se fosse feita, uma lei da natureza seria quebrada” – pode ser um comprometimento do compatibilista (p. 115). McKenna e Pereboom (2016) oferecem a seguinte formulação para as duas noções de “capacidade” distinguidas por Lewis:

Chame-as de capacidade causal (CC) e capacidade ampla (CA).

A [noção de] capacidade causal estabelece que:

CC: Uma pessoa tem a capacidade de provocar a ocorrência de algo, p, apenas no caso em que há um curso de ação tal que essa pessoa é capaz de realizar e, se ela fosse realizá-la, então ela causaria a ocorrência de p.

A [noção de] capacidade ampla estabelece que:

CA: Uma pessoa tem a capacidade de provocar a ocorrência de algo, p, apenas no caso em que há um curso de ação tal que essa pessoa é capaz de realizar e, se ela fosse realizá-la, então p ocorreria. (p. 79)

Para que essa distinção fique mais clara, considere o seguinte exemplo de Fischer (1994). Um velho marinheiro confere a previsão do tempo todas as manhãs, às 9 horas. Ele tem feito isso nos últimos quarenta anos, sempre para conferir como estará o tempo exatamente ao meio-dia. Se a previsão é de bom tempo, nesse horário, ele sai para navegar. Se é de mau tempo, ele nunca sai – prefere não se arriscar e detesta navegar se o céu está nublado. Imagine que seus padrões psicológicos e de comportamento são muito regulares; além disso, ele não está sujeito a coerção, manipulação, hipnose etc., e o mundo não é determinista. Agora é exatamente meio-dia, e o marinheiro soube essa manhã que a previsão era de chuvas intensas. Ele decide não sair para navegar, mesmo que o barco esteja pronto e todas as outras condições necessárias para a saída sejam satisfeitas. Do fato de que o marinheiro não saiu para navegar, não podemos inferir que ele não poderia ter saído – é claro que ele poderia. Mas o que teria ocorrido se o marinheiro tivesse exercido sua capacidade de sair para navegar exatamente agora? É muito razoável responder que, se ele tivesse ido navegar nesse momento, a previsão vista às 9 horas teria sido de bom tempo (MCKENNA & PEREBOOM, 2016).

Um compatibilista que pretenda objetar as premissas 4 e 6 do Argumento da Consequência poderia adotar a sugestão de Lewis e empreender duas estratégias. A primeira consiste em argumentar o seguinte: sendo o mundo determinista ou indeterminista, é falso que o marinheiro não tem, ou nunca teve, escolha sobre os estados de coisas no passado porque,

embora não pudesse causar a ocorrência de algum evento no passado diferente do atual (ele não é capaz de agir de modo tal que o passado seria diferente, no sentido de CC, que é a tese mais forte), ele era capaz de realizar uma ação – a de sair para navegar – que, se fosse realizada, alguma característica do passado seria diferente (ele é capaz de agir de modo tal que o passado seria diferente, no sentido de CA, que é a tese mais fraca). A noção de capacidade que o compatibilista precisa evocar, portanto, é CA em vez de CC.

A segunda estratégia envolve uma abordagem “humeana”⁵⁸ das leis da natureza, segundo a qual “uma lei da natureza é redutível a regularidades entre eventos” (MCKENNA & PEREBOOM, 2016, p. 81)⁵⁹. Segundo essa abordagem, se a história do mundo tivesse sido diferente (incluindo as ações humanas), então diferentes regularidades talvez emergissem, e a elas seriam redutíveis leis da natureza diferentes. A partir da mesma distinção entre noções de capacidade, o compatibilista pode dizer que, embora o marinheiro não seja capaz de causar uma variação nas leis da natureza realizando uma ação tal que as regularidades a partir das quais a ação ocorre seriam diferentes (CC), ele ainda é capaz de agir diferentemente do que age de modo que, se assim fizesse, as leis da natureza atuais não teriam sido as mesmas (CA)⁶⁰.

O Argumento da Consequência concluía que, a respeito de qualquer sentença verdadeira por exemplo sobre uma ação, e dado o determinismo, ninguém tem ou já teve qualquer escolha a respeito de essa sentença ser o caso. Vimos que um compatibilista pode contrapor as premissas “NM” e “NL” com base na intuição de que, mesmo em um mundo

⁵⁸ Esse tipo de abordagem deriva fundamentalmente da suspeita de David Hume sobre *conexões necessárias irreduzíveis* (BEEBEE & MELE, 2002). Hume famosamente argumentou que “não há nenhuma diferença entre a repetição de casos [de tipos de eventos] constantemente conjugados e um caso similar em particular, *exceto* que depois que experimentamos sua conjunção constante, o hábito nos determina a esperar o efeito quando a causa ocorre” (MORRIS & BROWN, 2022); e que, quando dizemos que um objeto ou um evento está necessariamente conectado a outro (como a conexão entre uma causa e seu efeito, por exemplo), o que queremos dizer de fato não é que estamos observando conexões necessárias (porque elas não podem ser observadas) ou que há entidades tais como conexões necessárias, mas que os objetos/eventos em questão “adquiriram uma conexão associativa em nosso pensamento que deu origem a essa inferência” (MORRIS & BROWN, 2022). Os “humeanos” sobre leis podem então sustentar que a existência de leis da natureza irreduzíveis à conjunção dos eventos que elas conectam não é passível de ser empiricamente observada, ou que é inadequado postular uma entidade tão duvidosa quanto uma lei da natureza que é uma conexão necessária irreduzível (BEEBEE & MELE, 2002).

⁵⁹ A abordagem humeana sobre leis terá como consequência um entendimento da própria tese do determinismo que difere do que estive supondo até aqui.

⁶⁰ Segundo Beebe e Mele (2002), uma visão humeana sobre as leis da natureza permite – diferentemente do que afirma Lewis – que agentes tenham também a capacidade de quebrar uma lei em sentido forte. Essa discussão, porém, não será explorada nesse trabalho. Também não discutirei os detalhes de um tipo de compatibilismo contemporâneo que argumenta que a verdade do determinismo é compatível com a capacidade de agir diferentemente relevante para a responsabilidade moral, o *Novo Disposicionalismo*. Para críticas dessas visões, cf. Clarke, 2009; e Franklin, 2017, capítulo 3.

determinista, o marinheiro poderia ter saído para navegar. E se ele poderia navegar, o passado ou as leis teriam sido ao menos ligeiramente diferentes. Para avaliar as possibilidades de estratégia disponíveis ao compatibilista que acabo de apresentar, porém, consideremos mais uma vez a capacidade de agir diferentemente possuída pelo marinheiro. Tomando “A” como “sair para navegar” (a ação alternativa que o marinheiro não realizou),

Aqui estão algumas coisas que talvez tenhamos em mente quando pensamos ou dizemos que alguém, S, tem a capacidade de (é capaz de, pode) fazer A:

(i) S tem uma capacidade geral de A. (S pode falar espanhol, andar de bicicleta, dirigir um carro etc.)

(ii) S tem uma capacidade geral de A, e as circunstâncias são favoráveis a que S exerça essa capacidade. (S é capaz de dirigir um carro, ela tem um carro em bom funcionamento à mão, o tempo está ameno etc.)

(iii) S tem uma capacidade geral de A, e está aberto a S (em algum instante especificado) exercer (em algum instante especificado) essa capacidade.

(iv) S tem uma capacidade geral de A, e depende de S (em algum instante especificado) se ela (em algum instante especificado) exerce essa capacidade.

(v) S tem uma capacidade geral de A, e S tem escolha (em algum instante especificado) sobre se ela (em algum instante especificado) exerce essa capacidade. (CLARKE, 2009, p. 336-337)

Em que sentidos o marinheiro do exemplo poderia ter feito A, se o mundo for determinista? O sentido (i), que poderíamos descrever como uma *propriedade disposicional* desse agente, claramente se aplica. Sabemos que ele possui todas as habilidades, conhecimentos, motivação etc. necessários para sair para navegar – afinal, ele tem trabalhado como marinheiro há quarenta anos, e nada em seus padrões psicológicos e de comportamento sofreu qualquer alteração. O marinheiro está em perfeito estado de saúde e, por todas essas razões, preserva sua capacidade de sair para navegar mesmo quando está dormindo, ou mesmo quando o barco ainda não está pronto. O sentido (ii), que é um pouco mais exigente do que (i), também se aplica em alguma medida – o barco está preparado para caso o marinheiro decida sair e todas as outras condições foram satisfeitas, exceto a da previsão de tempo favorável, que causou sua decisão de não sair naquele dia. Podemos dizer que o marinheiro ao menos tinha a oportunidade de sair para navegar, além da propriedade disposicional relevante.

Será que os sentidos (i) e (ii) em que o marinheiro poderia ter saído para navegar são tudo o que importa, porém, para a afirmação de que ele tinha escolha a respeito do fato de que não saiu para navegar – do modo apropriado para que fosse, por exemplo, culpado por alguma consequência moralmente relevante (um dano a alguém) que adviesse de sua ação? Uma das lições importantes a ser tirada dessa discussão é a seguinte: a resposta à pergunta sobre se a

capacidade de agir diferentemente é compatível com o determinismo depende do sentido de “capacidade” em questão, e alguns sentidos são claramente compatíveis segundo nossas intuições ordinárias. Por outro lado, talvez o marinheiro não tivesse, em um mundo determinista, a capacidade de sair para navegar nos sentidos (iii)-(v); e algo “do tipo caracterizado por (iv) e (v) é o que muitos dos que pesquisam o tema da liberdade têm em mente” (CLARKE, 2009, p. 339). Nas palavras de Pereboom,

O incompatibilista talvez alegue que, particularmente em contextos em que culpa e punição merecidas de maneira básica estão em questão, há uma razão muito boa para não aceitar simplesmente as intuições ordinárias sobre nossas capacidades de agir diferentemente. Imagine um assassino condenado à morte, sob a justificação de que ele *merece* a pena em virtude do que fez conscientemente. Suponha que ele proteste: “Minha ação foi acarretada pelo passado e pelas leis da natureza, então eu não podia ter evitado matar”. Seria adequado replicar “Sim, sua ação foi acarretada pelo passado e pelas leis, mas quando nós avaliamos a sentença ‘Você poderia ter evitado matar’, não mantemos o passado e as leis fixados, para acomodar nossas intuições ordinárias sobre a verdade desse tipo de sentença. E, nesse caso, a sentença acaba sendo verdadeira” [?] *Não é óbvio que essa réplica seja adequada (...)*. (2022, p. 25-26, grifo meu)⁶¹

Parece então que a estratégia compatibilista de apelar para os sentidos (i) e (ii) de “capacidade de agir diferentemente” (como objeção à solidez do Argumento da Consequência) não captura tudo aquilo que está em jogo na liberdade de agir diferentemente que parece necessária para a responsabilidade moral, no sentido associado a merecimento básico de responsabilização. A distinção de Lewis entre CC e CA apresentada acima, portanto, não seria capaz de oferecer uma resposta à altura da alegação do assassino. Nesse caso, o incompatibilista acerca do Problema da Compatibilidade aparentemente está em vantagem: por meio do Argumento da Consequência, teríamos uma razão substancial para acreditar na intuição de que o determinismo representa uma ameaça à liberdade e (consequentemente) à responsabilidade moral.

Para finalizar este capítulo, retomemos uma questão que deixei em aberto ao final da subseção 1.3.3: será que o Argumento da Consequência *de fato* desempenha o papel decisivo

⁶¹ No exemplo de Pereboom, o fato de o assassino ter sido condenado à morte pode levar algumas pessoas a pensarem que está em jogo o merecimento *dessa* punição em específico, ou então nenhum merecimento de punição. Não é esse o caso. Aqui, como ao longo de todo o debate sobre liberdade e responsabilidade moral, é importante distinguir duas questões: i) a natureza (e a existência) da liberdade que associamos à responsabilidade – e, portanto, a existência de merecimento básico de punição; e ii) a proporcionalidade, ou a adequação, entre a má ação e a punição. Da afirmação de que condenar uma pessoa à morte nunca seria moralmente justificável não se segue, é claro, que outra forma de punição para quem comete um assassinato não o seja. Quando nos referimos à noção de “punitivismo”, por exemplo em relação a Estados que massivamente encarceraram seus cidadãos por crimes de baixa gravidade, a discussão muitas vezes está em ii), ou ainda na questão adicional de se a ação realizada deveria de fato ser encarada como um crime.

que pretende desempenhar em favor do incompatibilismo? O ponto aqui não é acerca de sua solidez – podemos inclusive conceder que alguma de suas formulações seja sólida. Lembre-se, porém, de que esse argumento diz respeito à incompatibilidade entre o determinismo e a *liberdade de agir diferentemente*. Na subseção 1.3.2, introduzi também a noção de *liberdade da fonte genuína*. Não haveria então alguma estratégia disponível ao compatibilista que fizesse uso *dessa* noção de liberdade, em vez daquela⁶²? Essa pergunta será motivada e explorada nos próximos capítulos.

⁶² Esse questionamento foi historicamente introduzido pelos Casos Estilo-Frankfurt (discutidos no capítulo 3 deste trabalho).

3 CAPÍTULO 2: LIBERTISTAS E COMPATIBILISTAS ENTREGUES À PRÓPRIA SORTE

2.1 INTRODUÇÃO: O DILEMA DO DETERMINISMO

Considere mais uma vez a noção de liberdade da fonte genuína, apresentada no capítulo anterior:

Liberdade da fonte genuína. Uma agente é fonte genuína de sua ação apenas “se ela voluntariamente contribui com alguma condição C substantiva e necessária para as condições que de fato conduzem à sua ação, e não há condições suficientes para C que ocorram independentemente de sua agência.” (MCKENNA, PEREBOOM, 2016, p. 150)

No capítulo “Indeterminism and Control: An Approach to the Problem of Luck” (2012), Fischer atribui a William James (1907) o seguinte dilema:

1. Ou o determinismo é verdadeiro, ou é falso.
 2. Se o determinismo causal é verdadeiro, então eu *tenho* que agir do modo como ajo e, portanto, não sou moralmente responsável por minhas ações.
 3. Se o determinismo causal é falso (de maneira relevante), então como eu ajo é uma questão de sorte e, portanto, não sou moralmente responsável por minhas ações.
- Logo,
4. Não sou moralmente responsável por minhas ações. (FISCHER, 2012, p. 85)

Esse argumento é conhecido como *Dilema do Determinismo*. Uma outra formulação é a seguinte:

1. Ou o determinismo causal é verdadeiro, ou é falso.
 - 2*. Se o determinismo causal é verdadeiro, então eu não posso agir diferentemente, portanto não sou moralmente responsável pelas minhas ações.
 - 3*. Se o determinismo causal é falso (de maneira relevante, i.e., nas sequências levando ao meu comportamento), então minhas ações não estão apropriadamente conectadas aos meus estados [mentais] anteriores (i.e., “minhas ações” não são em sentido genuíno *minhas* ações), portanto não sou moralmente responsável pelas minhas ações.
- Logo,
4. Não sou moralmente responsável por minhas ações. (p. 85-86)

No capítulo 1, os desafios mencionados à liberdade de agir diferentemente e à liberdade da fonte genuína supunham que a tese do determinismo era verdadeira e faziam a suposição com base na possibilidade dessa descoberta. Se o determinismo é falso, porém, poderia parecer então que não há nenhuma ameaça (ao menos não-teológica) à existência de

liberdade. Mas não é esse o caso. Segundo Fischer, James chamou a atenção para o fato de que a verdade do *indeterminismo* também representa um desafio: se as conexões causais entre crenças, desejos, deliberação, escolha, movimentos corporais etc. são indeterministas, podemos nos perguntar se é *a agente* que faz a diferença na realização da ação. Parece que para uma decisão e a ação correspondente serem genuinamente da agente, elas precisam estar conectadas aos seus estados mentais anteriores da maneira apropriada, sem que essa conexão seja questão de sorte ou de acaso (FISCHER, 2012).

Para que fique mais claro o desafio à liberdade posto pelo indeterminismo, considere um exemplo apresentado por Peter van Inwagen (2000). Alice está passando por uma situação delicada e precisa decidir entre mentir e dizer a verdade. Ela acaba escolhendo contar a verdade, depois de avaliar seriamente cada uma das opções e pensar em suas prováveis consequências. Suponha que essa escolha foi indeterminada. Agora imagine que existe um deus que, imediatamente após Alice ter dito toda a verdade, faz com o que o tempo seja revertido até um instante logo anterior à decisão da agente, mantendo fixo o mesmo processo deliberativo. O que teria ocorrido nessa segunda sequência: a decisão de mentir ou de dizer a verdade? Considerando que a opção originalmente selecionada foi indeterminada e que a segunda escolha também o seria, parece que a questão sobre o que ocorre nessa última sequência não pode ser respondida – talvez ela minta, talvez diga a verdade. Suponha então que esse deus causa a reversão do tempo até o momento anterior à decisão mais 999 vezes. Ainda não é possível dizer o que exatamente ocorreria, mas, *provavelmente*, veríamos algo como o seguinte: algumas vezes Alice mente, algumas vezes ela diz a verdade. Não seria verdade que, “à medida em que vemos o número de repetições aumentar, deveríamos ficar convencidos de que aquilo que ocorrerá na *próxima* repetição é uma questão de acaso”?⁶³ (VAN INWAGEN, 2000, p. 15)

Vimos antes que o determinismo parece ameaçar a liberdade necessária para responsabilidade moral porque tem como consequência que as ações efetivamente realizadas por pessoas sejam as únicas ações fisicamente possíveis nos momentos em que ocorreram. Além disso, dado o determinismo, parece que nenhuma das condições para que uma certa ação seja realizada é uma contribuição da própria agente – esse é o desafio em relação à noção de liberdade da fonte genuína. O que é sugerido agora é que, também no indeterminismo, as condições para a especificação da escolha de Alice não dependem propriamente dela, mas “do acaso”, mantido o processo original de deliberação que não determina causalmente qual será sua decisão. Já que ela nem sempre toma a mesma decisão, e nada a respeito de sua agência

⁶³ Esse argumento é conhecido na literatura como *Rollback Argument*.

fornece uma explicação (em qualquer uma das mil seqüências) de por que Alice tomou uma decisão em vez de outra, pode-se questionar “a noção de que a *agente* tem o poder de fazer a seleção relevante” (FISCHER, 2012, p. 94).

Observe que não basta haver uma explicação coerente de por que, na primeira seqüência, Alice decidiu dizer a verdade com base em suas motivações, caráter, circunstâncias etc.; e uma outra explicação coerente de por que, em uma seqüência *n* entre as reversões, Alice decidiu mentir com base em suas motivações, caráter, circunstâncias etc. Nesse exemplo, o caráter e os estados mentais dessa agente claramente são compatíveis com ambas as decisões e podemos imaginar que, nos cenários em que mente, Alice está buscando evitar o mal maior (ou algo do gênero). Mas o que de fato importa para a afirmação de que essa agente possui a liberdade da fonte genuína é algo a respeito de sua agência que desempate a decisão – isto é, uma *explicação contrastiva* em termos “agenciais” (MELE, 2006). Temos então o seguinte desafio, colocado pelo indeterminismo, à liberdade da fonte genuína:

1. O indeterminismo é verdadeiro. [Suposição]
 2. Ao menos algumas pessoas [aparentemente] são moralmente responsáveis pelo que fazem.
 3. Responsabilidade moral exige liberdade.
 4. Liberdade exige a capacidade do agente de ser a fonte genuína de suas ações.
 5. A capacidade de ser a fonte genuína de ações é incompatível com o indeterminismo. (MCKENNA, PEREBOOM, 2016, p. 43).
- [Logo, ninguém é moralmente responsável pelo que faz.]

Supondo o indeterminismo, o passado, as leis da natureza e o acaso (utilizando nesse momento apenas noções intuitivas dessas coisas) fazem com que haja condições suficientes (quando vistas em retrospecto) para a ocorrência decisão de Alice. Isso é verdadeiro para qualquer uma das alternativas de ação.

Uma outra maneira de endossar a existência do Dilema do Determinismo está presente no livro *How Physics Make Us Free* (2016), de Jennan Ismael. A autora aponta que é fácil construir uma adaptação do Argumento da Consequência para o indeterminismo, tal como compreendido pela Física contemporânea. Considere uma versão simplificada do Argumento da Consequência:

1. O determinismo (...) implica que os fatos do passado, em conjunção com as leis da natureza, implicam todas as verdades sobre o futuro.
 2. O passado não está sob nosso controle.
 3. As leis da natureza não estão sob nosso controle.
 4. Nossas ações são implicadas pelo passado e pelas leis da natureza.
- Logo,
5. Nossas ações não estão sob nosso controle. (p. 88)

Adaptando essa versão para o indeterminismo, temos o seguinte argumento:

1. Nossa melhor Física atual implica que os fatos do passado, em conjunção com as leis da natureza e os fatos sobre resultados de processos quânticos aleatórios, implicam todas as verdades sobre o futuro.
 2. O passado não está sob nosso controle.
 3. Os resultados de processos quânticos aleatórios não estão sob nosso controle.
 4. As leis da natureza não estão sob nosso controle.
- Logo,
5. Nossas ações não estão sob nosso controle. (p. 98)

A formulação acima adiciona ao Argumento da Consequência original uma fonte da qual as ações humanas seriam consequências: processos quânticos aleatórios. A existência desses processos, se genuinamente indeterministas (em vez de meramente descritos por equações probabilísticas), faria com que o mundo fosse indeterminista. Essa adição, no entanto, não faz diferença para aquilo que está sob o poder da agente: trata-se de mais um elemento fora do controle de seres humanos (inclusive quanto a processos quânticos aleatórios que ocorram dentro do cérebro).

Uma leitura possível desse panorama de desafios à liberdade, que pretendo sugerir de modo breve, é a que se segue. Embora Fischer (2012) sustente que os problemas relacionados ao determinismo e ao indeterminismo são “imagens de espelho”, compartilhando uma estrutura, mas possuindo como alvos diferentes aspectos da liberdade necessária para responsabilidade moral, parece-me que há um enquadramento legítimo segundo o qual determinismo e indeterminismo ameaçam a liberdade da fonte genuína pela *mesma* razão: *sendo o determinismo verdadeiro ou falso, parece que não há uma condição C substantiva e necessária para a realização de uma ação em específico (em vez de qualquer outra) com a qual a agente contribua voluntariamente e para a qual não existam condições suficientes independentes de sua agência*. Se o determinismo for verdadeiro, parece que aquilo que fornece uma explicação contrastiva de uma decisão/ação qualquer de uma agente é, exclusivamente, o passado e as leis da natureza. Se o determinismo for falso, parece que aquilo que fornece uma explicação contrastiva de uma decisão/ação qualquer de uma agente é, exclusivamente, o passado, as leis da natureza e os resultados de processos quânticos aleatórios.

Podemos agora visualizar como o indeterminismo representa uma ameaça *também à liberdade de agir diferentemente*. No Argumento da Consequência, a explicação mais imediata para o fato de que o passado e as leis da natureza não estão sob o controle de agentes era a de que essas coisas são fixadas e, no caso das leis da natureza, fixadas por aspectos do mundo físico que nada têm a ver com a agência particular de indivíduos humanos. Se o determinismo é verdadeiro, não há um *mundo possível* (isto é, um cenário possível em que

operam as mesmas leis deterministas da natureza) com o mesmo passado em que agentes ajam diferentemente em relação a uma ação qualquer. Se o indeterminismo é verdadeiro, por outro lado, os resultados de processos quânticos aleatórios não são fixados: por definição, eles poderiam ter sido outros. Isso significa que, se uma agente faz x em um mundo indeterminista, existe um mundo possível em que ela faz y – agindo, portanto, de modo que os resultados de certos processos quânticos aleatórios seriam diferentes. Entretanto, ainda é verdade que o passado, as leis da natureza e os resultados de processos quânticos aleatórios não estão sob o poder de agentes.

Sendo o indeterminismo verdadeiro, existia uma possibilidade alternativa a essa agente fazendo x – fosse diferente o resultado de certos processos quânticos aleatórios, ela faria y . Mas parece que não há *liberdade* de agir diferentemente, porque a circunstância em que a agente faz y não se efetiva em virtude de uma seleção *da própria agente* de y em vez de x . Mais uma vez, são os resultados de certos processos quânticos aleatórios que parecem “selecionar” a decisão a ser tomada. Retornamos então ao desafio do indeterminismo à liberdade da fonte genuína. O movimento que acabo de traçar mostra o seguinte: a existência bruta de possibilidades alternativas não é condição suficiente para que um agente tenha o tipo importante de acesso a essas possibilidades – o acesso no sentido de ser a fonte da resolução entre duas escolhas fisicamente possíveis.

Essa ideia ficará mais clara com um exemplo. Imagine que Alice é agora uma agente que, sem o seu conhecimento, foi operada por uma neurocirurgiã, a Dra. Black. Black é uma cientista brilhante que desenvolveu um perigoso dispositivo capaz de monitorar a atividade cerebral e *controlar*, se ativado, as decisões e ações de pessoas. Durante a operação, esse dispositivo foi inserido no cérebro de Alice, com o objetivo de fazê-la cometer um assassinato. Suponha que não fazia parte do caráter dessa agente cogitar seriamente um assassinato. Além disso, o dispositivo possuía uma limitação importante: quando ativado, tudo o que ele poderia fazer seria gerar uma probabilidade de 50% de que Alice decidisse e agisse do modo pretendido pela neurocirurgiã. Ele é então ativado; e Alice decide e comete o assassinato – tendo passado a genuinamente desejá-lo como um objetivo seu quando o dispositivo começou a funcionar. Do modo como essa ação ocorre, não parece adequado dizer que a agente era livre e moralmente responsável. No entanto, havia uma possibilidade alternativa igualmente provável de que ela se abstinhasse de cometer o assassinato.

A conclusão que podemos extrair desse caso é a seguinte: a liberdade de agir diferentemente não é *liberdade* sem que haja também a liberdade da fonte genuína. Mas seria a existência de possibilidades alternativas *necessária* para que alguém possua a liberdade de

ser a fonte genuína da própria ação? Como veremos neste capítulo e no próximo, essa questão é altamente controversa. O que podemos estabelecer por ora é que tanto compatibilistas quanto libertistas *precisam oferecer uma resposta* à alegação de que sendo o determinismo verdadeiro ou falso, parece que não há uma condição C substantiva e necessária para a realização de uma ação em específico (em vez de qualquer outra) com a qual a agente contribua voluntariamente e para a qual não existam condições suficientes independentes de sua agência⁶⁴. Uma maneira de fazer isso é negar que uma decisão tomada em um mundo indeterminista seja questão de sorte/acaso, desde que satisfeitos certos requisitos. Essa alternativa endossa a formulação da liberdade da fonte genuína que explorei até agora. Outra resposta, porém, rejeita essa formulação e argumenta que uma outra, mais fraca, é tudo o que a responsabilidade moral exige.

Neste capítulo, meu objetivo será apresentar de que maneira alguns dos principais libertistas e compatibilistas contemporâneos empregam essas estratégias. Na seção 2.2, o foco estará nos compatibilismos baseados em controle. Apresentarei a princípio a proposta de P. F. Strawson ao Problema da Compatibilidade, que não é baseada no controle da agente, e uma razão para que rejeitemos o tipo de abordagem defendido pelo autor. Em seguida, apresentarei o compatibilismo das Visões do Eu Profundo e o Semicompatibilismo de John Fischer e Mark Ravizza, bem como algumas das principais críticas a essas teorias presentes na literatura sobre liberdade e responsabilidade moral. Já na seção 2.3, farei uma contextualização da posição libertista e passarei à exposição do libertismo da causalidade de eventos de Robert Kane, seguida de algumas das principais críticas direcionadas a essa teoria. Essa seção será finalizada com breves observações sobre o que ficou conhecido como o Problema da Sorte para o libertista.

2.2 COMPATIBILISMOS BASEADOS EM CONTROLE

No capítulo 1, vimos que é possível distinguir dois grupos de considerações sobre a agência de uma pessoa capazes de modificar, atenuar ou até causar a suspensão do nosso juízo de que ela é responsável por uma ação, o que ocorre particularmente na forma de uma alteração de nossas reações emocionais (indignação, raiva, rancor, ressentimento, gratidão etc.). Essa observação foi feita por P. F. Strawson no artigo “Liberdade e Ressentimento” (2008), publicado originalmente em 1962. O artigo de Strawson impactou profundamente o debate sobre liberdade e determinismo à época, além de continuar sendo uma grande

⁶⁴ Alguns incompatibilistas (céticos a respeito da liberdade e da responsabilidade moral no sentido associado a merecimento básico de responsabilização) endossariam precisamente essa ideia.

influência para as abordagens contemporâneas da responsabilidade moral. Para o autor, como também vimos muito brevemente, esse grupo de emoções que ele denomina de “atitudes reativas” é parte central das nossas práticas de responsabilização e constitui o fundamento da responsabilidade moral de um indivíduo. No Problema da Compatibilidade, a posição de Strawson é *compatibilista*.

Para compreender o modo como esse autor se coloca na discussão, o primeiro passo é estar ciente de quem eram, à época, os interlocutores de Strawson. A disputa entre compatibilistas e libertistas possuía contornos bastante restritos: enquanto o compatibilista típico sustentava que agir livremente significava essencialmente “agir da maneira desejada sem impedimentos” e enfatizava a utilidade social da responsabilização à maneira grosseiramente consequencialista e desvinculada de merecimento básico mencionada no capítulo 1, o libertista típico, embora justificasse a responsabilização com a noção de merecimento, oferecia uma abordagem da liberdade que é apontada como obscura mesmo por libertistas contemporâneos (como veremos na próxima seção)⁶⁵. Nesse contexto, Strawson chamou a atenção para como as nossas práticas de responsabilização estão fortemente atreladas à *qualidade da vontade* que percebemos nos agentes: nossa avaliação sobre se é adequado por exemplo censurar uma agente por uma má ação depende de qual era sua intenção; de se ela refletia boa vontade e consideração dos interesses dos outros, ou se refletia má vontade, indiferença, desprezo, malevolência etc.

McKenna e Pereboom (2016) identificam ao menos três argumentos distintos pelo compatibilismo no artigo de Strawson. A questão aqui é se a verdade do determinismo de fato fornece razão para suspensão *universal* de responsabilização. Essa seria a consequência óbvia do incompatibilismo, dado que, se o determinismo é verdadeiro, é verdadeiro de modo a constituir as ações de todos os agentes humanos⁶⁶. Os três argumentos de Strawson seriam então os seguintes⁶⁷:

Argumento da Desculpa

- 1) Analisando as formas como desculpamos ou aliviamos a censura a uma agente que agiu mal, há os dois grupos de circunstâncias mencionados no capítulo anterior: ou uma

⁶⁵ Nesse trabalho, não entrarei em detalhes sobre o compatibilismo e o incompatibilismo clássicos (aproximadamente anteriores à década de 1960). Para uma visão geral dessas posições e de seus problemas, cf. McKenna e Pereboom, 2016. É interessante observar, porém, que ao menos quanto a essas posições houve um claro progresso no debate do Problema da Compatibilidade.

⁶⁶ Partindo do pressuposto, é claro, de que o escopo das leis da natureza é universal ao menos quanto aos agentes.

⁶⁷ Os nomes desses argumentos são sugeridos por McKenna e Pereboom (2016).

agente em geral moralmente responsável não é censurável por aquela ação em particular porque *não agiu com uma qualidade de vontade moralmente objetável* (ela foi compelida, ou coagida, ou não sabia o que fazia etc.), ou a agente nem mesmo é moralmente responsável em geral porque é (ainda que temporariamente) *incapaz de relações interpessoais “normais”*.

2) A tese do determinismo não tem como consequência lógica que nenhum agente age com qualidade da vontade moralmente objetável ou que nenhum agente é capaz de relações interpessoais normais.

Logo,

3) A verdade da tese do determinismo não fornece razão para suspensão universal de responsabilização.

Argumento da Impossibilidade Psicológica

1) Não é possível para nós extinguir nossas atitudes reativas, já que elas são *parte crucial da natureza humana*. Ainda que às vezes adotemos “atitudes objetivas” perante certos agentes (aquelas que os tratam como alvo apropriado de tratamento, ou proteção, ou reclusão, em vez de atitudes de maior envolvimento interpessoal), não somos capazes de sustentá-las indefinidamente e, mesmo quando as adotamos, não o fazemos por uma convicção teórica (um "determinismo do caso concreto").

Logo,

2) Não é possível agir como o incompatibilista pensa que devemos agir se o determinismo for uma tese verdadeira.

Argumento da Racionalidade Prática

1) Ainda que fosse possível suspender nossas atitudes reativas, uma opção prática como essa deve ser feita à luz de uma consideração de perdas e ganhos para a experiência humana.

2) Muito plausivelmente, adotar apenas atitudes objetivas em relação a outras pessoas resultaria em um empobrecimento da experiência humana, tornando o mundo um lugar mais frio e impessoal. Suspender nossas atitudes reativas, então, traria mais perdas do que ganhos.

Logo,

3) Não devemos agir como o incompatibilista pensa que devemos se o determinismo for uma tese verdadeira.

Strawson conclui o artigo afirmando então que se as práticas de responsabilização moral fazem parte de nossa natureza, da maneira como fazem, então elas não são suscetíveis à justificação que tanto compatibilistas quanto incompatibilistas buscam. Estariam ambos “hiperintelectualizando os fatos”. Fischer (1994) reconstrói esse ponto da seguinte maneira:

Se alguém busca *fundamentar* a sustentação *dessas* atitudes [as atitudes reativas] (em vez de buscar o mero condicionamento de comportamento ou a terapia), essa pessoa deve buscar pela satisfação de alguma “exigência teórica”. E é plausível que a exigência teórica em questão seja que os agentes são “metafisicamente livres”. Mas essa liberdade precisa ser “libertista” (...), que é misteriosa e não é apelativa. Seria então mais plausível dizer que *nenhuma* exigência teórica precisa ser satisfeita para que um agente seja moralmente responsável. Em vez disso, a propensão da comunidade em sustentar atitudes reativas perante o agente *constitui* a responsabilidade moral desse agente. Strawson enfatiza que é um *erro* buscar alguma fundamentação teórica para a responsabilidade moral (...). (p. 211)

Mas seria a *rejeição de uma condição metafísica (ou de liberdade)* para a responsabilidade moral a melhor (ou mesmo a única) estratégia teórica disponível para o compatibilista? Fischer (1994) continua da seguinte forma: “se *agir livremente* é distinguido da liberdade de agir diferentemente, talvez haja um tipo de liberdade que é robusta o suficiente para fundamentar as atitudes reativas mas que não precisa ser entendida como libertista” (p. 212). Em outras palavras, o compatibilista teria a opção de trabalhar com a noção de liberdade da fonte genuína, preservando assim a busca por uma justificação das práticas de responsabilização. A alternativa proposta por Strawson seria a única plausível somente em um contexto insuficientemente matizado de posições teóricas.

Em *The Metaphysics of Free Will* (1994), Fischer reconhece os méritos da abordagem de Strawson, como a atenção dirigida ao papel das emoções na responsabilização moral. Ele oferece, porém, uma razão adicional para que o compatibilista não desista da busca de exigências teóricas satisfeitas por agentes responsáveis. Na perspectiva “strawsoniana”, não há algo como “ser moralmente responsável” e a prioridade explanatória é inversa: é o fato de uma agente ser “recipiente de atitudes reativas e participante nas práticas associadas” (p. 212) que explica que ela seja responsável, em vez de sua responsabilidade explicar que ela seja alvo apropriado de atitudes reativas. Mas essa conexão parece tornar a teoria incapaz de criticar práticas existentes.

Fischer sugere que imaginemos uma sociedade que dirige censura, ressentimento e punição a pessoas com graves deficiências cognitivas porque interpreta algumas de suas ações como provenientes de uma natureza má, ou da falta de consideração com os outros (isto é, como exibindo má qualidade da vontade). Deveríamos considerar que “o mero fato de seu

comprometimento com esses sentimentos elimina as preocupações com se a adoção dessas atitudes perante esses agentes é de fato justificada?” (p. 213). Parece que não. Sem um critério independente de nossas práticas, portanto, não poderíamos avaliar se são ou não adequadas⁶⁸ – e nós efetivamente *fazemos* esse tipo de avaliação. O critério independente mais intuitivo é, é claro, se a agente tinha ou não controle sobre sua ação. Por essa razão, o foco do restante desta seção será apresentar algumas propostas compatibilistas *baseadas em controle*.

2.2.1 As Visões do Eu Profundo

As abordagens que Wolf (1987) denominou de *Visões do Eu Profundo* afirmam que ações livres (da maneira necessária para a responsabilidade moral) são aquelas que emergem do “eu profundo” ou “eu real” da agente. A proposta mais conhecida desse tipo é a de Frankfurt (1971), segundo a qual a liberdade exige uma *coincidência* entre o desejo que motiva a ação e o desejo de segunda ordem a respeito do primeiro – mais precisamente, a respeito de possuir ou de tornar efetivo o desejo de primeira ordem. Um desejo de primeira ordem é identificado “por afirmações da forma “A deseja X”, nas quais o termo “X” refere-se a uma ação” (PEREBOOM, 2022, p. 17). Já um desejo de segunda ordem é identificado “por afirmações da forma “A deseja X”, nas quais o termo “X” refere-se a um desejo de primeira ordem” (PEREBOOM, 2022, p. 17). Para exemplificar a condição especificada por Frankfurt, imaginemos que Teresa é dependente química de uma droga *y*. A existência de responsabilidade moral pelo uso da droga nesse caso dependerá de se a agente desejava efetivar o desejo de usar *y*. Se Teresa faz uso de *y*, mas deseja honestamente que seu vício não existisse, então ela não é responsável. Já se Teresa não se incomoda em possuir um vício por *y* e, ao contrário, deseja mantê-lo, o uso de *y* está alinhado ao eu mais profundo dessa agente e é, portanto, uma ação pela qual Teresa é moralmente responsável. Considerando que a verdade da tese do determinismo não faria diferença alguma para se agentes endossam ou não os seus desejos de primeira ordem, essa abordagem das ações livres é compatibilista.

Segundo Wolf, as Visões do Eu Profundo capturam muito bem algumas de nossas intuições. Em primeiro lugar, elas “compartilham a ideia de que agência responsável envolve algo além de agência intencional” e que uma ação livre não emerge simplesmente de “estados psicológicos *em* nós, mas [de] expressões de caracteres que vêm *de* nós” (1987, p. 375). Além disso, parece que tanto uma pessoa cleptomaniaca quanto uma pessoa que age sob coerção,

⁶⁸ Para uma perspectiva recente que nega essa afirmação, cf. Shoemaker, 2011. Abordagens que rejeitam a condição metafísica da responsabilidade moral são minoritárias e não serão discutidas nesse trabalho.

por exemplo, estão alienadas de seu eu mais profundo; enquanto outros animais e talvez as inteligências artificiais não possuem “eus” reais na forma de desejos de segunda ordem. Wolf observa também que essas teorias possuem uma importante virtude teórica. Perante a preocupação ingênua de que se tudo é causalmente determinado, então o que acontece “simplesmente acontece” de modo independente de nossas vontades, o teórico do Eu Profundo responderia que o determinismo não impede a eficácia causal que os desejos de uma agente podem possuir sobre seus comportamentos.

Há, no entanto, ao menos três problemas com essas teorias. Frankfurt (1971) admite que o eu profundo de uma pessoa possa ser governado por um eu ainda mais profundo, sendo os desejos de segunda ordem alvo de desejos de terceira ordem, e assim por diante. Mas Wolf (1987) aponta que, embora não haja um limite lógico para o número de níveis de desejos, certamente há um limite *psicológico*: não conseguimos nem mesmo conceber o que seria um desejo de quarta ordem (e menos ainda de níveis mais altos) e, mesmo que conseguíssemos, ainda seria legítimo questionar o que é que governa o *último* nível. Aqui, a resposta precisaria apelar a algo externo a nós. Nesse caso, de que adiantaria a adição de “voltas” à corrente causal levando a uma ação? Se o último nível do eu de uma agente é ele próprio determinado por fatores externos, chamá-lo de seu “eu real” é de fato apropriado?

Embora levante essa preocupação, Wolf não a endossa completamente⁶⁹. A ressalva da autora é crucial para que compreendamos uma motivação geral do compatibilismo. Ela afirma o seguinte:

Do ponto de vista de quem possui esse medo, a responsabilidade parece exigir que sejamos um primeiro motor imóvel, cujo eu mais profundo não é nem aleatório nem determinado, mas é determinado por si mesmo – ele é, em outras palavras, *autocriado*. Nesse ponto, porém, os proponentes da Visão do Eu Profundo podem questionar se esse medo é legítimo. (...) (O) retorno ao ponto de vista interno do agente cuja responsabilidade está em questão tornará difícil enxergar que bem esse estatuto metafísico [a “autocriação”] deveria fornecer ou que mal sua ausência deveria impor. (1987, p. 378, grifo meu).

O ponto crucial para o qual Wolf chama a atenção aqui pode ser resumido da maneira que se segue. Quando nos atentamos ao desafio que o determinismo oferece à liberdade de uma agente ser a fonte genuína da própria ação, parece que a existência dessa liberdade é incompatível com o fato de essa ação ser, em última instância, causada por fatores externos à sua agência. Por outro lado, como apontei na introdução deste capítulo, há um problema

⁶⁹ Após a ressalva que indico a seguir, Wolf formula uma outra objeção às teorias do Eu Profundo que ajuda a motivar a abordagem das ações livres que ela própria oferece. O caso a partir do qual essa objeção é formulada ficou bastante conhecido na literatura sobre liberdade e responsabilidade moral e será explorado na subseção 2.2.3.

análogo quanto ao indeterminismo. Aplicando-o às teorias do Eu Profundo, se o determinismo é falso, ainda que o meu “eu real” não seja causalmente determinado por fatores externos a mim, ele ainda não será determinado *por* mim – mas será formado, digamos, aleatoriamente. Parece então que dessa “perspectiva externa” a respeito das sequências causais levando a ações, como diria Wolf, a única solução seria exigir que uma agente tenha criado a si mesma. Mas não apenas é difícil conceber como uma pessoa pode ter criado seu eu real (incluindo, de maneira relevante, seus desejos, preferências, hierarquias de valores etc.) como também é difícil visualizar que diferença, *do ponto de vista da agente*, essa exigência tão exagerada poderia fazer. Se nós, diferentemente de outros animais, somos já capazes de refletir sobre quem somos, de avaliar que marcas queremos deixar no mundo e, crucialmente, de revisar a nós mesmos frente a essas reflexões, que outro tipo de liberdade poderíamos querer? Essas capacidades não são já suficientes para que atribuamos responsabilidade moral a nós mesmos e aos outros?

Compatibilistas sustentam que caracterizações da liberdade que exigem tanto o indeterminismo quanto algum sentido de “autocriação” são excessos filosóficos a serem combatidos⁷⁰. Mas isso não significa, obviamente, que pensem que qualquer teoria compatibilista seja razoável. Quanto às Visões do Eu Profundo, há ainda dois problemas importantes a serem apontados. O primeiro motivará a construção de compatibilismos “históricos” como o Semicompatibilismo, que discutirei na próxima subseção. Observe que, na visão de Frankfurt, não importa como uma agente chegou a endossar seu desejo de primeira ordem para que ela seja responsável pela ação que resulta desse desejo: basta que ela o endosse. Mas imagine o exemplo a seguir⁷¹. Uma pessoa de pouco engajamento e convicção políticos e grande habilidade com armas é submetida a meses de lavagem cerebral por um grupo de opositores do presidente em exercício. Por meio da lavagem cerebral, o grupo altera radicalmente a personalidade desse agente, que passa a ter crenças políticas extremas e um ódio pessoal do presidente, chegando ao ponto de decidir assassiná-lo. Suponha então que essa pessoa o assassina e, quando isso ocorre, ela não apenas desejava assassiná-lo como também desejava que esse desejo fosse efetivado. Nesse caso, os desejos de primeira e de

⁷⁰ Cf. também Fischer, 2012.

⁷¹ Para casos desse tipo, cf. Mele, 1995; e Fischer e Ravizza, 1998.

segunda ordem coincidem, tal como exigido por Frankfurt⁷². É estranho, porém, dizer que esse agente é plenamente responsável por sua ação; afinal, ele foi severamente manipulado.

A conclusão que se pode extrair desse tipo de caso é que, independentemente de quais forem as condições necessárias e suficientes para a realização de uma ação livre estipuladas por um compatibilista, parece que elas precisam incluir alguma condição histórica, que garanta que os estados mentais da agente relevantes para sua liberdade e que operam no momento da ação tenham emergido de uma maneira que não ameace essa liberdade. O próximo problema para as Visões do Eu Profundo também oferecerá um *desideratum* para o compatibilista. Ele diz respeito a casos de *acrasia* – isto é, de um conflito de desejos em que aquilo que o agente chega a fazer contraria seus juízos sobre o melhor curso de ação naquele contexto⁷³.

Imagine que Teresa é uma atleta amadora que encontra a carteira de um vizinho na calçada do prédio⁷⁴. Ela sabe quem é o dono e pode facilmente devolvê-la. Mas, embora essa agente julgue que deve devolver a carteira, deseje fazer isso e deseje também que esse primeiro desejo seja efetivado, existe um conflito em seus estados motivacionais. Teresa percebe que há na carteira dinheiro suficiente para que ela compre um relógio esportivo novo, com muito mais funcionalidades para suas corridas. Ela então fica seriamente tentada a pegar o dinheiro para si e guarda a carteira por algumas horas, até tomar sua decisão. No fim do dia, Teresa não consegue conter o seu desejo pelo relógio e, mesmo sentindo-se culpada e reprovando o próprio comportamento, vai até a loja e faz a compra com o dinheiro do vizinho.

Se o exemplo anterior mostrava que a condição estipulada por Frankfurt não era suficiente para uma ação livre, o exemplo acima mostra que ela nem mesmo é necessária: o desejo de primeira ordem que motiva a ação de Teresa e seu desejo de segunda ordem sobre o que fazer com a carteira não coincidem e, ainda assim, julgamos que ela seja moralmente responsável por sua decisão. O compatibilista precisa, portanto, de uma abordagem das ações livres que *exclua* casos de manipulação, ou de mudança radical de personalidade/valores etc.⁷⁵

⁷² Esse exemplo pode ser facilmente adaptado para versões das teorias do Eu Profundo em que o que equivale ao eu real são os valores ou as crenças refletidas da agente em vez de meros desejos. O ponto aqui é que o fato de essas visões ignorarem o modo como o que quer que represente o eu mais profundo do sujeito foi formado/adquirido – isto é, o fato de ignorarem a história da agência desse sujeito – é um desafio teórico bastante sério.

⁷³ Adoto aqui uma noção mais ou menos aristotélica desse tipo de caso. Cf. Aristotle, 2009.

⁷⁴ Adaptado de Haji, 1998. Para outra objeção envolvendo um caso de *acrasia*, cf. Arpaly, 2003.

⁷⁵ Mais uma vez, vale mencionar o caso de Phineas Gage (DAMASIO, 1995). Um agente que não demonstra consideração com outras pessoas ao agir e que não se incomoda em ser desse modo, mas que só possui essas características em virtude de um dano neurológico que alterou relevantemente sua personalidade (e que pode ser menos dramático do que ter o cérebro atravessado por uma barra de

(além de casos de compulsão e coerção), e *inclua* casos de acrasia. Na próxima subseção, veremos como dois autores tentaram atender a essas exigências teóricas.

2.2.2 O Semicompatibilismo de Fischer e Ravizza

O Dilema do Determinismo estabelecia que se o determinismo é verdadeiro, então (2) eu tenho que agir do modo como ajo, ou (2*) eu não posso agir diferentemente; portanto, não sou moralmente responsável por minhas ações. Se, porém, o determinismo é falso, então (3) como eu ajo é uma questão de sorte, ou (3*) minhas ações não estão apropriadamente conectadas aos meus estados mentais anteriores; portanto, não sou moralmente responsável pelas minhas ações. Nesta subseção, veremos como o *Semicompatibilismo* de Fischer e Ravizza (1998) responde aos pares de premissas (2)-(2*) e (3)-(3*) e oferece uma teoria da condição metafísica da responsabilidade moral. O Semicompatibilismo é a posição no Problema da Compatibilidade que estabelece as três seguintes teses: (i) não sabemos se o determinismo é compatível com a liberdade de agir diferentemente – que Fischer chama de *controle regulador* (1994, 2012); (ii) a liberdade de agir diferentemente, porém, é irrelevante para a responsabilidade moral; e (iii) a liberdade de uma agente ser a fonte genuína da própria ação – aquilo que Fischer chama de *controle guia* (1994, 2012) – é compatível com o determinismo.

Se a responsabilidade moral exige a liberdade de agir diferentemente (em outras palavras, se o Princípio das Possibilidades Alternativas (PPA), apresentado no capítulo 1, é verdadeiro) e se o Argumento da Consequência é sólido, então as premissas (2)-(2*) são verdadeiras. No capítulo 1, vimos algumas estratégias pelas quais o compatibilista pode negar a solidez desse argumento. A preferência de Fischer (1994, 2012), no entanto, é *atacar o PPA*. O Semicompatibilismo é oficialmente agnóstico quanto à solidez do Argumento da Consequência, mas, se o PPA é falso, então não podemos imediatamente inferir de sua solidez que ninguém é moralmente responsável por suas ações caso o determinismo seja uma tese verdadeira. Por essa razão, seria possível “contornar” o Argumento da Consequência e construir uma teoria da liberdade necessária à responsabilidade moral, sem que seja necessário oferecer-lhe uma resposta. Em sua defesa da tese (ii), Fischer e Ravizza estão apoiados nos *Casos Estilo-Frankfurt*.

O capítulo 3 deste trabalho oferecerá uma abordagem detalhada, e uma defesa, dos Casos Estilo-Frankfurt. Nesse momento, porém, basta uma apresentação intuitiva de como

ferro) provavelmente não seria julgado por nós como moralmente responsável por suas ações de maneira plena.

esse modelo de casos, elaborado por Harry Frankfurt (1969), desafia o PPA. Alice e Dra. Black são novamente as personagens do exemplo: imagine que a neurocientista inseriu no cérebro de Alice um poderoso dispositivo capaz de monitorar a atividade cerebral e, caso seja engatilhado, de controlar as decisões e ações dessa agente. Black quer garantir que ela cometa um assassinato. Suponha dessa vez, porém, que Alice já está considerando cometê-lo por motivo de vingança; e que a neurocientista não deseja mostrar-se desnecessariamente – ela ativará o dispositivo somente na presença de algum sinal na atividade cerebral de Alice que indique que sua tendência é a de desistir de seus planos. Mas a agente acaba decidindo cometer o assassinato por suas próprias razões e, conseqüentemente, o dispositivo nunca é engatilhado.

O argumento formulado por Frankfurt a partir de casos desse tipo começa por observar que, embora o dispositivo da neurocirurgiã não interfira na sequência de eventos levando à ação (nem desempenho, portanto, qualquer papel na explicação dessa ação), sua presença elimina as possibilidades alternativas do cenário – afinal, Alice decidiria cometer o assassinato de qualquer modo. Poderíamos então concluir o seguinte: considerando que a agente não tinha possibilidades alternativas robustas de ação e, ainda assim, *parece livre da maneira relevante para ser moralmente responsável pelo assassinato*, então o PPA é falso (cf. FISCHER, 1994; SARTORIO 2016). Nossa intuição é a de que Alice é a fonte genuína da própria ação, sem que a falta da liberdade de agir diferentemente faça qualquer diferença para sua responsabilidade.

Mas e quanto à premissa (2), que estabelece que, se eu tenho que agir como ajo, então não sou moralmente responsável por minhas ações? As ideias presentes em *Responsibility and Control: a theory of moral responsibility* (1998) são uma resposta a essa preocupação. A tese central de Fischer e Ravizza nesse livro é a de que a liberdade de uma agente ser a fonte genuína da própria ação equivale ao “controle guia”. Esse tipo de controle possui duas componentes: a *responsividade moderada a razões* do “mecanismo” (o processo) que leva à ação e a *posse, por parte da agente, desse mecanismo*. O conceito de responsividade a razões desempenha papel fundamental aqui. Como apontam McKenna e Pereboom (2016), teorias da responsividade (ou sensibilidade) a razões “têm raízes na ideia aristotélica de que o ser humano é um animal racional que, quando maduro, pode compreender e ser sensível a boas razões sobre como viver bem e tomar decisões informadas” (p. 217). A visão de que a *flexibilidade* de ação de uma agente frente a considerações racionais é central à responsabilidade moral, endossada por essas teorias, também captura algumas de nossas intuições sobre casos paradigmáticos de ausência de liberdade, já que uma agente “agindo sob

transtornos compulsivos, fobias, vícios ou episódios psicóticos (...) talvez não sejam adequadamente sensíveis (...) às razões que possuem para agir” (p. 216).

Segundo Fischer e Ravizza (1998), a responsividade a razões é uma propriedade modal, ou contrafactual. Isso significa que “mundos possíveis não atualizados podem ser relevantes para a responsabilidade moral em virtude de ajudarem a especificar ou analisar propriedades (...) [presentes na] sequência [de eventos levando à ação no mundo atual]” (FISCHER, 2006, p. 17). Uma pessoa cujo mecanismo de ação é moderadamente responsivo a razões

(...) reconheceria razões (algumas delas morais) de tal modo que desse origem a um padrão compreensível (do ponto de vista de um terceiro que compreendesse os valores e crenças do agente) e reagiria a pelo menos uma razão suficiente para agir diferentemente (em algum cenário possível). (FISCHER; RAVIZZA, 1998, p. 243-244).

Em termos gerais, uma agente moralmente responsável por uma ação teria agido diferentemente, *mantendo fixo o tipo de mecanismo levando à ação* que opera no mundo atual, na presença de razões suficientes para agir de outro modo em pelo menos uma circunstância alternativa. A expressão “razão suficiente” significa aqui a razão mais forte da agente, ou sua melhor razão para ação.

Essa caracterização será mais bem explicada a seguir. Nesse momento, porém, podemos nos perguntar como é que uma agente de um Caso Estilo-Frankfurt poderia ser moralmente responsável segundo o Semicompatibilismo, considerando que a presença do dispositivo de Black impediria que ela agisse diferentemente em todos os cenários possíveis. A estratégia de Fischer e Ravizza é deslocar o foco da análise: não é *a agente*, mas *o mecanismo*, aquilo que precisa responder moderadamente a razões para agir de outro modo. Retomando o ponto de Frankfurt de que o dispositivo (apesar de eliminar possibilidades alternativas de ação) na verdade não interfere na decisão da agente, é plausível afirmar que sua presença na história *não faz parte* do mecanismo levando à decisão e à ação. No cenário contrafactual em que Alice tende a desistir de cometer o assassinato, o dispositivo é engatilhado e Black garante que essa decisão seja tomada, fazendo com que a agente não seja moralmente responsável. Nesse caso, a intervenção é parte do mecanismo e, mantendo-o fixo, não haveria qualquer mundo possível em que Alice agiria diferentemente. Mas aquilo que ocorre em um Caso Estilo-Frankfurt difere relevantemente desse cenário. Segundo Fischer e Ravizza, a razão pela qual uma agente como Alice parece ser moralmente responsável é o fato de que seu mecanismo de ação parece moderadamente responsivo a razões, tal como o de uma

agente que deliberasse sobre um assassinato sem que nenhum dispositivo tivesse sido implantado em seu cérebro – ambos os mecanismos seriam do mesmo tipo.

Mas o que significa dizer que o mecanismo de ação de uma agente é *moderadamente* responsivo a razões? Há um modo pelo qual poderíamos descrever a responsividade a razões em que a exigência de flexibilidade de ação é bastante forte: se o mecanismo efetivo de uma ação fosse operar e “houvesse razão suficiente para agir diferentemente, o agente reconheceria a razão suficiente para agir diferentemente e então escolheria (...) e faria diferentemente” (FISCHER; RAVIZZA, 1998, p. 41). Não é plausível, porém, que esse grau de exigência seja necessário para a responsabilidade: os casos de acrasia apontam justamente na direção contrária. Retomemos a história de Teresa, que encontra a carteira de um vizinho na calçada e gasta o dinheiro que há guardado nela em vez de devolvê-la. Se a responsividade *forte* a razões fosse condição necessária para a responsabilidade moral, precisaríamos dizer que Teresa não é responsável, dado que havia uma razão suficiente para que ela agisse de outro modo – uma razão moral para que a carteira fosse devolvida ao dono – e, apesar de reconhecê-la, a agente não é suficientemente motivada a agir com base nessa razão.

Além disso, se a responsividade forte a razões fosse exigida para a responsabilidade moral, nós nunca responsabilizaríamos alguém que realiza uma ação moralmente errada – afinal, para toda ação desse tipo há uma razão suficiente para não a realizar. Essa condição precisa, portanto, ser enfraquecida. Por outro lado, se exigirmos apenas a responsividade *fraca* a razões, demandando apenas que “exista [apenas] algum cenário possível (ou mundo possível) no qual haja razão suficiente para agir diferentemente, o agente reconheça essa razão, e aja diferentemente” (FISCHER; RAVIZZA, 1998, p. 44), existe o risco de que os padrões de responsividade de uma agente sejam tão irracionais que poderíamos questionar se de fato a agente é normativamente competente (um critério plausível para que, como explicado no capítulo 1, uma pessoa faça parte do grupo dos agentes morais).

Para que a exigência de responsividade a razões seja descrita da maneira adequada, nem forte nem fraca demais, Fischer e Ravizza sugerem que há dois elementos em jogo: a *receptividade* e a *reatividade* a razões. Receptividade é a capacidade de reconhecer e avaliar conjuntos de razões para agir e exige que o agente ao menos “não esteja substancialmente iludido sobre a natureza da realidade” (p. 73). Já a reatividade é a capacidade de “traduzir” essa avaliação em comportamento correspondente. A proposta dos autores envolve então uma assimetria: o controle guia exigiria receptividade *regular* do mecanismo a razões, mas apenas reatividade *fraca* a essas razões. Quanto à receptividade regular a razões,

é como se um “terceiro” (aquele avaliando a responsabilidade moral do agente) conduzisse uma “entrevista imaginária” com o agente. Nessa entrevista, ele pergunta sobre vários cenários efetivos e hipotéticos, e extrai visões do agente sobre o que constituiria razões suficientes. Mesmo se a pessoa afirmar que, dados seus efetivos valores (e preferências), apenas uma razão contaria como suficiente, o padrão de receptividade de seu mecanismo efetivo ainda poderia ser testado perguntando-a quais razões contariam como suficientes, dado um outro conjunto de valores (e preferências). O entrevistador então empregaria a informação (...) para buscar compreender o padrão no conjunto de reconhecimentos de razões. O que é exigido é que a configuração de respostas da entrevista imaginária possa (junto com informações de fundo) dar origem a um padrão compreensível, da perspectiva desse terceiro (...). (FISCHER; RAVIZZA, 1998, p. 71-72)

Essas exigências permitem que incluamos uma agente como Teresa no grupo de agentes moralmente responsáveis. Se pudermos supor que ela reconhece um amplo e razoavelmente coerente espectro de razões para agir, incluindo a razão segundo a qual ela deveria devolver a carteira ao dono, e que ela de fato *devolveria* a carteira em um cenário em que por exemplo soubesse que esse vizinho está passando por sérias dificuldades financeiras (correndo inclusive o risco de perder o próprio apartamento), então poderemos considerar que Teresa é moderadamente responsiva a razões (apesar de, no mundo atual, ela não ter sido motivada por uma razão suficiente para agir de outro modo). O Semicompatibilismo permite, portanto, que responsabilizemos agentes “acráticos”.

Já para evitar que agentes manipulados sejam incluídos no grupo dos agentes moralmente responsáveis, Fischer e Ravizza estipulam uma componente de posse do mecanismo de ação. Ela envolve um *assentimento de responsabilidade*. Crianças de qualquer sociedade são, enquanto crescem, submetidas a educação moral e gradativamente inseridas em práticas de responsabilização interpessoais. Em parte por essa razão, elas passam a enxergar a si mesmas como agentes, no sentido de que resultados no mundo dependem de suas decisões e comportamentos. Ocorre também a aquisição de algumas crenças, incluindo a de que elas são alvos apropriados de certas respostas como consequência do modo como exercem sua agência. Por meio desse processo, uma agente moralmente madura assente responsabilidade por tipos de mecanismo de ação, como o raciocínio prático e uma série de hábitos não-refletidos, a que pertencerão suas ações livremente realizadas. Esse assentimento é de caráter histórico: é adquirido em algum momento *anterior* à ação pela qual a agente é responsável, de modo tipicamente *não-voluntário*, sem que precise ser consciente ou refletido. Para que satisfaça a condição de posse do mecanismo de ação, a visão do indivíduo sobre si mesmo precisa ainda ser baseada, de maneira apropriada, em evidência.

Vejamos então como a adição dessa condição interfere na capacidade da teoria de lidar com casos de manipulação. Retomando dessa vez o exemplo da pessoa submetida a meses de lavagem cerebral, que passa então a sustentar crenças políticas extremas e decide assassinar o presidente da República, o diagnóstico de Fischer e Ravizza seria o seguinte: dado que o mecanismo levando a essa decisão não é um mecanismo pelo qual esse agente tenha naturalmente assentido responsabilidade, ele não pode ser considerado moralmente responsável. Aqui, o grupo de manipuladores “insere” no agente um processo decisório estranho, ou ao menos as condições para que esse processo estranho ocorra – o que explica que ele não seja responsável, ainda que cumpra a condição de responsividade moderada a razões.

A teoria semicompatibilista da responsabilidade moral é bastante sofisticada e influente no debate no Problema da Compatibilidade, a ponto de ser considerada um modelo para os compatibilismos contemporâneos (MELE, 2011). Fischer e Ravizza explicam a compatibilidade do controle guia com o determinismo do seguinte modo. Quanto ao “teste” de responsividade moderada a razões do mecanismo, é importante notar que “a existência do mesmo tipo de mecanismo não exige semelhança em todos os detalhes” e os cenários pertinentes “podem diferir tanto nos tipos de incentivos que o agente possui para agir diferentemente quanto nos detalhes particulares” (1998, p. 52). De maneira central, eles sustentam que *não é plausível exigir que todas as condições do passado sejam incluídas como parte do mecanismo*. Em suas palavras,

Se o determinismo causal é verdadeiro, então qualquer cenário possível (com as leis da natureza atuais) em que o agente aja diferentemente no instante T precisa diferir em *algum* aspecto do cenário atual anterior a T. A existência desses cenários possíveis é *tudo* o que é exigido pela nossa teoria da responsabilidade moral. Crucialmente, nossa abordagem *não* exige que o agente seja capaz de efetivar esse cenário, isto é, que ele tenha em seu poder, em T, agir de modo tal que o passado (relativo a T) teria sido diferente do que foi. E a existência desse tipo de cenário exigido é certamente compatível com o determinismo causal. Portanto, nossa abordagem (...) deixa espaço para a responsabilidade moral por ações mesmo em um mundo causalmente determinado. (p 52)

Isso significa que, se uma agente em um mundo determinista realiza a ação X com um mecanismo próprio, e se, entre os mundos possíveis próximos ao atual (que serão também deterministas), há pelo menos algum em que o mesmo tipo de mecanismo opera e a agente age diferentemente porque o passado é um pouco diferente do mundo atual (pelo menos quanto às razões suficientes disponíveis para que essa agente aja de outro modo ou quanto à motivação da agente perante o mesmo grupo de razões para agir), a existência *desse* mundo possível já seria suficiente para que consideremos (supondo que a agente seja regularmente

receptiva a razões) que o mecanismo *no mundo atual* é moderadamente responsivo a razões. Essa agente seria então moralmente responsável, mesmo que não possua a liberdade de agir de outro modo e mesmo que a ação X seja causalmente determinada.

Para finalizar esta subseção, vejamos como o Semicompatibilismo responde ao par de premissas (3)-(3*) do Dilema do Determinismo. O desafio do indeterminismo à liberdade de uma agente ser a fonte genuína da própria ação estabelece que se o determinismo é falso, então minhas ações não estão conectadas da maneira apropriada aos meus estados mentais anteriores. Fischer (2012) argumenta que há uma resposta a essa preocupação que emprega ingredientes similares aos de um Caso Estilo-Frankfurt. Imagine um mundo determinista M1 em que uma agente toma a decisão de levantar a mão em um instante t_2 , por suas próprias razões e da maneira como normalmente tomaríamos. Suponha que a “cola” entre estados mentais anteriores e decisão exigida para responsabilidade moral está presente. Suponha ainda um mundo possível M2 em que quase tudo é semelhante a M1, exceto pelo fato de que quando a agente começa a deliberar sobre levantar a mão, em t_1 , uma máquina aleatória, genuinamente indeterminista, começa a operar no intervalo até t_2 . Essa máquina tem 50% de chance de não interagir causalmente com o mundo (incluindo a agente), isto é, de não fazer nada. A outra possibilidade, também com 50% de chance de ocorrência, é a de que a máquina seja engatilhada e *cause* a decisão da agente de levantar a mão em t_2 . No mundo M2, o que ocorre é que a máquina acaba não fazendo nada. Assim, a agente decide levantar a mão por suas próprias razões. O ponto de Fischer é então o seguinte: se em M1 a “cola” necessária à responsabilidade moral está presente, por que razão diríamos que esse não é o caso em M2?

Parece, ao contrário, que não há diferença de controle da agente entre M1 e M2, dado que a máquina no segundo mundo não interage causalmente com a decisão tomada (tal como o dispositivo em um Caso Estilo-Frankfurt). No entanto, se um deus fizesse com o que o tempo fosse mil vezes revertido até um instante logo anterior à decisão da agente, mantendo fixo o mesmo processo deliberativo, talvez víssemos uma distribuição em que a agente decide por si própria em 467 reversões e tem sua decisão causada pela máquina nas outras 533. O argumento de Peter van Inwagen apresentado na introdução deste capítulo, portanto, aplica-se a esse caso. Ainda assim, responderíamos que, se supusermos que não há nada de errado com o controle da agente em M1, essa conclusão deveria se estender também a M2 (FISCHER, 2012).

Dada essa resposta à suposição indeterminista do Dilema do Determinismo, temos um quadro geral suficiente da posição semicompatibilista. Na próxima subseção, apresentarei

aquelas que entendo como as mais importantes críticas recebidas por essa teoria, seguidas de um problema que se aplica potencialmente a todos os compatibilismos.

2.2.3 Algumas críticas

Nesta subseção, tratarei brevemente dos problemas da Individuação de Mecanismos, dos Subsistemas⁷⁶ e da Superveniência, concluindo-a com o Argumento dos Quatro Casos. As três primeiras objeções aplicam-se especificamente ao Semicompatibilismo, ainda que a segunda possa também ser dirigida a outras abordagens baseadas no mecanismo de ação. Já a quarta expressa uma preocupação que deve ser levada em conta por todos os compatibilistas – a de que mesmo o acréscimo de uma condição de assentimento de responsabilidade como a de Fischer e Ravizza seja insuficiente para excluir casos de manipulação do conjunto de casos que satisfazem as exigências estipuladas para ações livres. A razão para essa conclusão seria a de que o compatibilista não é capaz de estabelecer um *princípio* que justifique por um lado a existência de responsabilidade em um mundo determinista e, por outro, a inexistência de responsabilidade de uma agente manipulada.

Um detalhe relevante da teoria semicompatibilista é que, como os próprios autores admitem⁷⁷, não são oferecidos critérios para a *individuação* de mecanismos – isto é, para que saibamos que quando mantemos fixo o mecanismo em outros mundos possíveis, trata-se de fato do *mesmo* tipo de mecanismo operando nesses mundos. Fischer e Ravizza (1998) acreditam que isso pode ser feito intuitivamente. Mas suponha que Teresa é causalmente determinada a gastar o dinheiro encontrado na carteira do vizinho com um novo relógio esportivo, ao mesmo tempo em que reconhece um amplo espectro de razões para agir. Imagine agora, como fizemos há pouco, que seu mecanismo de ação pareça cumprir a condição de responsividade moderada a razões porque há um mundo possível no qual o passado foi um pouco diferente e o vizinho está passando por sérias dificuldades financeiras, correndo o risco até mesmo de perder seu apartamento, e que, sabendo disso, Teresa devolve a carteira com o dinheiro intacto. Poderíamos questionar, porém, se o mecanismo levando à ação nesse outro mundo é o mesmo do mundo atual. Talvez o fato de a agente se refrear em seu desejo de comprar o relógio, tendo adquirido maior “energia ou foco” frente a uma razão tão forte para agir de outro modo⁷⁸, indique que um mecanismo diferente passou a operar naquele momento (MCKENNA & PEREBOOM, 2016).

⁷⁶ O nome “Problema dos Subsistemas” é uma sugestão minha.

⁷⁷ Cf. Fischer e Ravizza, 1998, p. 40.

⁷⁸ Cf. Fischer e Ravizza, 1998, p. 74.

O ponto aqui é que não é claro, para que o teste de responsividade moderada a razões do mecanismo seja feito, se todas as características da “estrutura motivacional” da agente nessa situação devem ser mantidas fixas. Se sim, temos então dois mecanismos diferentes e Teresa não é, afinal, moralmente responsável por sua ação. Se não, resta saber por quê; e que critérios devem ser utilizados para que façamos a avaliação. Esse é o *Problema da Individuação de Mecanismos*. Em “Reasons Responsiveness, Agents and Mechanisms” (2013), McKenna afirma que pode haver uma enorme variação de intuições nesses contextos altamente carregados do ponto de vista teórico: as intuições de neurofisiologistas, cientistas cognitivos ou terapeutas cognitivos acerca do que é agir com base no *mesmo* processo serão diferentes daquelas da psicologia popular, que por sua vez diferirão das intuições de filósofos comprometidos com posições distintas no debate do Problema da Compatibilidade.

Além disso, McKenna pensa que o Semicompatibilismo enfrenta um outro problema ainda mais profundo, que desafia qualquer abordagem da responsabilidade baseada no mecanismo de ação:

O problema geral, como o vejo, é que qualquer sistema complexo terá “subsistemas” desenhados para funcionar precisamente por meio do desligamento ou da permissão para que outros sistemas se sobreponham em alguns contextos, mas não em outros. Assumindo que uma pessoa operando como um agente prático pode ser entendida como um sistema complexo, (...) qualquer tentativa de delimitar quais elementos agenciais são aqueles implicados “no” mecanismo de ação é obrigada a restringir a flexibilidade do agente no que diz respeito à responsividade a razões. (2013, p. 164)

Essa objeção, o *Problema dos Subsistemas*, foi originalmente levantada por Ginet (2006). Considere uma pessoa dirigindo um carro por um caminho muito familiar, de modo que seu mecanismo de ação é do tipo *hábito não-refletido*. Essa pessoa segue o caminho habitual até uma saída, mas percebe que ela está bloqueada. Suponha agora que a agente não está certa sobre que outra saída poderia ser usada e, naturalmente, começa a deliberar sobre o assunto. Aqui há uma mudança relevante de tipo de mecanismo: do hábito não-refletido, a agente passa para um mecanismo de raciocínio prático. Mas isso gera uma tensão para a teoria semicompatibilista: segundo Fischer e Ravizza, precisaríamos manter fixo o mecanismo de hábito não-refletido para a avaliação de sua responsividade a razões. Ao mesmo tempo, parece que é parte do que significa “ser responsivo a razões” a capacidade de “mudança de marcha” entre um mecanismo habitual e outro refletido, caso algo no contexto de ação nos chame a atenção ou seja diferente do usual. O Semicompatibilismo não seria capaz de acomodar casos como esse.

Para que a seriedade desse ponto se torne mais vívida, imagine que uma mulher negra está entre os candidatos de uma comissão que oferece bolsas de estudo. No mundo atual, essa mulher passa pela entrevista que é parte do processo seletivo e, sem que se deem conta disso, os membros da comissão a eliminam motivados por uma mistura de racismo e machismo, contrariando a competência demonstrada pela candidata. De maneira habitual e não-refletida, eles julgam que uma mulher de pele negra não poderia ser tão competente em sua área de estudos quanto um homem de pele branca. Suponha agora que a comissão está ciente da possível interferência de preconceitos em seus juízos, entende as razões morais para agir envolvidas nesses contextos etc., e uma circunstância na qual os membros efetivamente agiriam de outro modo, dando-se conta da própria tendência preconceituosa e aprovando a candidata, seria se essa mulher discorresse durante a entrevista sobre alguma situação de racismo e machismo pela qual ela já passara. Precisaríamos então admitir a mudança de mecanismos de ação desses agentes, do preconceito implícito e habitual para a correção consciente de preconceito, para atribuir responsabilidade moral aos membros da comissão por seu juízo preconceituoso no mundo atual. O semicompatibilista, porém, teria que dizer que esses agentes não são responsáveis pelo modo como agiram, já que manteríamos fixo o mecanismo de tipo hábito não-refletido (que, por definição, opera de maneira padronizada)⁷⁹.

A terceira objeção importante ao Semicompatibilismo é o *Problema da Superveniência*, proposto por Carolina Sartorio (2016). Considere nossa agente do Caso Estilo-Frankfurt apresentado na subseção anterior, Alice. Segundo Fischer e Ravizza, Alice é moralmente responsável se o mecanismo levando à sua ação é moderadamente responsivo a razões, independentemente da presença do dispositivo da Dra. Black. Mas imagine agora uma “Alice Insensível”, que é

(...) insensível à maioria dos tipos de razões para se refrear (...) (e vamos assumir que não por causa de alguma decisão livre anterior de se tornar

⁷⁹ Outros problemas importantes poderiam ser levantados a partir desse exemplo. Em primeiro lugar, o Semicompatibilismo tem a implicação de que, se uma agente fosse amplamente receptiva a razões morais em contextos de operação de preconceitos, mas não agiria de outro modo em nenhum mundo possível suficientemente próximo porque seus preconceitos são muito enraizados e estão “blindados” por crenças como “eu não sou preconceituosa”, então essa agente não seria moralmente responsável por uma ação motivada por preconceito. Esse parece o resultado inadequado: trata-se de um caso em que nós atribuiríamos responsabilidade à agente. Ao mesmo tempo, a retirada da condição de reatividade a razões, mantendo apenas alguma versão da condição de receptividade (similarmente à sugestão de Fricker (2007) de que um homem que deslegitimasse a palavra de uma mulher em virtude de seu gênero precisaria somente de que a consciência crítica de esquemas de gênero fosse difundida em sua sociedade para que ele fosse responsável pelo modo como agiu) parece-me *arbitrária*. Por que exigiríamos que uma agente fosse receptiva às razões x e y para agir de outro modo, mas a responsabilizaríamos mesmo que ela não tivesse controle para agir segundo x e y? Essa discussão, porém, não será aprofundada aqui.

insensível desse modo): (...) [a agente] é tal que, na maioria dos tipos de cenário contrafactual em que essas razões estão presentes, (...) não é receptiv[a] a essas razões ou não é suficientemente motivad[a] por elas, e então seu desejo de vingança ainda prevalece.” (p. 119-120).

Fischer e Ravizza diriam que essa agente não possui responsabilidade moral. Mas o ponto de Sartorio é o seguinte: o que garante que a diferença entre Alice e Alice Insensível não é *meramente contrafactual*, de modo que o mecanismo levando à ação no mundo atual em cada caso seja o mesmo? O fato de a caracterização semicompatibilista da responsividade a razões dar-se em termos modais parece entrar em conflito com a intuição derivada dos Casos Estilo-Frankfurt: a de que devemos olhar para a sequência de eventos levando à ação no mundo atual, não para possibilidades alternativas, ao avaliar a responsabilidade moral de uma agente. Devemos nos atentar àquilo que é explanatoriamente relevante para a ação – e, no caso de Alice, vimos que a presença do dispositivo não desempenha nenhum papel naquilo que realmente ocorre. Dessa lição, Sartorio extrai um princípio de *Superveniência*: “a liberdade de uma agente em relação a X é superveniente à (...) sequência causal levando a X” (p. 122). Disso segue-se que, para haver diferença na responsabilidade de dois agentes por suas ações, deve haver diferença na sequência causal levando a essas ações. O Semicompatibilismo, porém, não é capaz de garantir que a diferença contrafactual entre Alice e Alice Insensível seja operacional nos mecanismos de ação de cada agente, de modo a justificar uma avaliação distinta de existência de responsabilidade em cada caso.

Por fim, vejamos uma versão dos argumentos de manipulação⁸⁰ contra o compatibilismo: o *Argumento dos Quatro Casos*. Argumentos de manipulação começam supondo que uma certa agente satisfaz a “estrutura de agência compatibilista” (EAC) da teoria que pretendem objetar – por exemplo, a posse de um mecanismo moderadamente responsivo a razões, ou a presença de desejos de segunda ordem coincidentes com os desejos que levaram à ação (MCKENNA & PEREBOOM, 2016). Eles procedem então segundo a forma geral a seguir⁸¹:

Argumento da Manipulação

Intuição de Não-responsabilidade: “Se uma agente, S, é manipulada da maneira X para realizar a ação A a partir de uma EAC, então S não realiza A livremente e, portanto, não é moralmente responsável por A”.

⁸⁰ Esse tipo de argumento é considerado uma das principais motivações para a posição incompatibilista entre determinismo e liberdade da fonte genuína. Cf. McKenna e Pereboom, 2016, capítulo 7.

⁸¹ Os nomes das premissas são sugestões minhas.

Suposição de Semelhança: “Qualquer agente manipulado da maneira X para fazer A é semelhante a qualquer agente normal causalmente determinado a fazer A a partir de uma EAC”.

Conclusão: “Logo, qualquer agente normal causalmente determinado a fazer A a partir de uma EAC não realiza A livremente e, portanto, não é moralmente responsável por A”. (MCKENNA & PEREBOOM, p. 163)

Vejam agora a formulação específica dos quatro casos, de Pereboom (2001, 2014). Suponha que o agente, Professor Plum, satisfaz todas as condições compatibilistas para liberdade apresentadas nesse capítulo ao realizar sua ação: assassinar a senhora White. Ele inclusive adquiriu as crenças necessárias para o assentimento de responsabilidade, como estipuladas por Fischer e Ravizza. Considere agora cada uma das variações de sua história, tais como resumidas em McKenna e Pereboom (2016):

Caso 1: Professor Plum foi criado por um time de neurocientistas que podem manipulá-lo diretamente por meio de uma tecnologia como a de rádio, mas esse agente é o mais parecido possível com um ser humano ordinário, dada sua história. Os cientistas o manipulam “localmente” para que ele inicie um processo de deliberação e diretamente produzem seus estados momento a momento, o que leva ao assassinato de White por razões egoístas. (...)

Caso 2: Plum é como um ser humano ordinário, exceto pelo fato de que ele foi criado por um time de neurocientistas que, embora não possam controlá-lo diretamente, programaram-no para pesar razões para agir de modo a ser frequentemente (mas não exclusivamente) egoísta, com a consequência de que naquelas circunstâncias ele era causalmente determinado a iniciar o processo de deliberação que resulta no assassinato de White. (...)

Caso 3: Plum é como um ser humano ordinário, exceto pelo fato de que ele foi determinado por rigorosas práticas de treinamento em sua casa e comunidade, de modo a ser frequentemente (mas não exclusivamente) egoísta em suas razões (tão egoísta quanto nos casos 1 e 2). Seu treinamento foi iniciado cedo demais para que ele tivesse a capacidade de prevenir ou alterar as práticas que determinaram seu caráter. Naquelas circunstâncias, Plum é então causalmente determinado a iniciar o ... processo ... que resulta no assassinato de White. (...)

Caso 4: O determinismo (...) é verdadeiro, e Plum é um ser humano ordinário, que foi gerado e cresceu sob circunstâncias normais, que é frequentemente (mas não exclusivamente) egoísta em suas razões (tão egoísta quanto nos casos 1-3). O assassinato de White por Plum ocorre como resultado de o agente iniciar o processo relevante. (p. 165-166)

Nesses casos, a determinação causal daquilo que o agente realiza assume quatro formas distintas. No primeiro, temos a intuição de que Plum não é moralmente responsável pelo assassinato de White. A estratégia de Pereboom é então sugerir que, a cada mudança gradual entre o contexto de manipulação radical e o contexto em que o determinismo é

verdadeiro, não há uma diferença relevante em relação ao caso anterior. Assim, transferiríamos nosso juízo de que Plum não é responsável de 1 para 2, de 2 para 3 e, finalmente, de 3 para 4 – concluindo então que as condições da EAC não são suficientes para que uma agente tenha a liberdade de ser a fonte genuína da própria ação.

A rigor, apenas o Caso 1 já poderia ser um contraexemplo para o compatibilismo. Pereboom (2014) afirma que se o compatibilista resistir à conclusão de que Plum, no Caso 1, de fato satisfaça a EAC por ele defendida em virtude da manipulação momento a momento, então poderíamos facilmente modificar o exemplo para que os neurocientistas, embora tenham essa capacidade, exercitem-na apenas moderadamente. Esse ponto tem sido bastante discutido e as respostas ao Argumento dos Quatro Casos (e a argumentos de manipulação em geral) costumam dividir-se entre ataques à Intuição de Não-responsabilidade e ataques à Suposição de Semelhança⁸². Aqui, porém, não entrarei em detalhes sobre essas respostas. Na próxima seção, passarei à apresentação de uma das principais vozes do *libertismo* contemporâneo no Problema da Compatibilidade.

2.3 UMA TEORIA LIBERTISTA CONTEMPORÂNEA

Em um artigo de 1987, Watson expressou a seguinte percepção:

(...) através da poeira de três séculos de debate, nós podemos ver claramente ao menos uma coisa: se nenhuma quantidade ou tipo de capacidade e complexidade cognitiva ou volitiva que poderia existir em um mundo determinista será suficiente para a agência livre, então a simples adição da exigência do indeterminismo não será suficiente idem. ([1987] 2004, p. 196)

A posição acima, de que o indeterminismo não aumenta controle, pretende objetar o *libertismo*. Como vimos, o libertismo é a tese incompatibilista acerca do Problema da Compatibilidade que afirma que a presença de indeterminismo é um ingrediente fundamental da natureza da liberdade de agentes humanos. Como observa Franklin (2017), uma defesa dessa tese não será possível “simplesmente reiterando [nosso] argumento favorito pelo incompatibilismo” (p. 155)⁸³. Na introdução deste capítulo, apontei que o libertista está sujeito a uma objeção de sorte: se uma agente possui razões concorrentes acerca do que fazer e finalmente decide fazer A em vez de B, o fato de o indeterminismo estar presente nessa

⁸² Respectivamente, as estratégias chamadas de *hardline* e *softline*. Algumas dessas respostas estão em Fischer, 2007; McKenna, 2014; e Nelkin, 2011. Cf. também Mele, 2006; e Sartorio, 2016.

⁸³ Alguns filósofos, como Pereboom (2014) e Caruso (2020), são incompatibilistas sem serem libertistas: eles sustentam o ceticismo sobre a liberdade e a responsabilidade moral associada a merecimento básico de responsabilização. Essa não é, porém, a única via para o ceticismo. Levy (2011), por exemplo, descreve a si mesmo como um compatibilista a respeito do determinismo, mas não acredita que haja qualquer teoria compatibilista da liberdade viável devido a problemas de sorte. Essas posições não serão exploradas neste trabalho.

escolha nos faria concluir que não há uma explicação contrastiva da decisão por A – assim, a decisão seria uma questão de acaso, ou de sorte. O ponto agora é o seguinte: ainda que o indeterminismo não diminua controle, dadas essas considerações, parece que ele não faz nada para *aumentá-lo* em comparação ao controle descrito pelo compatibilista. Esses são, respectivamente, o *Problema da Sorte* e o *Problema do Controle Não-aumentado*.

A existência desses problemas faz com que o libertista precise engajar-se em dois projetos em torno da exigência do indeterminismo (em adição à defesa do incompatibilismo): (i) explicar o papel do indeterminismo na natureza da liberdade, especificando *onde* ele precisa ocorrer, na sequência causal levando à ação, para que esse papel seja cumprido; e (ii) mostrar que a teoria resultante dessa explicação é capaz de resolver tanto o Problema da Sorte quanto o Problema do Controle Não-aumentado (FRANKLIN, 2017). É importante notar que as respostas a essas objeções são dependentes uma da outra em uma das direções: sem um argumento segundo o qual o indeterminismo não ameaça controle, não é possível argumentar que ele o aumente em ações intencionais que cumpram condições compatibilistas para a liberdade.

O libertismo contemporâneo divide-se em três grandes tipos: não-causais, da causalidade de eventos, e da causalidade do agente (CLARKE, CAPES & SWENSON, 2021). A disputa mais proeminente, porém, se dá entre os libertistas da causalidade de eventos e da causalidade do agente⁸⁴. Os teóricos da causalidade do agente sustentam que eventos do tipo “decisões livres” precisam ser causados por *agentes enquanto substâncias* – isto é, não por meio de eventos tais como o agente reconhecendo certas razões para agir, ou outros estados mentais (como sustentaria o teórico da causalidade de eventos) (MCKENNA & PEREBOOM, 2016; CLARKE, CAPES & SWENSON, 2021). Além disso, “não pode ser o caso que o que o agente causa ou sua causação daquele evento sejam causalmente determinados por eventos anteriores” (CLARKE, CAPES & SWENSON, 2021).

McKenna e Pereboom (2016) apontam que o libertismo da causalidade de eventos é geralmente considerado mais promissor por filósofos atraídos por uma “visão de mundo científica” (p. 233). Eles acreditam inclusive que as abordagens da causalidade do agente mais coerentes são aquelas que rejeitam o projeto de reconciliação com as melhores teorias físicas. No entanto, os autores sustentam que uma teoria da causalidade do agente está mais bem equipada para responder ao Problema da Sorte. Não entrarei em mais detalhes sobre o

⁸⁴ Não discutirei o libertismo não-causal aqui. Para um representante dessa posição, cf. Ginet, 1990.

libertismo resultante desse tipo de teoria aqui⁸⁵. Desejo apenas enfatizar dois pontos que justificam essa opção. Em primeiro lugar, se os autores têm razão quanto a essa abordagem precisar rejeitar a reconciliação com a ciência, esse é um custo teórico muito grande – eles próprios concordariam com isso. O segundo ponto é o seguinte: há pelo menos uma razão aparentemente forte (que explicarei mais adiante) para sustentar que a falta de explicação contrastiva em termos de elementos agenciais de uma decisão em um mundo indeterminista ocorre também em mundos deterministas, como argumentou Mirja Pérez de Calleja (2014). Isso reforçaria, afinal, a afirmação de que o Problema da Sorte também é uma dificuldade para o compatibilista, se o for para o libertista (e agora não somente em mundos indeterministas)⁸⁶. Por essas razões, apresentarei nesta seção os argumentos do influentíssimo libertismo da causalidade de eventos de Robert Kane, seguidos de algumas das principais críticas recebidas pela teoria. Por fim, passarei ao argumento de Pérez de Calleja sobre o Problema da Sorte e a breves observações sobre que consequências acredito que ele traga para o debate entre compatibilistas e libertistas.

2.3.1 Kane e o controle voluntário plural

Segundo Franklin (2017), Kane é o nome mais associado ao libertismo da causalidade de eventos, um tipo de teoria que abraça a tese do *reducionismo agencial*: “o papel causal do agente em todas as atividades agenciais (da deliberação à escolha à ação corporal) é redutível às atividades causais de estados mentais e eventos envolvendo o agente” (p. 2). O ponto aqui não é que não seja a agente a causar sua ação; o reducionismo responde à pergunta sobre o que é, para uma agente, causar sua ação⁸⁷. Para compreender a maneira como Kane defende o incompatibilismo e desenvolve os dois projetos para o libertista mencionados há pouco (a respeito do papel e local do indeterminismo e de como o Problema da Sorte e o Problema do Controle Não-aumentado devem ser respondidos), nos termos de uma teoria da causalidade de eventos, vejamos então uma interpretação das condições que ele estabelece como necessárias e suficientes para uma ação livre:

⁸⁵ O trabalho de Timothy O’Connor (2009) constitui um dos melhores representantes dessa abordagem. Para uma argumentação segundo a qual temos evidência empírica para pensar que a agência humana provavelmente não atende às exigências de uma teoria como a de O’Connor, cf. Pereboom, 2014.

⁸⁶ Nesse momento, o teórico da causalidade do agente poderia alegar que sua abordagem é então uma solução para o problema de sorte que ameaça *qualquer* descrição da liberdade em termos de causalidade de eventos, seja ela compatibilista ou incompatibilista. Deixarei, no entanto, a avaliação dessa possibilidade (incluindo de se esse tipo de abordagem é de fato irreconciliável com as ciências) para trabalhos futuros.

⁸⁷ Para uma discussão sobre esse ponto, cf. Franklin, 2017, capítulo 7.

A decisão X de uma agente S no instante t_2 foi diretamente livre, e é uma decisão pela qual a agente possui responsabilidade moral direta, se somente se: (i) no intervalo $t_1 - t_2$, S é normativamente competente em relação a X; (ii) em t_1 , S fez um esforço da vontade para X e fez também um esforço da vontade para Y; (iii) as razões de S que favoreciam X causaram (...) o esforço de S para X em t_1 e as razões de S que favoreciam Y causaram (...) o esforço de S para Y em t_1 ; (iv) cada esforço da vontade feito por S satisfaz exigências compatibilistas plausíveis; (v) o esforço da vontade de S para X (...) causou X em t_2 ; e (vi) era possível, dadas as leis da natureza e o passado até t_2 , que o esforço da vontade de S para Y tivesse causado Y em t_2 (caso em que seu esforço da vontade para X não teria causado X). (FRANKLIN, 2017, p. 100)

Começamos pelo ponto de uma ação ser diretamente livre. Kane (2016) destaca três tipos de ações em certo sentido “livres”: 1. aquelas realizadas “voluntariamente, de propósito e por razões e que não são coagidas, compelidas ou sujeitas ao controle de outros agentes”; 2. aquelas que cumprem as mesmas condições do tipo 1, mas que são também “feitas de nossa ‘livre vontade’ no sentido de uma vontade que nós somos em última instância responsáveis (...) por formar”; e por fim as 3. ações “*autoformadoras* (...) ou ações “formadoras da vontade” pelas quais nós formamos e reformamos a vontade a partir da qual agimos no sentido em 2” (p. 3, grifo meu). Ações do tipo 1 cumpririam a estrutura agencial compatibilista e poderiam, é claro, existir em mundos deterministas. Já as ações do tipo 2, embora possam ser determinadas, não podem existir em mundos deterministas – se determinadas, elas o são a partir das ações do tipo 3, sendo assim *livres de maneira derivada*. As ações do tipo 3 são por sua vez o foco da teoria. Segundo Kane, ações autoformadoras são livres em sentido básico porque, além de satisfazerem todas as características do tipo 1 (e algumas características adicionais que veremos a seguir), elas *exigem indeterminismo*.

Kane (2007) sustenta que, ao debater o Problema da Compatibilidade, não se deve enfocar apenas o ponto das possibilidades alternativas de ação (como fazem o Argumento da Consequência e as discussões sobre o sentido relevante da capacidade de agir diferentemente). Ele pensa que essa é uma base muito frágil para a posição incompatibilista e reconhece que o que está em jogo é a liberdade de uma agente ser a fonte genuína da própria ação – possuindo, assim, *responsabilidade última* por essa ação. Kane formula então um argumento segundo o qual a responsabilidade última exige indeterminismo. O argumento começa por assumir que

A ideia básica da RU [responsabilidade última] é a seguinte: Para ser ultimamente responsável por uma ação, um agente precisa ser responsável por qualquer coisa que é causa ou motivo suficiente da ocorrência da ação. Se, por exemplo, uma decisão é provocada e pode ser suficientemente explicada pelo caráter e pelos motivos do agente (...), então, para ser ultimamente responsável, o agente precisa ao menos ser parcialmente responsável por possuir o caráter e os motivos que possui, em virtude de escolhas e ações voluntariamente realizadas no passado. (p. 14)

Segundo essa suposição, precisamos ter formado nosso “arbítrio” ou “vontade” por ações voluntárias passadas pelas quais fomos responsáveis para que haja *livre*-arbítrio. Se havia *causas suficientes* para a realização de certa ação passada, precisamos ter sido responsáveis por essas causas por meio de uma ação ainda anterior. Surge então um potencial regresso infinito, que só pode ser parado pelo *indeterminismo* – afinal, o indeterminismo é a única circunstância em que não há causas suficientes para uma ação. Mas o mero indeterminismo não é suficiente. Se havia *motivos suficientes* para a tomada de uma decisão passada, então precisamos ter sido responsáveis por esses motivos em virtude de uma decisão ainda anterior. Surge aqui um outro potencial regresso infinito, que só pode ser parado pelas ações autoformadoras – ações em que o caráter, valores e estruturas motivacionais da agente, que não estavam ainda plenamente formados, são moldados por ela própria a partir de uma *seleção voluntária, intencional e racional*⁸⁸ entre possibilidades alternativas. O incompatibilismo seria, portanto, estabelecido pelo “regresso dual do livre-arbítrio” (p 22): a exigência de que sejamos ultimamente responsáveis por causas das nossas ações leva à exigência de que algumas ações em nossas histórias de vida sejam indeterminadas, enquanto a exigência de que sejamos ultimamente responsáveis por motivos das nossas ações leva à exigência de ações autoformadoras, que por sua vez implicam o acesso a possibilidades alternativas. A conexão entre responsabilidade última e possibilidades alternativas seria, portanto, indireta.

Como descrevem McKenna e Pereboom (2016), o tipo paradigmático de ação autoformadora para Kane é o de “conflito moral ou prudencial, em que há razões a favor e contrárias à ação em questão” (p. 235; cf. KANE, 1996). A produção desse tipo de ação é iniciada com o caráter atual e os motivos da agente e passa por um *duplo esforço da vontade*. Os esforços da vontade são “lutas” internas para que decidamos de um determinado modo, apesar das pressões, tendências ou tentações contrárias a essa decisão. Nas ações livres, esses esforços são indeterminados, “e, conseqüentemente, a escolha produzida pelo esforço será *indeterminada*” (MCKENNA & PEREBOOM, 2016, p. 235). Vejamos como o próprio Kane exemplifica esses conceitos:

Considere uma empresária que enfrenta um conflito como esse. Ela está a caminho de uma reunião importante, quando observa uma agressão acontecendo em um beco. Uma luta interior começa então a ocorrer entre sua consciência da necessidade de parar e pedir ajuda e suas ambições de carreira, que dizem a ela que a reunião não pode ser perdida. Ela precisa fazer um esforço da vontade para superar a tentação de seguir seu caminho. Se vencer essa tentação, isso será o resultado de seu esforço, mas se ela

⁸⁸ Isto é, tendo boas razões para agir e agindo segundo essas razões (KANE, 2007).

falhar, será porque ela não permitiu que esse esforço prevalecesse. Isso se deve ao fato de que, ao mesmo tempo em que desejava vencer a tentação, ela também desejava falhar em vencê-la, por razões diferentes e incomensuráveis. Quando nós, como a mulher, decidimos nessas circunstâncias, e os esforços indeterminados que fazemos tornam-se escolhas determinadas, nós *fazemos* com que um conjunto de razões concorrentes ou motivos prevaleçam sobre outros *por meio da decisão*. (2007, p. 26)

Chamemos essa empresária de Beatriz. Imagine que Beatriz (intencionalmente e por suas razões) decide parar e pedir ajuda. Suponha também que se trata de uma agente normativamente competente e que o mundo é indeterminista. Nesse caso, Kane (2007, 2016) argumenta que ao fazer um duplo esforço da vontade – digamos, tentar vencer a vontade de continuar andando e tentar ignorar sua consciência moral – causado por suas razões concorrentes, a agente desempenha papel causal na tomada de decisão e constitui a si mesma como uma pessoa mais propensa a renunciar a interesses egoístas e ajudar outras pessoas quando necessário, formando assim o próprio caráter. Não há um momento em que “o esforço é interrompido e o acaso “assume” o processo” (2007, p. 32), mesmo que o resultado do conflito entre esforços da vontade seja indeterminado – isto é, mesmo que em um mundo possível com um passado *idêntico* ao atual Beatriz tenha tomado a decisão alternativa de seguir seu caminho para a reunião.

Esse panorama nos dá boas pistas de como Kane lidará com o Problema da Sorte. Para afastar a objeção de que a decisão de parar e pedir ajuda em vez de continuar andando até o local da reunião foi uma questão de sorte, o autor enfatiza (2016) que essa decisão foi o resultado de um esforço da vontade que emergiu dos motivos da agente, com a capacidade e a oportunidade tanto de fazer quanto de não fazer com que ele fosse bem-sucedido, cumprindo condições compatibilistas razoáveis (o esforço era voluntário, intencional e racional). Crucialmente, a agente *endossa* a decisão tomada – e esse também seria o caso no mundo possível em que ela decidisse de outro modo. Em ambas as circunstâncias, a reação da agente não é a de que “houve um “erro” ou “acidente” (...) [em vez disso] a agente se identifica com a escolha como algo que ela (...) causou propositalmente e segundo a qual ela está agora pronta para agir” (p. 11). Esses pontos fornecem também a resposta de Kane ao Problema do Controle Não-aumentado: em uma ação autoformadora, a agente possui o que ele chama de *controle voluntário plural*: “capacidade de efetivar *qualquer* alternativa [entre ao menos duas] que o agente deseje ou mais deseje, quando desejar, pelas razões que deseja, sem ser coagido ou compelido” (1999, p. 237-238). A posse de controle voluntário plural significaria a posse de mais controle do que aquele possuído por uma agente que satisfaz condições compatibilistas em um mundo determinista.

No restante desta subseção, apresentarei brevemente algumas das principais críticas recebidas por Kane, que dirão respeito ao Problema da Sorte e ao Problema do Controle Não-aumentado. Considere em primeiro lugar as razões do autor para afirmar que a decisão da empresária não é questão de sorte. Além de considerações que poderiam ser evocadas por um compatibilista para sustentar que ações livres podem existir em mundos deterministas (PEREBOOM, 2001), ele acrescenta que essa agente possui a capacidade e oportunidade tanto de fazer quanto de não fazer com que seu esforço da vontade fosse bem-sucedido, constituindo assim o controle voluntário plural. Mas Pereboom (2001) é cético a respeito “do sentido em que a empresária possui o controle necessário para causar qualquer uma de suas duas opções para escolha” (p. 52). Ele argumenta que um cenário com um “manipulador randômico”⁸⁹ é capaz de nos fazer duvidar da existência de controle voluntário plural.

A *Objeção do Manipulador Randômico*⁹⁰ é construída da seguinte maneira. Imagine que a contraparte da empresária em outro mundo possível, a empresária*, possui capacidades, razões, esforços etc. iguais aos da primeira agente, além de também tomar a decisão de ajudar a vítima de agressão. Mas suponha que em seu cérebro existe um mecanismo que gira um ponteiro em um mostrador com duas regiões. Esse mecanismo é genuinamente indeterminista e o fato de o ponteiro parar em uma das regiões “é o componente indeterminista crucial da realização neurofisiológica da decisão [de ajudar a vítima]” (PEREBOOM, 2001, p. 53), substituindo o processo neurofisiológico normal da agente. Se o ponteiro cai na primeira região, os esforços da vontade da empresária* resultarão na escolha de seguir o caminho para a reunião; e, se o ponteiro cai na segunda região, seus esforços da vontade resultarão na decisão de parar para pedir ajuda para a vítima – como é o caso no cenário. Suponha que, para a empresária*, a experiência de primeira pessoa do conflito entre suas razões e esforços da vontade, resultando em uma decisão que ela endossa e identifica como sua, são idênticas às da primeira agente. A partir desse cenário, Pereboom (2001) aponta então que não parece que a empresária* seja a fonte genuína da própria ação, dada a atividade do manipulador randômico. Esse argumento é uma formulação específica do Problema da Sorte, em que se pode agora exigir que Kane estabeleça um princípio que diferencie a empresária original de sua contraparte. Pereboom pensa que isso não pode ser feito.

No texto “Hard Incompatibilism” (2007), Pereboom acrescenta ainda uma outra objeção a Kane, que chamarei de *Objeção do Continuum*. Ele afirma que a suposição de que

⁸⁹ Uma versão anterior desse ponto consta em van Inwagen, 1983.

⁹⁰ Os nomes dessa e das próximas duas objeções a Kane mencionadas nesta subseção são sugestões minhas.

“agentes ultimamente responsáveis precisam não apenas ser a fonte das próprias ações, mas também do *arbitrio* que deseja realizá-las” (KANE, 1996, p. 73) gera o seguinte desafio. Imagine eventos que são “verdadeiramente aleatórios”, em que isso significa que eles não são causados por nada. Se um desses eventos é uma decisão, então, segundo a suposição acima, a agente não pode ser responsável por ela porque não pode ser a fonte do *arbitrio* que a realiza. Entre decisões desse tipo e decisões causalmente determinadas por fatores que estão fora do controle da agente (e pelas quais, segundo o libertismo, ela também não será responsável),

(...) existe um leque de decisões para as quais fatores fora do controle da agente contribuem sem determiná-las causalmente, ao mesmo tempo em que não há nada que suplementa a contribuição causal desses fatores para produzi-la. (...) No entanto, dada a intuição do incompatibilista da fonte [a suposição de Kane acima], agentes não podem ser moralmente responsáveis por decisões se elas atendem a essas especificações. Se [é o caso que atendem] (...), então sua produção possui apenas uma combinação dos dois tipos de fatores que ameaçam a responsabilidade moral. (...) [S]e esses fatores, em vez de determinarem uma única decisão, simplesmente deixam em aberto mais de uma possibilidade, e a agente não desempenha nenhum papel *adicional* em determinar qual possibilidade será realizada, então não temos mais razão para acreditar que ela é moralmente responsável do que temos no caso determinista. (p. 108, grifo meu)

Pereboom aponta que a própria intuição incompatibilista a respeito da liberdade da fonte genuína implica que agentes não podem ser responsáveis por decisões que estejam em qualquer lugar desse *continuum*. Se o que contribui causalmente para uma decisão sem determiná-la são o caráter e os motivos da agente, que dão origem a seus esforços da vontade concorrentes, Kane poderia dizer que esse caráter e motivos não precisam estar fora do controle dessa agente. No entanto, dado que ela precisa ser ultimamente responsável por eles, haverá momentos no passado em que a agente fez as *primeiras* escolhas relevantes para formá-los – momentos em que os fatores que as influenciaram causalmente, moldando os esforços da vontade envolvidos, serão fatores fora de seu controle. Se a agente não era responsável por esses fatores, ela não pode ter sido responsável pelos esforços da vontade derivados e, conseqüentemente, não pode ter sido responsável pela decisão resultante desses esforços. O argumento termina apontando que a agente não pode então ser responsável pelas escolhas subsequentes, dado que não há responsabilidade pelas escolhas autoformadoras.

Vejam os por fim um dos contra-argumentos a Kane propostos por Franklin (2017), que é também um libertista da causalidade de eventos. Chamarei a crítica a seguir de *Objeção da Compensação*. Franklin sustenta que não é claro, *tudo considerado*, que o papel dos esforços da vontade nessa teoria aumenta o controle da agente em relação às abordagens compatibilistas – uma objeção que permanece ainda que haja boas respostas para os pontos de

Pereboom acima. Primeiramente, observe que há um sentido em que o indeterminismo de fato diminui controle, segundo o próprio Kane: no caso da empresária, “o indeterminismo está funcionando como um *obstáculo* (...) à realização de um de seus propósitos – (...) na forma de resistência em sua própria vontade que precisa ser superada pelo esforço” (2007, p. 178, grifo meu). Ele acrescenta que, não houvesse esse obstáculo, então a agente teria, em certo sentido, “controle completo” de *uma* de suas alternativas de escolha (isto é, ela poderia garantir que essa alternativa fosse realizada). No entanto, Kane argumenta que “ao ser um obstáculo para a realização de algum de nossos propósitos, o indeterminismo paradoxalmente abre a possibilidade genuína de que persigamos outros propósitos” (p. 178). O indeterminismo que impede o controle direto tem origem nos desejos e no esforço da agente para agir de outro modo e, ao superar esse obstáculo, ela pode tornar-se genuinamente autocriada e livre.

A configuração proposta por Kane para o controle voluntário plural é uma resposta ao Problema do Controle Não-aumentado. Franklin (2017) desconfia, porém, que a localização do indeterminismo nessa ação prévia à decisão – o esforço da vontade – seja justamente aquilo que impede a teoria de resolver esse problema. Ele aponta que os compatibilistas podem conceder que não há controle voluntário plural em um mundo determinista, mas que sua abordagem do controle, ao garantir o “controle completo” de uma alternativa de ação, pode “compensar” o duplo obstáculo à realização dos propósitos da agente colocado pelos esforços da vontade concorrentes. Se aceitarmos esse ponto e prestarmos atenção aos outros elementos que compõem o controle voluntário plural, veremos que eles são o tipo de coisa que é evocada pelo compatibilista para caracterizar a liberdade – a decisão emerge dos motivos da agente, por meio de um esforço voluntário, intencional e racional etc.⁹¹.

Muitas críticas promissoras ao libertismo da causalidade de eventos de Kane não serão mencionadas aqui⁹². Acredito, porém, que as objeções do Manipulador Randômico, do *Continuum* e da Compensação são representativas de parte dos pontos mais sensíveis dessa teoria. Para finalizar este capítulo, sugeri que o argumento de Pérez de Calleja a ser apresentado na próxima subseção nos auxilia a mapear quais dessas objeções podem ser coerentemente empregadas pelo compatibilista, em adição à sua abordagem positiva da natureza da liberdade necessária para a responsabilidade moral. O resultado será o seguinte: enquanto a Objeção do *Continuum* poderá ser utilizada, já que emprega uma intuição

⁹¹ Cf. também Pereboom, 2001.

⁹² Assim como tentativas de tornar o libertismo da causalidade de eventos mais simples e resistente às críticas. Para apenas duas dessas tentativas, cf. Balaguer (2010) e Franklin (2017).

incompatibilista como pressuposto, as objeções do Manipulador Randômico e da Compensação não poderão ser utilizadas.

2.3.2 Sorte entre mundos e aumento de controle

Em um artigo de 2014, Pérez de Calleja argumenta que o que Alfred Mele (2006) chamou de *sorte presente* (a sorte que ocorre próximo ao momento da decisão/ação, como pode ser o caso⁹³ em relação a agentes que cumprem condições libertistas de controle) é um problema também para os compatibilistas. O ponto do Problema da Sorte pode ser colocado em termos de uma diferença entre mundos possíveis: há uma diferença na decisão tomada em diferentes mundos possíveis e essa diferença não possui uma explicação em termos de agência, isto é, não há nada sobre as capacidades, estados mentais, caráter etc. do agente que explique a diferença no resultado da deliberação (MELE, 2006). No entanto, como aponta Pérez de Calleja, isso é verdadeiro para *qualquer* decisão em que a agente esteja dividida entre alternativas de ação que pensa possuir, com razões concorrentes para cada uma dessas ações. E esse tipo de decisão, é claro, pode ser tomada em um mundo determinista.

Considere mais uma vez o exemplo da empresária Beatriz. Imagine que sua contraparte, Beatriz*, vive em mundo determinista. O processo deliberativo, as razões e disposições e a decisão final dessas duas agentes são as mesmas; e nem a deliberação nem a decisão de Beatriz* de parar para pedir ajuda para a vítima da agressão refletem um “compromisso inequívoco” de fazer a coisa certa em vez de agir motivada por ambição, ou uma “incapacidade psicológica” de resistir ao ímpeto de fazer a coisa certa. Pérez de Calleja argumenta então que “parece claro que o fato de [a agente] ser determinada (...) *não* está fundamentado nas atitudes, tendências, capacidades e características da circunstância relevantes para aquilo que [ela] está motivada e disposta a fazer naquela circunstância” (2014, p. 115).

Esses fatores não podem ser aquilo que determina causalmente que Beatriz* decida como decidiu, já que eles são os mesmos de Beatriz. Pérez de Calleja acrescenta ainda uma outra razão para essa conclusão: é possível conceber uma contraparte de Beatriz*, Beatriz**, que vive também em um mundo determinista e, a partir do mesmo processo deliberativo,

⁹³ O uso da expressão “pode ser o caso” aqui merece algumas considerações. Essa expressão faz sentido se, ao dizer que há sorte em uma decisão, supusermos que a presença de sorte implica ausência de controle. Mas há um outro sentido, mais fraco, em que claramente *há* sorte em casos como o da empresária: embora Kane argumente que ela possui controle voluntário plural, é fato que não há uma explicação contrastiva em termos agenciais da decisão tomada no mundo atual. Talvez a sorte e a presença de controle sejam compatíveis, se satisfeitas certas condições (cf. Hartman, 2017).

razões e disposições, toma a decisão de seguir seu caminho para a reunião. É claro que “o mesmo conjunto de fatores causais não pode determinar causalmente resultados mutuamente excludentes” (2014, p. 115). Para dissociar tanto a existência de sorte entre mundos do mero indeterminismo quanto o controle do determinismo, a autora compara esse grupo de casos com o de alguém que joga golfe pela primeira vez. Suponha que Beatriz* por sorte acerta em um buraco de dificuldade alta. Aqui, sua pouca habilidade em jogar golfe não é o que determina que ela acerte o buraco, mesmo que esse acerto seja causalmente determinado. Diferentemente de um jogador muito habilidoso, o fato de o acerto ser determinado não faz com que ela tenha a capacidade de garantir (ou quase garantir) o acerto em um buraco daquele nível de dificuldade.

Pérez de Calleja afirma que há, em casos como os das agentes de mundos deterministas acima, uma explicação de por que razão elas tomaram uma decisão em vez de outra – afinal, as leis da natureza e todo o passado até aquele momento eram compatíveis apenas com uma decisão. A presença dessa explicação, porém, não faz “nenhum bem” a essas agentes “no que diz respeito à liberdade ou ao controle racional que [elas] possuem (...), já que essas explicações contrastivas não apresentam causas que distintivamente racionalizem, ou mesmo causem de forma saliente, a decisão tomada em vez de uma outra” (2014, p. 116).

Depois da exposição desse argumento, a autora observa que há uma clara tensão para o compatibilista que tente alegar que há mais sorte em mundos indeterministas porque, afinal, agentes nesses mundos são capazes de decidir de outro modo. Além disso, compatibilistas contemporâneos (como Fischer e Ravizza) sustentam que o fato de o mundo ser determinista ou indeterminista é irrelevante para a liberdade. Dado que decisões sujeitas à sorte entre mundos são muito comuns e muitas vezes relevantes do ponto de vista moral, o compatibilista que pretenda evocar o Problema da Sorte contra o libertista teria que admitir que nenhuma delas é livre em sua própria abordagem se o mundo for indeterminista, contradizendo a posição de que o modo como são as leis da natureza não faz diferença para a responsabilidade moral dos agentes. (PÉREZ DE CALLEJA, 2014).

Neste trabalho, não explorarei a questão de se a sorte de fato elimina controle, como é muitas vezes assumido no debate do Problema da Compatibilidade. De qualquer forma, o que penso que se pode concluir da discussão que abordei acima para as teorias apresentadas neste capítulo é o seguinte: se o compatibilista tiver uma boa resposta para o Argumento da Manipulação⁹⁴, sua oposição ao libertismo como proposta do mais forte sentido de liberdade necessário para a responsabilidade moral será suplementada não por um endosso do Problema

⁹⁴ E para o Argumento Direto, que não foi discutido neste trabalho.

da Sorte, mas pela estratégia de mostrar que o indeterminismo não aumenta a liberdade, ou que, se aumenta, essa “liberdade adicional” *não* é a liberdade em jogo para a responsabilidade moral. Considerando a observação de Pérez de Calleja de que o determinismo não faz “nenhum bem” ao controle dos agentes, Kane não precisa conceder, como concede, que há um sentido em que o indeterminismo diminui controle por funcionar como um obstáculo à realização de um propósito. Consequentemente, a sugestão de Franklin de que, tudo considerado, o libertismo de Kane não fornece mais controle à agente (a Objeção da Compensação) não poderá ser utilizada pelo compatibilista para endossar o Problema do Controle Não-aumentado. A Objeção do Manipulador Randômico também não poderá ser coerentemente empregada, já que se trata de uma formulação específica do Problema da Sorte.⁹⁵

Por outro lado, a Objeção do *Continuum* (segundo a qual a intuição incompatibilista sobre a liberdade da fonte genuína implica que agentes não podem ser responsáveis por decisões que estejam em qualquer lugar do *continuum* que vai de decisões verdadeiramente aleatórias a decisões causalmente determinadas) poderá ser coerentemente empregada pelo Compatibilista, dado que pode ser vista como uma formulação específica do Problema do Controle Não-aumentado⁹⁶. Outro caminho possível, para o qual eu gostaria de chamar a atenção, foi sugerido por Fischer (2007): embora Kane se esforce em argumentar por (em vez de simplesmente assumir) uma conexão entre responsabilidade última e possibilidades alternativas de ação, inserindo o conceito de ações autoformadoras, *não há uma razão pela qual não possamos “frankfurtar” uma ação autoformadora* – isto é, torná-la um Caso Estilo-Frankfurt. Se concluirmos que a agente é moralmente responsável pela decisão tomada (e se de fato não há possibilidades deixadas em aberto no cenário), então é falso que ações desse tipo exigem possibilidades alternativas. Além disso, sugiro, teremos uma razão para pensar que a própria exigência de responsabilidade última para que uma agente seja a fonte genuína da própria ação é, afinal, exagerada.

⁹⁵ O libertismo da causalidade do agente parece uma solução sob medida para o Problema do Controle Não-aumentado, como aponta Franklin (2017), ao sustentar que “o indeterminismo é necessário para aumentar o controle porque é necessário para a existência do poder de causalidade do agente” (p. 156). O autor argumenta, no entanto, que a influente abordagem de O’Connor (2009) admite o exercício desse poder mesmo em mundos deterministas – o que faz com que o libertismo da causalidade do agente não se saia melhor do que o libertismo da causalidade de eventos. Para outra argumentação de que a causalidade do agente é compatível com o determinismo, cf. Nelkin, 2011.

⁹⁶ Em outros termos, esse argumento estabelece a impossibilidade de que uma agente seja a fonte última da própria ação (o que equivale à fonte genuína, segundo o incompatibilista). Por essa razão, é claro que ele também pode ser utilizado, como faz o próprio Pereboom, no contexto de uma argumentação cética sobre a existência da liberdade necessária para a responsabilidade moral.

No próximo capítulo, farei uma apresentação detalhada dos Casos Estilo-Frankfurt e algumas das principais objeções a esse argumento, seguida de minha própria defesa de que os Casos são bem-sucedidos em objetar o Princípio das Possibilidades Alternativas.

4 CAPÍTULO 3: CASOS ESTILO-FRANKFURT E A IRRELEVÂNCIA DE POSSIBILIDADES ALTERNATIVAS DE AÇÃO PARA A RESPONSABILIDADE MORAL

3.1 INTRODUÇÃO: OS CASOS ESTILO-FRANKFURT

No capítulo 1, apresentei um princípio bastante intuitivo que vinculava a responsabilidade moral à liberdade de agir diferentemente, o Princípio das Possibilidades Alternativas. Retomemos seu conteúdo:

PPA. “uma pessoa é moralmente responsável pelo que faz apenas se podia ter agido diferentemente.” (FRANKFURT, 1969)

O PPA é o alvo de um dos artigos que ajudou a delinear o debate contemporâneo do Problema da Compatibilidade. Em “Alternate Possibilities and Moral Responsibility” (1969), Harry Frankfurt ofereceu um modelo de contraexemplos ao PPA que tem sido discutido desde a sua proposição⁹⁷ e cuja “moral”, segundo Fischer (2012), é bastante simples: “se o determinismo elimina a responsabilidade moral, não é em virtude de eliminar possibilidades alternativas” (p. 36). Estabelecendo a distinção entre a liberdade de agir diferentemente e a liberdade de ser a fonte genuína da própria ação, os Casos Estilo-Frankfurt pretendem mostrar que (ao menos quanto a uma ação em particular) a primeira *não é necessária* para a responsabilidade moral – o que enfatiza, tanto para compatibilistas quanto para incompatibilistas, a importância da questão sobre o que significa ser “fonte” da própria ação (FISCHER, 1982).

Quanto à relevância desse resultado, podemos apontar o seguinte: além de contrariarem as intuições comuns que dão origem ao PPA, se os Casos de fato objetam esse princípio, então nós podemos simplesmente “contornar” o Argumento da Consequência, no sentido de que sua possível solidez não interferirá na noção de liberdade que importa para o Problema da Compatibilidade (FISCHER, 1994, 2007, 2012). Por essa razão, e dada a possibilidade de “frankfurtar” as ações autoformadoras paradigmáticas para o libertismo, os Casos podem ser úteis particularmente para os compatibilistas⁹⁸. Vejamos então uma de suas formulações mais básicas:

⁹⁷ A literatura em torno desses casos é extremamente vasta. Para uma visão geral e recente da discussão, cf. Nelkin e Pereboom (eds.), 2022, capítulo 8.

⁹⁸ A exemplo de Fischer e Ravizza (1998) e Sartorio (2016). Observe, porém, que um incompatibilista não necessariamente rejeita os Casos Estilo-Frankfurt.

Caso Estilo-Frankfurt (tradicional). Uma neurocientista chamada Black tem monitorado secretamente os processos cerebrais de um agente, Jones, que está deliberando sobre se fará certa escolha E. Black pode prever confiavelmente as escolhas que Jones está prestes a fazer observando a atividade em seu cérebro, além de poder manipulá-la de um modo que garanta que o agente fará a escolha E. A neurocientista planeja intervir somente se a previsão for de que Jones não decidirá E por si mesmo. Mas Jones acaba fazendo a escolha E motivado por suas próprias razões, sem a intervenção de Black (que corretamente prevê que a escolha seria feita).⁹⁹

Observemos o papel desempenhado por cada elemento da história. Está em aberto para Jones se ele decidirá ou não por E; e suponha que há boas razões para ambas as alternativas de que esse agente pensa dispor. Aqui está o ponto mais importante da construção desse caso: o agente na verdade *não* dispõe de possibilidade alternativa à decisão tomada. A presença de Black faz com que Jones não tenha a liberdade de agir diferentemente porque a escolha E é inevitável – ela teria ocorrido ou pelas próprias razões do agente (como ocorre na sequência atual¹⁰⁰), ou pela intervenção da neurocientista; podemos imaginar que ela instalou um dispositivo no cérebro de Jones e que esse dispositivo seria engatilhado caso a tendência a desistir de E fosse detectada. Ainda assim, *nós julgamos que Jones é moralmente responsável pela escolha E* (assumindo que E seja moralmente relevante). Segundo Frankfurt (1969), esse exemplo aponta para uma distinção entre fatores que tornam uma escolha/ação inevitável e fatores que são relevantes do ponto de vista explanatório em relação a essa escolha/ação. A presença de Black, embora seja um fator do primeiro tipo, não é um fator do segundo tipo: ela *não* desempenha qualquer papel no que acaba ocorrendo entre a deliberação e a escolha de Jones. Ao contrário, a escolha E é feita do modo exato como seria feita na ausência de Black. Segundo Frankfurt, consideramos Jones responsável justamente por essa razão. A consequência é que possibilidades alternativas não seriam necessárias para responsabilidade moral.

A ausência de possibilidades alternativas aqui está apoiada no fato de os Casos pertencerem a uma subcategoria de casos de *sobredeterminação*, o terceiro tipo distinguido por E. J. Lowe (2002)¹⁰¹. Assuma que um evento C é a causa de um evento E. Em um Caso

⁹⁹ Adaptado de Sartorio, 2016, p. 13.

¹⁰⁰ “Atual”, aqui, é sinônimo de “efetivo”.

¹⁰¹ Lowe refere-se a três categorias desses casos, chamadas de (1) *actual overdetermination*, (2) *pre-emption* e (3) *fail-safe cases*. Esses casos são muito importantes para discussões sobre causalidade.

Estilo-Frankfurt, ocorre sobredeterminação do seguinte tipo: “se C não tivesse ocorrido, então outro evento, D – que não ocorre efetivamente – teria ocorrido e então causado E” (LOWE, 2002, p. 179). Nesse caso, E significa a escolha de Jones, C significa o processo de decisão por suas próprias razões e D significa a intervenção da neurocientista (que, embora esteja presente monitorando os processos cerebrais desse agente, não chega a agir no mundo atual).

Tipicamente, mesmo os defensores do PPA não discordam de que a figura de Jones em um Caso Estilo-Frankfurt pareça moralmente responsável. Eles tendem a argumentar, porém, que “apesar das impressões iniciais, o exemplo no fim das contas apresenta uma possibilidade alternativa relevante” (MCKENNA & PEREBOOM, 2016, p. 104). A forma mais imediata que essa objeção pode tomar é a de que Jones pode ao menos começar a decidir por não-E, já que seria com base em alguma tendência de escolha que o dispositivo de Black seria capaz de prevê-la¹⁰². Mas Fischer (1994) argumentou que se trata de uma alternativa tão exígua que não pode ser aquela que o defensor do PPA tem em mente como uma condição necessária para a responsabilidade moral. Para que uma possibilidade alternativa componha a liberdade de agir diferentemente de uma agente, ela precisa ser *robusta*. As principais críticas e defesas dos Casos Estilo-Frankfurt são, como veremos a seguir, versões refinadas desses dois pontos.

Nesse capítulo, meus objetivos serão (i) apresentar uma parte dos principais argumentos da discussão sobre os Casos Estilo-Frankfurt e (ii) defender que há um estilo de formulação – os casos Zona-Tampão¹⁰³ *Modificados* – aparentemente capaz de resistir a objeções. O capítulo contará, portanto, com uma tentativa minha de contribuir com o debate da natureza da liberdade necessária para a responsabilidade moral. Embora um caso Zona-Tampão Modificado tenha sido anteriormente proposto por McKenna (2018), as razões que apresentarei para endossá-lo e alguns detalhes do caso que proporei como um candidato promissor a contraexemplo ao PPA, o caso *Decepção*, serão originais. Minha defesa desse caso consistirá em oferecer respostas a três objeções recentes a Casos Estilo-Frankfurt que acredito serem representativas das estratégias de defensores do PPA: os argumentos de Franklin (2017), Elzein (2017) e Capes (2022). Também defenderei, de modo breve, que Casos Estilo-Frankfurt devem ser construídos sempre assumindo o indeterminismo (ao contrário do que sugere Fischer (2012) em sua defesa da viabilidade de um Caso com relações causais deterministas).

¹⁰² Essa estratégia ficou conhecida como *flicker of freedom* (cf. FISCHER, 1994, p. 134-147)

¹⁰³ No original, *buffer cases*.

3.2 A DEFESA DO DILEMA E ALGUMAS RÉPLICAS

O principal desafio aos Casos Estilo-Frankfurt mais simples, semelhantes à formulação que apresentei acima, assume a forma de um dilema (FISCHER, 2012; PEREBOOM, 2014). A *Defesa do Dilema* foi proposta de modos ligeiramente distintos por vários autores (cf. por exemplo KANE, 1996; GINET, 1996; e GOETZ, 2005). A seguinte reconstrução de sua forma geral está em Fischer (2012). Suponha que a escolha E de Jones fosse votar no candidato democrata no segundo turno das eleições, em vez de no candidato de extrema direita. A escolha ocorreria em um instante $t2$. Black está ciente de que, se Jones for decidir votar no democrata, ele involuntariamente levantará a sobrancelha esquerda em $t1$. Por outro lado, se a decisão for pelo candidato da extrema direita, Jones ficará ruborizado em $t1$. O sinal involuntário em $t1$ servirá como o gatilho do dispositivo: caso haja o rubor, Black intervirá e fará com que Jones decida votar pelo candidato democrata em $t2$. O que ocorre é que Jones levanta a sobrancelha esquerda e toma a decisão desejada por Black, sem que ela tenha que intervir.

Observe agora que um Caso Estilo-Frankfurt não faz qualquer suposição explícita sobre se a tese do determinismo é verdadeira. Assumamos então que não é; a tese do indeterminismo é verdadeira. A sequência causal que leva do sinal involuntário exibido em $t1$ à escolha em $t2$, portanto, é indeterminista – o que significa que Jones teria ao menos alguma probabilidade de optar pelo candidato da extrema direita ainda que levantasse a sobrancelha esquerda em $t1$. Podemos então questionar se não havia aqui uma possibilidade alternativa relevante, uma vez que o dispositivo de Black já não seria engatilhado após a detecção do sinal involuntário desejado. Sob a hipótese do indeterminismo, a presença de Black não seria plenamente capaz de privar Jones de possibilidades alternativas. Assuma agora que a tese do determinismo é verdadeira. Há, pois, uma relação determinista entre o sinal exibido em $t1$ e a escolha realizada em $t2$, o que garante o poder do dispositivo em eliminar possibilidades alternativas. Mas essa hipótese também é problemática: o incompatibilista diria que, já que o que está em questão no Problema da Compatibilidade é justamente se um agente como Jones é moralmente responsável, a expectativa de uma resposta afirmativa intuitiva seria uma petição de princípio em favor do compatibilista. (FISCHER, 2012)

Stewart Goetz (2005) apresentou um ponto adicional à suposição determinista desse dilema. Primeiramente, ele assume que a verdade do determinismo elimina o acesso de agentes a possibilidades alternativas; o que é algo que alguns compatibilistas (como o próprio Fischer) também estariam dispostos a aceitar. Em seguida, ele afirma que o que de fato

elimina as possibilidades alternativas de Jones não é a presença Black, mas o próprio determinismo. Em suas palavras,

[O Caso Estilo-Frankfurt] cria a aparência que é o dispositivo de Black, que está na sequência alternativa dos eventos, que faz com que Jones não seja livre para decidir diferentemente. Essa aparência é ilusória porque, sem a ocorrência do determinismo causal na sequência atual dos eventos, o dispositivo não pode impedir Jones de fazer uma escolha alternativa; e com o determinismo causal na sequência atual dos eventos, não é o dispositivo que impede Jones de fazer uma escolha alternativa. Em resumo, se Jones não é livre para decidir de outro modo, é por causa da ocorrência do determinismo causal na sequência atual dos eventos, e não por causa do dispositivo de Black (...). Portanto, é errado concluir que Jones seja moralmente responsável embora não seja livre para escolher de outro modo. (p. 85)

Esta seção possuirá quatro etapas. Na primeira, apresento a resposta de Fischer em “The Frankfurt Cases: The Moral of the Stories” (2012) às objeções da suposição determinista do Dilema. Fischer pretende defender que, embora um Caso Estilo-Frankfurt não precise ser construído supondo o determinismo, não há nenhuma razão pela qual isso não possa ser feito. Em seguida, apresento uma crítica a essa resposta e defendo que os Casos não devem ser construídos assumindo o determinismo¹⁰⁴. Concluo que o sucesso dos Casos Estilo-Frankfurt depende da existência de uma boa resposta à objeção da suposição indeterminista do Dilema. Essas duas etapas compõem a primeira subseção. Já a terceira etapa será a apresentação de um estilo de Caso que tenta justamente oferecer essa resposta: o Caso Zona-Tampão de Pereboom (2001, 2014). Por fim, na última subseção, apresento a principal crítica recebida pela primeira versão desse caso, assim como a resposta de Pereboom e sua nova versão do argumento.

3.2.1 Casos Estilo-Frankfurt deterministas

A resposta de Fischer (2012) às objeções da suposição determinista começa por reiterar que não precisamos, nem deveríamos, afirmar que Jones é de fato moralmente responsável; isso afasta a acusação de petição de princípio. O que um defensor dos Casos Estilo-Frankfurt deveria dizer é que “Eu não sei ainda se o agente é moralmente responsável por seu comportamento, mas se ele não é, isso não se dá porque lhe faltaram possibilidades alternativas” (FISCHER, 2006, p. 128). Outro ponto chave para o argumento de Fischer será assumir a verdade do determinismo, sem, no entanto, tomar posição sobre se ele elimina o acesso de agentes a possibilidades alternativas (um agnosticismo por exemplo quanto ao Argumento da Consequência). Quando Black vê que Jones levanta a sobrancelha esquerda, ela relaxa; afinal, a relação entre o sinal em $t1$ e a escolha em $t2$ é determinista e, não fosse

¹⁰⁴ A argumentação que apresentarei neste ponto foi esboçada em meu trabalho de conclusão de curso, na Universidade Federal de Uberlândia (UFU). Cf. Rodrigues, 2020.

esse o sinal, o dispositivo garantiria a decisão de Jones pelo voto no candidato democrata. Fischer afirma então que esse fato *mais* a verdade do determinismo é que privam Jones de possibilidades alternativas.

Ele formula o argumento da seguinte maneira:

1. Assumamos que o determinismo causal ocorre e que o caso estilo-Frankfurt se desenrola do modo a seguir.
2. Nesse ponto do argumento, o determinismo causal não é pensado como, em si mesmo, eliminando o acesso de agentes a possibilidades alternativas (nem é suposto que a presença de Black e do dispositivo, em si mesma, elimina esse acesso);
3. O determinismo causal mais a presença de Black eliminam a liberdade de Jones de escolher diferentemente em t_2 ;
4. Se Jones não é moralmente responsável por escolher votar no democrata em t_2 , isso não é pelo mero fato de que ele não era livre em t_2 para escolher diferentemente;
5. Então, se o determinismo causal exclui a responsabilidade moral de Jones por sua escolha, não é em virtude de eliminar possibilidades alternativas (se de fato as elimina). (p. 44-45)

Já que o determinismo não é pensado como excluindo possibilidades alternativas para agentes, haveria um papel substancial a ser desempenhado por Black nessa eliminação. Fischer explica que Jones poderia exibir o sinal involuntário correspondente ao voto no candidato da extrema direita (o rubor) em t_1 e, não fosse o dispositivo, não haveria nada no caso capaz de eliminar a liberdade de agir diferentemente desse agente. Simetricamente, a presença de Black sem o determinismo também não eliminaria possibilidades alternativas; Jones poderia levantar a sobrancelha esquerda em t_1 e, ainda assim, optar pelo candidato da extrema direita. Essas observações justificam a premissa 3. Já a respeito da transição entre 3 e 4, Fischer sustenta as três seguintes afirmações:

- A. O dispositivo de Black mais o determinismo causal eliminam a capacidade de Jones decidir diferentemente, mesmo que eu não esteja supondo (nesse estágio do argumento) que cada um o faria por si mesmo.
- B. O dispositivo e as disposições de Black são irrelevantes para a responsabilidade moral de Jones.
- C. O fato de que o dispositivo e as disposições de Black em um contexto causalmente determinista eliminam a capacidade de Jones de escolher e agir diferentemente é irrelevante para a responsabilidade moral de Jones. (p. 46)

A mais controversa é C: considerando que não é nada claro que o determinismo seja irrelevante para a responsabilidade moral de Jones, parece questionável que algo de que o próprio determinismo é uma causa – a eliminação das alternativas – seja irrelevante para a responsabilidade desse agente. Fischer apresenta então a seguinte analogia. Suponha que uma pessoa com a doença D precisa de dois medicamentos, M1 e M2, para curá-la. Ambos são necessários e sua ação conjunta provoca um aumento colateral na temperatura corporal da

pessoa em tratamento. O medicamento M2, por exemplo, é necessário para esse aumento. O fato de que M1 e M2 causem o aumento de temperatura, no entanto, é irrelevante para a cura de D. De maneira similar, o fato de que a verdade do determinismo mais a presença de Black eliminem a liberdade de agir diferentemente de Jones é irrelevante para sua responsabilidade moral. Dada essa argumentação, poderíamos agora afrouxar a posição agnóstica e construir uma defesa similar de um Caso Estilo-Frankfurt determinista supondo que o determinismo é incompatível com o acesso de agentes a possibilidades alternativas, o que se justifica pelo fato de que nada no argumento exploraria essa incompatibilidade. (FISCHER, 2012)

É duvidoso, porém, que a estratégia de Fischer sustente a possibilidade de construção de um Caso Estilo-Frankfurt explicitamente determinista. Parece-me, ao contrário, que *não deveríamos* construí-lo desse modo, independentemente da suposição agnóstica sobre o determinismo eliminar possibilidades alternativas. Vejamos novamente a premissa 3: “O determinismo causal mais a presença de Black eliminam a liberdade de Jones de escolher diferentemente em t_2 ”. Nesse ponto, concedamos a Fischer a suposição agnóstica. Ainda assim, penso que não faça sentido afirmar que “já que [o argumento] explicitamente não pressupõe que o determinismo causal exclua possibilidades alternativas, há um distinto e substantivo papel para Black desempenhar” (FISCHER, 2012, p. 45). Uma primeira razão é a seguinte: como do agnosticismo (uma posição epistemológica) pode emergir um papel substantivo a ser desempenhado por Black? Esse papel talvez existisse como resultado de uma suposição *compatibilista* entre o determinismo e possibilidades alternativas; mas essa suposição seria tão ou mais controversa do que os próprios casos Estilo-Frankfurt. Não podemos depender dela para o argumento.

Feita essa observação, há ainda uma segunda razão pela qual a premissa 3 não se sustenta. Assumamos por um momento que o determinismo exclui possibilidades alternativas (uma tese ao menos plausível) e que ele ocorre no mundo de Jones. Consequentemente, o passado e as leis da natureza explicam suficientemente a escolha desse agente. Mas o que parece estar escapando a Fischer (e que corresponde, em diferentes termos, ao traço crucial do caso para o qual Goetz chama a atenção) é que a mesma coisa é verdadeira para as intenções e a ação de Black. O passado e as leis da natureza explicariam suficientemente que ela instale o dispositivo no cérebro de Jones, deseje que ele decida votar no candidato democrata e que sua presença não seja desnecessariamente revelada etc. Ainda que Jones tivesse ruborizado em t_1 , o próprio rubor, a ativação do dispositivo e seus efeitos já estariam determinados. Nessa sequência alternativa de eventos, poderíamos dizer: “Jones votou no candidato democrata por causa do dispositivo de Black”. Essa afirmação, porém, está em um registro diferente da frase

“Jones votou no candidato democrata porque estava determinado a agir desse modo”; e não faz sentido dizer que a primeira *mais* a segunda explicam por que Jones votou no candidato democrata.

Para que o ponto fique mais claro, considere o seguinte exemplo. Clarice rapidamente dá xeque-mate em João em um jogo de xadrez. Imagine que uma explicação correta para esse fato é que Clarice é uma jogadora muito habilidosa. Outra explicação correta é que a movimentação da Dama preta de Clarice, junto à movimentação de um de seus cavalos, causou o xeque-mate no Rei branco de João. Não há sentido em dizer que o fato de Clarice ser uma jogadora muito habilidosa *mais* a movimentação da Dama e do Cavalo de Clarice explicam o xeque-mate; podemos supor que as respectivas movimentações são elas próprias explicadas pela habilidade da jogadora. Analogamente, se o curso de ação de Black e do dispositivo na sequência alternativa é acarretado pelo determinismo, sua presença “sobra” em um Caso Estilo-Frankfurt deterministicamente construído. Esse ainda é o caso se assumirmos que o determinismo não elimina possibilidades alternativas. Basta considerar o exemplo do xeque-mate, no qual não havia nenhuma suposição sobre a habilidade de Clarice fazer ser o caso que ela necessariamente daria xeque-mate. Assim, acredito que não haja aqui um argumento convincente pela tese de que a presença de Black elimina as possibilidades alternativas de Jones em um Caso Estilo-Frankfurt determinista. Essa afirmação seria, como apontam os proponentes do Dilema, uma petição de princípio para o defensor dos Casos. Como descrevi anteriormente (RODRIGUES, 2020),

Supondo que não sabemos se o determinismo é compatível com o acesso dos agentes a alternativas – a incompatibilidade é plausivelmente sugerida pelo argumento (...) da consequência – e que a “moral dos casos de Frankfurt” pretende ser consistente com a solidez do argumento da consequência: i. se o próprio determinismo exclui possibilidades, então Black “sobra”, como foi anteriormente apontado; e ii. se o determinismo não fecha alternativas, então o chifre determinístico do dilema pode até funcionar como o indeterminístico, mas nesse caso a dependência de que haja compatibilidade com a presença de alternativas para assegurarmos um papel a Black tornaria a construção desse caso de Frankfurt inconsistente com a aceitação do argumento da consequência. Não me parece haver vantagem em construir uma versão explicitamente determinista desses casos (...). (p. 62)

Em adição a essas considerações, penso que um Caso Estilo-Frankfurt deva ser construído em um cenário *indeterminista* devido ao próprio papel que ele pretende desempenhar na discussão do Problema da Compatibilidade. A figura de Black é um *substituto* do determinismo, não um elemento que possa ser acrescido a ele, em *um* dos aspectos que nos parece ameaçador em leis da natureza deterministas: a ausência de possibilidades alternativas. Esse substituto pretende ser teoricamente neutro quanto ao

Problema da Compatibilidade, possibilitando que um traço do determinismo seja isolado e que as consequências da decisão de Jones se desenrolando nesse modelo sejam analisadas. Não fosse assim, Frankfurt estaria de fato cometendo uma petição de princípio. Em “Alternate Possibilities and Moral Responsibility”, ele apontava que ninguém parecia disposto a negar o PPA e que a controvérsia entre filósofos trabalhando com o Problema da Compatibilidade era sobre se aceitar esse princípio implicava um comprometimento com o incompatibilismo. Se a presença de Black priva Jones da liberdade de agir diferentemente, podemos agora nos perguntar, sem a preocupação com possibilidades alternativas de ação, o que há na agência desse e de outros agentes que os torna moralmente responsáveis.¹⁰⁵

3.2.2 Os casos Zona-Tampão

Temos então que o sucesso de um Caso Estilo-Frankfurt será decidido pela suposição indeterminista do Dilema. Nesta subseção, apresentarei uma estrutura de formulação dos Casos chamada de “Zona-Tampão” e desenhada para responder à objeção que assume o indeterminismo¹⁰⁶. Nessa estrutura, segundo Fischer (2012), é como se fosse estabelecida uma “terra de ninguém” na fronteira entre dois países A e B. Para que alguém vá do país A ao país B, precisa entrar nessa porção de terra; a entrada, no entanto, não garante que essa pessoa chegará ao outro lado: ela ainda pode evitar ou ser impedida de entrar no país B. Analogamente, um Caso Estilo-Frankfurt com essa estrutura contará com uma *condição necessária* para que a agente tenha a possibilidade de decidir diferentemente que poderá ser obtida de maneira voluntária, mas que não contará ela própria como uma possibilidade alternativa robusta porque engatilhará o dispositivo de Black ou, na ausência do dispositivo, não será *suficiente* para que a agente decida diferentemente. Veremos esses pontos a seguir. Considere o seguinte caso:

Caso Estilo-Frankfurt “Sonegação Fiscal”. “Joe está considerando solicitar uma dedução fiscal para a taxa de registro que ele pagou quando comprou uma casa. Ele sabe que solicitar essa dedução é ilegal, mas que provavelmente não será pego, e que se

¹⁰⁵ Esse ponto sobre o papel dos Casos Estilo-Frankfurt na discussão do Problema da Compatibilidade será retomado mais adiante, ao final da seção 3.3.

¹⁰⁶ Outros autores tentaram construir um Caso Estilo-Frankfurt explicitamente indeterminista que escapasse a essa objeção; cf., por exemplo, Mele e Robb, 1998. Mas alguns defensores tanto do compatibilismo (cf. Fischer, 2012) quanto do incompatibilismo (cf. Franklin, 2017) sugerem que alguma versão dos casos Zona-Tampão seja a estratégia mais promissora para mostrar que a presença de Black elimina as possibilidades alternativas de Jones e que, ainda assim, julgamos que ele seja moralmente responsável por sua escolha. Por essa razão, explorarei apenas essa estratégia daqui em diante no capítulo.

ele fosse, poderia convincentemente alegar ignorância. Suponha que ele tenha um forte, mas nem sempre prevalecente desejo de promover o próprio interesse, independentemente dos custos para terceiros e mesmo que envolva atividade ilícita. Além disso, a única maneira pela qual ele poderia deixar de decidir sonegar impostos nessa situação é por razões morais, das quais ele está ciente. Ele não poderia, por exemplo, deixar de escolher sonegar impostos sem motivo ou simplesmente porque teve vontade. É também causalmente necessário para que ele decida não sonegar impostos que um certo nível de atenção às razões morais seja atingido. Joe pode garantir esse nível de atenção voluntariamente. No entanto, atingir esse nível de atenção não é causalmente suficiente para que ele deixe de sonegar impostos. Se alcançasse esse nível de atenção, ele ainda poderia, exercendo seu livre-arbítrio à maneira do libertista, optar por sonegar impostos ou abster-se de escolher (sem o dispositivo do interventor instalado). Para garantir que ele decida sonegar impostos, uma neurocientista, sem o conhecimento de Joe, implantou em seu cérebro um dispositivo que, se detectasse o nível necessário de atenção a razões morais, estimularia eletronicamente os centros neurais certos de modo a inevitavelmente resultar na escolha desejada. O que ocorre é que Joe não atinge esse nível de atenção às suas razões morais e decide sonegar impostos por si próprio, enquanto o dispositivo permanece ocioso.” (PEREBOOM, 2014, p. 15)¹⁰⁷

A objeção do defensor do Dilema para a suposição indeterminista era a seguinte: se a relação entre o sinal involuntário que Jones exhibe em $t1$ e a decisão que ele toma em $t2$ é indeterminista, o dispositivo de Black não é capaz de privar esse agente de possibilidades alternativas pelo fato de que, mesmo tendo exibido em $t1$ o sinal correspondente à decisão que Black deseja, Jones ainda poderia decidir diferentemente em $t2$. Vimos que uma possibilidade alternativa precisa ser robusta para contar como tal; é o que parece ocorrer quando estipulamos sinais involuntários indeterministicamente relacionados a cada escolha e esclarecemos um Caso Estilo-Frankfurt tradicional. Para avaliar se a versão Sonegação Fiscal apresentada acima (um caso do tipo Zona-Tampão) é capaz de responder a essa objeção, vejamos a princípio como seu proponente, Pereboom, descreve a noção de robustez.

Avançando a sugestão em termos intuitivos de Fischer (1994), Pereboom (2014) buscou esclarecer o que seria, para uma possibilidade alternativa, a robustez relevante para avaliações de responsabilidade moral. Como explicam McKenna e Pereboom (2016), se a

¹⁰⁷ Cf. Pereboom, 2001.

posse de alternativas é importante para a responsabilidade de uma agente por uma ação, ela precisa ser robusta ao menos “no sentido de que, como resultado de sua efetivação, a agente teria assim se engajado em um empreendimento voluntário por meio do qual teria evitado a responsabilidade moral que possui por sua ação no mundo atual” (p. 107). No caso em que Jones ou levantaria a sobrelha esquerda ou ruborizaria, se ele não levanta a sobrelha esquerda, isso não significa que Jones esteja voluntariamente evitando sua responsabilidade por decidir votar no candidato democrata. Afinal, o sinal é involuntário e, portanto, não há uma possibilidade alternativa robusta.

A robustez possui também uma componente epistêmica: se a única maneira pela qual Joe teria evitado sua decisão de sonegar impostos era (em vez de dar atenção às razões morais contrárias a essa decisão) dando um gole em sua xícara de café contendo, sem que ele saiba, uma substância que induz conformidade com as leis sobre impostos, essa não é uma possibilidade alternativa robusta porque “ele precisa em algum sentido entender que, ou como, [a alternativa] estava disponível” (MCKENNA & PEREBOOM, 2016). A proposta de Pereboom para uma condição necessária (embora talvez não suficiente) substancial da robustez é então a seguinte:

Robustez (B): Para que uma agente tenha uma alternativa robusta à sua ação imoral A, isto é, uma alternativa *per se* relevante para explicar porque ela é culpada por realizar A, precisa ser o caso que

- (i) ela poderia ter voluntariamente agido ou se refreado de agir, resultando na ausência de culpa, e
- (ii) para ao menos uma dessas coisas que ela poderia ter feito, ela era cognitivamente sensível ao fato de que poderia voluntariamente agir ou se refrear desse modo, e ao fato de que se ela o fizesse, então não seria, ou provavelmente não seria, culpada. (2014, p. 13)¹⁰⁸

Retomemos agora o caso Sonegação Fiscal. Observe que ele é construído de maneira indeterminista: *existe uma condição necessária para decidir diferentemente* – a atenção a razões morais – *na ausência da qual, porém, Joe não é causalmente determinado a escolher sonegar impostos*. Para qualquer instante específico em que o nível necessário de atenção estivesse ausente, “Joe ainda poderia fazê-lo ocorrer em um momento posterior, então ele não está causalmente determinado a decidir sonegar impostos por essa ausência do momento anterior” (PEREBOOM, 2014, p. 15). Isso permitiria a neutralidade teórica do cenário, já que a exigência libertista geral para uma decisão livre é satisfeita. Mas o libertista apontará agora que o caso conta com uma possibilidade alternativa – novamente, dar a atenção devida às

¹⁰⁸ A expressão “cognitivamente sensível” desempenha aqui o papel de uma condição epistêmica mais fraca do que “saber” ou “compreender”, já que essas duas últimas noções talvez sejam exigentes demais nesse contexto. Cf. Pereboom, 2014, p. 12.

razões morais – cuja realização, dada a presença da neurocientista, seria uma ação voluntária de Joe que resultaria na prevenção de sua culpa por sonegar impostos (em virtude da intervenção do dispositivo). Pereboom responde então que essa alternativa não é robusta porque Joe não está, nem tem razão para estar, cognitivamente sensível ao fato de que evitaria sua culpa ao voluntariamente atingir o nível necessário de atenção. Embora ele saiba essa é uma condição necessária para sua desistência da sonegação, Joe “acredita que atingir esse nível de atenção é compatível com sua livre decisão de sonegar impostos de qualquer jeito (o que seria verdade se o dispositivo da interventora não estivesse presente)” (2014, p. 16). Ainda assim, parece que esse agente é moralmente responsável por sua decisão e ação.

3.2.3 A Defesa *Timing*

Carl Ginet (2002) formulou uma importante objeção aos Casos Estilo-Frankfurt indeterministas precisamente com foco na versão Sonegação Fiscal: a Defesa *Timing*. Estipulemos que a decisão de Joe de sonegar impostos foi tomada em um instante $t1$. Esse agente não poderia ter postergado sua decisão indefinidamente: havia um prazo, até $t2$, a ser seguido. Se logo antes de $t2$ o dispositivo da neurocirurgiã tivesse detectado o nível necessário de atenção às razões morais para não sonegar impostos, sua ação teria sido engatilhada em $t2$, fazendo com que Joe decidisse pela sonegação em $t3$. Ginet argumenta então que Joe *não é* culpado por decidir sonegar impostos até $t3$, mas que ele talvez seja culpado por decidir agir desse modo em $t1$ e que, se o for, a explicação é a existência de uma *possibilidade alternativa robusta anterior a $t1$* – a de dar o nível necessário de atenção às razões morais. Ele afirma que, se Joe tivesse dado atenção às razões morais antes de $t1$, “ele estaria naquele momento se abstendo de uma vontade [a de decidir sonegar impostos *naquele* momento] (...) tal que, restando-se desse modo, ele teria evitado a responsabilidade [pela decisão de sonegação] (...) naquele momento e teria estado ciente [desse fato]” (GINET, 2002).

Reconstruindo a Defesa *Timing*, Pereboom (2014) aponta que podemos falhar em distinguir “Joe decidindo pela sonegação *em $t1$* ” de “Joe decidindo pela sonegação *até $t3$* ” devido ao fato de essas circunstâncias serem moralmente equivalentes, no sentido de que, na ausência da neurocirurgiã, o agente teria sido igualmente culpado em ambas. Mas a análise correta da culpa de Joe por decidir sonegar impostos até $t3$ seria, segundo Ginet, que esse agente teve uma possibilidade alternativa a essa decisão em $t1$; e aquilo pelo que Joe realmente é culpado é a escolha pela sonegação *em $t1$* (isto é, pela obtenção desse estado de

coisas temporalmente mais específico). Frente a essa objeção, Pereboom modifica alguns aspectos de Sonegação Fiscal e apresenta a seguinte versão de um Caso Zona-Tampão¹⁰⁹:

Caso Estilo-Frankfurt “Corte de Imposto”. “Jones pode votar a favor ou contra um corte de impostos modesto para aqueles em seu grupo de alta renda, pressionando o botão de sim ou de não na cabine de votação. Assim que ele entrar na cabine, terá exatamente dois minutos para votar e um cronômetro de contagem regressiva será exibido. Se Jones não votar, terá de pagar uma multa cara o suficiente para que, em sua situação, ele se comprometa com certeza a votar (a favor ou contra); isso é corroborado pelo fato de que a perspectiva da multa, juntamente às condições de fundo, o determina causalmente a votar. Jones concluiu que votar a favor do corte é moralmente errado, já que acredita que isso não estimularia a economia de forma desejável, acrescentando riqueza aos já ricos sem ajudar pessoas de baixa renda (apesar de essa conclusão contrariar o modo como o corte foi anunciado). Ele é receptivo e reativo a esses tipos de razões morais: ele votaria contra um corte de impostos substancialmente maior para seu grupo de renda por razões desse tipo, e de fato o fez no passado. Ele gasta algum tempo na cabine de votação ensaiando as razões tanto morais quanto de interesse próprio relevantes. Mas o que seria necessário para Jones decidir votar contra o corte de impostos é que ele imaginasse vividamente sua chefe descobrindo um voto favorável ao corte; e então, devido às suas inclinações políticas, ela o puniria deixando de promovê-lo a uma posição melhor. Nessa situação, é causalmente necessário para sua decisão de não votar a favor do corte de impostos, e em vez disso votar contra, que ele imagine vividamente a descoberta da chefe e sua própria não-promoção – o que pode ocorrer involuntariamente ou voluntariamente, por seu livre-arbítrio à maneira do libertista. Jones entende que imaginar o cenário de punição o colocará em uma posição motivacional apropriada para votar contra o corte. Mas essa imaginação não é

¹⁰⁹ Outra objeção importante a Sonegação Fiscal está em Widerker (2006). Ele argumenta que uma explicação plausível para nossa impressão de que Joe é moralmente responsável por sua decisão é que se trata de responsabilidade *derivada* de uma omissão – não ter feito um esforço razoável para dar atenção a razões morais – pela qual o agente é diretamente culpado, em virtude de seu acesso a possibilidades alternativas. Se essa é a explicação correta da nossa intuição sobre o caso, então o PPA não é objetado. Widerker concede, porém, que já não estaria tão certo de seu ponto se “a condição necessária fosse modificada para algo que não fosse um ato voluntário moralmente impressionante” (PEREBOOM, 2014, p. 20). A versão de um Caso Estilo-Frankfurt que apresentarei a seguir, além de tentar responder à Defesa *Timing*, incorpora a sugestão de Widerker (embora Pereboom ofereça também uma outra resposta a esse autor; cf. 2014, p. 18-20). Na próxima seção, porém, veremos que essa substituição não parece fazer tanta diferença para proteger o Caso de objeções.

causalmente suficiente para que ele decida votar contra; mesmo assim ele ainda poderia, por seu livre-arbítrio à maneira do libertista, decidir votar ou a favor ou contra (sem o dispositivo da interventora em jogo). No entanto, uma neurocientista, sem o conhecimento de Jones, implantou um dispositivo em seu cérebro que, se detectasse a imaginação vívida do cenário de punição, estimularia seu cérebro para determinar causalmente a decisão de votar a favor do corte de impostos. A imaginação de Jones não é exercitada dessa maneira e ele decide votar a favor do corte por si próprio, enquanto o dispositivo permanece ocioso.” (PEREBOOM, 2014, p. 23)

Imagine que os dois minutos do cronômetro se encerram pouco depois de t_3 , sendo então esse o último instante em que Jones pode tomar sua decisão. Ele decide votar a favor do corte em t_1 . Pereboom (2014) afirma então que “Jones é culpado por decidir votar a favor do corte até t_3 apesar do fato de não ter uma possibilidade alternativa robusta” (p. 23). Primeiramente, uma preocupação em relação a Sonegação Fiscal poderia ser que o fato de Joe compreender que poderia a qualquer momento, estendendo-se indefinidamente no futuro, evitar a responsabilidade por decidir sonegar impostos (dando o nível de atenção necessário a razões morais) lhe daria uma possibilidade alternativa robusta para cada momento no futuro, indefinidamente (apesar da presença do dispositivo). Em Corte de Imposto, Jones está comprometido com a tomada de alguma decisão até t_3 , o que elimina essa preocupação. Ainda é verdade que Jones está ciente de que pode evitar a responsabilidade pela decisão de votar a favor do corte imaginando vividamente o cenário de punição. Mas ele não teria uma possibilidade alternativa robusta relativa a todo o intervalo de dois minutos até t_3 e incluindo t_3 porque, segundo Pereboom, “o compromisso de Jones de tomar a decisão do voto até t_3 elimina a possibilidade de que ele não tome qualquer decisão durante esse intervalo em virtude de ter ficado vividamente imaginando o cenário de punição” (p. 24).

Para Ginet, Jones não é culpado pelo estado de coisas de tomar a decisão de votar em favor do corte até t_3 porque ele não dispõe de possibilidades alternativas a esse estado de coisas; mas ele seria culpado por tomar essa decisão em t_1 porque dispunha de uma possibilidade alternativa para esse estado de coisas mais específico. Pereboom (2014) oferece então o seguinte contra-argumento. Suponha que Adão está causalmente determinando por fatores fora de seu controle a atirar em uma pessoa com seu rifle e a tomar essa decisão em t_1 . O defensor do PPA dirá que ele não é moralmente responsável pela decisão/ação. Mas imagine agora que, apesar de Adão estar determinado a tomar essa decisão, depende dele o momento específico no qual a decisão será tomada, em um intervalo de dois minutos entre t_0

e $t3$ (incluindo $t3$). Seu desejo de assassinar a vítima é muito forte; e Adão entende que não há diferença moral entre tomar a decisão em um instante ou em outro, sendo então indiferente a essa questão. Ele decide em $t1$. Dada a posição incompatibilista de maneira geral, Adão não será culpado por tomar essa decisão em $t1$ e a justificativa precisará ser que esse agente não tem possibilidade alternativa robusta a tomar a decisão até $t3$. Esse agente seria, no máximo, responsável em um sentido moralmente neutro (mas não culpado) pelo momento da decisão. Nas palavras de Pereboom,

O (...) incompatibilista [especificamente o defensor do PPA] não será capaz de defender a afirmação de que a decisão de Jones em $t1$ de votar a favor do corte de imposto, na presença do dispositivo (...), é tão culpável, e pelas mesmas razões, quanto seria sua decisão de votar a favor até $t3$ sem o dispositivo em jogo. Com o dispositivo, o (...) incompatibilista não pode explicar a culpa de Jones por essa decisão em $t1$, mas apenas sua responsabilidade em um sentido neutro por tomar a decisão em $t1$ em vez de em outro instante. Embora, como na situação de Adão, Jones tenha alternativa a decidir em $t1$ – por exemplo, continuando a deliberar em $t1$ e decidindo em $t2$ – isso seria insuficiente para explicar a culpa de Jones por tomar essa decisão em $t1$. Do modo como o caso é construído, Jones não tem alternativa a tomar sua decisão até $t3$, e como mostra a situação de Adão, isso tem que ser suficiente, da perspectiva do (...) incompatibilista, para que Jones não seja culpado por tomar sua decisão em $t1$. Tanto Adão quanto Jones podem tomar suas decisões em $t2$ em vez de em $t1$, mas, como no caso de Adão, o fato de Jones não ter possibilidade alternativa robusta a tomar sua decisão até $t3$ faz com que a decisão em $t2$ não seja uma alternativa robusta. Ainda assim, temos a forte intuição de que Jones é culpável por decidir votar a favor em $t1$ – fato para o qual o (...) incompatibilista [defensor do PPA] agora não tem uma explicação. (2014, p. 25-26)

Pereboom conclui então que o Caso Estilo-Frankfurt Corte de Imposto é uma séria objeção ao PPA. Em sua reconstrução do argumento acima, Franklin (2017) comenta que, para que Jones tivesse uma oportunidade de agir diferentemente *per se* relevante para explicar sua responsabilidade moral, seria necessário que ele antes agisse em uma oportunidade de agir diferentemente que não é *per se* relevante – imaginar a descoberta da chefe. Um caso Zona-Tampão colocaria um espaço entre oportunidades irrelevantes e relevantes do ponto de vista explanatório.

3.3 UMA DEFESA DOS CASOS ESTILO-FRANKFURT

Nesta seção, farei uma defesa de um estilo de formulação dos Casos Estilo-Frankfurt chamados de casos *Zona-Tampão Modificados*. Na subseção 3.3.1, apresentarei uma objeção a Corte de Imposto levantada por McKenna e desenvolvida de modo independente por Franklin. Argumentarei que a versão de Franklin indica a necessidade de uma revisão no critério de Robustez para possibilidades alternativas – um enfraquecimento da componente

epistêmica. Na subseção 3.3.2, introduzirei o caso Zona-Tampão Modificado originalmente proposto por McKenna e apontarei suas vantagens em relação a Corte de Imposto. Penso que esse caso, no entanto, ainda necessita de revisão. Argumentarei então que meu próprio Caso Estilo-Frankfurt, *Decepção*, é um candidato mais promissor a contraexemplo ao Princípio das Possibilidades Alternativas. Em seguida, apresentarei uma resposta à objeção de Franklin. Por fim, na subseção 3.3.3, tornarei a defesa de *Decepção* mais robusta por meio do enfrentamento de outras duas objeções importantes¹¹⁰.

3.3.1 Outro desafio a “Corte de Imposto” e uma revisão de Robustez

No artigo “A Critical Assessment of Pereboom’s Frankfurt-style example” (2018), McKenna levanta uma objeção ao caso Corte de Imposto. Endossando um ponto anteriormente formulado por Widerker (2006), ele sustenta que a responsabilidade moral de Jones é apenas *derivada*: a “deixa” para intervenção do dispositivo (a imaginação da chefe) é afinal algo que o agente pode fazer, com uma possibilidade alternativa robusta de ação, em virtude da qual – o libertista poderia alegar – esse agente é responsável de maneira derivada pelo voto. Nas palavras de McKenna,

Jones está ciente de que, exceto se decidir imaginar vividamente a reação de sua chefe (...), uma decisão que é livre para tomar, ele certamente fará algo moralmente errado (...). Apesar de ser verdadeiro que, mesmo com a imaginação da chefe, ele ainda correria risco de agir mal (e está ciente disso), (...) [é] como se, ao livremente (...) omitir-se em tomar certa decisão, ele estivesse conscientemente colocando a si mesmo no caminho da má ação. (p. 3121-3122)

Ao final de *A Minimal Libertarianism* (2017), Franklin oferece uma versão mais elaborada dessa objeção. Ele afirma que, uma vez que aspectos pouco esclarecidos de Corte de Imposto são mais bem descritos, “ficará claro ou que o agente em questão não é moralmente responsável, ou que ele tem uma oportunidade razoável de agir diferentemente” (2017, p. 202). Franklin começa estipulando que Joe tem uma obrigação moral de votar contra o corte de imposto. A questão então é a seguinte: esse agente está ciente de que o *único* modo pelo qual ele pode cumprir sua obrigação moral é imaginando a descoberta da chefe? Pereboom não deixa isso claro. Suponhamos primeiro que Jones “não é cognitivamente sensível a sua psicologia um tanto estranha” (p. 204). Para que o alvo de Corte de Imposto seja a responsabilidade moral direta pela ação que esse agente realiza, precisamos assumir que Jones não é culpado pela psicologia estranha (aquela em que o único estado mental suficiente

¹¹⁰ Agradeço a Pedro Merluzzi pelo incentivo a que eu desenvolvesse melhor essa seção. Agradeço também aos grupos Ignorantes Culpáveis e GEMF pelas objeções e sugestões.

motivador da ação correta é a imaginação da chefe) que possui. Mas nesse caso, argumenta Franklin, já não é claro que Jones seja culpado pela decisão de votar pelo corte. Ele oferece a seguinte situação análoga:

Salva-vidas: Jeff é o salva-vidas da praia local. Na noite passada, ele foi sequestrado e manipulado por um ansioso time de neurocientistas que desejava testar os poderes da tecnologia científica atual. Eles então tentaram rearranjar a estrutura motivacional de Jeff de modo que ele resgataria pessoas se afogando no mar apenas se, primeiro, se imaginasse sendo parabenizado e elogiado pelos outros banhistas. Os neurocientistas se asseguram de que Jeff retenha sua capacidade e oportunidade de imaginar esse cenário; e que a imaginação é apenas uma condição necessária para o resgate. Exceto se imaginar os elogios, ele não terá qualquer motivação e então (suponha) não terá a oportunidade de resgatar ninguém. No entanto, mesmo que Jeff imagine esse cenário, ele ainda será capaz e terá a oportunidade de se abster de fazer o resgate. Os neurocientistas são bem-sucedidos em fazer essas mudanças na psicologia de Jeff. No dia seguinte, ele acorda pensando que tudo está como era antes e vai trabalhar. Durante sua patrulha, ele percebe alguém no mar pedindo por ajuda. Jeff não se imagina sendo elogiado por ninguém e simplesmente assiste a vítima se afogar, surpreso pelo fato de que ele simplesmente não deseja ajudá-la. (p. 204-205)

Assumindo que Jeff não é culpado por deixar de decidir salvar o banhista, Franklin afirma que a explicação é que esse agente não teve uma oportunidade razoável de agir diferentemente – isto é, não teve uma possibilidade alternativa robusta e, portanto, não foi livre para agir diferentemente. Ele não tinha ideia, nem deveria ter, da condição necessária para que ele decidisse resgatar a pessoa que se afogava. Analogamente, Jones não seria culpado por decidir votar a favor do corte, pela mesma razão do caso de Jeff. Esse ponto é facilmente negligenciado, afirma Franklin, pelo fato de que “a maioria das pessoas na posição de Jones teria a oportunidade de decidir contra o corte de imposto diretamente” (p. 205), sem precisar imaginar o próprio chefe descobrindo um voto favorável. A suspeita aqui seria de que parte da plausibilidade inicial da atribuição de culpa a Jones resulta de não termos em mente esse aspecto de sua psicologia.

Suponha agora que Jones está ciente de qual é o único modo de ele se sentir motivado a decidir votar contrariamente ao corte. Considerando que ele pode imaginar a descoberta da chefe voluntariamente, julgamos que existe culpa pela decisão tomada. Mas agora há de fato uma possibilidade alternativa robusta. Como coloca Franklin,

(...) seria uma desculpa muito ruim da parte de Jones apontar que imaginar a chefe não garantiria a decisão de votar contra o corte. A resposta óbvia é: “mas era necessário, você entendia isso e, portanto, deveria ter imaginado a descoberta da chefe; e então você deveria ter decidido votar contra o corte de imposto (...)”. (...) Afinal, se a oportunidade de agir diferentemente [nesse caso] era irrelevante do ponto de vista explanatório, então o fato de Jones ser

ou não cognitivamente sensível acerca dessa oportunidade (...) deveria também ser irrelevante. Mas, como vimos, isso não é irrelevante (...)" (p. 206)

A objeção de Franklin é precisa e representa mais um desafio importante aos Casos Estilo-Frankfurt indeterministas. Acredito, porém, que uma versão modificada de um Caso Zona-Tampão possa ser satisfatoriamente defendida – o que indicaria uma resiliência dos Casos frente aos defensores do PPA. Antes de apresentá-la, vejamos uma lição que penso que pode ser extraída da objeção de Franklin: a de que a componente epistêmica de Robustez precisa ser revisada.

Em *Sonegação Fiscal*, Pereboom sustentava que a possibilidade alternativa de voluntariamente dar o nível necessário de atenção a razões morais não era robusta devido à condição (ii) de Robustez: Joe não estaria cognitivamente sensível ao fato de que evitaria ser culpado por decidir sonegar impostos ao atingir esse nível de atenção porque, embora saiba que essa é uma condição necessária para decidir diferentemente, ele “acredita que atingir esse nível de atenção é compatível com sua livre decisão de sonegar impostos de qualquer jeito (o que seria verdade se o dispositivo da interventora não estivesse presente)” (2014, p. 16). A respeito de Corte de Imposto, Pereboom ofereceria a mesma razão para negar que a imaginação voluntária do cenário de punição seja uma possibilidade alternativa robusta. Franklin, no entanto, poderá insistir no ponto da suposição de que Jones estava ciente de essa ser o único modo pelo qual ele poderia cumprir sua obrigação – a de que fazer a alegação acima seria uma “desculpa ruim”.

Para melhor visualizar o que está em jogo, considere o seguinte exemplo. Se Sabina deve dinheiro a Teresa, e se, para pagar a dívida, ela deve evitar gastar o dinheiro que possui bebendo drinks em um bar na noite anterior, a alegação de que evitar gastar o dinheiro não era condição suficiente para o pagamento da dívida não fará com que Teresa deixe de culpar Sabina por não a ter pago, já que passou a noite bebendo drinks¹¹¹ – mesmo na presença de um dispositivo que, caso Sabina simplesmente passasse em frente ao bar sem fazer menção de entrar, engatilharia sua decisão de dar meia-volta e entrar. Suponha que o único caminho do trabalho de Sabina até a sua casa é passando em frente ao bar. Aqui, a agente tem uma possibilidade alternativa aparentemente robusta: a condição necessária é uma ação (passar em frente ao bar sem entrar) que ela pode realizar voluntariamente para ao menos tentar evitar a própria culpa por não pagar a dívida e, suponha, Sabina está ciente de que pode agir voluntariamente desse modo e de que *há uma probabilidade não-desprezível* de que ela evite

¹¹¹ Agradeço a Gabriel Maruchi pela discussão sobre esse exemplo.

ser culpada por não pagar a dívida se assim o fizer. Essa informação é verdadeira independentemente da presença do dispositivo; e estar ciente dela já parece suficiente para que uma possibilidade alternativa seja robusta, embora a exigência (ii) de Robustez fosse que a agente estivesse cognitivamente sensível ao fato de que, se cumprisse a condição necessária voluntariamente, “então não seria, ou provavelmente não seria, culpada” (PEREBOOM, 2014, p. 13). Penso que esse exemplo, porém, nos dá razões para enfraquecer Robustez do seguinte modo:

Robustez Revisada. “Para que uma agente tenha uma alternativa robusta à sua ação imoral A, isto é, uma alternativa per se relevante para explicar por que ela é culpada por realizar A, precisa ser o caso que

(i) ela poderia ter voluntariamente agido ou se refreado de agir, resultando na ausência de culpa,” (PEREBOOM, 2014, p. 13) e

(ii) para ao menos uma dessas coisas que ela poderia ter feito, ela era cognitivamente sensível ao fato de que poderia voluntariamente agir ou se refrear desse modo, e ao fato de que se ela o fizesse, então teria uma *probabilidade não-desprezível* de não ser culpada.

Vejamos como essa revisão gera o resultado correto acerca do caso de Sabina e, conseqüentemente, de Jones em Corte de Imposto. Independentemente de se o poderoso dispositivo da neurocientista está ou não presente em seu cérebro, é verdade que ela poderia ter voluntariamente agido de modo a evitar ser culpada: se o dispositivo está presente, ela poderia passar em frente ao bar sem fazer menção de entrar e isso engatilharia a ação do dispositivo, de modo que a decisão de dar meia-volta e entrar seria causada por ele, e não tomada pelas razões da agente; por outro lado, se não há dispositivo, Sabina poderia passar em frente ao bar sem entrar e, embora essa seja apenas uma condição necessária para sua ausência de culpa, haveria ainda outras ações que, se realizadas, resultariam em ausência de culpa (por exemplo, evitar fazer compras *online* assim que chegasse em casa etc.). O contexto de agência de Sabina, como dito acima, satisfaz a condição (i) estabelecida por Pereboom. O mesmo ocorre em Corte de Imposto: sabemos que se há dispositivo, Jones poderia ter imaginado o cenário em que a chefe o pune e então ele não seria culpado em virtude da interferência do dispositivo; se não há, podemos supor que, somadas à imaginação, haveria ainda outras coisas que o agente poderia voluntariamente fazer (como evitar voltar a pensar

nas razões auto-interessadas que enfraqueceriam sua motivação a votar contra o corte) e cuja consequência seria a ausência de culpa por um voto favorável ao corte.

O ponto importante aqui é a condição (ii): tanto Jones quanto Sabina satisfazem a condição epistêmica revisada. Na presença do dispositivo, Sabina não será culpada caso passe em frente ao bar sem entrar, mas ela não tem razões para estar ciente disso. Ela sabe, porém, que há uma probabilidade (apenas) não-desprezível de não ser culpada por deixar de pagar Teresa ao fazer isso porque ainda poderia gastar seu dinheiro fazendo compras *online* (e essa possibilidade faz com que não seja o caso que ela saiba que *provavelmente* não será culpada ao cumprir a condição necessária), embora deixar de entrar no bar seja o primeiro passo na direção de evitar a própria culpa. Estar ciente disso já é, como apontei, suficiente para que haja uma possibilidade alternativa robusta. O mesmo ocorre com Jones em relação à imaginação da chefe. Na presença do dispositivo, ele não está ciente de que não será culpado ao cumprir a condição necessária e, mesmo sem o dispositivo, também não pensa que *provavelmente* não será culpado ao cumpri-la porque (podemos supor) acredita que logo em seguida pode voltar a dar atenção às razões autointeressadas, enfraquecendo a motivação recém-adquirida de votar contra o corte. Jones pensa, porém, que há uma chance não-desprezível de que ele evite ser culpado por um voto favorável ao corte se imaginar sua chefe punindo-o. Portanto, Jones tinha uma possibilidade alternativa robusta e o caso Corte de Imposto, de Pereboom, não é um contraexemplo bem-sucedido ao Princípio das Possibilidades Alternativas. Quanto à suposição de que Jones está ciente da condição necessária, penso então que a objeção de Franklin está correta.

3.3.2 Um caso Zona-Tampão Modificado

Com o objetivo de revisar o caso Zona-Tampão de Pereboom perante o tipo de objeção que explorei acima, McKenna (2018) propôs que a condição necessária para a decisão correta seja estipulada como um evento *involuntário* – não como algo que pode ocorrer também voluntariamente. Seguindo Capes (2022), farei referência aos Casos com essa estrutura com o termo “casos *Zona-Tampão Modificados*”. Como descreve McKenna, podemos imaginar situações bastante realistas nas quais uma decisão em específico não é causalmente determinada e, se a possibilidade alternativa de decisão fosse efetivada, sua ocorrência exigiria o surgimento prévio de algum estado motivacional por meio de um processo não-voluntário e indeterminista. Ele enfatiza que sempre há alguma sorte nas razões para agir que nos vêm à mente ou nas motivações que se revelam mais fortes quando estamos deliberando sobre o que fazer, e isso pode depender de coisas tão triviais quanto a quantidade

de café que tomamos mais cedo. Além disso, ele afirma que “não deveria haver nada de problemático em supor que tal processo é (...) compatível com exercícios de (...) liberdade” (2018, p. 3125).

Se a imaginação da chefe só pudesse ocorrer a Jones involuntariamente, como resultado de um processo indeterminista, então esse novo Caso *não satisfaz a condição (i) de Robustez* (e de Robustez Revisada): não há nada que o agente pudesse ter feito voluntariamente que resultaria em sua ausência de culpa. Essa é a razão pela qual um Caso Zona-Tampão Modificado é mais promissor contra o PPA do que os casos Zona-Tampão de Pereboom - não precisamos apelar à condição epistêmica de Robustez (ou de Robustez Revisada), como faz Pereboom em Corte de Imposto, para justificar a afirmação de que a deixa para a intervenção do dispositivo não é uma possibilidade alternativa robusta.

No entanto, penso que essa proposta simples de revisão ainda é vulnerável a uma objeção importante, que considero motivo suficiente para refinamentos adicionais. Poderíamos agora perguntar a McKenna: as probabilidades objetivas envolvidas nesse processo indeterminista e involuntário *resultam de algum traço* (ou traços) da personalidade de Jones por cuja formação ele é responsável *em virtude de ter tido possibilidades alternativas naquela ocasião?* Se algo desse tipo puder ser afirmado, parece que a responsabilidade pela decisão favorável ao corte de imposto seria, mais uma vez, apenas derivada. Para que um Caso seja bem-sucedido contra o PPA, devemos então garantir que a responsabilidade em questão seja básica, não-derivada.

Para garantir que não estamos misturando intuições de responsabilidade básica com intuições de responsabilidade derivada por ações, decisões etc.¹¹², minha proposta é de que construamos *um caso Zona-Tampão Modificado que também é uma decisão formadora de caráter*. Aqui o agente (ainda) não tem responsabilidade por alguma parcela de seu caráter. Tomando emprestada alguma inspiração de Kane (mas sem endossar o libertismo), o que desejo sugerir para esse Caso refinado é que o agente seja agora uma pessoa jovem avaliando diferentes razões para a ação (como razões morais e razões de interesse próprio) e cuja decisão ajudará a *formar* tendências e disposições de seu futuro caráter moral.

Em especial, suponhamos que não existe nenhuma situação anterior em que esse agente, Jones, tenha decidido entre a ação moralmente correta para o contexto e uma outra ação apenas *um pouco errada* - uma vez que o corte de impostos é apenas modesto, como o próprio Pereboom estabelece em sua descrição de Corte de Imposto. Acredito que enfatizar este ponto seja relevante para mitigar suspeitas de que a decisão aqui em jogo é uma questão

¹¹² Agradeço a Gabriel Maruchi pela discussão sobre esse ponto.

de sorte, apesar das considerações de McKenna. Essa afirmação será explicada a seguir. Vejamos o caso Zona-Tampão Modificado que proponho como contraexemplo ao PPA:

Caso Estilo-Frankfurt *Decepção*. Tudo é igual a Corte de Imposto, exceto por dois fatos: (i) Jones é uma pessoa jovem formando sua agência moral; e (ii) a condição causalmente necessária (mas não suficiente) para a decisão de votar contra o corte de impostos (agindo então de maneira moralmente correta, como ele próprio reconhece) é que, ao considerar sobrepor seu interesse pessoal às razões morais, já que se trata afinal de um corte apenas modesto (portanto, menos errado do que um corte substancial), ele se sinta *decepcionado* consigo mesmo. Esse sentimento só pode ocorrer involuntariamente, como resultado de um processo indeterminista (embora causado, caso ocorra, pela sua receptividade e reatividade a esses tipos gerais de razões morais). A decisão que Jones está prestes a tomar ajudará ou a desenvolver ou a bloquear sua tolerância em relação às próprias ações que sejam ligeiramente erradas, uma vez que o corte de impostos em questão é apenas um pouco errado e essa é a primeira situação em que ele pode fazer algo ligeiramente errado sem o risco de ser descoberto (sozinho na cabine de votação). No caso de uma decisão favorável ao corte, ele desenvolveria então alguma tolerância a erros morais modestos pelo mesmo processo que ocorre com muitas pessoas: o privilégio não inteiramente voluntário ou deliberado de razões de interesse próprio sobre razões morais em sua economia mental, dadas circunstâncias favoráveis. Jones entende que o sentimento de desapontamento o colocará em uma posição motivacional apropriada para votar contra o corte. Sem que ele soubesse, porém, uma neurocientista implantou em seu cérebro um dispositivo que, caso detectasse a decepção, determinaria causalmente a decisão de votar a favor do corte de impostos. Mas Jones não chega a se sentir dessa maneira e decide votar a favor do corte por si próprio, enquanto o dispositivo permanece ocioso.

Em *Decepção*, não há nenhuma possibilidade alternativa robusta em aberto para Jones: não há nada que o agente pudesse ter feito voluntariamente e que evitasse sua culpa pelo voto a favor do corte. No entanto, ele ainda parece ser moralmente responsável, ou culpado, pela sua decisão. Portanto, o PPA é falso. Penso que não há nada de problemático em supor um cenário como *Decepção*; seus elementos compõem um conjunto tanto metafisicamente possível quanto, com exceção do poderoso dispositivo da neurocientista, razoavelmente realista e representativo de experiências humanas nas quais julgamos que existe

responsabilidade moral. Consequentemente, não deveria haver nada de problemático em supor que Jones é moralmente responsável por sua decisão nesse Caso.

Em sua crítica a Corte de Imposto, Franklin (2017) afirma que “se a oportunidade de agir diferentemente fosse realmente irrelevante do ponto de vista explanatório, então a sensibilidade cognitiva de Jones... [isto é, se ele está ciente da condição necessária] também deveria ser irrelevante. Mas, como vimos, não é irrelevante” (p. 206). Em Decepção, por outro lado, o estatuto epistêmico de Jones é irrelevante. Estipulando que ele está ciente de que se sentir decepcionado consigo mesmo era a única maneira pela qual ele poderia estar motivado o suficiente para decidir votar contra o corte de imposto, ainda não poderíamos *exigir* dele que se sentisse dessa maneira (já que não era algo que ele poderia realizar voluntariamente).

A suposição de que Jones não era cognitivamente sensível à condição necessária para fazer a escolha moral também seria irrelevante para a sua responsabilidade moral. Como as considerações de McKenna que mencionei acima ajudam a mostrar, a psicologia de Jones (tanto em Corte de Imposto quanto em Decepção) não é tão estranha como Franklin supõe. O facto de o corte em questão ser apenas modesto pode passar despercebido no caso de Pereboom, mas torna a decisão de o apoiar menos errada do que o apoio a um corte substancialmente maior. Este ponto contribui crucialmente, acredito, para que sua psicologia “estranha” não seja afinal tão estranha assim. Não parece incomum que uma pessoa disposta a agir de forma ligeiramente errada, com base em razões egoístas, só possa ser desmotivada por alguma razão prudencial ou associação de ideias e sentimentos bastante peculiares à sua psicologia, sem que tenha consciência disso. Ainda assim, nós tipicamente a responsabilizaríamos por essa ação, asseguradas outras condições necessárias para a responsabilidade moral. Ou, se o problema aqui é o fato de existir uma condição necessária única, não vejo razão para que não adicionemos mais estados motivacionais, estabelecendo então uma condição necessária disjuntiva para decidir favoravelmente ao corte. Dada essa possibilidade, e por uma questão de simplicidade, essa modificação não precisa ser efetivamente feita.

Nesta subsecção, apresentei meu próprio Caso Estilo-Frankfurt, Decepção, e argumentei que esse contraexemplo (ao carregar as características de um Caso Zona-Tampão Modificado e refiná-lo de modo a evitar preocupações com responsabilidade derivada) é capaz de responder à sofisticada objeção de Franklin. Há ainda, no entanto, ao menos outros dois trabalhos recentes e diretamente relevantes para uma defesa bem-sucedida desse Caso contra o PPA: a objeção de Justin Capes (2022) precisamente aos casos do tipo Zona-Tampão

Modificados e a ambiciosa objeção de Nadine Elzein (2017) a *qualquer* Caso Estilo-Frankfurt. Na próxima e última subseção, apresentarei esses argumentos e oferecerei respostas a ambos, obtendo assim o que espero ser uma defesa robusta de Decepção.

3.3.3 A irrelevância de possibilidades alternativas de ação para a responsabilidade moral

A partir do *insight* de Fischer de que uma possibilidade alternativa relevante deve ser robusta, Leon e Tognazzinni (2010) contrastaram uma definição tradicional de sucesso para um Caso Estilo-Frankfurt com uma definição mais fraca e que, segundo os autores, seria já suficiente. De acordo com a primeira, um Caso bem-sucedido descreve um cenário metafisicamente possível em que o agente é responsável, mas não poderia ter agido de outro modo (ou seja, um contraexemplo clássico). No entanto, poderíamos também aceitar como bem-sucedido um Caso em que, se o agente podia ter agido diferentemente, então essa capacidade é irrelevante para a sua responsabilidade moral. Em Decepção, por exemplo, Jones poderia ter ficado desapontado consigo mesmo; mas, dada a presença da neurocientista, essa possibilidade é irrelevante para sua responsabilidade pela decisão tomada. Esse cenário contaria então como um Caso bem-sucedido contra o PPA.

Elzein (2017) acredita, porém, que “não há exemplos [Estilo-Frankfurt], e provavelmente *não poderia* haver exemplos, que tenham sucesso, mesmo nesse sentido fraco” (p. 171). Ela argumenta que a aparência de força de qualquer Caso depende inteiramente de uma obscuridade sobre se as possibilidades alternativas envolvidas são de fato relevantes. Mas como a relevância do sentimento de decepção de Jones seria *obscura*? Elzein afirma que é fácil construir casos simples em que as alternativas deixadas em aberto são obviamente irrelevantes para a responsabilidade dos agentes, mas a consequência é que “não é nem um pouco óbvio que o agente possa ser responsabilizado” (p. 171). Seu argumento geral é que, para preservar a intuição de que existe responsabilidade, um defensor dos Casos deve construir um cenário contando com possibilidades alternativas *de natureza diferente* daquelas de um caso simples; mas então parece que essas alternativas já não são irrelevantes e, portanto, o PAP não é objetado. Vejamos um dos casos simples apresentados pela autora:

Caso Simples 2: Suponhamos que Brenda viva em um mundo completamente determinado, exceto por um evento indeterminado: em t , Brenda poderia ser atingida por um raio enviado pelos Deuses (...). Brenda definitivamente cortará os pneus de seu chefe em um ataque de ódio em $t1$, a menos que os Deuses a atinjam com um raio em t (o que irá feri-la gravemente, garantindo que ela nunca comece a deliberar sobre isso). Esse é o único modo pelo qual Brenda poderia evitar a decisão de cortar os pneus de seu chefe em $t1$. (ELZEIN, 2017, p. 173)

Elzein ressalta que dado que as decisões de Brenda são determinadas, exceto por aquela em que há a chance de que essa agente seja atingida por um raio (algo, é claro, fora de seu controle), não parece nada óbvio que ela seja responsável por cortar os pneus de seu chefe.

Aplicando o argumento de Elzein a Decepção, teríamos o seguinte dilema¹¹³. Ser atingida por um raio é claramente uma alternativa inacessível à agente, no sentido de ter controle sobre essa alternativa; e, se isso ocorrer, será, é claro, contra a vontade da agente. Na minha versão de um Caso Zona-Tampão Modificado, no entanto, Elzein diria que a alternativa do agente não é tão obviamente inacessível, dado que ter um certo sentimento pode não ser descrito como algo que meramente ocorre a uma pessoa contra sua vontade. Se descrevêssemos o exemplo para que a decepção fizesse parte do decidir ou do querer de Jones, não ficaria claro se esse sentimento de fato só poderia ocorrer involuntariamente¹¹⁴. Mas sob essa descrição, embora Jones pareça ser moralmente responsável por sua decisão, o caso não elimina as possibilidades alternativas acessíveis ao agente.

Poderíamos então estabelecer um cenário na direção oposta. Suponha que a decisão de Jones em favor do corte de impostos é inteiramente determinada, exceto por um evento indeterminado que talvez lhe ocorra: pode ser que, durante a sua deliberação na cabine de votação, ele sofra uma convulsão que perturbe sua atividade neural e o faça optar por votar contra o corte. O dispositivo da neurocientista ficaria então ocioso, a menos que ocorresse a convulsão. Mas isso não chega a acontecer. A respeito desse tipo de descrição de um Caso, Elzein argumenta então que “está longe de ser óbvio que Jones possa ser responsabilizado pela sua escolha”, na medida em que sua decisão foi “inteiramente determinada (tal como a escolha de Brenda em nosso Caso Simples 2), exceto pela chance de que algo que está inteiramente fora de seu controle lhe aconteça” (2017, p. 183-184).

Abordando especificamente o caso Zona-Tampão Modificado de McKenna, Capes (2022) também levantou a preocupação acima. Ele oferece um caso simples semelhante ao de Elzein e argumenta que, na versão de Corte de Imposto em que a imaginação é involuntária, Jones está na mesma situação. Embora McKenna (2018) insista que podemos tratá-lo como um agente responsável porque sua decisão flui de “recursos de agência típicos” (p. 3125) e porque ele poderia ter agido de outro modo fora do contexto de um Caso Estilo-Frankfurt, Capes vai ainda mais longe do que Elzein e responde que Jones não tinha controle sobre sua

¹¹³ Alguns dos Casos discutidos por Elzein no artigo são os de Fischer (2007) e Stump (1999), ambos com deixas para intervenção do dispositivo que são involuntárias para o agente. Minha aplicação da objeção a Decepção será baseada na discussão desses casos.

¹¹⁴ Elzein afirma que uma condição como essa poderia ser como um *esforço da vontade*, para utilizar a terminologia de Kane.

decisão em virtude de não ter controle sobre se o evento relevante de aumento de motivação ocorreria ou não. Isso seria verdade independentemente da presença da neurocientista. No restante desta seção, oferecerei respostas tanto a Elzein quanto a Capes.

Em primeiro lugar, é preciso retomar alguns elementos gerais sobre o papel dialético de um Caso Estilo-Frankfurt, no sentido dos passos que ele representa no debate do Problema da Compatibilidade. Como Fischer (1982) e outros autores reconhecem há bastante tempo, os Casos não são suficientes para afirmar que a liberdade e a responsabilidade moral são compatíveis com o determinismo causal. Ele afirma, como mencionei anteriormente, que a “moral da história” é que *se* o determinismo exclui a responsabilidade moral, não é por eliminar possibilidades alternativas. Levando a sério essa afirmação, eu suspeitaria de qualquer tentativa de construir Casos explicitamente deterministas pela razão apresentada na subseção 3.2.1: quando especificamos um exemplo que resulte nessa “moral”, a presença da neurocientista e seu poderoso dispositivo no cérebro do agente está desempenhando o papel do que parece ser uma variável isolada, ou aspecto, do determinismo - o fato (plausível) de que ele elimina possibilidades alternativas. Esse é, como argumentei, o único aspecto considerado pelos Casos. É importante enfatizar esse ponto porque há alguns argumentos, como o Argumento Direto (VAN INWAGEN, 1983), segundo os quais o determinismo elimina a responsabilidade moral *diretamente*, sem o passo intermediário de eliminar possibilidades alternativas. Nos termos com os quais estivemos trabalhando desde o primeiro capítulo, há argumentos segundo os quais o determinismo ameaça a liberdade de uma agente ser a fonte genuína da própria ação, sem que seja considerada sua liberdade de agir diferentemente. Não faria sentido, portanto, assumir a verdade do determinismo em um experimento mental cujo papel é testar nossa intuição de um único elemento dessa tese, adequadamente substituído pela figura da neurocientista e, portanto, representado de uma maneira teoricamente neutra.

Agora podemos avaliar o que há de inadequado no argumento de Elzein. Mesmo sem levar em conta a capacidade de Jones de decidir diferentemente, mesmo considerando a capacidade do agente de agir de outra forma, poderíamos gerar a intuição de que ele não é moralmente responsável por sua ação chamando a atenção para a fonte (talvez) “contaminada” de sua ação caso o determinismo seja verdadeiro: dele se seguiria que qualquer ação é uma consequência do passado remoto e das leis da natureza, pelos quais ninguém é moralmente responsável; ninguém seria, então, moralmente responsável por qualquer ação (como estabelece o Argumento Direto). Quando questionamos uma pessoa não-familiarizada com o debate da Compatibilidade sobre se a responsabilidade moral é ou

não compatível com o determinismo, essa intuição é parte do que explica uma possível resposta negativa. Mas observe que Elzein faz uma suposição crucial em seu Caso Simples 2: ela supõe que, anteriormente à alternativa irrelevante em questão, o mundo de Brenda era causalmente determinado. Isso significaria que, embora sua decisão específica de cortar os pneus de seu chefe não tenha sido determinada, suas motivações, traços de personalidade e crenças *foram* determinadas. Por essa razão, penso que a comparação entre Jones em Decepção (ou o agente em qualquer outro caso Zona-Tampão Modificado) e Brenda no Caso Simples 2 é inadequada.

Para que a relevância desse ponto seja mais bem visualizada, considere a inferência na qual Elzein diz estar se apoiando: “1. Mudar a natureza das alternativas em jogo (e nada mais) [de um Caso Simples para um Caso Estilo-Frankfurt] faz toda a diferença no que diz respeito à responsabilidade do agente. [Portanto,] 2. As alternativas em jogo provavelmente são relevantes de alguma forma para a responsabilidade do agente.” (p. 186). O que estou argumentando aqui é que a premissa dessa inferência é falsa. Algo mais difere Decepção de Caso Simples 2 – a suposição de determinação anterior. Considere a seguinte versão da história de Brenda:

Caso Simples Modificado. Brenda vive em um mundo como parece ser o nosso, no qual o determinismo causal é *falso*. Em t , Brenda poderia ser atingida por um raio enviado pelos Deuses. Ela está tentada a cortar os pneus de seu chefe em $t1$. Sua raiva é tão intensa e as circunstâncias são tão favoráveis (Brenda pode agir sem ser descoberta e punida) que ela definitivamente tomará essa decisão, *mantendo fixo o mesmo conjunto de razões para agir* (incluindo as circunstâncias favoráveis), a menos que os Deuses a atinjam com um raio em t . Dado que os deuses não a atingem com um raio, Brenda corta os pneus do chefe.

Em Caso Simples Modificado, torna-se muito mais fácil ter a intuição de que Brenda é moralmente responsável. Não há nada (como o determinismo) “contaminando” a fonte de suas ações, seus motivos e disposições, de modo que não negaríamos que essas coisas são genuinamente dela. Em especial, não parece haver compulsão envolvida. Dado que Brenda definitivamente tomará a decisão de cortar os pneus do chefe apenas com base no exato conjunto de razões para agir (que envolvem sua raiva e as circunstâncias muito favoráveis) existente no mundo atual, não há nenhuma razão pela qual não poderíamos afirmar que ela é moderadamente responsiva a razões: a condição de responsividade a razões de Fischer e

Ravizza estabelecia que o agente precisava reagir diferentemente em pelo menos um mundo possível em que *outra* razão suficiente para agir de outro modo estivesse presente¹¹⁵. Essa agente, portanto, parece satisfazer ao menos uma condição compatibilista plausível para a responsabilidade moral, além de tomar sua decisão em um mundo indeterminista.

Por essas razões, a afirmação de que ela seja responsável é ao menos tão plausível quanto a afirmação de que Jones é responsável por sua decisão em Decepção; e não é enfraquecida pelo fato de que a única alternativa à sua decisão é uma possibilidade irrelevante. Sua motivação era incrivelmente forte, mas era também *sua*. Consequentemente, temos um caso simples em que a possibilidade alternativa deixada em aberto é obviamente irrelevante, não robusta, no qual existe responsabilidade moral. Esse tipo de caso só poderia ser utilizado contra os Casos Estilo-Frankfurt caso fizéssemos, como Elzein, uma suposição inadequada em sua construção.

Passemos agora à objeção de Capes (2022). Ele sustenta que um agente em um caso Zona-Tampão Modificado não possui controle (e, portanto, não é responsável), em vez de simplesmente apontar obscuridade a esse respeito. A razão era que o agente “não tem em seu poder trazer essa alternativa à tona de maneira ativa” (p. 719), dado que a condição necessária para sair do “caminho da má ação” não é acessível voluntariamente. Curiosamente, Capes menciona em uma nota de rodapé que, segundo um parecerista, sua noção de controle é tão forte que tem a consequência implausível de que nós frequentemente não temos o controle suficiente para responsabilidade, apesar das aparências. Capes responde que esse resultado é muito menos implausível do que parece, mas que fundamentar essa afirmação exigiria um outro artigo. O autor admite, então, que está empregando uma noção de controle mais forte do que a intuitiva. Este ponto será crucial para a minha resposta.

Capes afirma que um agente moralmente responsável deve ter tido controle sobre se uma condição necessária substancial para sua decisão/ação ocorreria. No entanto, como apontei anteriormente, não parece muito incomum que uma pessoa disposta a agir de maneira ligeiramente errada só possa ser desmotivada por alguma razão particularmente atraente para sua psicologia - uma razão cuja ocorrência pode ser involuntária e indeterminada, ou uma questão de sorte. Mesmo nessas condições, nós a responsabilizariamos por sua ação. Admitindo então que a noção de controle de Capes seja realmente forte, penso que ela não pode ser empregada em uma discussão sobre os Casos Estilo-Frankfurt sem a argumentação adicional que Capes sabe estar ausente. Retomando minhas considerações sobre o papel dialético de um Caso, Capes precisaria admitir que ao menos as nossas avaliações sobre a

¹¹⁵ Agradeço a Beatriz Sorrentino Marques pela discussão sobre esse exemplo.

presença de responsabilidade num caso como esse (independente daquilo que pensamos que a fundamenta) são um tanto “pré-teóricas”. Se basearmos conscientemente o nosso julgamento sobre se Jones é culpado em uma noção de controle que difere de outra mais intuitiva e que exige argumentos adicionais, então estaremos misturando inadequadamente o nível pré-teórico dessa discussão a compromissos teóricos não-argumentados.

Além disso, se (i) estou certa sobre o quão realista é a situação de Jones em Decepção, se (ii) a responsabilidade moral realmente exige, como afirma Capes, o controle sobre a ocorrência das condições necessárias substanciais para a decisão tomada e, por fim, se (iii) nós de fato temos a intuição de que Jones é moralmente responsável, então o defensor do PPA teria dificuldade em explicar como Jones pode ser livre e responsável se eliminarmos o neurocientista do cenário. Mas isso não deveria ser um problema para essa posição: Jones teria então uma possibilidade alternativa robusta, sendo capaz de agir voluntariamente de modo a evitar a própria culpa (isto é, a capacidade de votar contra o corte de impostos; aqui a probabilidade de que ele obtenha esse resultado, podemos supor, é alta). Se a sorte na ocorrência do sentimento necessário para essa decisão é um problema para os casos Zona-Tampão Modificados, acredito que seja um problema também para os defensores do PPA. Não é nada claro, afinal, como seria possível uma agente decidir de maneira direta, “puramente” escolhendo “de um modo ou de outro dada sua configuração motivacional precisa” (MCKENNA, 2018, p. 3126).

Nesse capítulo, discuti o papel dos Casos Estilo-Frankfurt no contexto do Problema da Compatibilidade e defendi uma formulação específica contra diferentes objeções. Em primeiro lugar, argumentei que um Caso bem-sucedido precisa ser construído assumindo um mundo possível indeterminista. Em seguida, considerei a objeção segundo a qual a responsabilidade moral que pensamos existir em um Caso Zona-Tampão é apenas derivada, em virtude de possibilidades alternativas deixadas em aberto. Argumentei que Franklin tem o mérito de apontar uma possibilidade alternativa realmente existente em Corte de Imposto; e sugeri uma revisão para Robustez. Em Robustez Revisada, a componente epistêmica é mais fraca do que era assumido por Pereboom. Endossei então a proposta de McKenna de um caso Zona-Tampão Modificado, em que a condição necessária para a decisão moralmente correta é um evento involuntário. Meu próprio caso, Decepção, é construído para lidar com um potencial problema para McKenna quanto à garantia de que estamos tratando de responsabilidade básica. Por fim, argumentei que as objeções de Elzein e Capes, cada uma a seu modo, violam a dialética em torno dos Casos Estilo-Frankfurt. Acredito que Decepção seja um contraexemplo promissor ao Princípio das Possibilidades Alternativas, à luz do

poderoso *insight* de Harry Frankfurt. Assim, concluo que a existência de possibilidades alternativas de ação disponíveis a uma agente é irrelevante para sua responsabilidade moral pela ação efetivamente realizada.

5 CONCLUSÃO

Neste trabalho, meus objetivos eram os seguintes:

- (1) apresentar as noções de liberdade e de responsabilidade moral, bem como o vínculo entre elas, tais como trabalhadas na Metafísica do Livre-Arbítrio e Filosofia da Ação contemporâneas;
- (2) apresentar o Problema da Compatibilidade em sua formulação contemporânea e as principais posições em disputa, o compatibilismo e o incompatibilismo, oferecendo um panorama de argumentos influentes e alguns comentários sobre a discussão;
- (3) defender um Caso Estilo-Frankfurt do tipo Zona-Tampão Modificado que chamei de *Decepção*, tendo em vista objeções recentes às formulações mais sofisticadas dos Casos.

No capítulo 1, *O vínculo entre liberdade e responsabilidade moral*, comecei por apresentar distinções e pressupostos da discussão filosófica sobre aquilo que era minha preocupação fundamental: a natureza da liberdade que parece estar associada à responsabilidade moral de uma agente por uma ação, bem como os fatores que talvez representem uma ameaça à existência dessa liberdade. Na seção 1.2, fiz uma apresentação ampla das condições necessárias e variedades da responsabilidade, considerando circunstâncias em que atenuamos ou suspendemos nossos juízos sobre sua existência. Já na seção 1.3, discuti a noção de liberdade e a investigação de sua natureza sob o recorte do Problema da Compatibilidade. Ao final dessa seção, tratei de uma importante contribuição para o incompatibilismo, o Argumento da Consequência, e de algumas estratégias compatibilistas para objetá-lo. O capítulo foi concluído com o questionamento sobre se o Argumento da Consequência de fato desempenha um papel decisivo em favor do incompatibilismo. Apontei então para uma estratégia disponível ao compatibilista (introduzida historicamente, embora não logicamente, pelo argumento dos Casos Estilo-Frankfurt): fazer uso da noção de liberdade da fonte genuína em vez da liberdade de agir diferentemente.

No capítulo 2, *Compatibilistas e libertistas entregues à própria sorte*, apresentei o modo como o indeterminismo também ameaça a liberdade. Sugeri brevemente a seguinte interpretação do problema colocado pelo Dilema do Determinismo: sendo o determinismo verdadeiro ou falso, parece que não há uma condição C substantiva e necessária para a realização de uma ação em específico (em vez de qualquer outra) com a qual a agente contribua voluntariamente e para a qual não existam condições suficientes independentes de sua agência (isto é, não há a liberdade de uma agente ser a fonte genuína da própria ação).

Aponte que tanto compatibilistas quanto libertistas precisam oferecer uma resposta a essa alegação e listei duas maneiras de realizar essa tarefa. A primeira era negar que uma decisão tomada em um mundo indeterminista seja questão de sorte/acaso, desde que satisfeitos certos requisitos – essa é a alternativa disponível ao libertista. A segunda, disponível ao compatibilista, era rejeitar a formulação da noção de liberdade da fonte genuína utilizada até esse ponto do trabalho e argumentar que uma outra, mais fraca, é tudo o que a responsabilidade moral exige.

Em seguida, apresentei de que maneira alguns dos principais libertistas e compatibilistas contemporâneos empregam essas estratégias. Na seção 2.2, explorei algumas versões do compatibilismo baseado em controle, as Visões do Eu Profundo e o Semicompatibilismo de John Fischer e Mark Ravizza, e mencionei aquelas que acredito serem as principais críticas já feitas a essas teorias. Já na seção 2.3, ofereci algum contexto a respeito da posição libertista e explorei em maior profundidade o libertismo da causalidade de eventos de Robert Kane, incluindo também algumas das principais críticas direcionadas à teoria.

Ao final desse capítulo, ofereci um comentário sobre o que poderíamos concluir do debate apresentado. A partir da importante contribuição de Mirja Pérez de Calleja, segundo a qual a sorte é também um problema para o compatibilista (apesar de ser amplamente considerada o principal problema para os libertistas), argumentei que (i) se o compatibilista tiver uma boa resposta para o Argumento da Manipulação, sua oposição ao libertismo será então suplementada não por um endosso do Problema da Sorte (incluindo a Objeção do Manipulador Randômico), mas pelo endosso do Problema do Controle Não-aumentado; e que (ii) a Objeção da Compensação não poderá ser utilizada pelo compatibilista para a defesa do Problema do Controle Não-aumentado, mas a Objeção do Continuum é uma boa opção para satisfazer essa finalidade. Também chamei a atenção para a alternativa, com a qual concordo, sugerida por Fischer contra o libertismo de Kane: não há uma razão pela qual não possamos transformar uma ação autoformadora em um Caso Estilo-Frankfurt.

No terceiro e último capítulo, *Casos Estilo-Frankfurt e a irrelevância de possibilidades alternativas de ação para a responsabilidade moral*, meus objetivos eram apresentar uma parte dos principais argumentos da discussão sobre os Casos Estilo-Frankfurt e defender que há um estilo de formulação, os casos Zona-Tampão *Modificados*, capaz de resistir a várias objeções importantes. Embora esse estilo de Caso tenha sido anteriormente proposto por Michael McKenna (como uma revisão dos casos Zona-Tampão de Derk Pereboom), as razões que apresentei para endossá-lo e alguns detalhes do caso que propus, *Decepção*, são contribuições originais para o debate.

Os Casos Estilo-Frankfurt são uma classe de contraexemplos ao Princípio das Possibilidades Alternativas propostos em diferentes graus de complexidade e vastamente discutidos na literatura contemporânea. Na seção 3.2, apresentei aquela que é a mais importante objeção aos Casos, dada sua completude: a Defesa do Dilema. Em primeiro lugar, discuti a suposição determinista do dilema e critiquei a posição de Fischer de que, embora um Caso não precise ser construído supondo o determinismo, não há nenhuma razão pela qual isso não possa ser feito. Defendi que, na verdade, um Caso bem-sucedido precisa ser construído assumindo o indeterminismo. Em seguida, apresentei os casos Zona-Tampão de Pereboom e uma das críticas recebidas por essa proposta.

Na seção 3.3, comecei por considerar uma outra objeção, segundo a qual a responsabilidade moral que pensamos existir em um Caso Zona-Tampão é apenas derivada. Essa objeção, originalmente proposta por Widerker, foi endossada por McKenna e independentemente desenvolvida por Christopher Franklin. Argumentei que a versão de Franklin tem o mérito de apontar uma possibilidade alternativa de fato deixada em aberto naquele tipo de Caso e que ela nos leva a uma revisão para a condição de Robustez (para uma possibilidade alternativa) proposta por Pereboom. Endossei então a proposta de McKenna de um caso Zona-Tampão Modificado, em que a condição necessária para a decisão moralmente correta é um evento involuntário. Meu caso Decepção, a parte mais importante do capítulo, é construído para lidar com um potencial problema para a versão de McKenna, além de não ser vulnerável à objeção de Franklin.

Para que a defesa de Decepção fosse mais robusta, lidei ainda com outras duas objeções, propostas por Nadine Elzein e Justin Capes, a Casos Estilo-Frankfurt semelhantes ao meu. Argumentei por fim que as objeções desses autores violam a dialética em torno dos Casos. Pelas razões expostas ao longo da seção final do último capítulo, acredito que Decepção seja um contraexemplo promissor ao Princípio das Possibilidades Alternativas. Assim, a conclusão mais substancial deste trabalho é a seguinte: a existência de possibilidades alternativas disponíveis a uma agente é irrelevante para sua responsabilidade moral pela ação realizada.

REFERÊNCIAS

- ARISTOTLE. **The Nicomachean Ethics**. Tradução: David Ross. New York: Oxford University Press, 2009.
- ARPALY, N. **Unprincipled Virtue**. Oxford: Oxford University Press, 2003.
- BALAGUER, M. **Free Will as an Open Scientific Problem**. The MIT Press: 2010.
- BALAGUER, M. Why there are no good arguments for any interesting version of determinism. **Synthese**, 168. 2009. p. 1-21.
- BEEBEE, H. & MELE, A. Humean Compatibilism. **Mind**, 111(442), 2002. p. 201–23.
- BOURGET, D. & CHALMERS, D. Philosophers on Philosophy: The 2020 PhilPapers Survey. **Philosophers' Imprint**, 23(0), 2023.
- BRINK, D. & NELKIN, D. Fairness and the Architecture of Responsibility. In: SHOEMAKER, D. (ed.). **Oxford Studies in Agency and Responsibility**, vol. 1. Oxford: Oxford University Press, 2013. p. 284–313.
- CAPEL, J. Against (modified) buffer cases. **Philosophical Studies**, 179, 2022. p. 711-723.
- CAROL, J. W. MARKOSIAN, N. Freedom and Determinism. In: ____ **An Introduction to Metaphysics**. Cambridge University Press: 2010. p. 45-79.
- CARUSO, G. Justice without retribution: An epistemic argument against retributive criminal punishment. **Neuroethics**, v. 13, n. 1, 2020.
- CARTER, I. Positive and Negative Liberty. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Spring 2022 Edition). Disponível em: <<https://plato.stanford.edu/archives/spr2022/entries/liberty-positive-negative/>>.
- CLARKE, R. Dispositions, Abilities to Act, and Free Will: The New Dispositionalism. **Mind**, 118, 2009. p. 323–51.
- CLARKE, R. CAPEL, J. SWENSON, P. Incompatibilist (Nondeterministic) Theories of Free Will. **The Stanford Encyclopedia of Philosophy** (Fall 2021 Edition). In: ZALTA, Edward N. (Ed.). Disponível em: <<https://plato.stanford.edu/archives/fall2021/entries/incompatibilism-theories/>>. Acesso em: ago. 2022.
- COATES, J. MCKENNA, M. Compatibilism. **The Stanford Encyclopedia of Philosophy** (Winter 2018 Edition). In: ZALTA, Edward N. (Ed.). Disponível em: <<https://plato.stanford.edu/archives/win2018/entries/compatibilism/>>. Acesso em: jan. 2019.
- DAMASIO, A. **O erro de Descartes: emoção, razão e o cérebro humano**. Tradução: Dora Vicente, Georgina Segurado. 3a ed. São Paulo : Companhia das Letras, 2012.
- EARMAN, J. **A primer on determinism**. D. Reidel Publishing Company: 1986. p. 235-251.
- ELZEIN, N. Frankfurt-style counterexamples and the importance of alternative Possibilities. **Acta Analytica**, 32, 2017. p. 169–191.
- FISCHER, J. **The Metaphysics of Free Will: A Study of Control**. Oxford: Blackwell, 1994.
- FISCHER, J. **Deep Control: Essays on Free Will and Value**. New York: Oxford University Press, 2012.
- FISCHER, J. Compatibilism. In: FISCHER, J. KANE, R. PEREBOOM, D. VARGAS, M. **Four Views on Free Will**. Blackwell Publishing, 2007.

- FISCHER, J. Responsibility and Control. **The Journal of Philosophy**, vol. 79, n. 1, 1982.
- FISCHER, J. KANE, R. PEREBOOM, D. VARGAS, M. **Four Views on Free Will**. Blackwell Publishing, 2007.
- FISCHER, J. & RAVIZZA, M. **Responsibility and Control: a theory of moral responsibility**. Cambridge University Press: 1998.
- FRANKLIN, C. **A Minimal Libertarianism: Free Will and the Promise of Reduction**. New York: Oxford University Press, 2017.
- FRANKLIN, C. Farewell to the Luck (and Mind) Argument. **Philosophical Studies**, 156, 2011. p. 199–230.
- FRANKFURT, H. Alternate Possibilities and Moral Responsibility. **Journal of Philosophy**, 66, 1969. p. 829–39.
- FRANKFURT, H. Freedom of the Will and the Concept of a Person. **Journal of Philosophy**, 68. 1971. p. 5–20.
- FRICKER, M. **Epistemic Injustice: Power and Ethics of Knowing**. Oxford: Oxford University Press, 2007.
- GINET, C. In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing. **Philosophical Perspectives**, v. 10, 1996, p. 403–417.
- GINET, C. Might we have no choice? In: LEHRER, K. (ed.). **Freedom and Determinism**. New York: Random House, 1966. p. 87-104.
- GINET, C. Review of Living without Free Will. **Journal of Ethics**, 6, 2002. p. 305–9.
- GINET, C. Working with Fischer and Ravizza's Account of Moral Responsibility. **Journal of Ethics**, 10, 2006.
- GOETZ, Stewart. Frankfurt-Style Counterexamples and Begging the Question. **Midwest Studies in Philosophy**, v. 29, n. 1, 2005. p. 83-105.
- HAJI, I. **Moral Appraisability**. New York: Oxford University Press, 1998.
- HARTMAN, R. **In Defense of Moral Luck: Why Luck Often Affects Praiseworthiness and Blameworthiness**. New York: Routledge, 2017.
- HOBBS, T. **The English Works of Thomas Hobbes of Malmesbury**, Vol. 3: Leviathan. London: John Bohn, 1839.
- HOEFER, C. Causal Determinism. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Spring 2016 Edition). Disponível em: <<https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>>.
- HUME, D. **Tratado da Natureza Humana: Uma tentativa de introduzir o método experimental de raciocínio nos assuntos morais**. São Paulo: UNESP, 2009.
- ISMAEL, J. **How Physics Makes Us Free**. Oxford: Oxford University Press, 2016.
- JAMES, W. The Dilemma of Determinism. In: _____. **The Will to Believe and Other Essays in Popular Philosophy**. New York: Longmans Green and Co., 1907.
- JEPPSSON, S. Accountability, Answerability, and Attributability: On Different Kinds of Responsibility. In: NELKIN, D. & PEREBOOM, D. **The Oxford Handbook of Moral Responsibility**. New York: Oxford University Press, 2022.

- JEPPSSON, S. Accountability, Answerability and Freedom. **Social Theory and Practice**, 42, 2016. p. 681–705.
- KANE, R. Libertarianism. In: FISCHER, J. KANE, R. PEREBOOM, D. VARGAS, M. **Four Views on Free Will**. Blackwell Publishing, 2007.
- KANE, R. On the Role of Indeterminism in Libertarian Free Will. **Philosophical Explorations**, 19, 2016. p. 2–16.
- KANE, R. **The Significance of Free Will**. New York: Oxford University Press, 1996.
- KANT, I. **Crítica da Razão Prática**. São Paulo: Martins Fontes, 2016.
- LAPLACE, P. Determinism, Ignorance, and Probability. In: MCGREW, T.; ALSPECTOR-KELLY, M.; ALLHOF, F. (eds.). **The Philosophy of Science: An Historical Anthology**. Oxford: Wiley-Blackwell, 2009, p. 250-254.
- LEON, F. TOGNAZZINI, N. Why Frankfurt-examples don't need to succeed to succeed. **Philosophy and Phenomenological Research**, 80, 2010. p. 551–565.
- LEVY, N. **Hard Luck: How Luck Undermines Free Will and Moral Responsibility**. New York: Oxford University Press, 2011.
- LEWIS, D. Are We Free to Break the Laws? **Theoria**, 47, 1983. p. 113–21.
- LOCKE, J. **Ensaio acerca do entendimento humano**. Tradução de Anoar Aiex. São Paulo: Editora Nova Cultural, 1999.
- LOWE, E. J. **A Survey on Metaphysics**. Oxford University Press, 2002.
- MCKAY, T. & JOHNSON, D. A Reconsideration of an Argument against Incompatibilism. **Philosophical Topics**, 24, 1996. p. 113–22.
- MCKENNA, M. A Critical Assessment of Pereboom's Frankfurt-Style Example. **Philosophical Studies**. 175, 2018. p. 3117-29.
- MCKENNA, M. **Conversation and Responsibility**. Oxford University Press: 2012.
- MCKENNA, M. Reasons-Responsiveness, Agents and Mechanisms. In: SHOEMAKER, D. (ed.). **Oxford Studies in Agency and Responsibility**. v. 1. Oxford: Oxford University Press, 2013. p. 151–83.
- MCKENNA, Michael. Resisting the Manipulation Argument: A Hard-Liner Takes It on the Chin. **Philosophy and Phenomenological Research**, v. 89, n. 2, 2014.
- MCKENNA, M. & PEREBOOM, D. **Free will: a contemporary introduction**. New York: Routledge, 2016.
- MELE, A. **Free Will and Luck**. Oxford: Oxford University Press, 2006.
- MELE, A. Moral Responsibility for Actions: Epistemic and Freedom Conditions. **Philosophical Explorations**. v. 13, n. 2, 2011. p. 101–11.
- MELE, A. Motivation and Intention. **Journal of Philosophical Research**, 21, 1996. p. 51–67.
- MELE, A. (org.). **Surrounding Free Will**. New York: Oxford University Press, 2015. p. 255-273.
- MELE, A. & ROBB, D. Rescuing Frankfurt-Style Cases. **The Philosophical Review**, v. 107, n. 1, 1998, p. 97-112.
- MERLUSSI, P. **O problema do livre-arbítrio e do determinismo: uma defesa do incompatibilismo**. Dissertação submetida ao Programa de Pós-Graduação em Filosofia da

Universidade Federal de Santa Catarina para a obtenção do grau de Mestre em Filosofia. Florianópolis, 2013.

MORRIS, W. & BROWN, C. David Hume. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Summer 2022 Edition). Disponível em: <<https://plato.stanford.edu/archives/sum2022/entries/hume/>>.

MURCHO, D. **Lógica Elementar**: Raciocínio, linguagem e realidade. Lisboa: Edições 70, 2019.

NELKIN, D. **Making Sense of Freedom and Moral Responsibility**. Oxford: Oxford University Press, 2011.

NELKIN, D. Moral Luck. In: ZALTA, E. (ed.) **The Stanford Encyclopedia of Philosophy**, 2013. Disponível em: <<http://plato.stanford.edu/archives/win2013/entries/moral-luck/>>.

NELKIN, D. & RICKLESS, S. Moral Responsibility for Unwitting Omissions: A New Tracing View. In: ____ (eds.). **The Ethics and Law of Omissions**. Oxford: Oxford University Press, 2017.

NELKIN, D. & PEREBOOM, D. (eds.). **The Oxford Handbook of Moral Responsibility**. New York: Oxford University Press, 2022.

O'CONNOR, T. Agent-Causal Power. In: HANFIELD, T. (ed.). **Dispositions and Causes**. Oxford: Clarendon Press, 2009, p. 189–214.

PEREBOOM, D. A Notion of Moral Responsibility Immune to the Threat from Causal Determination. In: CLARKE, R. MCKENNA, M. SMITH, A. **The Nature of Moral Responsibility**. Oxford University Press: 2015.

PEREBOOM, D. **Free Will** (Elements in Philosophy of Mind). Cambridge: Cambridge University Press, 2022.

PEREBOOM, D. **Free Will, Agency, and Meaning in Life**. Oxford: Oxford University Press, 2014.

PEREBOOM, D. Hard Incompatibilism. In: FISCHER, J. KANE, R. PEREBOOM, D. VARGAS, M. **Four Views on Free Will**. Blackwell Publishing, 2007.

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

PÉREZ DE CALLEJA, M. Cross-world luck at the time of decision is a problem for compatibilists as well. **Philosophical Explorations**, v. 17, n. 2, 2014.

PRATES, V. Pai se revolta com bloqueio de bolsonaristas que impedia cirurgia do filho. **Correio Braziliense**, 2022. Disponível em: <https://www.correio braziliense.com.br/brasil/2022/11/5053900-pai-se-revolta-com-bloqueio-d-e-bolsonaristas-que-impedia-cirurgia-do-filho.html>.

RODRIGUES, L. **O semicompatibilismo como resposta ao problema do determinismo e da liberdade**. Trabalho de conclusão do curso apresentado ao Instituto de Filosofia da Universidade Federal de Uberlândia como requisito parcial para obtenção dos títulos de Bacharela e Licenciada em Filosofia. Uberlândia, 2020.

RUDY-HILLER, F. The Epistemic Condition for Moral Responsibility. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Fall 2018 Edition).

SARTORIO, C. **Causation and Free Will**. Oxford University Press: 2016.

- SCHLICK, M. When is a Man Responsible? In: BEROFSKY, B. (ed.). **Free Will and Determinism**. New York: Harper & Row, 1966. p. 54–63.
- SMITH, A. Attributability, answerability, and accountability: in defense of a unified account. **Ethics**, v. 122, n.3. 2012.
- SHOEMAKER, D. Attributability, answerability, and accountability: toward a wider theory of moral responsibility. **Ethics**, v. 121, n. 3, 2011. p. 602-632.
- STRAWSON, G. The Impossibility of Moral Responsibility. **Philosophical Studies**, 75, 1994. p. 5–24.
- STRAWSON, P. F. **Freedom and Resentment and Other Essays**. New York: Routledge, 2008.
- STRAWSON, P. F. Liberdade e ressentimento. In: CONTE, J. (org.). GELAIN, I. (org.). **Ensaio sobre a filosofia de Strawson**. Florianópolis: Editora da UFSC, 2015.
- STUMP, E. Alternative possibilities and moral responsibility: the flicker of freedom. **The Journal of Ethics**, 1999, p. 299–324.
- TALBERT, M. Moral Responsibility. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Fall 2022 Edition).
- TALBERT, M. Unwitting Wrongdoers and the Role of Moral Disagreement in Blame. In: SHOEMAKER, D. (ed.). **Oxford Studies in Agency and Responsibility**, Vol. 1. Oxford: Oxford University Press, 2013.
- VAN INWAGEN, P. **An Essay on Free Will**, Oxford: Clarendon Press, 1983.
- VAN INWAGEN, P. Free Will Remains a Mystery. In: TOMBERLIN, J. (ed.). **Action and Freedom** (Philosophical Perspectives, a supplement to *Noûs*, vol. 14). Malden: Blackwell Publishing, 2000.
- VARGAS, M. **Building Better Beings: A Theory of Moral Responsibility**. New York: Oxford University Press, 2013.
- VIHVELIN, K. Arguments for Incompatibilism. In: ZALTA, E. (ed.). **The Stanford Encyclopedia of Philosophy** (Fall 2022 Edition). Disponível em: <<https://plato.stanford.edu/archives/fall2022/entries/incompatibilism-arguments/>>.
- WATSON, G. **Agency and Answerability: Selected Essays**. New York: Oxford University Press, 2004.
- WATSON, G. Two Faces of Responsibility. **Philosophical Topics**, v. 24, n. 2, 1996, p. 227–248.
- WIDERKER, D. Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. **The Philosophical Review**, v. 104, n. 2, 1995, p. 247-261.
- WIDERKER, D. Libertarianism and the philosophical significance of Frankfurt scenarios. **Journal of Philosophy**, 103, 2006, p. 163–187.
- WOLF, S. **Freedom within Reason**. New York: Oxford University Press, 1990.
- WOLF, S. Sanity and the Metaphysics of Responsibility. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 372-387.
- ZIMMERMAN, M. Moral Responsibility and Ignorance. **Ethics**, 107, 3, 1997.