

**UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA
NÍVEL DOUTORADO**

TANIA SCHNEIDER DA FONSECA

**UMA TEORIA SEMICOMPATIBILISTA SOBRE RESPONSABILIDADE MORAL:
John Fischer e o controle de direcionamento**

São Leopoldo

2018

TANIA SCHNEIDER DA FONSECA

**UMA TEORIA SEMICOMPATIBILISTA SOBRE RESPONSABILIDADE MORAL:
John Fischer e o controle de direcionamento**

Tese apresentada como requisito parcial para a obtenção do título de Doutora em Filosofia, pelo Programa de Pós-Graduação em Filosofia da Universidade do Vale do Rio dos Sinos – UNISINOS.

Orientador: Prof. Dr. Denis Coitinho

São Leopoldo

2018

F676t Fonseca, Tania Schneider da.
Uma teoria semicompatibilista sobre responsabilidade moral : John Fischer e o controle de direcionamento / Tania Schneider da Fonseca. – 2018.
163 f. ; 30 cm.

Tese (doutorado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Filosofia, 2018.
“Orientador: Prof. Dr. Denis Coitinho.”

1. Responsabilidade – Aspectos morais e éticos. 2. Fischer, John Martin, 1952- 3. Liberdade – Filosofia. 4. Livre arbítrio e determinismo. I. Título.

CDU 141.5

Dados Internacionais de Catalogação na Publicação (CIP)
(Bibliotecário: Flávio Nunes – CRB 10/1298)

TANIA SCHNEIDER DA FONSECA

**UMA TEORIA SEMICOMPATIBILISTA SOBRE RESPONSABILIDADE MORAL:
John Fischer e o controle de direcionamento**

Tese apresentada como requisito parcial para
obtenção do título de Doutora em Filosofia, pelo
Programa de Pós-Graduação em Filosofia da
Universidade do Vale do Rio dos Sinos -
UNISINOS

Aprovada em 27 de agosto de 2018

BANCA EXAMINADORA

Denis Coitinho Silveira (Orientador) – Universidade do Vale do Rio dos Sinos

Alfredo Culleton – Universidade do Vale do Rio dos Sinos

João Hobuss – Universidade Federal de Pelotas

Leonardo Ribeiro – Universidade Federal de Minas Gerais

Marco Azevedo – Universidade do Vale do Rio dos Sinos

À minha mãe,
por ser a minha grande inspiração.

In Memoriam

AGRADECIMENTOS

À Capes (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pelo auxílio financeiro durante todo o período do doutorado. Ao programa de Pós-Graduação em Filosofia da UNISINOS (Universidade do Vale do Rio dos Sinos) por ter me possibilitado a realização deste trabalho. À secretária Luciane Silva, por resolver as questões burocráticas de forma eficiente e sempre muito gentil. Aos professores Dr. Inácio Helfer e Dr. Eduardo Vicentini, pelas sugestões e comentários na qualificação da tese. Aos professores Dr. Alfredo Culleton, Dr. João Hobuss, Dr. Leonardo Ribeiro e Dr. Marco Azevedo, pelos comentários construtivos e críticos durante a defesa desta tese. Ao meu orientador Dr. Denis Coitinho, por ter acompanhado o meu trabalho durante os quatro anos do doutorado, pelo apoio, pelas inestimáveis sugestões, por me incentivar a ir além e por ter contribuído com a minha formação desde a graduação. Ao professor Dr. João Hobuss, meu orientador no mestrado e iniciação científica, pelo enriquecimento proporcionado à minha carreira. Em especial ao meu pai Nelson e à minha mãe Ceni, pelo apoio nos momentos difíceis, por me incentivarem a seguir em frente, pela compreensão nas ausências, e por tudo. Aos meus irmãos, aos meus sobrinhos Taís, Lucas e Rafael, pela compreensão nas minhas ausências. Aos meus familiares e amigos, que de uma forma ou outra me auxiliaram. E ao Tiaraju, meu esposo, pelas valiosas discussões filosóficas, por ter revisado atentamente o meu texto e pela parceria na vida pessoal.

RESUMO

Quais são as condições necessárias e suficientes para responsabilizar moralmente o comportamento do agente? Várias respostas a essa pergunta têm sido dadas na história recente da filosofia. De um lado, diversos filósofos acreditam que o critério especificando essas condições poderia ser compatível com a tese determinista. Por outro lado, há aqueles que negam que esse critério possa conciliar-se com a visão de um mundo determinista. Tradicionalmente, a liberdade enquanto uma capacidade para agir de outro modo é defendida como uma condição necessária para a responsabilidade moral. Com o seu artigo de 1969, “Alternate Possibilities and Moral Responsibility”, Harry Frankfurt mudou o curso do debate sobre o problema da vontade livre. Ele forneceu exemplos hipotéticos, por meio de experimentos de pensamento, de agentes que, conforme ele argumentou, embora não pudessem ter agido de outro modo, ainda assim seriam moralmente responsáveis pelas suas ações. O artigo de Frankfurt entusiasmou muitos filósofos, destacadamente John Fischer, a repensar o problema da responsabilidade moral. Para Fischer, Frankfurt teria mostrado que o debate não diz mais respeito ao problema de demonstrar a compatibilidade entre liberdade e determinismo, mas, sim, à questão da compatibilidade da responsabilidade moral com o determinismo. Para lidar com essa questão, e qualificar a posição de Frankfurt, Fischer desenvolve o que ele denomina de uma posição semicompatibilista. Essa posição responderia às objeções incompatibilistas, assim mostrando a compatibilidade da responsabilidade moral com o determinismo. Esse trabalho é dedicado a um estudo dessa posição. A tese defendida é a de que o semicompatibilismo proposto por Fischer de fato responde às principais objeções dos incompatibilistas, e é mais vantajosa se comparada à posição compatibilista tradicional, que defende que a capacidade para agir de outro modo seria uma condição necessária para a responsabilidade moral.

Palavras-chave: Responsabilidade Moral. Liberdade. Determinismo. John Fischer. Compatibilismo. Incompatibilismo. Semicompatibilismo.

ABSTRACT

What are the necessary and sufficient conditions for the moral responsible agency? Many answers to this question have been given in the recent history of philosophy. On the one side, some philosophers believe that the criterion specifying these conditions could be compatible with the determinist's worldview. On the other side, there are those who deny that this criterion and the determinist's position could be ever reconciliated. Traditionally, freedom as a capacity to do otherwise has been defended as a necessary condition for moral responsibility. In 1969, when "Alternative Possibilities and Moral Responsibility" (1969) was published, Harry Frankfurt changed the course of the discussion about the free will problem. Frankfurt showed, through some thought experiments, agents that, he argued, were moral responsible for their actions even though they could not have acted otherwise. Frankfurt's essay pushed several philosophers, remarkably John Fischer, to rethink the problem of moral responsibility. For Fischer, Frankfurt showed that the debate should not be about the compatibility between freedom and determinism, but rather should address the question of whether moral responsibility is compatible with determinism. To deal with this problem, and to further qualify Frankfurt's position, Fischer develops, as he calls it, a semicompatibilist position. This position, he claims, successfully address the objections from the incompatibilists, thus showing that moral responsibility is compatible with determinism. This study is an investigation of this position. The thesis defended is that the semicompatibilism proposed by Fischer does indeed answers the main incompatibilist's objections, and it is better than the traditional compatibilist position, which argues that the capacity to do otherwise is a necessary condition for moral responsibility.

Key-words: Moral Responsibility. Freedom. Determinism. John Fischer. Compatibilism. Incompatibilism. Semicompatibilism.

SUMÁRIO

Introdução	10
1 O debate em torno do problema da responsabilidade moral	20
1.1 Determinismo, fatalismo e fatalismo teológico	20
1.2 Compatibilismo clássico: Hume	29
1.3 Libertismo: Kant, Inwagen e Clarke.....	34
1.4 Compatibilismo contemporâneo: Strawson e as atitudes reativas	40
2 Responsabilidade moral e possibilidades alternativas	46
2.1 A refutação de Frankfurt do princípio das possibilidades alternativas	46
2.2 A estratégia átimo de liberdade [<i>flicker of freedom</i>]	51
2.3 A rejeição da abordagem átimo de liberdade	56
2.4 A defesa do dilema	60
2.5 Réplica à defesa do dilema	63
3 O semicompatibilismo de Fischer	66
3.1 Controle de direcionamento e regulativo.....	66
3.2 Sensibilidade às razões: forte, fraca e moderada.....	72
3.3 Uma noção não histórica de responsabilidade	75
3.4 Mecanismo do próprio agente.....	80
4 Teorias alternativas	88
4.1 A posição libertista de Robert Kane	88
4.2 O problema da sorte para a posição libertista de Kane e a sua resposta às críticas.....	95
4.3 Desacordo entre Kane e Fischer	99
4.4 A posição internista de Frankfurt.....	102
4.5 A proposta de Mele para o compatibilismo.....	105
4.6 A distinção entre a posição de Mele e Fischer	113
5 Mecanismos e manipulação.....	119
5.1 A teoria de Fischer baseada em mecanismos	119
5.2 Objeções à teoria baseada em mecanismos	123
5.3 Um modelo de liberdade baseado na sequência real.....	135
5.4 O argumento da manipulação	137
Considerações Finais	147
Referências	153

INTRODUÇÃO

O debate filosófico sobre a responsabilidade moral tem uma longa história. Uma das razões para o interesse nessa questão parece estar conectada com a concepção de nós mesmos como pessoas. Poderia se dizer que uma característica distintiva das pessoas é a sua posição de agentes enquanto moralmente responsáveis. Nesse sentido, é possível afirmar que unicamente as pessoas deteriam um tipo especial de controle. Muito tem se discutido sobre se uma pessoa poderia ser moralmente responsável pelo seu comportamento se ele fosse explicado unicamente por meio de um processo causalmente determinístico.

Aristóteles parece ter sido o primeiro a construir uma teoria sobre a responsabilidade moral. Ele trata sobre as condições que seriam apropriadas para se poder louvar ou censurar alguma ação particular ou traço de caráter do agente. Sua estratégia consistiu em apontar que um candidato será apto para louvor ou censura se e apenas se sua ação ou disposição for voluntária. Uma ação ou uma disposição de caráter voluntária possuiria para Aristóteles duas características. A primeira seria a condição de controle, isto é, a ação ou traço de caráter deve ter sua origem no agente. Quer dizer, a ação ou o traço de caráter deve estar no agente significa que ele não poderia ser compelido externamente. A segunda seria a condição epistêmica, ou seja, o agente deveria estar consciente do que está fazendo. Uma ação involuntária é aquela que ocorre por compulsão ou por ignorância. É compulsório ou forçado aquilo cujo princípio motor se encontra fora do agente, como, por exemplo, se a pessoa foi empurrada pelo vento ou por outro indivíduo.

Aristóteles trata também de casos em que é discutível se os atos são voluntários ou involuntários, que são as chamadas ações mistas. O exemplo que ele discute é o de alguém que durante uma tempestade, a fim de evitar o naufrágio, joga a carga no mar. Em teoria ninguém voluntariamente joga fora seus bens, porém qualquer pessoa sensata o fará se a sua própria segurança e a da tripulação assim exigirem. Tais ações mistas assemelham-se mais aos atos voluntários, e isso se deve ao fato de serem escolhidos no momento da ação, levando em consideração as circunstâncias particulares. Ambos os termos voluntário e involuntário devem, portanto, ser utilizados com relação ao momento da ação. O homem age voluntariamente se o princípio que move as partes apropriadas do seu corpo se encontra nele. E aquelas coisas cujo princípio motor está em nós, está em nós igualmente o fazê-las ou não as fazê-las. As ações mistas são voluntárias, mas em abstrato é provável que sejam involuntárias, pois por si mesmas

ninguém as escolheria. Nesse caso, sendo nobre a finalidade as pessoas poderiam ser louvadas. Mas, se o objetivo é insignificante, então poderiam ser censuradas.

Tem sido muito discutido se o conceito de vontade livre está conectado com o conceito de responsabilidade moral. Para Aristóteles, por exemplo, a liberdade significa uma capacidade para agir diferentemente, podendo o agente tanto fazer a ação A como o seu contrário. O agente seria moralmente responsável apenas se possuísse essa capacidade¹.

Aristóteles não tratou especificamente do problema do livre arbítrio, mas, sim, da questão da responsabilidade moral. Sobre essa questão não há consenso. Sauvé Meyer, por exemplo, afirma que ele introduziu o tópico da voluntariedade das ações como condição necessária para a responsabilidade moral (2006, p. 137). Segundo a interpretação de David Ross, “em linhas gerais, devemos dizer que (Aristóteles) compartilhava a crença do homem no livre arbítrio, mas que não analisou o problema muito bem, não tendo se expressado com perfeita consistência” (1995, p. 211). Para Pamela Huby partindo da evidência fragmentada e a falta de clareza sobre a história dos estoicos e Epicuro, é possível afirmar que é mais provável que Epicuro tenha sido o primeiro a reconhecer o problema do livre arbítrio, do conflito entre a vontade livre e o determinismo. E que foi nos estoicos, particularmente em Crisipo, que o problema foi tomado com mais entusiasmo. Diferentemente de Aristóteles, ele compreendeu que havia um problema (HUBY, 1967, p. 357-358).

Apesar disso, penso que Carlo Natali acerta ao afirmar que Aristóteles introduziu a base tanto no plano conceitual como no plano terminológico para o debate posterior em torno do problema do livre arbítrio, razão pela qual a sua obra permanece pertinente para o debate contemporâneo em torno do problema (NATALI, 2004, p. 65).

O problema do livre-arbítrio diz respeito ao possível conflito existente entre a liberdade e o determinismo. Há diferentes formulações do problema e historicamente a formulação clássica é a seguinte:

1. Alguns agentes poderiam ter agido diferentemente do modo que agiram.
2. Ações humanas são eventos.
3. Todo evento tem uma causa.
4. Se um evento é causado, então é causalmente determinado.

¹ Na literatura não há consenso sobre se Aristóteles pode ser interpretado como um libertista ou um determinista. Ver os capítulos II, III, V e VII da *EN* sobre a possibilidade do agente poder ou não agir diferentemente uma vez adquirida uma disposição de caráter.

5. Se um evento é um ato que é causalmente determinado, então o agente que fez o ato não poderia ter agido de outro modo.

O problema do livre-arbítrio surge quando refletimos sobre essas cinco proposições. Elas formam um conjunto inconsistente, ou seja, não é possível afirmar a verdade de todas essas proposições. (1) trata a liberdade como uma capacidade para selecionar entre diferentes cursos possíveis de ação. (2) identifica as ações com os eventos, que são fenômenos naturais. (3) e (4) são pressuposições da ciência natural. (5) é baseada no nosso senso comum sobre o que significa dizer que um evento é causalmente determinado, isto é, dadas as condições causais antecedentes, então unicamente um futuro é possível.

A partir da formulação do problema surgiram diferentes posições filosóficas. Alguns acreditam na falsidade da proposição (5), assim defendendo uma posição compatibilista entre a liberdade e o determinismo, enquanto outros defendem a verdade da proposição, nesse caso se comprometendo com uma posição incompatibilista entre a liberdade e o determinismo. Há também a posição libertista, que é incompatibilista, que defende a existência da vontade livre e sustenta a falsidade do determinismo. O libertista nega a proposição (4) e defende a existência de um mundo causalmente indeterminístico.

É mais útil pensar em termos de diferentes formulações do problema do livre-arbítrio. A razão para se pensar assim é porque cada formulação tem usado noções diferenciadas de liberdade como condições para a aplicação das práticas de responsabilidade moral, louvor e censura. Duas formulações do problema podem ser apresentadas como sendo dois argumentos para o incompatibilismo. Esses argumentos partem de dois modelos diferenciados de controle. Um deles faz referência à liberdade como abertura às possibilidades alternativas, e o outro às origens das ações.

O argumento incompatibilista clássico pode ser formulado no seguinte modo:

1. Se uma pessoa age com a sua própria vontade livre, então ela poderia ter feito de outro modo.
2. Se determinismo é verdadeiro, então ninguém pode fazer de outro modo do que, na realidade, faz.
3. Portanto, se determinismo é verdadeiro, ninguém age com a sua própria vontade livre.

Dito de outro modo, determinismo implicaria um único futuro possível dados os fatos do passado e as leis da natureza. E liberdade corresponderia a mais de um futuro possível. Nesse caso, se determinismo fosse verdadeiro ninguém teria acesso a alternativas.

Os compatibilistas têm duas opções: negar a premissa 1 ou a premissa 2. Ao rejeitar a premissa 1 eles não estariam aceitando o modelo de liberdade como capacidade para agir diferentemente e estariam assumindo um modelo de controle baseado na origem das ações. Ao rejeitar a premissa 2 eles estariam sustentando que o determinismo não seria uma ameaça à liberdade, e que, portanto, poderíamos ter feito de outro modo mesmo em um mundo determinista.

Já o argumento incompatibilista da origem, que é um ataque ao compatibilista que nega a premissa 1 do argumento acima, pode ser apresentado da seguinte forma:

1. Uma pessoa age com a sua própria vontade livre apenas se ela é a sua origem última.
2. Se determinismo é verdadeiro, ninguém é a origem última das suas ações.
3. Portanto, se determinismo é verdadeiro, ninguém age com a sua própria vontade livre.

Para o agente ter o controle sobre suas ações é preciso que ele desempenhe um papel fundamental na produção dessas ações. Esse ponto pode ser explicado em termos de uma distinção entre aqueles eventos que são produtos da própria agência do agente e aqueles que são meros acontecimentos corporais. Por exemplo, por um lado temos a escolha de alguém para pegar uma xícara de café e, por outro, os eventos como os batimentos cardíacos e a circulação sanguínea. No primeiro caso, um agente é a origem e o produtor do acontecimento em questão, e no segundo caso pode se dizer que eventos estariam meramente acontecendo.

Como que o determinismo seria uma ameaça a esse modelo particular de vontade livre? Se o determinismo é verdadeiro, então as ações fariam parte de uma conexão causal determinística que remeteria ao passado e às leis da natureza. Mesmo que o agente fosse a origem das suas ações, ele não seria livre porque a origem última das ações estariam fora dele. Nesse caso as suas ações seriam necessitadas ou determinadas pelos eventos do passado e pelas leis da natureza.

Uma das alternativas para o compatibilista, e a maioria faz uso dessa estratégia, consiste em negar a premissa 1 do argumento acima, segundo a qual uma pessoa age da sua própria vontade livre apenas se ela é a sua origem última.

Refletir sobre a liberdade é um dos caminhos para tratar sobre o problema do livre-arbítrio. Outro caminho é refletir sobre a noção de responsabilidade. A liberdade está diretamente conectada, dentre outras, às noções de louvor, censura e atitudes reativas, pois quando censuramos alguém ou ficamos ressentidos, ou com raiva, pressupomos que estava em seu poder não ter feito o que fez.

Começo com um exemplo. Digamos que assistimos o julgamento de um jovem que esteja sendo acusado de assassinato. No início nossa atitude em relação ao réu é de raiva, pensamos que o que ele fez foi horrível. Porém, na medida em que o seu advogado começa a expor a sua defesa, tomamos conhecimento de histórico triste de abuso infantil e de negligência por parte de seus pais. Começamos então a questionar: até que ponto esse jovem é responsável por ter se tornado o tipo de pessoa que ele é hoje? Ele escolheu ser o tipo de pessoa que ele é ou o seu caráter foi determinado pela educação recebida e pelo ambiente social no qual ele estava inserido? Nós sabemos que a genética, a educação e a sociedade na qual convivemos têm uma influência sobre o que nós somos. Porém, a questão é se essas influências externas completamente determinam aquilo o que nós somos, se para o bem ou para o mal, ou, se essas influências não são tão determinantes, se podemos dizer que somos responsáveis por nos tornarmos o tipo de pessoas que somos apesar desses condicionantes sociais. Em resumo, a questão sobre o jovem assassino é se o seu caráter é o resultado de más circunstâncias, o que nos levaria a considerá-lo uma vítima, ou se ele tem alguma responsabilidade por ser o que ele é, o que nos levaria a culpá-lo. Parece que a questão de se o caráter do jovem foi formado a partir da sua própria vontade livre depende de saber se esses fatores externos foram ou não completamente determinantes.

Para entender as discussões contemporâneas sobre o tema é útil retomar, em linhas gerais, às origens de toda a problemática. Podemos pensar em termos de três estágios. O primeiro envolve a ideia defendida na modernidade. O segundo abrange três distintas contribuições na década de 1960 que desafiaram muitos dos pressupostos do compatibilismo clássico. E o terceiro estágio remete às formas contemporâneas de compatibilismo que divergem da forma clássica, mas que refletem de alguma maneira as contribuições encontradas no segundo estágio.

Na Filosofia Antiga, podemos encontrar raízes da posição compatibilista, por exemplo, nos estoicos e em Aristóteles (conforme algumas interpretações). Porém, essa posição se popularizou no século XVII, na era moderna, na obra de filósofos como Thomas Hobbes, David Hume e John Stuart Mill, provavelmente porque forneceu uma solução simples para o problema do livre-arbítrio, do conflito entre a liberdade e o determinismo. Se o compatibilismo estiver

correto então não precisaríamos nos preocupar com a possibilidade da ciência futuramente vir a enfraquecer a nossa crença de que somos livres e responsáveis, pois haveria uma reconciliação entre a nossa experiência de liberdade e a visão científica de um mundo determinista.

Para os compatibilistas clássicos ser livre implica possuir um poder ou uma capacidade para fazer o que se deseja fazer, havendo uma ausência de constrangimentos ou obstáculos que nos impeçam de fazer o que queremos. Tais obstáculos podem ser físicos (estar preso ou amarrado), coercivos (se alguém estivesse apontando uma arma), fobia a lugares fechados que me levassem a evitá-los (compulsão). Mesmo se o que a pessoa quer ou deseja seja determinado pelo passado, ela é livre se não houver nenhum obstáculo que a impeça de fazer o que ela quer.

O compatibilismo clássico foi criticado por não lidar com os casos de obstáculos internos. Por exemplo, várias doenças mentais podem fazer com que uma pessoa aja como ela deseja, sem impedimentos externos, porém intuitivamente parece que ela não age de sua própria vontade livre nesses casos. Às vezes os próprios desejos que explicam as ações são origem da falta de liberdade dos agentes, como os casos de vícios ou neuroses.

A explicação compatibilista clássica responde ao argumento incompatibilista da origem afirmando que se a ação do agente surge a partir dos seus desejos sem impedimentos externos, então ele seria a origem genuína da sua ação, e esse tipo de liberdade seria suficiente para a responsabilidade moral. Certamente ele não seria a origem última da sua ação, pois isso exigiria a falsidade do determinismo.

A partir de 1960, três contribuições mudaram radicalmente o modo como o problema da vontade livre passou a ser abordado nas discussões contemporâneas.

A primeira foi o argumento incompatibilista, conhecido como o argumento da consequência, que procurou mostrar que um agente determinado carece de controle sobre as possibilidades alternativas. A outra contribuição remete ao argumento contra o princípio das possibilidades alternativas, elaborado por Harry Frankfurt (1969). Frankfurt procurou mostrar a falsidade do princípio que afirma que um agente é moralmente responsável apenas se ele poderia ter feito de outro modo. E, por último, Peter Strawson (originalmente publicado em 1962, republicado em 2008), adotando uma posição compatibilista, destacando para ambos os oponentes do debate (compatibilistas e incompatibilistas) a importância do papel das atitudes reativas nas nossas relações interpessoais para entendermos o conceito de responsabilidade moral. Para Strawson, a suposta ameaça que o determinismo coloca para a vontade livre e a responsabilidade moral desapareceria quando o papel das atitudes reativas, tais como ressentimento, gratidão, culpa e remorso etc., é devidamente compreendido.

Na literatura contemporânea as explicações compatibilistas foram influenciadas por pelo menos uma dessas contribuições. Algumas estratégias tentaram interpretar a liberdade em termos de possibilidades alternativas, enquanto outras não.

John Fischer aceita o pressuposto incompatibilista, que é o argumento de que a liberdade (no sentido que exige possibilidades alternativas) é incompatível com o determinismo. Sua posição é chamada de semicompatibilismo, pois, por um lado, ele aceita a incompatibilidade da liberdade com o determinismo e, por outro, propõe a compatibilidade da responsabilidade moral com o determinismo. Em *Responsibility and Control* (1998), uma obra escrita com Mark Ravizza, ele distingue entre dois tipos de controle: o de direcionamento (*guidance control*) e o regulativo (*regulative control*). Influenciado pela contribuição de Frankfurt, de que não precisaríamos da capacidade para agir de outro modo para sermos moralmente responsáveis, Fischer argumenta que o único controle necessário para a responsabilidade moral seria o de direcionamento. O ataque de Frankfurt ao princípio das possibilidades alternativas levou muitos compatibilistas a desenvolver explicações de liberdade que recusam o modelo do controle regulativo, isto é, enquanto capacidade para agir diferentemente. As explicações de controle de direcionamento estariam fixadas apenas sobre modelos de controle da origem das ações.

Os casos de obstáculos internos à nossa vontade (vícios, fobias, obsessões, neuroses e outros tipos de comportamento compulsivos) não apenas afetam a nossa liberdade para fazer o que nós queremos, mas afetam também a nossa liberdade para *querer* o que nós queremos. Para lidar com essa falha Frankfurt definiu a liberdade como um poder que teríamos para determinar quais serão as nossas vontades. Para esclarecer esse ponto Frankfurt fez uma distinção entre a liberdade da ação e a liberdade da vontade. Liberdade da ação é a liberdade de escolher um ou outro curso de ação, e liberdade da vontade é a liberdade de poder determinar aquilo que se quer. Para podermos dizer que escolhemos livremente uma ação temos de pressupor que essa foi a minha vontade.

A teoria de Frankfurt recebeu várias objeções. Uma delas diz respeito ao fato de que sua explicação de liberdade seria insuficiente. Por exemplo, um agente pode sofrer um processo de manipulação, por meio de lavagem cerebral ou hipnose, fazendo com que ele tenha desejos distintos daqueles que originalmente teria.

A proposta de Fischer é importante porque ela procura aperfeiçoar a noção de liberdade encontrada em Frankfurt. Para Fischer, o controle de direcionamento exige mais do que apenas agir de acordo com as razões e motivos e guiar o comportamento em conformidade com essas razões e motivos. Agentes compulsivos e neuróticos também guiam o seu comportamento em termos de razões e motivos. Fischer argumenta contra a estratégia de Frankfurt que pensa que

as condições históricas de responsabilidade (como as pessoas vieram a ter as razões ou os motivos que elas têm) não importam para avaliar se as pessoas são responsáveis.

Fischer também acredita que os agentes manipulados ou controlados por outros agentes são distintos de agentes meramente determinados. A responsabilidade moral seria eliminada apenas quando há um controle completo por outros agentes. Além do controle de direcionamento, e capacidade para responder às razões, é exigido que os agentes assumam a responsabilidade por agir com base nos motivos e razões que eles agem e, assim, vejam a si mesmos como candidatos justos de atitudes reativas e práticas associadas (recompensa e punição). Aqui Fischer estaria aceitando a proposta de Strawson de pensar a responsabilidade moral em termos de atitudes reativas.

O elemento chave para entender a reconciliação da responsabilidade moral com o determinismo consiste na ideia do agente assumir a responsabilidade pelo seu comportamento. Mesmo sendo o determinismo verdadeiro, ele ainda vê a si mesmo como um candidato justo ou apropriado de atitudes reativas.

O tema desta tese é o problema da relação entre liberdade, determinismo e responsabilidade que vem sendo debatido pelos filósofos por mais de dois milênios. O determinismo é incompatível com a responsabilidade moral do agente? Se sim, por quê? Se não, qual tipo de liberdade é exigida para se responsabilizar moralmente o comportamento de alguém? Diferentes respostas foram dadas a essas questões na história da filosofia, e persistem a ser dadas contemporaneamente, formando uma literatura marcada pela ausência de consenso. O debate, em parte, tem se concentrado no seguinte ponto: nós somos moralmente responsáveis pelas nossas ações se e somente se nós somos livres. O que significa liberdade é uma das discordâncias entre os teóricos. Existem também os chamados céticos que pensam que a responsabilidade moral é impossível tanto em um mundo determinista como em um mundo indeterminista. Além disso, com o desenvolvimento da ciência, pesquisas na física, na biologia, na psicologia e na neurociência, têm cada vez mais desafiado a tese de que o agente tem, de fato, algum controle sobre o seu comportamento. As influências externas e internas às quais ele estaria sujeito são determinantes ou deixam algum espaço para a sua liberdade?

A tese defendida aqui será a de que a teoria semicompatibilista, desenvolvida e defendida por Fischer, responde às objeções dos filósofos incompatibilistas, mostrando que

seria possível compatibilizar a responsabilidade moral com o determinismo². Acredito que a proposta de Fischer apresenta uma solução satisfatória para o problema do livre-arbítrio, pois nos permite pensar a responsabilidade moral em termos da capacidade do agente para responder às razões morais, sem a necessidade de invocar uma concepção metafísica de liberdade como capacidade para agir diferentemente dadas as mesmas circunstâncias passadas. O presente estudo pretende abordar os dois sentidos de liberdade que são abordados na literatura como necessários para a aplicação da noção de responsabilidade moral, quais sejam, o controle regulativo e o de direcionamento. Eu analiso essas duas interpretações sobre a liberdade para defender a tese de que seria suficiente para a responsabilidade moral a noção de controle de direcionamento, isto é, sem a necessidade de apelar para um controle regulativo. Acredito que o controle de direcionamento seria capaz de lidar com as críticas dirigidas aos modelos incompatibilistas, em particular, as objeções de que a noção de liberdade adotada por eles não seria inteligível, de que as suas explicações seriam obscuras e misteriosas e que as ações, quando em um mundo indeterminista, ocorreriam por acaso ou sorte.

A estrutura da tese está elaborada em cinco capítulos.

No primeiro capítulo apresentarei uma definição de determinismo, fatalismo, compatibilismo, incompatibilismo e libertismo.

No segundo capítulo tratarei do argumento compatibilista de Frankfurt contra o princípio das possibilidades alternativas. Partindo de situações hipotéticas, Frankfurt defende que um agente não necessitaria da capacidade para agir diferentemente para ser moralmente responsável. Depois tratarei das objeções à tese de Frankfurt. Essas objeções reformulam os próprios casos apresentados por Frankfurt com o intuito de mostrar que o princípio das possibilidades alternativas não teria sido falsificado. Em seguida, eu apontarei que todas essas objeções estariam cometendo o mesmo equívoco que seria pressupor uma possibilidade alternativa não robusta. A seguir, discutirei a estratégia incompatibilista de reformular o argumento em forma de um dilema com o objetivo de preservar a capacidade para agir de outro modo seja num mundo determinista seja num mundo indeterminista. E, por último, eu apresentarei a ideia de que para a atribuição de responsabilidade moral não haveria a necessidade de contar com uma noção de liberdade metafísica, mudando assim o foco do debate.

² Apesar deste trabalho conter referências ao livro *Responsibility and Control* (1998), escrito conjuntamente por Fischer e Ravizza, ele se concentrará especificamente na teoria de Fischer tal como ela pode ser encontrada nesse livro e em outros escritos seus.

No terceiro capítulo irei expandir a ideia de que não precisaríamos depender do conceito de liberdade enquanto capacidade para agir diferentemente. Nesse capítulo tratarei detalhadamente da teoria semicompatibilista de Fischer. Apontarei a distinção entre o controle de direcionamento e o controle regulativo. Depois apresentarei o conceito de capacidade para responder apropriadamente às razões. Em seguida tratarei da abordagem hierárquica de Frankfurt, ressaltando as suas diferenças em relação à teoria de Fischer. Depois retorno à teoria de Fischer discutindo o conceito de mecanismo do próprio agente.

No quarto capítulo eu apresentarei a crítica que Robert Kane faz a toda e qualquer teoria compatibilista, incluindo a de Frankfurt e Fischer. Para Kane, pelo menos algumas das nossas ações deveriam ser indeterminadas para que possamos ser moralmente responsáveis. Kane faz um esboço de um argumento que mais tarde na literatura será conhecido como o argumento da manipulação. Para ele haveria mais semelhanças do que diferenças entre um agente determinado e um agente manipulado. Também apresentarei a réplica de Alfred Mele a esse argumento. Para Mele, os compatibilistas teriam os recursos necessários para distinguir os casos de manipulação dos casos de mera determinação causal.

No quinto e último capítulo retomarei o conceito de capacidade para responder apropriadamente às razões, proposto por Fischer. Em seguida tratarei da exigência que a teoria semicompatibilista faz do conceito de mundos possíveis ou cenários contrafactuais para testar se o mecanismo sobre o qual o agente está agindo seria responsivo às razões. Também apresentarei um outro modelo baseado na sequência real dos eventos muito semelhante àquele de Fischer, exceto que não invocaria a necessidade de se verificar como que o processo pelo qual o agente está agindo se comporta em mundos possíveis. Ainda tratarei também do argumento da manipulação que Derk Pereboom elabora contra todo e qualquer critério compatibilista de agência livre e responsável. Para Pereboom, nenhum critério seria suficiente para explicar a responsabilidade moral em um contexto determinista. E, por fim, apontarei que esse argumento incompatibilista estaria partindo de pressuposições e premissas inadequadas.

1 O DEBATE EM TORNO DO PROBLEMA DA RESPONSABILIDADE MORAL

Neste capítulo será feita uma caracterização das principais posições que compõem o debate, a saber, determinismo, indeterminismo, compatibilismo, incompatibilismo e fatalismo. O objetivo deste capítulo é oportunizar uma compreensão geral do quadro conceitual que norteará as discussões nos capítulos seguintes.

1.1 Determinismo, fatalismo e fatalismo teológico

Tanto o determinismo como o fatalismo foram apresentados na história da filosofia como uma ameaça à liberdade humana, e, conseqüentemente, à responsabilidade moral. Apesar disso, há uma diferença entre o determinismo e o fatalismo. De acordo com a explicação fatalista, a deliberação, a escolha e a ação são completamente ociosas. O que está fadado a acontecer acontecerá, não importa o que se escolha fazer. Por outro lado, para o determinismo, a deliberação, a escolha e a ação dependem do agente apesar de serem resultado de causas anteriores.

Desde os estoicos, a tese do determinismo causal (científico) tem estado no centro do debate sobre a discussão da responsabilidade moral³ (SHARPLES, 1996). No período medieval, em Agostinho, por exemplo, reflexões sobre a liberdade e a responsabilidade envolveram a seguinte questão: dado a presciência divina, nós somos livres e moralmente responsáveis (AGOSTINHO, 1996)? Já no período moderno, devido ao mecanicismo motivado pela física de Newton, houve uma retomada de interesse no determinismo científico. Foi dada uma explicação de todo aspecto do universo, incluindo as ações humanas, em termos de causas físicas (LAPLACE, 2009).

A partir dessa explicação determinista do universo e das ações humanas duas posições surgiram no debate sobre o problema da vontade livre. Por um lado, os incompatibilistas afirmam que as pessoas não poderiam ser livres nem moralmente responsáveis se essa explicação determinista da ação humana fosse verdadeira. Por outro lado, os compatibilistas argumentam que as pessoas poderiam ser livres e moralmente responsáveis mesmo se a tese determinista fosse verdadeira⁴.

³ A diferença entre o determinismo científico e o fatalismo parece não ter sido claramente reconhecida até o desenvolvimento da filosofia estoica.

⁴ Na Grécia Antiga essas posições foram representadas pelo pensamento de Epicuro e dos estoicos, respectivamente. Para uma apresentação dessas posições helenísticas, ver Sharples (1996, p. 59-82).

Interessante observar a existência de duas interpretações do conceito de responsabilidade moral. Louvor e censura podem ser apropriadas em dois sentidos: (i) o agente merece uma tal reação dado o seu comportamento ou caráter e (ii) uma tal reação provavelmente trará um aperfeiçoamento no seu comportamento ou caráter. Esses dois sentidos caracterizam duas interpretações do conceito de responsabilidade moral: (i) uma visão baseada no mérito, de acordo com a qual uma tal reação seria apropriada em relação ao agente se e apenas ele merece, e (ii) uma visão consequencialista, de acordo com a qual louvor ou censura seriam apropriados se e apenas uma tal reação tivesse uma alta probabilidade de conduzir uma mudança desejada no comportamento do agente. As pesquisas nos últimos 50 anos têm se concentrado em oferecer versões alternativas de (i). Derk Pereboom, por exemplo, defende que a responsabilidade moral estaria conectada com a ideia de que o agente merece ser louvado ou censurado e não com considerações consequencialistas (PEREBOOM, 2009, p. 2).

A distinção entre esses dois conceitos de responsabilidade moral é importante porque, por exemplo, filósofos que aceitam uma concepção de responsabilidade moral baseada no mérito tendem a ser incompatibilistas. Para o agente merecer punição, censura ou condenação moral ele deveria possuir um tipo de liberdade que é incompatível com a verdade da tese determinista. Aqueles que aceitam uma concepção consequencialista de responsabilidade moral tendem a ser compatibilistas. Para os compatibilistas não haveria a necessidade da falsidade da tese determinista, pois a eficácia das práticas de punição e censura para regular o comportamento em modos socialmente desejáveis seria uma base adequada para essas práticas. Mesmo em um mundo determinista, louvor ou censura poderiam ainda ser meios eficazes de influenciar o comportamento do agente (STRAWSON, 2008, p. 2-3)⁵.

Pesquisas feitas na área da psicologia chegam à uma conclusão perturbadora. Os estudos mostram que quanto mais conhecimento temos sobre as origens do comportamento humano, menos espaço parece haver para a escolha livre. Nesse sentido, é interessante analisar as repercussões de um crime ocorrido em 1924, em Chicago nos EUA. Dois adolescentes, Richard Loeb e Nathan Leopold, raptaram e assassinaram Bobby Franks. O fato curioso sobre esse caso é que a despeito da brutalidade do ato cometido, os dois adolescentes não pareciam ser

⁵ A tendência geral de vincular, de um lado, a concepção consequencialista de responsabilidade moral com o compatibilismo sobre determinismo causal e, de outro, a concepção baseada no mérito com o incompatibilismo, persistiu até o fim da primeira metade do século XX.

perversos. Eram estudantes brilhantes e vinham de famílias ricas⁶. Como explicar esse assassinato? (RACHELS, 2009, p. 180).

Os familiares contrataram o famoso advogado Clarence Darrow para os defender⁷. Em defesa dos réus o advogado não apelou ao fato de que os réus eram loucos. Mesmo assim, defendeu que eles não eram responsáveis pelo o que tinham feito. Darrow fundamentou a sua defesa em pesquisas conduzidas por psicólogos, especificamente na ideia de que “o caráter humano é moldado pelos genes do indivíduo e pelo ambiente” (RACHELS, 2009, p. 157). Darrow também se baseou nas conclusões dos psiquiatras que atestaram que os réus não tinham sentimentos normais, pois não demonstravam qualquer reação emocional ao crime, além do histórico familiar dos dois adolescentes. Loeb foi privado durante a sua infância de afeto, passava os dias estudando, e as noites lendo secretamente histórias sobre crimes, fantasiando como praticar o crime perfeito e enganar a polícia, enquanto Leopold não tinha amigos, cresceu desprezando as outras pessoas e tentando a todo custo mostrar-se superior em relação aos outros. Eles acabaram se conhecendo e juntos cometeram o assassinato, que nenhum deles poderia ter praticado sozinho. Darrow estava convicto que os criminosos não deveriam ser enforcados tampouco censurados, pois causas que remontam ao começo, antes mesmo deles terem nascido, podem ter influenciado a mente desses adolescentes tendo os levado a realizar esse ato. Eles já nasceram sem o tipo correto de emoções. Dito de outro modo, eles já nasceram maus, sem os sentimentos de piedade e empatia. Após um mês de deliberações sobre o caso, o juiz deu a sentença pretendida pelo advogado e os livrou da pena de morte, condenando-os à prisão perpétua⁸ (RACHELS, 2009, p. 158 e 177).

Esse julgamento criminal foi o primeiro em que se usou a ideia moderna de que a nossa personalidade é resultado da hereditariedade e do ambiente para justificar o fato de que os réus não eram responsáveis pelo crime cometido. Contudo, Darrow não foi o primeiro a lançar

⁶ Leopold, aos dezoito anos, era o mais jovem licenciado na história da Universidade de Chicago, e estava prestes a entrar em Harvard, no curso de Direito. Loeb, aos dezenove anos, era o mais novo licenciado na Universidade de Michigan (RACHELS, 2009, p. 155-156).

⁷ Darrow era conhecido como o defensor das causas impossíveis e como um opositor da pena de morte. Leopold e Loeb já tinham admitido a culpa, então a tarefa do advogado era apenas evitar a pena de morte. Caberia ao juiz ouvir os argumentos dos advogados e decidir se eles seriam enforcados, pois não haveria um júri (RACHELS, 2009, p. 156-157).

⁸ Loeb, passado doze anos na prisão, foi morto numa discórdia com outro prisioneiro. Leopold ficou trinta e quatro anos na prisão, e durante esse tempo deu aulas, serviu como cobaia para experimentos e pesquisas médicas sobre a malária, trabalhou na biblioteca e no hospital da prisão. Quando conseguiu a liberdade condicional foi morar em Porto Rico, onde continuou se esforçando para se tornar uma pessoa melhor, realizando trabalhos que envolviam ajudar os outros. Ele morreu em 1971 (RACHELS, 2009, p. 158-159).

dúvidas sobre a ideia de que o destino está sob nosso controle. Já na antiguidade Aristóteles ocupou-se da possibilidade das leis da lógica significarem que nós carecemos de controle sobre o que fazemos⁹. Para Aristóteles, se toda a afirmação ou negação for verdadeira ou falsa, então tudo ocorrerá por necessidade. Parece que Aristóteles estaria antecipando o argumento preguiçoso, pois se tudo ocorre por necessidade então não haveria sentido em deliberar na medida em que nada dependeria de nós¹⁰.

Vejamos o exemplo da batalha naval, que ele discute em “De Interpretatione”, no capítulo IX (1984b). Adotando o princípio da bivalência (toda proposição p é ou verdadeira ou falsa), se a proposição “ocorrerá uma batalha naval amanhã” for verdadeira, necessariamente ocorrerá uma batalha naval amanhã. Se for falsa, necessariamente não ocorrerá. Se o princípio da bivalência fosse adotado para proposições passadas, presentes e futuras, logo tudo ocorreria por necessidade. Conforme a passagem:

Estas são, dentre outras, as consequências absurdas às quais chegamos ao admitir que para todo conjunto constituído de uma afirmação e de uma negação contraditórias (sobre universais falados universalmente ou sobre particulares), necessariamente uma é verdadeira e a outra falsa, e que nada nas coisas em devir é segundo qualquer um dos dois membros da alternativa, e que tudo é e ocorre por necessidade. Em consequência, *não haveria necessidade de deliberar*, nem de se preocupar, uma vez admitido que se nós fazemos isto, isto ocorrerá, e que se nós não fazemos isto, isto não ocorrerá (ARISTÓTELES, 1984b, *Int* 18 b 25-35)¹¹.

Aristóteles parece ter percebido o problema com essa conclusão, pois ele negou a validade do princípio da bivalência para proposições futuras. Acontecimentos futuros não são necessários, sendo possível que ocorram ou não ocorram, pois do contrário não haveria sentido deliberar.

Se o futuro está determinado ou predestinado, então nós não temos poder para mudá-lo. Isso ficou conhecido, na literatura, como o problema do fatalismo. Na idade média, sobretudo a partir de Agostinho, a onisciência divina gera uma dificuldade semelhante para os teólogos. Pressupondo que Deus sabe tudo, então Ele sabe o que vamos fazer no futuro. Porém, se Deus já sabe o que vamos fazer, então como podemos agir diferentemente?

⁹ Para a discussão aristotélica sobre a questão da responsabilidade moral, ver *Nicomachean Ethics* (1984a, *EN* 1110 a 1112 a).

¹⁰ Sobre esse ponto ver o artigo de João Hobuss (2013, p. 291-313).

¹¹ Essa é a tradução feita por Hobuss (2013, p. 307).

Conforme discutido por Michael Dummett, na Segunda Guerra Mundial, durante o bombardeamento de Londres, o seguinte argumento fatalista se tornou popular¹²:

A sirene toca e eu corro para o abrigo anti-aéreo para evitar ser morto por uma bomba. O fatalista argumenta que “ou você será morto por uma bomba ou você não será. Se você for, então quaisquer precauções que você tomar serão ineficazes. Se você não for, todas as precauções que você tomar serão supérfluas. Portanto, é inútil tomar precauções.” Essa crença foi aplicada até mesmo para bombas particulares. Se a bomba fosse me matar, então “ela tinha o meu número”, e não havia sentido em tomar precauções para evitar ser morto por *aquela* bomba; se ela não tivesse o meu número, então, é claro, precauções seriam sem sentido também (DUMMETT, 1964, p. 345).

O fatalismo é definido como a tese de que as ações humanas estão fadadas a acontecer não importa o que façamos. Conforme identificado por Dummett, haveria uma falha evidente nesse argumento assim formulado. Parece falacioso argumentar que se você não será morto por uma bomba, então quaisquer precauções que você tome serão supérfluas, pois poderia ser o caso que você não morreu simplesmente por ter se precavido (DUMMETT, 1964, p. 348). Por exemplo, digamos que tenha chovido e eu não tenha me molhado porque estava com meu guarda-chuva. Não seria razoável dizer que eu não teria me molhado mesmo se não tivesse levado o meu guarda-chuva.

Outra versão do fatalismo é o teológico, discutido na Idade Média. Há uma tensão entre a onisciência divina e o livre-arbítrio¹³. Como pode alguma ação humana ser livre se Deus sabe tudo o que o ser humano fará? Se admitirmos a onisciência divina então devemos aceitar que se eu realizo uma ação em um determinado tempo, t_2 , por exemplo, poderia ser dito, levando em consideração a onisciência divina, que Deus sabe, em t_1 , que eu exercerei a ação em t_2 . Nesse caso o fatalista dirá que eu não poderia me abster de realizar a ação porque a ação foi determinada pela presciência divina.

Para Alfred Ayer, o fatalismo não apresentaria dificuldades à liberdade humana, pois o conhecimento prévio de uma ação não a tornaria inevitável. Isso fica evidente na seguinte passagem:

¹² Para uma discussão do argumento fatalista no contexto clássico na filosofia antiga, ver Bobzien (1998, p. 182-199).

¹³ Ver a discussão do problema em Agostinho, na sua obra *A Cidade de Deus* (1996, p. 483-495), especialmente os capítulos 8-10, do livro V, onde Agostinho argumenta que a liberdade e a graça divina não se opõem.

Se o fato de alguém saber o que eu farei amanhã não torna necessário que eu o farei, então o fato de alguém saber o que eu farei, não somente amanhã, mas em todos os dias da minha vida, tampouco torna necessárias essas ações (...). Nem faz qualquer diferença se a pessoa a quem o conhecimento prévio é atribuído é considerada humana ou divina. Se existe algum sentido de que nós somos livres, então nós somos livres, não importa se existe alguém que tem algum conhecimento do que nós de fato faremos, nem importa o quanto de conhecimento ele tem, e não importa que tipo de pessoa ele é. Todas essas considerações são irrelevantes (AYER, 1963, p. 252-253).

Se Ayer estiver correto, a implicação é que o fatalismo teológico não constituiria uma ameaça real à liberdade humana.

O desafio mais preocupante atualmente à liberdade humana não decorre do fatalismo, mas do determinismo, que já era corrente no mundo antigo, mas que se tornou notável na modernidade, com o avanço da ciência moderna. O determinismo pode ser definido como a tese de que todos os eventos, incluindo as vontades, as escolhas e as ações humanas, são causadas por acontecimentos anteriores. Pressupondo que as ações humanas possam ser explicadas pelas leis físicas, isso implicaria que dado as leis da natureza e o estado do universo em qualquer tempo passado, é fisicamente impossível, antes e depois de t , ser diferente do que é. Existem dois modos de colocar a ameaça do determinismo ao comportamento humano. Primeiro, se uma explicação determinista for verdadeira, então o agente não seria a origem das suas ações. Seriam as condições antecedentes e não o agente que determinaria o seu comportamento. Segundo, o determinismo implicaria a ausência de possibilidades alternativas. Nessa segunda formulação, o agente não seria livre porque careceria da liberdade para agir diferentemente (WATSON, 2013, p. 2).

Diferentemente do fatalismo, que depende de premissas lógicas ou *a priori*, a tese determinista envolve premissas empíricas que remetem às leis físicas do mundo tal como o conhecemos.

O que significa dizer que um sistema é determinista? Nas palavras de Rachels:

Dizer que um sistema é determinista é afirmar que tudo o que nele acontece resulta de causas anteriores, e que, logo que as causas ocorreram, os efeitos têm de se seguir inevitavelmente, dadas as circunstâncias circundantes e as leis da natureza. Provavelmente, vemos o edifício em que vivemos como um sistema determinista. Se as luzes se apagam, pensamos que isso tem de ter uma causa. Supomos que, logo que a causa ocorreu, o efeito tinha de se seguir. Se o eletricitista nos dissesse que “isso pura e simplesmente aconteceu”, sem qualquer razão, essa afirmação violaria a nossa concepção do funcionamento das coisas (RACHELS, 2009, p. 160).

Na modernidade, com a ascensão da ciência mecanicista, o universo foi concebido como um sistema determinista. Tudo o que acontece na natureza é explicado pelas leis causais da física. Em 1819 o matemático francês Pierre-Simon Laplace (1749-1827) declarou que se um observador conhecesse todas as leis da física então ele seria capaz de prever todos os estados futuros do universo. Esse observador saberia tudo antes do acontecimento desses eventos. (LAPLACE, 2009, p. 250-254).

Mas a questão que poderia ser colocada é esta: qual é a relação desses eventos naturais, como a causa da luz estar apagada, com os nossos movimentos corporais e as nossas ações ou atitudes racionais? A relação seria que nós somos considerados como parte da natureza no sentido que o nosso comportamento está sujeito às mesmas leis físicas dos acontecimentos naturais¹⁴. Tanto os nossos estados mentais como os desejos e as escolhas, como os movimentos do nosso corpo, braços e pernas, são causados por acontecimentos neurológicos. Seria possível para o observador de Laplace prever as nossas ações do mesmo modo que prevê os eventos futuros do universo. Recuando na cadeia causal ele poderia prever, antes mesmo de termos nascido, que ação nós faremos. Para Laplace, a nossa liberdade é uma ilusão (LAPLACE, 2009, p. 251).

Há experimentos que mostram que a atividade cerebral se inicia um segundo e meio antes de decidirmos o que faremos. Esses experimentos foram conduzidos pela primeira vez nos anos 70 do século passado, pelo cientista alemão H. H. Kornhuber (1976). Digamos que daqui a um minuto moveremos o dedo espontaneamente. Podemos ter certeza que a decisão de mover o dedo está no nosso controle. Porém, vamos imaginar que se conecte algum dispositivo em nossas cabeças a fim de observar nossa atividade cerebral, e nos peçam para repetir a ação. Um técnico monitorando a nossa atividade cerebral, através de uma electroencefalografia, é capaz de constatar que, quando movemos o dedo, a atividade cerebral inicia-se antes de movermos o dedo¹⁵. O observador, nesse caso, poderia saber o que nós faremos antes mesmo de tomarmos uma decisão.

¹⁴ Como Darrow observou, as causas do nosso comportamento podem estar distantes, antes mesmo de termos nascido, denominadas causas últimas. No entanto, as causas imediatas são “acontecimentos que ocorrem no nosso cérebro” (RACHELS, 2009, p. 161).

¹⁵ Esse técnico desempenha a mesma função do observador de Laplace. É possível fazer mais do que apenas monitorar a atividade cerebral. É concebível induzir ações por meio de uma estimulação elétrica cerebral. José Delgado conduziu um experimento na Yale University, e chegou à conclusão que ao estimular algumas regiões do cérebro era possível causar certos tipos de movimentos corporais, como o franzir da sobrancelha. Primeiro ele fez o experimento em animais, macacos e gatos, e percebeu que os animais consideram seus movimentos como se estes estivessem no seu controle e fossem voluntários. Não se mostravam surpreendidos. Filósofos podem dizer que esse experimento não causa ações, mas

O argumento contra o livre-arbítrio até aqui apresentado partiu de princípios da física. Mas o argumento pode também ser formulado a partir da psicologia, afinal a psicologia estuda o comportamento humano. No século XX houve discórdias na psicologia entre a psicanálise e o behaviorismo. Na psicanálise, Freud enfatizou o inconsciente como sendo o motivador do comportamento humano, enquanto o behaviorismo defendeu que o comportamento é uma resposta aos estímulos externos. Para Freud, os acontecimentos na nossa infância criam desejos e impulsos inconscientes que determinam as nossas decisões na vida adulta. Por exemplo, a conduta de uma mulher que tem vários relacionamentos com homens abusivos pode ser explicada pela sua personalidade formada na infância quando conviveu com um pai abusivo que a espancava (FREUD, 1961). Para os behavioristas, não são os pensamentos inconscientes que explicam o comportamento, mas o processo de condicionamento. As nossas ações podem ser explicadas como sendo uma resposta a um condicionamento prévio. Tendemos a responder quando somos recompensados ou castigados. Por exemplo, digamos que uma criança possa jogar vídeo game quando tem uma boa relação com os seus amigos, porém seja proibida de jogar quando briga com eles. Por meio da recompensa e do castigo, que são meios de condicionamento, a criança começará a tratar bem os seus amigos (SKINNER, 1971).

O debate filosófico tradicional discorda sobre o problema da compatibilidade da liberdade com o determinismo. De um lado estão os compatibilistas, como Hobbes (1839) e Hume (2009), que defenderam que o determinismo não comprometeria a nossa liberdade. De outro lado estão os incompatibilistas, que sustentaram que a vontade livre e o determinismo são incompatíveis. Os incompatibilistas se dividem em dois grupos: os libertistas e os deterministas:

- (i) Libertistas: acreditam que a tese determinista deveria ser falsa, pois o determinismo enfraqueceria a nossa liberdade. Nós somos livres.
- (ii) Deterministas: defendem que a tese determinista é verdadeira, e que nós carecemos de vontade livre. Nós não somos livres.

apenas movimentos corporais, na medida em que ações não implicariam apenas movimentos, mas razões e decisões. O interessante é que Delgado fez o experimento também em seres humanos. E, do mesmo modo que os animais, os seres humanos não se mostraram surpreendidos. Além disso, deram razões para o comportamento exercido. Em um dos voluntários produziu o movimento de virar a cabeça como se estivesse procurando alguma coisa. Para esse voluntário, o comportamento resultante era espontâneo, o qual ele justificava por meio de razões. Ao ser questionado sobre o que estava fazendo, ele respondia que estava à procura dos seus chinelos (DELGADO, 1969, p. 115-116).

Outros defenderam que os compatibilistas e os libertistas estão ambos errados porque a liberdade é incompatível tanto com o determinismo quanto com o indeterminismo. Galen Strawson, por exemplo, defendeu que vontade livre é logicamente impossível. O conceito de vontade livre, seja o compatibilista ou libertista, seria incoerente. A vontade livre é, para Strawson, uma ilusão, e uma vez que a responsabilidade moral pressupõe a vontade livre, ela também seria uma ilusão (STRAWSON, 1994, p. 5-7).

Agora digamos que a teoria física da mecânica quântica dissesse que o nosso universo não é determinístico. As concepções libertistas da vontade livre estariam sujeitas às seguintes questões: se a ação do agente é resultante de um processo indeterminístico, como explicar por que ele fez uma ação ao invés de outra? Supondo que a tese determinista seja falsa, a questão passa a ser: como explicar a responsabilidade moral de uma ação resultante de um processo indeterminístico sobre o qual o agente pode não ter controle algum? O agente que faz uma ação correta não seria meramente um sortudo? (WATSON, 2013, p. 9).

Vejamos o argumento incompatibilista típico. Suponha que eu faça a ação *A* livremente (acene meu braço para cumprimentar alguém) no contexto de um mundo determinista. Para o incompatibilista, se *A* é livre eu poderia ter me absterido de fazer *A*. Porém, de acordo com a definição de determinismo dada anteriormente, as leis da natureza e o estado do universo antes de eu fazer *A* implicam que o futuro não poderia ser diferente do que ele é na realidade. Dito de outro modo, eu não poderia ter me absterido de fazer *A*. Se a tese determinista é verdadeira, então, de acordo com o incompatibilista, eu não fiz *A* livremente, pois não havia possibilidades alternativas (GARRETT, 2006, p. 113).

Uma das réplicas possíveis ao argumento incompatibilista afirma que as ações humanas não são governadas por leis causais, mas pertencem a uma esfera diferente: elas seriam explicadas pela atribuição de razões, e razões não são causas. Já os movimentos corporais são explicados pela lei causal. O ponto seria que a verdade ou a falsidade da tese determinista seria irrelevante para o entendimento da ação humana livre.

No entanto, parece haver um problema com essa réplica. Mesmo que as ações humanas e os movimentos corporais façam parte de categorias lógicas distintas, eles ainda estão conectados. Por exemplo, a grande maioria das ações humanas, como o aceno de mãos, envolve movimentos corporais. E qualquer restrição sobre o movimento corporal acarretaria uma restrição sobre a ação humana. Portanto, a ideia de que razões não são causas não produziria uma réplica convincente ao argumento incompatibilista (GARRETT, 2006, p. 113).

O que o compatibilista teria a dizer sobre o argumento do incompatibilista? A resposta padrão consiste em negar que haja alguma incompatibilidade entre a vontade livre e o

determinismo. Para o compatibilista, a vontade livre deve ser contrastada com restrições de vários tipos, não com a causação universal ou com o determinismo. Um agente não é livre quando certas restrições estão presentes. Por exemplo, quando ele é ameaçado com uma arma.

Há duas razões possíveis para a insatisfação com a resposta compatibilista. Primeiro, o uso ordinário das palavras ‘livre’ e ‘não livre’ não corresponde exatamente à definição compatibilista. A minha ação é ainda livre mesmo se alguém aponte uma arma para a minha cabeça, e, em resposta, eu abra o cofre e entregue o dinheiro. A tendência em descrever esse ato como um caso de ação não livre consiste talvez na confusão entre vontade livre e responsabilidade moral. Certamente eu não seria moralmente responsável pelo roubo do dinheiro nas circunstâncias descritas, porém o fato de eu não ser moralmente responsável não significa que a minha ação não foi livre. A minha ação foi livre, pois eu poderia ter escolhido não consentir com os desejos do assaltante. O propósito desse esclarecimento é mostrar que os casos de restrição (ação restritiva) com que a ação livre deveria ser contrastada são aqueles em que alguém é literalmente privado do poder de escolha ou ação, como restrição física, cleptomania e hipnose (GARRETT, 2006, p. 114).

Segundo, a acusação de arbitrariedade seria a segunda razão para a insatisfação com a resposta compatibilista. O compatibilista estabelece uma distinção dentro do âmbito das causas, entre as causas que são compatíveis com a vontade livre (causas determinantes) e as causas incompatíveis com a vontade livre (causas restritivas). Essa não seria uma distinção arbitrária? Um agente que estivesse sujeito a ambos os tipos de causas, seja determinante ou restritiva, não poderia ter feito de outro modo. Então por que as causas determinantes são consistentes com a ação livre e as causas compelidas não são?

Há uma outra réplica ao argumento incompatibilista. Frankfurt, por exemplo, aponta uma diferente concepção compatibilista de vontade livre. A incompatibilidade da vontade livre e determinismo não se sustentaria pois a vontade livre não exigiria a capacidade para agir diferentemente. Vontade livre requereria apenas que a nossa ação estivesse em conformidade com os nossos desejos e crenças (FRANKFURT, 1998, p. 1-10).

David Hume é reconhecidamente um defensor da posição compatibilista clássica, que defende que a liberdade e a responsabilidade moral podem ser reconciliadas com o determinismo causal. Esse será o tema do item a seguir.

1.2 Compatibilismo clássico: Hume

Vejamos o que Hume tem a dizer sobre liberdade e necessidade no *Tratado sobre a Natureza Humana* e nas *Investigações sobre o Entendimento Humano*. No *Tratado*, Hume

chama a atenção para uma distinção entre dois tipos de liberdade, a liberdade da espontaneidade e a liberdade da indiferença:

Poucos são capazes de fazer uma distinção entre a liberdade de espontaneidade, como é chamada na escolástica, e a liberdade de indiferença, ou seja, entre aquilo que se opõe à violência e aquilo que significa uma negação da necessidade e das causas. O primeiro sentido da palavra é o mais comum; e, uma vez que é somente essa espécie de liberdade que nos interessa preservar, nossos pensamentos têm-se voltado sobretudo para ela, confundindo-a quase sempre com a outra (HUME, 2009, p. 443).

Para Hume, as ações livres são aquelas causadas pelos desejos do agente, sendo, portanto, voluntárias. O agente é responsável pela ação se os seus desejos forem as causas determinantes da ação. Nas circunstâncias em que o agente é forçado ou compelido a agir, isto é, quando sua ação resulta da violência ou constrangimento de algum tipo, mesmo que essa ação possa ser dita ser necessitada ou causada, não é resultante da vontade ou dos desejos do agente e, portanto, o agente não é responsável por ela.

Para Hume, o que definirá uma ação livre e responsável não será a ausência de causa, pois uma ação pode ser causada e ainda carecer de liberdade e responsabilidade, mas o que irá distinguir uma ação livre e responsável de uma que carece de liberdade e responsabilidade será um tipo diferente de causa. Ações livres e responsáveis são causadas pelos nossos próprios desejos. Ações não livres ou não responsáveis são causadas ou necessitadas por causas externas ao agente, que é o caso da ação forçada.

Podemos encontrar a continuação da definição de Hume de liberdade no contexto da *Investigação*. Aqui define-se liberdade hipotética, uma explicação de liberdade distinta, mas relacionada à liberdade da espontaneidade:

Por liberdade, então, só nos é possível entender um poder de agir ou não agir, de acordo com as determinações da vontade; isto é, se escolhermos ficar parados, podemos ficar assim, e se escolhermos nos mover, também podemos fazê-lo. Ora, essa liberdade hipotética é universalmente admitida como pertencente a todo aquele que não esteja preso e acorrentado. Não há aqui, portanto, matéria para disputas (HUME, 2003, p. 136-137).

Hume está dizendo que não devemos confundir a liberdade hipotética com a liberdade da espontaneidade. Uma pessoa pode desfrutar de liberdade da espontaneidade, mas carecer de liberdade hipotética. Por exemplo, eu posso escolher ficar na sala ao invés de sair, porém pode ser o caso que se eu tivesse escolhido sair eu não poderia ter saído porque a porta estava trancada. É importante distinguir a liberdade da espontaneidade da liberdade hipotética no

debate da vontade livre, uma vez que a liberdade hipotética representaria a explicação compatibilista de possibilidades alternativas.

A posição compatibilista de Hume consiste não apenas na apresentação de uma definição de liberdade da espontaneidade. Hume também elabora um argumento anti-libertista contra a posição rival. Esse argumento é desenvolvido usando características da sua própria definição de liberdade da espontaneidade com o objetivo de destacar uma falha fatal na posição libertista.

Conforme citado anteriormente, a partir do *Tratado* a liberdade da indiferença significa uma negação da necessidade e causa. Para o libertista, é uma condição necessária da responsabilidade moral que ação não seja necessitada ou causada. Para Hume, em contraste, se a necessidade é removida das ações humanas a causalidade também é eliminada. E a eliminação da necessidade e da causalidade das ações acabaria por reduzir a liberdade ao acaso (HUME, 2009, p. 443). Conforme Hume, se a explicação das ações não procede de alguma causa no caráter e na disposição da pessoa que as realizou, então as ações não podem ser honradas, quando boas, ou censuradas, quando más. Não seria razoável responsabilizar um agente pelas ações resultantes da indiferença ou da violência externa.

Há muitas objeções incompatibilistas ao argumento da espontaneidade. A crítica consiste em apontar a completa insuficiência da liberdade da espontaneidade. Kant, por exemplo, afirmou que se nossa vontade é determinada por causas naturais antecedentes, então a nossa vontade não seria diferente de qualquer outro objeto mecânico cujos movimentos são internamente condicionados. Agentes que desfrutam de uma liberdade desse tipo são como marionetes sujeitos ao destino e à natureza. O próximo subcapítulo deste capítulo tratará da concepção kantiana de liberdade, mas antes disso eu gostaria de ressaltar algumas objeções endereçadas ao argumento da espontaneidade.

O incompatibilista diz que se as nossas escolhas são determinadas por causas antecedentes, então não poderíamos ter escolhido outra coisa a não ser o que nós de fato escolhemos. Não poderíamos, assim, ser responsáveis pela nossa conduta, pois nessa explicação nós não teríamos possibilidades alternativas. A noção de liberdade hipotética, segundo o incompatibilista, não pode lidar com essa objeção. A liberdade hipotética implicaria a ideia de que nós poderíamos ter agido de outro modo *se* tivéssemos escolhido agir diferentemente. No entanto, há casos em que o agente não poderia ter feito de outro modo dadas as circunstâncias reais. Para o incompatibilista, a liberdade hipotética não é suficiente, responsabilidade requer liberdade categórica para escolher de outro modo nas *mesmas circunstâncias*. Dito de outro modo, para o incompatibilista responsabilidade exige mais do que

apenas liberdade da ação, nós também precisamos de liberdade da vontade, isto é, um poder para escolher entre alternativas. Apenas liberdade da ação não garantiria que nós tivéssemos um controle último sobre a nossa conduta. Nós devemos ser a origem não apenas das nossas ações, mas também da nossa vontade.

Esse ponto ficará mais claro a partir de uma concessão feita pela própria posição compatibilista. Como visto anteriormente, Hume distinguiu as ações livres das ações não livres (isto é, as ações compelidas) em termos da diferença entre as causas internas e as externas. O problema é que essa distinção aparentemente não se sustenta, pois há casos em que o agente age conforme a sua vontade, porém carece de liberdade. Por exemplo, há possibilidade dos próprios desejos serem compulsivos, como é o caso de um drogadito ou um cleptomaníaco. Um desejo desse tipo parece enfraquecer a liberdade tanto quanto a força externa. O agente estaria agindo de acordo com os seus próprios desejos, porém esse agente não seria livre tampouco responsável pelo seu comportamento. Nesse sentido, se poderia dizer que há algumas causas internas ao agente que seriam compelidas. A dificuldade decorrente dessa concessão é que não é evidente quais causas internas deveriam ser consideradas como compulsivas e quais não deveriam. Por trás dessa objeção está a preocupação de que o argumento da espontaneidade pressupõe um entendimento inadequado da natureza das desculpas (RUSSELL, 2016).

A distinção de Hume entre dois tipos de liberdades é central para a sua estratégia compatibilista. O argumento da liberdade (que engloba a definição de Hume de liberdade da espontaneidade e a liberdade anti-libertista) não traz nada novo em relação ao que já foi dito por Hobbes, por exemplo. A novidade de Hume estaria em apresentar uma nova definição de necessidade.

Na *Investigação sobre o Entendimento Humano*, Hume define a doutrina da necessidade como se segue:

Nessas duas circunstâncias, esgota-se toda a necessidade que atribuímos à matéria. Fora da *conjunção* constante de objetos semelhantes, e da conseqüente *inferência* de um ao outro, não temos a menor ideia de qualquer necessidade ou conexão. Se verificamos, portanto, que toda a humanidade sempre reconheceu, sem nenhuma dúvida ou hesitação, que essas duas circunstâncias ocorrem nas ações voluntárias dos homens e nas operações da mente, seguir-se-á que toda a humanidade sempre esteve de acordo quanto à doutrina da necessidade, e que têm polemizado até agora simplesmente por não se entenderem uns aos outros (HUME, 2003, p. 122).

Devemos concluir que a mesma necessidade ocorre na esfera moral. Além disso, Hume revela que universalmente se admite a existência de uma grande uniformidade nas ações humanas. A regularidade das nossas ações fica evidente a partir da seguinte passagem:

Parece, então, não apenas que a conjunção entre motivos e ações voluntárias é tão regular e uniforme como a que existe entre a causa e o efeito em qualquer parte da natureza, mas também que essa conjunção regular tem sido universalmente reconhecida pela humanidade, e nunca foi objeto de disputa, seja na filosofia, seja na vida ordinária (HUME, 2003, p. 128-129).

No *Tratado* Hume reitera que é uma questão de fato que existe uma conjunção constante na esfera moral bem como na esfera natural. Hume admite a possibilidade de irregularidades aparentes e incertezas em ambas as esferas natural e moral, mas que “o acaso e a indiferença se deve ao nosso conhecimento imperfeito, e está apenas no nosso julgamento, não nas próprias coisas” (HUME, 2009, p. 440).

Para Hume, assim como a união entre os motivos e as ações tem a mesma constância que a união entre os objetos naturais, a inferência na esfera moral é mesma “*determinando-nos a inferir a existência de uns da existência dos outros*” (HUME, 2009, p. 440). A explicação de Hume sobre a uniformidade das ações humanas se baseia sobre dois fatos incontestáveis sobre a humanidade: “que os homens sempre buscam a sociedade”, e os homens consideram as outras pessoas como objetos de louvor ou censura (HUME, 2009, p. 438).

Para as pessoas viverem em sociedade, e responsabilizarem umas as outras, elas devem ser capazes de prever as ações do outro a partir do seu caráter, e o seu caráter a partir da suas ações. Hume afirma que é da experiência passada que deduzimos o comportamento futuro:

A dependência mútua entre os homens é tão grande em todas as sociedades que dificilmente haverá uma ação humana inteiramente completa em si mesma, ou realizada sem alguma referência às ações de outros que são requeridas para fazê-la corresponder plenamente à intenção do agente. (...) Em todas essas conclusões, do mesmo modo que em seus raciocínios sobre objetos externos, eles extraem seus padrões da experiência passada, e crêem firmemente que os homens, assim como todos os elementos, devem continuar, em suas operações, a portar-se como sempre se observou (HUME, 2003, p. 129-130).

Para Hume a sociedade não seria possível se a esfera moral estivesse sujeita à liberdade da indiferença, ou seja, tivesse ausência de regularidade e inferência.

As observações de Hume sugerem que as teorias metafísicas tradicionais de causalção têm confundido a noção de um evento sendo causado com a de um evento sendo forçado. Partindo de uma concepção equivocada de causalção, muitos filósofos têm defendido uma

conclusão igualmente equivocada de que deveria haver uma incompatibilidade entre o determinismo e a liberdade. Para Hume, ação livre e responsável deve ser causada pelos desejos do agente e não deve ser forçada.

O argumento da espontaneidade apresentado por Hume é interpretado como contendo as sementes de uma explicação utilitarista de responsabilidade moral. Seguindo Hobbes, Hume apontou que recompensas e punições serviriam para fazer com que as pessoas agissem corretamente, e não de forma errada, tendo uma preocupação em regular o comportamento futuro, sendo essa uma questão de utilidade social.

Essa explicação utilitarista de responsabilidade foi criticada por não tratar de questões de merecimento, portanto, carecendo do elemento retributivo retrospectivo, isto é, não possuindo uma visão voltada ao comportamento já realizado, no passado (RUSSELL, 2016). Mais sobre essa questão será tratado no item que trata sobre o projeto de Peter Strawson de reconciliação entre os compatibilistas e os incompatibilistas. Mas, antes disso, gostaria de apresentar uma abordagem libertista de liberdade, com o intuito de esclarecer melhor o problema a que ela também estaria sujeita.

1.3 Libertismo: Kant, Inwagen e Clarke

Neste subcapítulo, Immanuel Kant e Randolph Clarke serão apresentados como representantes do libertismo¹⁶, enquanto Peter van Inwagen como um representante do incompatibilismo.

Kant associa a liberdade ao conceito de causalidade. Na terceira seção da *Fundamentação da Metafísica dos Costumes*, Kant afirma que a vontade é “uma espécie de causalidade dos seres vivos, enquanto racionais, e *liberdade* seria a propriedade dessa causalidade” (KANT, 2007, p. 93). Em comparação com o que Hume disse acerca da causalidade das ações, para Kant a questão não é sobre se a causa da ação é interna ou externa. O que importa é se a ação está no controle do agente. Para esclarecer esse ponto podemos considerar o exemplo, oferecido pelo próprio Kant, de um homem que cometeu um roubo. Avaliar se essa ação foi moralmente errada consistiria para Kant verificar, dentre outras coisas, se essa ação estava sob seu controle. Estar sob o seu controle significa que estava em poder do agente no dado momento não ter realizado o roubo. Se a ação não estivesse em poder do agente, então não seria apropriado dizer que a ação foi moralmente errada e passível de punição (KANT, 2016, p. 155).

¹⁶ Roderick Chisholm (2013) também é um representante do libertismo.

Para o compatibilista a ação do ladrão é livre porque a causa está nele e porque o roubo foi uma ação voluntária. O ladrão deliberou e decidiu cometer o roubo, portanto a ação seguiu da sua decisão prévia. O ponto que gostaria de chamar atenção é que para Kant se a decisão do ladrão é vista como um fenômeno natural, que ocorre no tempo, então deve ser o efeito de alguma causa anterior. Pressupondo que todos os eventos naturais ocorram no tempo, e são determinados por cadeias causais que remontam ao passado distante, não haveria lugar para a liberdade. O problema estaria na ideia de tempo. Se a escolha do ladrão para cometer o roubo é um evento natural no tempo, então é o efeito de uma cadeia causal que remete ao passado. O passado está fora de nosso controle agora no presente. Não podemos mudar o passado. Se as nossas ações fossem determinadas por eventos no passado, então as ações não estariam, no presente, sob o nosso controle. Se a decisão do ladrão para cometer o roubo é um evento natural que ocorre no tempo, então não está sob o controle do agente agir diferentemente, isto é, ele não poderia ter feito outra coisa a não ser roubar. Se a minha ação hoje é determinada por uma decisão que eu fiz ontem, ou se a minha ação está determinada pelo meu caráter que eu formei na minha infância, então a ação não estaria sob o meu controle agora.

É a partir da distinção entre o mundo como numênico e o mundo como fenômeno que é possível, segundo Kant, explicar como que a causa da minha ação pode estar sob o meu controle. O idealismo transcendental explicaria o tipo de liberdade que a moralidade exige. A partir do idealismo transcendental seria possível localizar a causa da minha ação fora do tempo, isto é, no eu enquanto numênico, que é livre porque existe em um mundo do entendimento, não estando sujeito às leis da natureza e da causalidade necessitada pelos eventos passados (KANT, 2016, p.186-187). Para Kant o tipo de caráter que eu tenha desenvolvido ou as influências externas que tenham contribuído para a formação do meu caráter seria irrelevante, pois o que importaria seriam as ações voluntárias e intencionais vistas a partir da perspectiva do eu enquanto numênico, sendo causalmente indeterminadas.

Uma das objeções ao compatibilismo e que teve grande influência na literatura contemporânea foi o chamado argumento da consequência. Inwagen procurou, através desse argumento, demonstrar a incompatibilidade da vontade livre e o determinismo (INWAGEN, 2013, p. 38-57).

De acordo com o argumento da consequência, se o determinismo é verdadeiro, então nenhuma pessoa pode agir a não ser do modo como agiu, pois existiria apenas um futuro possível. O argumento pode ser esboçado da seguinte forma:

1. Não há algo que possamos fazer para mudar o passado e as leis da natureza.

2. Nossas ações presentes são consequências necessárias do passado e das leis da natureza.
3. Portanto, nós não podemos agir a não ser do modo como na realidade agimos.

O que estaria errado com o argumento da consequência? Muitos críticos costumam focar na expressão “não há algo que possamos fazer para mudar...”. Podemos observar que essa expressão contém a palavra ‘pode’, uma mais difíceis de qualquer linguagem para interpretar. Em linhas gerais, falar sobre o que uma pessoa pode ou não fazer significa falar sobre os seus poderes ou capacidades. Compatibilistas afirmam que se os termos “poder” e “capacidade” foram interpretados em um sentido hipotético, então o argumento da consequência seria falho. Para os compatibilistas clássicos, como visto anteriormente, dizer que uma pessoa tem o poder ou capacidade para fazer alguma coisa significa que ela teria feito se, na ausência de obstáculos externos, ela tivesse escolhido ou desejado.

Para os compatibilistas, as premissas do argumento da consequência são verdadeiras, pois nenhuma pessoa poderia mudar o passado ou as leis da natureza mesmo se escolhesse ou desejasse, já que não está sob o poder de qualquer pessoa mudar o passado ou as leis da natureza. O que estaria errado com o argumento da consequência seria a conclusão “nós não podemos agir a não ser do modo como na realidade agimos”. A conclusão sendo falsa o argumento da consequência seria inválido. Para mostrar a falsidade da conclusão podemos considerar uma ação simples como o fato de eu levantar a mão. Dizer que eu poderia ter feito outra coisa que não levantar a mão (ou seja, não levantar a mão) significa, a partir da análise hipotética e condicional, que eu teria feito de outro modo se eu tivesse escolhido ou desejado fazer diferentemente. A afirmação hipotética pode ser verdadeira mesmo se o meu ato foi determinado. Isso significa que a análise hipotética teria a implicação de que eu teria feito de outro modo *se o passado tivesse sido diferente em algum modo* (isto é, se eu tivesse escolhido ou desejado diferentemente). Portanto, nós podemos agir diferentemente do modo como na realidade agimos. Eu posso fazer o contrário de levantar a mão porque nada teria me impedido externamente (por exemplo, minha mão não está amarrada) se eu tivesse desejado não levantar a mão.

A expressão “poderia ter feito de outro modo” admite pelo menos dois significados:

1. A possibilidade de agir diferentemente, todas as outras condições sendo as mesmas.
2. A possibilidade de agir diferentemente, todas as outras condições sendo um pouco diversas.

A opção 1 estaria representando o argumento da consequência de Inwagen. Portanto, interpretando nesse sentido não seria possível uma reconciliação entre a liberdade e o determinismo. Já na opção 2 poderia haver a possibilidade dessa compatibilidade. Para explicar esse ponto eu cito o exemplo de Costa:

Um médico culpa-se a si mesmo por não ter pedido um certo exame, o que lhe teria levado o diagnóstico correto. Ele pensa: “Eu poderia ter agido de outro modo”. Ele não está pensando, porém, que ele poderia ter agido de outro modo, todas as outras condições permanecendo as mesmas; o que ele considera é que, se as condições internas que o levaram a agir tivessem sido um pouco diversas, ou seja, se ele tivesse seguido os procedimentos recomendados pela presença de certos sintomas, só esse cuidado já o teria levado ao diagnóstico correto; aqui a condição de que ele fosse um pouco mais responsável ao agir é que teria sido diferente. Com efeito, se a diferença requerida nas condições determinadoras da decisão fosse externa ao agente, ou se ela fosse muito grande, não dizemos que poderíamos ter agido de outro modo (COSTA, 2000, p. 28).

A expressão “poderia ter feito de outro modo” tem um sentido retrospectivo e prospectivo. Retrospectivo porque remete ao passado, quando o médico diz que ele “poderia ter agido de outro modo” significa que se ele tivesse tomado as precauções necessárias ele teria descoberto o diagnóstico correto. Prospectivo porque se dirige ao futuro, à ideia de reforma do comportamento futuro. Quando o médico se culpa ou pede desculpas isso significa que a correção do seu erro está ao seu alcance, e que ele é capaz de melhorar a sua conduta futuramente.

O argumento da consequência, mesmo se consistente, não mostra que nós teríamos vontade livre. Um argumento incompatibilista serve apenas para mostrar que a vontade livre e o determinismo não podem ser ambos compatíveis. O libertista pode ser definido como aquele que aceita as seguintes sentenças:

1. Vontade livre e determinismo são incompatíveis (incompatibilismo).
2. Vontade livre existe.
3. E determinismo é falso.

O libertista além de mostrar a incompatibilidade da vontade livre com o determinismo, deve também mostrar que nós podemos ter uma vontade livre que é incompatível com o determinismo.

Mas como que o libertista explica as ações livres? Para mostrar como as ações livres podem ser explicadas sem invocar uma causação física ou natural, o caminho inicial consiste em apontar para formas incomuns de agência ou causação¹⁷. Por exemplo, somos motores imóveis, causas não causadas, eu numênico, fora do espaço e tempo, entre outras (KANE, 2005, p. 33). A ideia de que nós precisaríamos acreditar na liberdade libertista para fazer sentido tanto da moralidade como da responsabilidade, como visto antes, foi defendida por Kant, e é importante enfatizar que ele admite que nós não podemos entender completamente uma tal liberdade em termos teóricos ou científicos.

¹⁷ O artigo de Terry Horgan (2015) fornece uma explicação sobre o significado do conceito de agência livre. Como é a experiência de si mesmo como um agente? Dito de outro modo, quais são as características da fenomenologia da agência? O agente tem a experiência do seu comportamento sendo causado por ele mesmo. Os libertistas, em geral, denominam a agência nesse sentido como causação pelo agente. Essa experiência diferencia-se daquela dos movimentos corporais sendo causados pelos seus próprios estados mentais. Por exemplo, quando alguém sente arrepio e sente a experiência desse arrepio como causado por um estado de medo. Além disso, a fenomenologia inclui também o aspecto das possibilidades alternativas, pois quando o agente tem a experiência de si mesmo como livremente realizando a ação, “está no agente” se ou não ele realiza a ação. Em resumo, o agente tem a experiência de si mesmo como causando a ação intencionalmente e também tem a experiência de causar a ação de modo que ele poderia ter feito de outro modo (HORGAN, 2015, p. 34-36). Ambos os aspectos da fenomenologia estão conectados, o da capacidade de agir de outro modo e o do agente como causa da sua ação, da seguinte forma: “ao experimentar o próprio comportamento como derivando de si mesmo como sua origem o agente experencia a si mesmo como sendo capaz de se abster – ou como sendo capaz de se abster intencionalmente de produzir tal comportamento” (HORGAN, 2015, p. 37). Até aqui tratou-se apenas do caráter introspectivo da fenomenologia da agência livre. Para Horgan, é preciso explorar a relevância dessa fenomenologia para o debate da vontade livre discutindo suas conexões com as seguintes dimensões: conceitual, metafísica, epistêmica e moral. Sobre a dimensão conceitual, integra o conteúdo do conceito de agência livre com o conteúdo da fenomenologia da agência livre (isto é, o aspecto da experiência e a capacidade para agir de outro modo). Assume que a noção libertista de agência é incoerente. Sobre a dimensão metafísica, reconhece a crença defendida pela grande maioria dos filósofos, de que a agência livre é um fenômeno genuíno e compatível com a hipótese científica de que as nossas escolhas e ações fazem parte de um processo determinístico. Sobre a dimensão epistêmica, reconhece o fato de que as pessoas acreditam que elas exercem agência livre em razão do caráter experiencial da fenomenologia da própria agência livre. Além disso, assume que as outras pessoas normalmente exercem esse mesmo tipo de agência livre. Seria essa fenomenologia que tornaria justificada a crença na agência livre. E, por último, a dimensão moral: assume que para alguém ser moralmente responsável por um ato, e para as atitudes reativas serem moralmente justificadas, a pessoa deveria exercer o tipo de agência moral que nós humanos experienciamos. Dito de outro modo, a responsabilidade moral e a liberdade estariam conectadas com a capacidade para agir diferentemente (HORGAN, 2015, p. 42-44). Para Horgan, a posição compatibilista seria mais vantajosa do que a do libertista pelas seguintes razões: o libertista falha em dar uma explicação satisfatória do agente como origem das suas escolhas e ações e, portanto, a agência livre acaba sendo conceitualmente misteriosa e incoerente. Apesar de reconhecer que agência livre é um fenômeno genuíno, essa explicação entra em conflito com a hipótese científica de que nossas escolhas e ações resultam de um processo determinístico. Dado que a noção libertista de agência é misteriosa e a possibilidade da hipótese científica determinista, a noção de que os seres humanos exercem agência livre careceria de justificção epistêmica (HORGAN, 2015, p. 47-49).

Críticos da liberdade libertista argumentam que os eventos que são indeterminados acontecem meramente por acaso, e, como não estão sob o controle de alguma coisa, não estão, portanto, também sob o controle do agente. Ademais, afirmam também que a noção de liberdade libertista não é nem mesmo inteligível, ou possível, e suas explicações de causação são obscuras e misteriosas (KANE, 2005, p. 33 e 35). Em síntese, a crítica é que as ações indeterminadas seriam arbitrárias (por exemplo, alguém indeciso sobre o que fazer pode lançar para cima uma moeda e decidir com base nisso), irracionais ou uma questão de sorte. O que estaria por trás de todas essas objeções é o problema de como reconciliar as ações livres com o indeterminismo, isto é, como é possível o agente ser capaz de escolher (ou agir) diferentemente *dadas as mesmas circunstâncias passadas e leis da natureza* (KANE, 2005, p. 38). Dito de outro modo, parece irracional, inexplicável e arbitrário eu escolher passar as minhas férias no Brasil, especificamente no Tocantins, que é calor durante todo o ano, dada a mesma deliberação anterior (o fato de eu gostar do frio) que me levou a escolher passar as férias na Rússia.

Os libertistas tentaram responder essa objeção argumentando que a capacidade do agente para agir diferentemente não pode ser explicada unicamente pelas circunstâncias anteriores à sua ação, pois não há nenhuma diferença nessas circunstâncias anteriores. Para que a ação ou escolha não seja meramente por acaso, arbitrária ou inexplicável, algum fator extra, não incluído entre as circunstâncias passadas, deve explicar por que o agente age ou escolhe de um modo ao invés de outro. Os libertistas, ao longo da história, invocaram diferentes fatores extras para explicar como a vontade livre seria possível nesse sentido. Por exemplo, para Kant o eu numérico fora do espaço e tempo, para René Descartes o dualismo da mente e corpo. Se a mente fosse distinta do corpo ela estaria fora do mundo físico e a sua atividade não seria governada pelas leis da natureza que governam os eventos físicos. Seriam esses fatores extras que explicariam por que as ações ou escolhas livres não meramente ocorrem por acaso ou arbitrariamente. Porém, essa explicação também foi criticada pelo fato de exigir suposições metafísicas, transformando o “eu” num mistério, sem questionamento.

Há uma outra estratégia libertista moderna que se tornou popular entre os filósofos contemporâneos. Essa estratégia é conhecida como a teoria da causação pelo agente. Chisholm é um defensor dessa posição:

Se nós consideramos unicamente objetos naturais inanimados, podemos dizer que causação, se ela ocorre, é uma relação entre *eventos e estados de coisas*. O rompimento da barragem foi um evento que foi causado por um conjunto de outros eventos, a barragem sendo fraca, a inundação sendo forte, e assim por diante. Mas, se o homem é responsável por uma ação particular, então (...)

existe algum evento ou conjunto de eventos que é causado, não por outros eventos ou estados de coisas, *mas pelo agente* (CHISHOLM, 2013, p. 30).

Podemos dizer, segundo Chisholm, que as ações livres são causadas. Mas essas ações livres não são causadas por circunstâncias anteriores, eventos ou estados de coisas. O agente foi a causa da sua ação, ele age com a capacidade para agir diferentemente. Nesse sentido, o agente foi uma causa não causada do seu próprio agir. Tal explicação de vontade livre é considerada hoje desacreditada.

Discordando da explicação de Chisholm, Clarke propõe uma nova compreensão de liberdade da vontade. Ele reconcilia elementos de causação pelo agente com causação por eventos. A explicação de liberdade, nesse sentido proposta por Clarke, pode ser vista a partir de um exemplo. Digamos que eu me deparo com duas possibilidades de escolha mutuamente exclusivas, pois os dois eventos ocorrem no mesmo horário: (a) ir a uma palestra de um conhecido ou (b) ir ao cinema com um amigo. Supondo que eu escolha ir à palestra, o que explicaria essa escolha seria o fato de eu ter razões e agir a partir delas. As razões para eu escolher ir à palestra são: eu estou interessada na palestra e eu conheço o palestrante. Eu seria a causa do meu agir quando eu ajo a partir de um conjunto de razões ao invés de um diferente conjunto de razões. O meu comportamento (escolha ou ação) é causado por mim (causado pelo agente) e é causado por eventos anteriores (como crenças, desejos, preferências, valores e razões). A noção de causação pelo agente é o que explicaria eu ter o controle sobre a escolha que fiz. Eu deliberei a partir das razões e escolhi com base nas razões mais significantes pra mim (CLARKE, 2013, p. 285-298).

Eu penso que a estratégia libertista de Clarke, quando invoca a noção de causação pelo agente, estaria sujeita às mesmas críticas que as outras teorias de causação pelo agente. Isso quer dizer que essa estratégia não explicaria como que o agente controla ou escolhe uma ação ao invés da outra.

A seguir apresentarei uma teoria compatibilista contemporânea que tenta reconciliar o incompatibilismo com o compatibilismo.

1.4 Compatibilismo contemporâneo: Strawson e as atitudes reativas

O famoso artigo “Freedom and Resentment” (2008), de Peter Strawson, trouxe uma contribuição importante para o debate da vontade livre, na medida em que procurou reconciliar duas posições antagônicas: o compatibilismo e o incompatibilismo. Para os incompatibilistas (a quem Strawson chamou de pessimistas), se a tese determinista fosse verdadeira, os conceitos de obrigação e responsabilidade moral não teriam aplicação, e as práticas de punição e culpa,

condenação e aprovação moral, seriam injustificadas. Já os compatibilistas (os otimistas) afirmam que esses conceitos e práticas são justificadas mesmo sendo verdadeira a tese determinista. Há ainda a posição do cético moral, que afirma que as noções de culpa moral, de censura e de responsabilidade moral são confusas, sendo o determinismo verdadeiro ou falso (STRAWSON, 2008, p. 1).

Como reconciliar posições tão opostas como a dos pessimistas e a dos otimistas? Uma possibilidade seria uma parte recuar de um lado, e a outra parte fazer uma concessão de outro lado. Conforme Strawson:

Alguns dos otimistas sobre o determinismo assinalam a eficácia das práticas de castigo, e da condenação e aprovação moral, na regulação do comportamento de formas socialmente desejáveis. Pelo fato de sua eficácia, sugerem, há uma base adequada para essas práticas; e esse fato certamente não mostra que o determinismo seja falso. A isso o pessimista replica, e de imediato, que o castigo *justo* e a condenação *moral* implicam culpa moral, e que a culpa moral implica a responsabilidade moral, e a responsabilidade moral implica a liberdade, e a liberdade implica a falsidade do determinismo (STRAWSON, 2008, p. 2-3)¹⁸.

Os otimistas costumam responder que essas práticas exigem liberdade, e que a existência de liberdade, nesse sentido, significa a ausência de certas condições como coerção, incapacidade inata, insanidade, ou outras formas menos extremas de desordem psicológica. A essa lista acrescenta-se outros fatores que, sem ser limitações de liberdade, podem tornar a condenação ou castigo moral inapropriados como, por exemplo, algumas formas de ignorância, erro ou acidente (STRAWSON, 2008, p. 3).

Os pessimistas concordam com a definição de liberdade no sentido negativo que os otimistas reconhecem. Eles também admitem que a existência da liberdade nesse sentido é compatível com o determinismo. Mas os pessimistas acrescentariam que o agente deve merecer o castigo, a culpa ou a condenação moral. E no caso em que o agente é culpado por uma ação que ele fez, para que ele realmente mereça ser censurado é exigido um sentido de liberdade que é incompatível com o determinismo (STRAWSON, 2008, p. 3).

Para os otimistas, a verdade da tese determinista não implica que ninguém decide fazer nada, ou que não se faz nada intencionalmente, ou que é falso que as pessoas sabem o que estão fazendo. Agora os pessimistas podem questionar: por que liberdade, nesse sentido, justifica a culpa? A razão que os otimistas fornecem para as práticas de condenação e de castigo moral,

¹⁸ Essa tradução do texto de Strawson é de Jaimir Conte (2015, p. 246).

nos casos em que a liberdade negativa está presente, é a eficácia das práticas sociais para regular o comportamento de modo socialmente desejável. Quanto a isso os pessimistas diriam que a admissibilidade dessas práticas exige um tipo de liberdade que é incompatível com a tese determinista (STRAWSON, 2008, p. 4).

Strawson está interessado na grande importância que nós atribuímos às atitudes e intenções que as outras pessoas adotam em relação a nós e o quanto os nossos sentimentos e reações pessoais envolvem nossas crenças acerca dessas atitudes e intenções. O ponto é destacar o quanto valorizamos que as ações das outras pessoas reflitam para nós atitudes de boa vontade ou afeto, por um lado, ou desprezo e indiferença, por outro. Digamos que alguém pise na minha mão acidentalmente, pois ele estava tentando me ajudar e não foi a sua intenção me ferir. Nesse caso eu irei desculpá-lo. Porém, se ele pisa na minha mão de forma desrespeitosa, com um desejo de me ferir, então eu sentirei ressentimento (STRAWSON, 2008, p. 5-6).

Muitos são os tipos de relações que podemos ter com as outras pessoas: como participantes de um interesse comum, como membros de uma mesma família, como amigos, como colegas, como amantes, por exemplo. Em geral, esperamos algum grau de boa vontade ou estima das pessoas que mantém relações interpessoais conosco (STRAWSON, 2008, p. 6).

Na presença de determinadas condições as atitudes reativas como o ressentimento, por exemplo, podem ser suspensas. Essas condições podem ser divididas em dois grupos. No primeiro grupo estariam expressões como “Ele não teve intenção”, “Não sabia”, “Ele não pôde evitá-lo”, “Ele foi compelido” ou “Não lhe deixaram alternativa”. Quando essas condições estão presentes elas nos levam a considerar a ofensa em questão como inapropriada (STRAWSON, 2008, p. 7-8).

O segundo grupo de considerações divide-se em dois subgrupos. No primeiro subgrupo encontram-se expressões como “Ele agiu sob o efeito da hipnose” ou “Ele não era o mesmo”. No segundo subgrupo considerações como “Ele é uma criança”, “Ele é um esquizofrênico” ou “Seu comportamento é compulsivo”. Essas desculpas, diferentemente daquelas do primeiro grupo, nos convidam a ver o agente como um candidato inapropriado para as atitudes reativas. Nesse segundo subgrupo permite-se que as circunstâncias sejam normais, mas o agente é apresentado como anormal, psicologicamente imaturo. Nessas circunstâncias nossas atitudes reativas em relação a essa pessoa tendem a modificar-se. O que Strawson pretende destacar aqui é a distinção entre a atitude de envolvimento ou participação nas relações pessoais, por um lado e a atitude objetiva, por outro. Strawson define a atitude objetiva como se segue:

A adoção da atitude objetiva em relação a outro ser humano consiste em vê-lo, talvez, como objeto de política social, como sujeito a quem, em um sentido muito amplo, caberia submeter a um tratamento; como alguém que certamente deve ser levado em conta, talvez tomando medidas preventivas; a ser controlado, tratado ou treinado (STRAWSON, 2008, p. 9)¹⁹.

A atitude objetiva pode incluir repulsa ou medo, piedade e inclusive amor, ainda que nem todos os tipos de amor. Porém, não se pode incluir as atitudes e os sentimentos como ressentimento, gratidão, perdão e o amor que dois adultos sentem um pelo outro, que são próprios à participação nas relações interpessoais. Ver alguém como um ser pervertido, ou transtornado, ou compulsivo no seu comportamento, implica adotar uma atitude objetiva em relação a essa pessoa. Também pode se adotar uma atitude objetiva em relação à conduta de uma criança pequena com o intuito de tratá-la ou educá-la. Além disso, podemos inclusive adotar essa atitude objetiva com relação ao comportamento do ser humano normal e maduro. Claro que não podemos fazer isso por muito tempo, em uma situação normal. Se a tensão continuar por muito tempo o melhor é romper a relação (STRAWSON, 2018, p. 10).

Já as atitudes reativas, Strawson define como “reações humanas naturais diante da boa ou má vontade ou diante da indiferença dos demais, conforme se manifestam em *suas* atitudes e ações” (STRAWSON, 2018, p. 10-11). A questão que teríamos que fazer, segundo Strawson, é que efeito teria sobre essas atitudes reativas caso nós aceitássemos a tese determinista? Dito de outro modo, a aceitação da tese determinista levaria à rejeição das atitudes reativas? Caso admitíssemos sua verdade, isso significaria o fim do ressentimento, da gratidão e do perdão? A verdade da tese determinista não implicaria em nós adotarmos uma atitude objetiva nas nossas relações interpessoais (STRAWSON, 2008, p. 11).

As atitudes reativas pessoais tendem a dar lugar às atitudes objetivas nas circunstâncias em que o agente é considerado excluído das relações humanas adultas por ter alguma anormalidade psicológica, ou por ser imaturo. Porém, o ponto é que a tese determinista não deveria implicar que a anormalidade fosse universalizada. Adotar sempre uma atitude objetiva causaria um isolamento, algo que os seres humanos não seriam capazes de fazer. Nesse sentido, uma convicção teórica sobre a verdade do determinismo não implicaria em nós adotarmos em relação aos outros uma atitude de completa objetividade. E mesmo que empregássemos tal atitude em um caso particular a razão de por que o faríamos não se deve ao fato de acreditarmos na possibilidade de um mundo determinista, mas ao fato de que quando abandonamos as

¹⁹ Essa tradução do texto de Strawson é de Jaimir Conte (2015, p. 253).

atitudes reativas nas relações interpessoais é devido à presença de condições particulares, tal como o comportamento anormal (STRAWSON, 2008, p. 12-14).

As atitudes reativas discutidas até aqui, em particular o ressentimento, são reações à qualidade da vontade dos demais em relação a nós mesmos. O ressentimento expressa uma reação nossa ou à ofensa ou à indiferença manifestada no comportamento de uma outra pessoa em relação a nós mesmos. Dito de outro modo, aqui trata-se do ponto de vista daquele cujo interesse está diretamente implicado. Strawson também discute nesse texto atitudes reativas que não envolvem a relação entre os ofendidos e os beneficiados, mas são semelhantes às atitudes reativas pessoais. Deixando de lado a atitude reativa do ressentimento, Strawson colocará no centro da discussão agora a indignação moral, ou a desaprovação moral, isto é, atitudes reativas vicárias, impessoais. Essas são reações à qualidade da vontade das pessoas, não em relação a nós, mas às outras pessoas. Aqui trata-se do ponto de vista daqueles cujos interesses não estão diretamente envolvidos (STRAWSON, 2008, p. 14-15)²⁰.

A suspensão das atitudes vicárias, assim como ocorre com as atitudes reativas pessoais, não é uma consequência da verdade da tese determinista. Nem segue a partir dessa tese que ninguém saberia o que está fazendo ou que ninguém seria capaz de adotar atitudes impessoais ou autorreativas (STRAWSON, 2008, p. 19).

Tendo dito isso, agora é o momento de tentar preencher a lacuna que o pessimista encontra na explicação do otimista do conceito de responsabilidade moral. Para o otimista, a verdade do determinismo não significaria a inexistência do conceito de responsabilidade moral e das práticas de condenação moral. Strawson acredita que a lacuna na explicação do otimista pode ser preenchida levando em conta as atitudes e sentimentos que são partes da vida moral. Nesse sentido, não haveria a necessidade de apontar a falsidade da tese determinista como faz o pessimista. É prestando a atenção a essa gama de atitudes que podemos recuperar os fatos como o conhecemos, não temos de ir além dos fatos. Dado que o otimista negligencia essas atitudes, o pessimista está correto quando aponta que há uma lacuna na explicação do otimista. Essa lacuna pode ser preenchida, mas em troca o pessimista deveria renunciar à sua metafísica (STRAWSON, 2008, p. 22-24).

O que há em comum entre o otimista e o pessimista é que ambos procuram, de modos diferentes, super-intelectualizar os fatos. Dentro da gama de atitudes e sentimentos humanos,

²⁰ Há também as atitudes autorreativas que dizem respeito às exigências que nós fazemos a nós mesmos, e aqui inclui-se fenômenos tais como sentir-se constrangido ou obrigado, remorso, culpa ou arrependimento (STRAWSON, 2008, p. 16).

vistos acima, há espaço para modificação, críticas e justificação. Porém, as questões de justificação são internas à estrutura geral das atitudes. Ambos, o otimista e o pessimista, se mostram incapazes de aceitar isso. O estilo do otimista de super-intelectualizar os fatos consiste na busca de uma fundamentação para suas práticas sociais com base nas suas consequências, e perde de vista as atitudes humanas das quais essas práticas são a expressão. Já o pessimista não perde de vista essas atitudes, porém é incapaz de aceitar o fato de que são essas atitudes que preenchem a lacuna na explicação do otimista. O pessimista pensa que a lacuna só pode ser preenchida se a tese determinista for falsa (STRAWSON, 2008, p. 25).

Não está errado o otimista enfatizar a eficácia das práticas sociais, ao regular o nosso comportamento do modo desejável. O que é errado é esquecer que essas práticas são expressões de nossas atitudes morais, e não meramente artifícios que empregamos de modo calculado para fins regulativos. Ao modificar a posição do otimista simultaneamente corrige-se sua posição sem recorrer à obscura metafísica do libertista (STRAWSON, 2008, p. 27).

Eu penso que o problema da vontade livre poderia ser beneficiado a partir de um aperfeiçoamento de uma posição compatibilista. Eu acredito que não haveria necessidade, como disse Strawson, de super-intelectualizar os fatos.

2 RESPONSABILIDADE MORAL E POSSIBILIDADES ALTERNATIVAS

Neste capítulo tratarei das posições compatibilistas e incompatibilistas sobre o determinismo e a liberdade. Primeiro apresentarei a posição de Frankfurt sobre a possibilidade do agente ser moralmente responsável apesar de não poder agir diferentemente. Depois, diferentes versões da teoria átimo de liberdade, proposta por filósofos que não ficaram convencidos pelo argumento de Frankfurt. Em seguida, será analisada a rejeição da abordagem átimo de liberdade por Fischer. Depois, uma réplica apresentada por filósofos conhecida como a defesa do dilema. E, por fim, a resposta de Fischer a essa nova tentativa de preservar a possibilidade do agente poder agir diferentemente para poder ser moralmente responsável pelo seu comportamento.

2.1 A refutação de Frankfurt do princípio das possibilidades alternativas

Existem vários desafios céticos à ideia intuitiva de que nós teríamos vontade livre no sentido que envolveria possibilidades alternativas. Um dos argumentos para a incompatibilidade da vontade livre e determinismo é conhecido como o “argumento da consequência”. Segundo esse argumento, se determinismo causal é verdadeiro, então nossas ações são consequências das leis da natureza e eventos no passado remoto. Mas não está em nós o que se passou antes de nós nascermos; e nem está em nós o que são as leis da natureza. Portanto, as consequências dessas coisas (incluindo nossas próprias ações) não estão em nós (INWAGEN, 1983, p. 15).

Esse argumento emprega ingredientes do senso comum, como a fixidez do passado e das leis da natureza, para mostrar que nós não teríamos o tipo de controle que abrangeria cursos alternativos de ações. Um dos mais importantes desafios a essa ideia é oferecido pela doutrina do determinismo causal. Determinismo causal é “a tese que todo evento é deterministicamente causado por algum evento no passado; assim, toda a escolha e todo o comportamento é o resultado de uma cadeia causal” (FISCHER, 2006, p. 183). Nossas ações, escolhas e deliberações seriam eventos mecânicos naturais, como o são os movimentos das águas e das pedras.

Assim, supondo que o determinismo causal seja verdadeiro e que eu faça uma escolha C no (tempo) t_2 :

Segue a partir da suposição do determinismo causal que o estado do mundo no t_1 juntamente com as leis da natureza implica que eu faço a escolha C no t_2 . Assim, em todos os mundos possíveis com as mesmas leis do mundo real

(...) eu faço a escolha *C* no t_2 . É evidente que se determinismo causal é verdadeiro, eu não posso fazer qualquer outra escolha a não ser a que eu faço na realidade (FISCHER, 2006, p. 183).

Nem todos têm aceitado que esse argumento é sólido. Vários filósofos têm argumentado que nós poderíamos ter possibilidades alternativas mesmo em um mundo causalmente determinístico²¹. Porém, outros estão inclinados a aceitar o argumento que estabelece que se determinismo causal é verdadeiro, então nós não teríamos o tipo de controle que envolveria possibilidades alternativas. Esses filósofos acreditam que responsabilidade moral exigiria o poder para escolher ou agir diferentemente, além da necessidade da falsidade da tese determinista.

O argumento da consequência é defendido por incompatibilistas que não foram convencidos pelo argumento de Frankfurt contra o que ele denomina, em “Alternate Possibilities and Moral Responsibility” (1969), de princípio das possibilidades alternativas (PPA). Esse princípio pode ser formulado como se segue: “uma pessoa é moralmente responsável pelo o que fez apenas se ela poderia ter feito diferentemente” (FRANKFURT, 1998, p. 1). Seu significado é assunto de controvérsia, no entanto, escreveu Frankfurt, “praticamente ninguém parece inclinado a negar ou a questionar que o princípio das possibilidades alternativas (construído em alguma forma ou outra) é verdadeiro” (FRANKFURT, 1998, p. 1). Alguns filósofos têm o caracterizado como “uma verdade *a priori*”.

Teorias podem fazer uso de exemplos de situações que envolvem coerção para anular o PPA. Elas podem afirmar que uma pessoa que tem sido coagida a fazer alguma coisa não poderia ter feito de outro modo; que ela não agiu livremente, e portanto, não poderia ser moralmente responsável pelo o que fez. Porém, para Frankfurt, a ausência de responsabilidade moral não estaria conectada com o fato de que a pessoa não poderia ter agido diferentemente. Ele argumenta que poderia existir alguma circunstância que não deixa alternativa para a pessoa, porém essa circunstância, na realidade, não desempenharia nenhum papel ou não conduziria ela a fazer o que ela faz. Frankfurt apresenta uma série de experimentos mentais, nos quais as situações são alteradas até que o agente não possua possibilidades alternativas, mas ainda assim age livremente, por sua própria vontade e poderia ser moralmente responsável pela ação exercida.

²¹ Ver Horgan (1979, 1985), Lewis (1981) e Bok (1998).

Exemplifico o ponto citando o personagem Jones, realizando a mesma ação, em quatro diferentes situações (FRANKFURT, 1998, p. 3-6). Vamos supor que alguém ameace Jones com certas penalidades caso ele não praticasse a ação exigida. Por exemplo, o oficial nazista entra na sua casa e ameaça prendê-lo e torturar a sua família caso ele não entregue o paradeiro do judeu, escondido no porão da sua casa. Jones poderia ser considerado responsável pela ação exercida nessas circunstâncias? Isso dependeria, segundo Frankfurt, do que realmente o levou a agir, se foi a sua decisão original ou a ameaça do oficial nazista. Imagine que Jones₁ seja um homem não razoável, que não se importa com as consequências das suas ações, pratica ação por sua própria vontade, faz o que já havia decidido, e uma vez que ele tenha tomado uma decisão nunca volta atrás, custe o que custar. Ele agiu sem levar em consideração a ameaça. Sendo assim, esse não seria um caso propriamente dito de coerção, visto que a ameaça feita não teria surtido efeito na tomada de decisão de Jones₁. Esse exemplo não representaria uma situação na qual o agente estaria sem alternativas para sua ação, pois é como se a ameaça não tivesse sido feita. Se ele tivesse decidido não revelar o paradeiro do judeu, a ameaça sofrida não o teria forçado a fazê-lo. Nesse caso, esse exemplo não seria um contraexemplo ao PPA.

Outra hipótese seria pensar a ação de Jones₂ sendo resultado da coerção. Aterrorizado pela ameaça do oficial nazista, e em virtude dela, ele fez o que lhe foi exigido e entregou o paradeiro do judeu, independentemente da sua decisão tomada anteriormente à ameaça. O fato de que no tempo anterior Jones₂ tenha decidido, por razões próprias, agir naquele modo poderia ser relevante para a avaliação do seu caráter, e “ele pode ser completamente responsável moralmente por ter feito *essa* decisão” (FRANKFURT, 1998, p. 3). Porém, nesse contexto dificilmente poderia ser dito que ele é moralmente responsável pela sua ação, considerando que ele a teria realizado simplesmente como resultado da coerção.

A terceira possibilidade seria aquela em que Jones₃ não agiria motivado pela ameaça (Jones₂) nem seria indiferente à ela (Jones₁). Assim como as suas contrapartes hipotéticas, Jones₃ realiza a mesma ação e entrega o paradeiro do judeu, porém Jones₃ agiu com base na sua decisão tomada antes da ameaça. Ele não estaria agindo motivado pelo medo da penalidade. A ameaça de fato não teria influenciado a realização da sua ação e o forçado a entregar o paradeiro do judeu, mas ela teria feito isso se ele já não tivesse motivo suficiente para realizar a ação em questão.

Esse caso mostra que o fato de que uma ameaça irresistível é feita não necessariamente implicaria que a pessoa que a recebe é coagida a fazer o que ela faz. A irresistibilidade da ameaça à qual Jones₃ está sujeito pode significar que não havia alternativa para ele a não ser exercer a ação que ele fez. Quer dizer, apesar da ameaça, isto é, embora ele não tivesse

alternativas disponíveis, parece natural afirmar que ele é moralmente responsável pelo que fez, na medida em que ele teria agido a partir da sua própria vontade e não por causa da ameaça (FRANKFURT, 1998, p. 4).

Apesar do caso do Jones₃ apresentar uma conclusão intuitiva, esse exemplo não convenceu a todos. A seguinte objeção pode ser feita: mesmo Jones₃ estando suscetível a ameaças, lhe restaria ainda uma possibilidade alternativa, pois ele poderia resistir à ameaça e aceitar a penalidade. Portanto, esse é um caso que seria difícil afirmar o que teria motivado a ação de Jones₃: se a sua decisão anterior à ameaça ou a própria ameaça e o medo da penalidade.

Frankfurt acredita que essa não é uma boa objeção à sua defesa da falsidade do PPA. Para mostrar por que, ele altera a estrutura do exemplo:

Suponha que alguém, Black, quer que Jones₄ realize uma determinada ação. Black está preparado para intervir, mas ele prefere evitar agir desnecessariamente. Então, ele espera até que Jones pense o que fazer, e ele não faz nada a não ser que seja claro para ele (Black é um excelente juiz de tais coisas) que Jones₄ está decidindo fazer alguma coisa *diferente* do que ele quer que ele faça. Se se torna claro que Jones₄ está decidindo fazer alguma coisa diferente, Black tomaria as medidas eficazes para assegurar que Jones₄ decidiria fazer e que ele faria o que ele (Black) quer que Jones faça (FRANKFURT, 1998, p. 6).

Para ilustrar esse caso, imaginemos que Black, trabalhando no partido dos democratas, quer que Jones₄ escolha e vote em Hillary Clinton, nas eleições de novembro de 2016 nos EUA, e não em Donald Trump, candidato do partido republicano. Black fará o que for necessário para garantir que Jones₄ decida votar e vote em Clinton. Vamos imaginar que Black manipule o processo cerebral de Jones₄ (por meio de algum dispositivo instalado em seu cérebro) de modo a determinar a sua escolha e o seu voto. Dado essa condição, Jones₄ não pode fazer de outro modo. Agora suponha que Black não tenha precisado intervir porque Jones₄ decidiu e realizou, por razões próprias, a ação que Black desejava. Sendo assim, parece claro que Jones₄ manterá a mesma responsabilidade pela sua ação que teria se tivesse feito o mesmo sem a presença do dispositivo de Black. Não seria razoável desculpar Jones₄ pela sua ação alegando que ele não poderia ter feito de outro modo, pois esse fato não influenciou a sua ação. Ele teria agido do mesmo modo sem a presença do dispositivo instalado em seu cérebro por Black (FRANKFURT, 1998, p. 7).

A ideia central, segundo Frankfurt, é observar a existência de condições suficientes para a realização da ação de Jones₄. Significa que Jones₄ não teria possibilidades alternativas porque ele exerce a ação desejada por Black. A atribuição de responsabilidade dependeria de qual teria

sido a motivação da ação (se se trata de sua própria predisposição ou de uma intervenção de Black). Agindo por razões próprias ou agindo por influência de Black, a ação seria a mesma. Se ele age por razões próprias “sua responsabilidade moral por essa ação não é afetada pelo fato de Black estar à espera com planos sinistros, visto que essa intenção não teve nenhum papel na realização da ação” (FRANKFURT, 1998, p. 8).

Os exemplos discutidos por Frankfurt não são exatamente morais, então podemos rerepresentá-los de uma maneira a evidenciar a responsabilidade especificamente moral dos agentes envolvidos. Podemos imaginar um cenário envolvendo um assassino. Reusando os personagens Jones e Black, digamos que Black quer que Jones mate o prefeito da cidade. Black espera Jones tomar a sua decisão. Se Black percebe que Jones está decidindo não matar o prefeito, ele intervirá para assegurar que Jones decida e execute a ação pretendida. Porém, Jones decidiu, por razões próprias, matar o prefeito, e Black não precisou intervir.

O fato de que uma pessoa não poderia ter evitado fazer uma determinada ação é uma condição suficiente para ela ter feito. Porém, como mostra esse último exemplo, esse fato (ela não poderia ter feito de outro modo) pode não desempenhar qualquer papel na explicação do por que ela agiu. Quer dizer, “embora a pessoa fosse incapaz de fazer de outro modo, pode não ser o caso de que ela agiu *porque* não poderia ter feito de outro modo” (FRANKFURT, 1998, p. 8). Agora se uma pessoa não tinha alternativa para realizar uma certa ação, mas não a realizou porque era incapaz de fazer de outro modo, então ela teria realizado a mesma ação mesmo se ela poderia ter feito de outro modo.

O equívoco do princípio das possibilidades alternativas estaria na pressuposição de que uma pessoa não é moralmente responsável por ter realizado uma ação se havia circunstâncias que tornaram impossível para ela evitar de realizá-la. Desse modo, o princípio das possibilidades alternativas poderia ser revisado de modo a declarar “que uma pessoa não é moralmente responsável pelo o que ela fez se ela o fez porque ela não poderia ter feito de outro modo” (FRANKFURT, 1998, p. 9).

Pode ser observado que essa revisão do princípio não necessariamente afeta os argumentos daqueles que contam com o princípio original nos seus esforços para defender o incompatibilismo da responsabilidade moral com o determinismo. Se foi causalmente determinado que uma pessoa realiza uma certa ação, então será verdadeiro que ela a realizou por causa desses determinantes causais. E se o fato de que foi causalmente determinado que ela realizou uma determinada ação significa que ela não poderia ter feito de outro modo (como supõem os argumentos incompatibilistas), então o fato de que foi causalmente determinado que uma pessoa realizou uma certa ação significará que a pessoa agiu porque não poderia ter feito

de outro modo. O princípio das possibilidades alternativas “revisado” implicaria, nesse caso, que uma pessoa não é moralmente responsável pelo o que ela fez se o que ela fez foi causalmente determinado. Frankfurt não acredita que essa revisão do princípio seja aceitável:

Imaginemos que uma pessoa nos diga que ela fez o que ela fez porque ela era incapaz de fazer de outro modo; ou que essa pessoa faça uma afirmação similar, a de que fez o que fez porque ela tinha de fazer o que ela fez. Nós frequentemente aceitamos essas justificativas como desculpas válidas (se nós acreditamos nelas), e elas podem bem parecer, à primeira vista, invocar o princípio revisado das possibilidades alternativas. Porém, eu penso que quando nós aceitamos tais afirmações como desculpas válidas é porque nós assumimos que o que está sendo dito é mais do que o que as afirmações, literalmente, dizem (FRANKFURT, 1998, p. 9-10).

Quer dizer, quando a pessoa oferece tais justificativas, nós entendemos que ela quer significar que ela fez o que ela fez *apenas porque* ela era incapaz para fazer de outro modo, ou *apenas porque* ela tinha de fazer o que ela fez. Além disso, entendemos, mais particularmente, que ela quer significar que ela fez o que ela fez não porque isso era o que ela, na realidade, queria fazer. Assim, o princípio deveria ser substituído pelo seguinte princípio “uma pessoa não é moralmente responsável pelo o que ela fez se ela o fez unicamente porque ela não poderia ter feito diferentemente” (FRANKFURT, 1998, p. 9-10). Esse novo princípio não pareceria entrar em conflito com a visão de que responsabilidade moral é compatível com o determinismo.

A questão que imediatamente surge é se Frankfurt refutou, de modo satisfatório, o PPA. Enquanto alguns filósofos foram persuadidos de que nós deveríamos rejeitar o PPA, muitos não foram convencidos, seja pelos argumentos de Frankfurt, seja pelos defensores do compatibilismo de tipo Frankfurt (FISCHER, 2010a, p. 317).

2.2 A estratégia átimo de liberdade [*flicker of freedom*]

Alguns céticos rejeitam a ideia de que Frankfurt tenha fornecido casos em que o agente seria moralmente responsável apesar de não poder agir diferentemente. Eles insistem que nós não somos forçados pelos casos de tipo Frankfurt a desistir do PPA. Outros céticos estão dispostos a abandonar o PPA, mas argumentam que o determinismo causal eliminaria a responsabilidade moral por alguma outra razão. O primeiro grupo de céticos, aqueles que insistem que nos exemplos de tipo Frankfurt há possibilidades alternativas, defendem uma estratégia conhecida como átimo de liberdade. Embora existam várias versões distintas dessa estratégia, elas defendem o mesmo argumento, a saber, de que apesar da impressão inicial de que nos casos de tipo Frankfurt não haveria possibilidades alternativas, de fato alguém poderia

perceber que existem nesses casos tais possibilidades alternativas, ainda que mínimas. Embora o interventor contrafactual, Black, eliminaria a maioria dos cursos alternativos de ações, ele não excluiria todos. A crítica que estaria por trás dessa estratégia seria a de que Frankfurt não teria justificado adequadamente a compatibilidade da responsabilidade com o determinismo. Subjacente a essa crítica estaria a visão tradicional de que se o determinismo ameaçaria a responsabilidade seria em virtude de enfraquecer o tipo de controle correspondente às possibilidades alternativas. Irei apresentar três versões dessa estratégia, e depois uma réplica que procura identificar um problema que seria comum a todas elas.

(a) Caso A

Um dos pontos de interesse no debate sobre os casos de Frankfurt foi a investigação de como Black saberia como Jones irá agir. A literatura posterior à publicação do artigo de Frankfurt introduziu a possibilidade da existência de um “sinal anterior” que poderia ser lido pelo interventor contrafactual²², o qual o auxiliaria na sua atividade (FISCHER, 2002, p. 282). Com base nos exemplos de tipo Frankfurt, nós podemos imaginar um caso em que Jones teria uma inclinação para mostrar algum *signal* que indicaria em quem ele irá votar, antes mesmo de começar a fazer uma escolha ou antes de formar uma intenção. Suponha que Jones levantasse a sobancelha antes de começar qualquer processo de tomada de decisão se e somente se ele estivesse prestes a escolher votar em Trump. Supondo que isso seja assim, então Black poderia, ao ler o sinal, impedir Jones de *começar* a fazer tal escolha ou decisão. Nesse caso, um átimo de liberdade emerge, pois Jones teria o poder para mostrar o sinal antes de começar a escolher, como o franzir da sobancelha²³. Sem um átimo de liberdade, Black não poderia prever o que Jones faria.

(b) Caso B

Para Inwagen, mesmo que o argumento de Frankfurt contra o PPA seja sólido, existem ainda três princípios que se não são versões desse princípio, são muito similares a ele. O ponto básico seria observar que os defensores dos exemplos de tipo Frankfurt não são precisos em especificar *pelo o quê* o agente é moralmente responsável. O agente seria moralmente responsável, segundo a posição do autor, não apenas por atos realizados, mas também por atos não realizados (omissões) e pelas consequências do que ele fez ou deixou de fazer.

²² O termo “interventor contrafactual” foi introduzido por Fischer (1982).

²³ David Blumenfeld desenvolveu este exemplo de tipo Frankfurt (1971, p. 339-345).

O argumento pretende mostrar que os seguintes três princípios são imunes aos contraexemplos de tipo Frankfurt ao PPA. São esses:

Princípio da Ação Possível (PAP): Uma pessoa seria moralmente responsável por deixar de realizar um dado ato apenas se ela poderia ter realizado esse ato²⁴.

Princípio da Prevenção Possível (PPP1): Uma pessoa seria moralmente responsável por um certo evento (particular) apenas se ela poderia tê-lo evitado (INWAGEN, 1978, p. 206).

Princípio da Prevenção Possível (PPP2): Uma pessoa seria moralmente responsável por um certo estado de coisas apenas se (esse estado de coisas ocorre e) ela poderia ter o impedido de ocorrer (INWAGEN, 1978, p. 210, 220).

PAP diz respeito aos atos não realizados, PPP1 às consequências do que o agente fez e PPP2 às consequências do que o agente deixou de fazer. A diferença entre PPP1 e PPP2 está no estatuto ontológico: o primeiro é sobre eventos e o segundo é sobre estado de coisas. PPP1 se aplicaria às consequências consideradas como eventos particulares e PPP2 às consequências consideradas como eventos universais²⁵.

Sobre o PPP2, Inwagen desenvolve um caso em que o agente não poderia ter impedido um estado de coisas de ocorrer, o que implicaria que ele não poderia ser moralmente responsável. Por exemplo, suponha que o cavalo de Ryder, Dobbin, saiu em disparada. Dobbin responde ao freio, e em cada encruzilhada Ryder pode guiar Dobbin ou para a direita ou para a esquerda. Sem que Ryder saiba, ambas as ruas em uma encruzilhada levam para Roma. Ryder poderia, nesse caso, ser responsável por guiar Dobbin para a esquerda ou para a direita, mas

²⁴ Sobre PAP, Inwagen nos pede para imaginar uma situação em que ele está na janela da sua casa quando testemunha a cena de um crime. Diante dessa situação, talvez seja melhor chamar a polícia. Então, ele pega o telefone, mas coloca de volta no gancho. Ele pensa por um momento e percebe que se ligar para a polícia, o criminoso talvez o ouça e queira se vingar. De qualquer forma, provavelmente o policial pedirá para ele se deslocar até a delegacia para fazer uma ocorrência, e como já é quase meia noite e no outro dia ele precisa levantar cedo, ele decide “não se envolver”. Agora, suponha que, sem o seu conhecimento, houve algum tipo de interrupção nos serviços da central telefônica e nenhum telefone da cidade está funcionando e permanecerá assim por várias horas. O ponto é que ele não seria responsável por não ter comunicado a polícia sobre o ocorrido, pois ele *não poderia* ter feito. Ele poderia ser responsável por *não ter tentado* chamar a polícia, pois isso ele poderia ter feito (INWAGEN, 1978, p. 204-205).

²⁵ Entre os filósofos que acreditam que nós responsabilizamos os agentes por eventos de certos tipos, existe um desacordo em relação ao status ontológico desses eventos. Eles devem ser construídos como universais ou como particulares? (FISHER, 2006, p. 28) Para evitar tomar partido nesse debate, Inwagen usa a seguinte estratégia: indica um certo princípio para eventos ou estado de coisas pelo qual nós responsabilizamos as pessoas como particular e outro princípio para eventos ou estado de coisas como universal (INWAGEN, 1978, p. 206).

não seria razoável dizer que ele seria responsável pelo fato de que ele passa por Roma (INWAGEN, 1983, p. 177; 1978, p. 216-217). Na intuição de Inwagen, “parece óbvio que ele não é responsável por esse estado de coisas *apenas porque* ele não poderia ter impedido” (INWAGEN, 1978, p. 217). Dobbin e Ryder teriam passado por Roma, independentemente de ir pela esquerda ou pela direita.

Agora sobre o princípio das possibilidades alternativas aplicado às consequências-particulares:

PPP1: Uma pessoa é moralmente responsável por um certo evento (particular) apenas se ela poderia tê-lo evitado.

Inwagen questiona *se* o seguinte contraexemplo de tipo Frankfurt poderia ser construído contra esse princípio:

Black atira e mata Jones (intencionalmente), portanto, produzindo sobre a morte de Jones um certo evento. Mas existe algum fator F que (i) desempenhou nenhum papel causal na morte de Jones, e que (ii) teria causado a morte de Jones *se* Black não tivesse atirado nele (visto que o fator F pode ter causado a morte de Jones ao causar Black a atirar nele, talvez nós deveríamos dizer “se Black não tivesse decidido atirar nele”), e (iii) é tal que Black não poderia ter impedido de causar a morte de Jones exceto por matar (ou decidir matar) ele mesmo a Jones. Pareceria que Black é responsável pela morte de Jones, embora ele não poderia ter impedido a morte de Jones (INWAGEN, 1978, p. 210)²⁶.

É facilmente perceptível, na perspectiva do autor, que essa história é inconsistente (INWAGEN, 1978, p. 210). Podemos adaptar esse contraexemplo de tipo Frankfurt para ver com mais clareza o que estaria por trás dessa versão da estratégia apresentada por Inwagen. Assim, teríamos que, na sequência real, Jones escolhe votar e vota em Clinton. Agora, imagine que Jones comece a mostrar uma inclinação (envolvendo um sinal anterior diferente) de escolher votar em Trump e isso provocaria a intervenção de Black, e a subsequente escolha para votar (e a concretização dessa escolha) em Clinton. Sob tais hipóteses, Jones teria votado em Clinton.

Partindo de uma distinção entre sequência real e sequência alternativa, se Jones não tivesse votado em Clinton, e como resultado o fator F tivesse causado Jones a votar em Clinton, então teria ocorrido um evento denotado “votar em Clinton” que teria o fator F como uma de suas causas. Esse evento teria sido diferente do evento “votar em Clinton” na sequência real,

²⁶ Eu adaptei os personagens de Inwagen ao caso original de Frankfurt.

resultante das próprias deliberações de Jones²⁷. No final das contas, o evento na sequência real não teria ocorrido (INWAGEN, 1978, p. 210).

Mais especificamente, essa abordagem adota o seguinte princípio para identificar e individualizar eventos particulares: “ x é o mesmo evento y se e apenas se x e y têm as mesmas causas”²⁸. Usando esse critério, Jones teria causado um evento particular *diferente* quando ele vota em Clinton, na medida em que a causa seria a intervenção de Black. Supondo que Black não desempenhou qualquer papel no evento da sequência real, mas exerceu um tal papel no cenário alternativo, dado o princípio da individualização, o evento real particular não poderia ser idêntico ao evento hipotético da sequência alternativa. Portanto, Jones possuiria o poder para produzir um evento particular diferente, ele poderia *evitar* que o evento particular, na sequência real, ocorresse. Aí estaria o átimo de liberdade (INWAGEN, 1978, p. 206-208). Se responsabilidade pela ação, de fato, envolve responsabilidade por provocar um evento particular concreto, então responsabilidade pela ação envolveria possibilidades alternativas.

(c) Caso C

A terceira versão da estratégia do átimo de liberdade está associada à posição libertista²⁹. O que distinguiria uma ação de um mero evento seria que a ação é precedida por uma volição³⁰. Partindo da intuição básica dessa explicação libertista, a volição deveria ser causada pelo agente³¹. O ponto é que quando um agente causa uma volição por esse tipo de relação causal, isso significaria que nenhum fator externo faz com que o agente cause a volição (FISCHER, 2006, p. 43-44).

²⁷ Aqui pode ser adotado a seguinte suposição: quando o agente realiza uma ação, ele causa a ocorrência de algum evento.

²⁸ O ponto é encontrar um modo para saber, de um dado evento, se esse mesmo evento teria ocorrido se as coisas tivessem sido diferentes (INWAGEN, 1978, p. 207). O princípio para individualizar eventos deve fornecer um modo de identificar eventos por meio de mundos possíveis, isto é, individualizar que um evento A no *mundo alfa* é o mesmo evento particular B no *mundo beta* apenas se A e B tem os mesmos antecedentes causais (FISCHER, 1982, p. 29). Inwagen parte de critério de Donald Davidson, com alguma modificação, para individualizar eventos. Para Davidson, x e y são os mesmos eventos se e apenas se x e y tem as mesmas causas e *efeitos* (INWAGEN, 1978, p. 207).

²⁹ Existem diferentes versões da posição libertista que procuram dar explicações de como preencher o espaço aberto pelo indeterminismo. Muitos libertistas modernos, por exemplo, argumentam que nós alcançamos autodeterminação, em um mundo indeterminístico, ao sermos nós mesmos os determinantes de nossas ações (CHISHOLM, 2013, p. 26-37). Outros, como David Wiggins e Robert Kane, defendem uma posição libertista sem apelar para a teoria de que a causa da ação é o agente. Ver, Wiggins (2013, p. 94-121) e Kane (2013, p. 299-321).

³⁰ Esta afirmação não necessita conduzir à visão libertista (FISCHER, 2006, p. 43).

³¹ A ideia básica que está por trás da posição que defende o agente como sendo a causa da sua ação seria de que ele poderia iniciar novas cadeias causais que não são pré-determinadas por eventos anteriores do passado ou pelas leis da natureza.

Adotando essa explicação no caso de tipo Frankfurt, teríamos que Jones delibera, escolhe e age no modo normal. Vamos supor que Jones cause a sua volição para votar em Clinton. Agora imagine a sequência alternativa. Suponha que tão logo Jones começasse a mostrar uma inclinação para escolher votar em Trump, Black interferiria e asseguraria que ele escolhesse votar em Clinton. Sob tais circunstâncias, Jones teria algum tipo de estado mental consistindo em uma escolha para votar em Clinton, mas ele *não* teria causado uma volição para votar nesse candidato, na medida em que sua volição foi causada por um fator externo. Assim, de acordo com a análise libertista, embora Jones não tivesse o poder para causar uma volição para votar em Trump, ele teria um átimo de liberdade, pois ele possuiria o poder de abster-se de causar uma volição para votar em Clinton. Nessa abordagem, não é pressuposto que Jones deveria ter o poder para agir de outro modo ou mesmo formar um tipo diferente de volição para ser considerado moralmente responsável pelo o que faz, mas Jones deveria ter o poder de *não* formar o tipo de volição que ele, na realidade, forma³².

2.3 A rejeição da abordagem átimo de liberdade

O argumento que será apresentado a seguir contra a estratégia átimo de liberdade pode ser aplicado a todas as versões apresentadas anteriormente. É mais conveniente começar com a posição de Inwagen, e então estender a crítica às outras visões. Porém, antes de apontar o suposto equívoco, eu gostaria de esclarecer sobre o tipo de controle que tem sido tradicionalmente adotado como condição necessária para atribuições de responsabilidade moral.

O problema com a estratégia átimo de liberdade, em todas as suas variantes, seria que ela baseia-se em uma incompreensão do PPA. Observe que na visão clássica, liberdade implicaria que o agente teria uma escolha entre dois ou mais cenários. Em um destes cenários, o agente deliberaria e formaria uma intenção para realizar um ato de um certo tipo e, então, realizaria essa intenção de um modo apropriado. No outro cenário possível, ele deliberaria e formaria uma intenção para realizar um diferente tipo de ato e, então, efetuaria essa intenção apropriadamente. Isso é o que estaria envolvido em ter possibilidades alternativas robustas, e esse seria certamente o modo natural de se pensar sobre o tipo de alternativas no qual a

³² Para Fischer não há necessidade de postular volições ou teorias baseadas no agente como sendo causa da ação para obter o tipo de resultado já esboçado. Assim, alguém poderia dizer que, embora Jones não tivesse o poder para escolher de outro modo, ele teria de fato o poder para abster-se de escolher votar em Clinton (FISCHER, 2006, p. 44).

responsabilidade moral estaria baseada. É evidente que esse tipo de liberdade não está presente nos contraexemplos de tipo Frankfurt (FISCHER, 2006, p. 47).

Agora sobre os casos apresentados anteriormente. Embora possa existir uma possibilidade alternativa, *ela não seria* suficientemente robusta para basear atribuições de responsabilidade moral. Desse modo, é implausível supor que seria em virtude da existência de uma tal possibilidade alternativa que um agente poderia ser moralmente responsável pelo o que faz. Seria preciso que a estratégia átimo de liberdade demonstrasse que essa possibilidade alternativa desempenha um certo papel no entendimento apropriado dos casos.

Retomando a ideia presente no princípio da prevenção possível, defendida por Inwagen, um agente seria moralmente responsável por um certo evento (particular) apenas se ele poderia tê-lo evitado. Mais especificamente, responsabilidade moral estaria conectada à ideia da possibilidade da ocorrência de um diferente evento. Partindo da distinção entre sequência real e sequência alternativa, dois eventos distintos resultariam, um tendo como causa um fator externo e outro tendo como causa as próprias deliberações do agente.

Mesmo que seja admitida a ocorrência, na sequência alternativa, de um evento diferente daquele que teria ocorrido na sequência real, é evidente que o agente, no cenário alternativo, *não estaria agindo livremente*. Parece estranho que o acréscimo de um cenário alternativo, em que a causa da ação é exterior e não resultante da escolha do agente, seria suficiente para tornar verdadeiro que ele possui o tipo de controle relevante necessário para responsabilizá-lo moralmente pelo dado evento (FISCHER, 2006, p. 45-46).

Suponha que aceitemos o seguinte princípio de Inwagen: “ x é o mesmo evento y se e apenas se x e y têm as mesmas causas”. Então, poderia ser estabelecido que um diferente evento teria ocorrido na sequência alternativa, visto que esse evento possui uma causa diferente. Mas o problema é que Inwagen não teria justificado adequadamente que o agente teria controle deliberativo. Ele estaria confundindo a “capacidade” de agir diferentemente com a “possibilidade” de alguma coisa diferente ocorrer³³. O equívoco cometido seria pensar que a única forma que o contraexemplo de Frankfurt poderia ameaçar o PPA seria apresentar um cenário alternativo em que o evento particular ocorrido é o mesmo da sequência real. Assim, “os casos de tipo Frankfurt não seriam melhor construídos como casos de acesso a um diferente

³³ Inwagen insiste que Frankfurt não teria dissociado responsabilidade de controle e, sendo assim, não teria demonstrado que a responsabilidade é compatível com determinismo. O que estaria por trás dessa crítica é a visão tradicional de que, se determinismo ameaçaria a responsabilidade, seria em virtude de enfraquecer a liberdade enquanto capacidade para agir diferentemente (FISCHER, 1982, p. 32-33).

mundo com o mesmo evento-particular, mas da *falta de acesso* ao tipo *relevante* de mundo alternativo” (FISCHER, 2002, p. 288).

A mesma objeção poderia ser feita à versão libertista (C). Nesta perspectiva, para que o agente possa ser moralmente responsável ele deveria ter o poder *de não causar* a volição que ele, na realidade, causa. Suponha que Jones tenha o poder de não causar a volição para votar em Clinton, como visto anteriormente. Observe que Jones, mesmo nesse caso, não formaria na sequência alternativa uma intenção para abster-se de causar uma volição para votar em Clinton e, então, procederia para realizar essa intenção. Embora possa ser admitido que Jones tenha o poder de não causar a volição para votar nesse candidato, o ponto é que ao não causar uma tal volição, *ele não estaria agindo livremente*. Por causa da intervenção de Black, Jones não teria se absterido de causar essa volição livremente. Mesmo se existisse alguma possibilidade alternativa aqui, de novo, ela não pareceria ser suficientemente robusta para basear atribuições de responsabilidade moral (FISCHER, 2006, p. 47).

E, por último, a primeira estratégia poderia insistir que apesar do agente, na sequência alternativa, não poder escolher ou agir livremente, existe alguma coisa, nesse cenário, que seria feita livremente, isto é, Jones poderia *começar a iniciar* um processo para fazer uma escolha. Ele poderia, por exemplo, começar a escolher votar em Trump, ou poderia ter o poder de exibir algum sinal ou padrão neurológico diferente. Nesse sentido, a possibilidade de começar a escolher, a qual poderia ser considerada como sendo feita livremente, seria o que basearia as atribuições de responsabilidade moral.

Vejamos essa versão da estratégia para que possamos ver qual seria o problema que ela também apresentaria. Suponha que Jones esteja na cabine de votação deliberando em qual dos partidos ele deveria votar, se nos republicanos ou nos democratas. Caso Jones esteja prestes a votar em Trump, um dispositivo instalado no seu cérebro seria ativado fazendo com ele escolhesse votar e votasse em Clinton (FISCHER, 1999, p. 109-110). A questão que imediatamente surge é esta: como o dispositivo informaria a Black em qual candidato Jones estaria prestes a escolher votar? Um papel crucial seria desempenhado por algum tipo de sinal involuntário que indicaria as escolhas futuras do agente. Por exemplo, um franzir a testa, que meramente ocorreria nele e que não estaria sobre seu controle voluntário. John Fischer, um dos defensores dos casos de tipo Frankfurt, faz a seguinte observação:

Pode ser objetado que apesar da aparência inicial Jones tem *alguma* possibilidade alternativa. Embora Jones não pode escolher ou votar diferentemente, ele pode exibir um diferente padrão neurológico no seu cérebro. Eu tenho chamado uma tal possibilidade alternativa um átimo de

liberdade (...). Mas eu afirmo que a mera exibição involuntária de algum sinal – tal como um padrão neurológico no cérebro, um corar, ou um franzir de testa, é insuficiente como uma base para a responsabilidade moral. O poder involuntário para exibir um diferente sinal parece para mim ser insuficientemente robusto para basear nossas atribuições de responsabilidade moral (FISCHER, 2002, p. 288-289).

Dito de outra forma, se em t_2 Jones estivesse prestes a escolher votar em Clinton em t_3 , então um sinal involuntário teria sido exibido (por exemplo, o franzir da sobrancelha esquerda, ou algum padrão neurológico no seu cérebro) em t_1 . Nesse caso, o dispositivo não seria ativado porque esse é o candidato no qual Black queria que Jones votasse. Mas se Jones tivesse mostrado outro sinal involuntário, então esse padrão neurológico (indicativo de que ele votaria em Trump) teria ativado o dispositivo de Black para intervir e fazer com que Jones escolha em t_2 votar em Clinton e com que ele vote em Clinton em t_3 (FISCHER, 1999, p. 110).

O ponto é que na sequência alternativa, na qual Jones mostraria um distinto padrão neurológico como um indicativo de uma escolha para votar em Trump, o sinal exibido seria completamente *involuntário* e a subsequente escolha e ação seriam produzidas eletronicamente. Sendo assim, não poderia ser dito de modo razoável que no cenário alternativo Jones escolheria e agiria livremente, pois tanto a sua escolha quanto a sua ação não seriam resultantes do seu próprio raciocínio prático. Além disso, ele não poderia ser moralmente responsável de maneira apropriada por esse comportamento (FISCHER, 1999, p. 110-111). Ademais, “o demonstrativo de um tal sinal não é nem mesmo uma ação e certamente não plausivelmente pensado ser robusto suficiente para basear atribuições de responsabilidade moral” (FISCHER, 2006, p. 48).

Os exemplos de tipo Frankfurt, embora possam parecer bizarros ou incomuns, pretendem ser construídos para a finalidade de apontar que a responsabilidade moral depende do que acontece na sequência real e não do que acontece na sequência alternativa (FISCHER, 2006, p. 58-59). Para que o argumento de Frankfurt contra o PPA seja refutado, seria necessário mostrar que o átimo de liberdade, presente nos exemplos vistos acima, seja um ato voluntário suficientemente robusto.

Obviamente que nem todos ficaram convencidos de que nós poderíamos abandonar a visão tradicional de liberdade enquanto capacidade para agir diferentemente. Uma outra objeção, conhecida como a defesa do dilema³⁴, pode ser vista como uma das ferramentas mais potentes dos céticos que negam que os casos de tipo Frankfurt tenham refutado o PPA de modo satisfatório.

³⁴ Esta objeção foi desenvolvida por Robert Kane (1985), David Widerker (1995), Carl Ginet (1996) e Keith Wyma (1997).

2.4 A defesa do dilema

Como vimos anteriormente, a estratégia átimo de liberdade é um argumento em defesa da presença de possibilidades alternativas nos casos de tipo Frankfurt. Portanto, de acordo com essa visão, Frankfurt não teria dissociado responsabilidade de liberdade. Entretanto, a característica distintiva para o argumento que será tratado agora, a defesa do dilema, remete à relação entre o sinal anterior, exibido em t_1 , pelo dispositivo neurológico, e a escolha ou ação, em t_2 . Essa relação seria problemática, pois ela seria ou determinística ou indeterminística.

Vamos supor que a relação entre o sinal anterior e a escolha é indeterminista e Black somente irá interceder se for necessário. Digamos que Black percebe o franzir da sobrancelha esquerda de Jones, em t_1 , e aguarda ele escolher, em t_2 , votar em Clinton por razões próprias. Poderia ser pensado que se houvesse um pequeno espaço de tempo entre o sinal e a escolha de Jones, Black poderia interferir para impedir a concretização da escolha. Porém, o problema dessa versão do dilema é que embora o sinal anterior exibido pelo dispositivo neurológico seja confiável, pressupondo o indeterminismo seria possível que Jones começasse a escolher votar em Trump mesmo que ele tenha exibido o sinal relevante (de que iria escolher votar em Clinton) (TIMPE, 2008, p. 34-35; FISCHER, 2010a, p. 319-320). Dado a hipótese da possibilidade do indeterminismo causal, parece que Black não poderia eliminar todas as possibilidades alternativas, pois Jones teria a possibilidade de “começar a fazer uma escolha”, o que seria mais robusta do que um mero átimo de liberdade (FISCHER, 2002, p. 291). Além disso, embora a possibilidade de começar a escolher diferentemente possa ser bloqueada antes de ser concluída, seria, no entanto, um episódio voluntário. Sendo assim, ainda não teríamos um contraexemplo ao PPA. Seria a existência da possibilidade de começar a escolher de outro modo que tornaria Jones moralmente responsável tanto pela sua escolha quanto pela sua ação.

Agora sobre a relação determinística entre o sinal prévio e a escolha ou ação. Nesse caso, parece que Jones não poderia escolher ou agir diferentemente. Digamos que Jones, em t_1 , levantasse a sobrancelha direita, como indicativo de que ele escolheria votar em Trump em t_2 , e de que ele votaria em Trump em t_3 . Black, ao tomar conhecimento desse padrão neurológico em t_1 poderia impedir que Jones escolhesse ou votasse em Trump. Partindo da pressuposição sobre a existência do determinismo, o cético insistirá que nós não poderíamos afirmar que Jones é moralmente responsável por sua escolha e ação (FISCHER, 2002, p. 290). A compatibilidade do determinismo com a responsabilidade não seria justificada de modo adequado. De novo, ainda não teríamos um caso em que poderíamos dizer que o agente é moralmente responsável apesar de não poder agir de outro modo.

Apresentarei, a seguir, as objeções feitas por Robert Kane, David Widerker e Stewart Goetz ao contraexemplo de Frankfurt ao PPA.

(a) Kane, Widerker e Goetz

Robert Kane desenvolveu a seguinte versão do dilema:

Imagine o que deve ser feito para exercer o controle de Frankfurt sobre a escolha de uma pessoa. O controlador, Black, planeja que Jones faça A. Mas se ele espera para ver se Jones está indo fazer A por razões próprias e unicamente intervém se Jones está prestes a fazer B...[a não ser que nós assumimos que a ação em questão é causalmente determinada], o controlador não pode saber qual delas vai ocorrer antes a não ser que ele predetermina uma delas a ocorrer. Ele pode, portanto, esperar até que ele descubra se o agente fará A ou B, mas então é tarde demais para controlar a escolha. Ou ele pode intervir no cérebro desligando a indeterminação ou seus efeitos antes que qualquer escolha ocorra, determinando, portanto, o resultado que ele deseja. No último caso, a escolha será determinada pelo controlador e o controlador, e não o agente, será responsável por isso (...) Por contraste, se o controlar não intervém para predeterminar o resultado, e a indeterminação permanece no lugar até que a escolha seja feita (...) então o agente, e não o controlador, é responsável por isso. Mas então é também o caso de que o agente poderia ter feito de outro modo (KANE, 1998, p. 142).

David Widerker fez uma objeção similar. Segundo Widerker, “o sucesso do caso de Frankfurt contra PPA depende da sua capacidade para nos convencer da plausibilidade” do seguinte princípio: podem existir circunstâncias que não deixam alternativa de ação para a pessoa, mas que apesar disso não a forçam a agir de uma determinada maneira (WIDERKER, 1995, p. 248).

Widerker nos pede para considerar o seguinte caso de tipo Frankfurt. Jones está deliberando se deveria ou não matar Smith em t_3 . Black estaria preparado para coagir Jones a matar Smith, caso ele decidisse, em t_2 , não o matar por razões próprias. Black deveria observar o sinal prévio, presente em t_1 , para saber se ele precisará interferir ou não. Para melhor compreender o cenário de Frankfurt, Widerker o descreve em termos mais concretos:

- (1) Se Jones está corando em t_1 , então desde que ninguém intervenha, Jones decidirá, em t_2 , matar Smith.
- (2) Se Jones não está corando em t_1 , então desde de que ninguém interfira, ele não decidirá, em t_2 , matar Smith.
- (3) Se Black percebe que Jones mostra o sinal de que não decidirá, em t_2 , matar Smith, isto é, se ele percebe que Jones não está corando em t_1 , então Black força Jones a decidir, em t_2 , matar Smith; mas se ele vê que ele está corando em t_1 , ele não faz nada.

Finalmente, suponha que Black não intervenha, porque:

(4) Jones está corando em t_1 e decide, em t_2 , matar Smith por razões próprias (WIDERKER, 1995, p. 249-250).

O problema estaria na premissa (1), segundo Widerker. Os defensores dos exemplos de tipo Frankfurt precisariam enfrentar o desafio do dilema. Por um lado, supondo a verdade da tese determinista, a relação entre o sinal neurológico prévio, em t_1 , e a decisão, em t_2 , seria causalmente suficiente. Então, o princípio acima falharia visto que a decisão de Jones seria determinada, portanto, ele não estaria, em t_2 , decidindo livremente. Por outro lado, pressupondo o indeterminismo, não existindo a relação de causalidade entre o sinal, em t_1 , e a decisão, em t_2 , então seria “difícil ver como a decisão de Jones é inevitável” (WIDERKER, 1995, p. 251). Sendo assim, seria preservado o poder de Jones de decidir não matar Smith. Para esse raciocínio, os contraexemplos de tipo Frankfurt não teriam sucesso em mostrar a falsidade do PPA quando aplicado a estados mentais tais como escolhas, decisões ou volições: “nós podemos dizer que um interventor contrafactual de tipo Frankfurt poderia privar um agente da liberdade de executar um determinado desejo ou intenção, mas ele não poderia impedi-lo da sua liberdade de formá-lo” (WIDERKER, 1995, p. 255).

Já Stewart Goetz sustenta que nós estaríamos sendo distraídos pelo mecanismo filosófico (tanto de Frankfurt como dos defensores do compatibilismo de tipo Frankfurt) quando é dito que Black estaria desempenhando a tarefa de eliminar as possibilidades alternativas. Porém, isso seria uma ilusão porque seria o determinismo causal, e não Black e seu dispositivo, que suprimiria o tipo de controle metafísico em discussão. Vejamos esse ponto mais especificamente:

[O exemplo de tipo 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 fazer de outro modo. Esta aparência é ilusória porque sem a ocorrência do determinismo causal na sequência real dos eventos, o dispositivo não pode impedir Jones de fazer uma escolha alternativa, e com determinismo causal na sequência real dos eventos, não é o dispositivo que impede Jones de fazer uma escolha alternativa. Em resumo, se Jones não é livre para fazer de outro modo, é por causa da ocorrência do determinismo causal na sequência real dos eventos, e não por causa do dispositivo de Black na sequência alternativa. Portanto, é errado concluir que Jones é moralmente responsável embora ele não seja livre para escolher de outro modo (GOETZ, 2005, p. 85).

O que faz com que Jones fosse incapaz de escolher de outro modo, em t_2 , seria o estado anterior do universo, em t_1 , e as leis da natureza. Black e o seu dispositivo são nesse caso impotentes. Dito de outro modo, seria unicamente por causa da suposição de uma relação

causalmente determinista, presente na sequência real, que a possibilidade de Jones de *começar* a escolher diferentemente seria eliminada. Se a tese do determinismo causal fosse falsa, então Black e o seu dispositivo não poderiam impedir que Jones escolhesse de outro modo.

O próximo subcapítulo será para procurar responder às críticas levantadas acima.

2.5 Réplica à defesa do dilema

Terei por foco o dilema que pressupõe o determinismo causal. Eu concordo com a crítica de que, sendo verdadeira a tese determinista, não seria razoável ou justificável assumir que o agente é moralmente responsável pelo seu comportamento, pois seria problemático concluir, somente a partir da análise dos exemplos (de tipo Frankfurt), que o agente é moralmente responsável. Esses casos poderiam ser usados como o primeiro passo de um argumento mais complexo para a compatibilidade do determinismo causal com a responsabilidade. A partir da reflexão sobre eles, seria plausível chegar à conclusão preliminar de que *se* Jones não é moralmente responsável pela sua escolha e ação, a razão não é simplesmente porque ele careceria de possibilidades alternativas. O argumento pode ser apresentado como segue:

Isto é, supondo que nós explicitamente assumimos que determinismo causal é verdadeiro, nesses casos é importante primeiro observar que eu não proponho que nós precipitadamente devemos concluir, a partir da mera reflexão sobre os casos, que Jones é moralmente responsável pela sua escolha e ação. Em vez disso, a conclusão inicial é que, se ele não é moralmente responsável, não é porque ele carece de possibilidades alternativas apropriadas (FISCHER, 2010a, p. 323).

Nesse sentido, seria intuitivamente razoável afirmar, como primeiro passo do argumento, que as possibilidades alternativas são irrelevantes para atribuições de responsabilidade moral. Em outras palavras, esses casos forneceriam fortes razões para pensar que responsabilidade não exige a capacidade para agir diferentemente. Se isso se confirmar como justificável, então o segundo passo do argumento consistiria em questionar se dado a verdade do determinismo causal, existiria algum outro fator que desempenharia algum papel, na sequência real, e que ameaçaria a responsabilidade moral (FISCHER, 2002, p. 291-292)³⁵. O desafio agora é outro, não mais sobre a compatibilidade do determinismo com o tipo de

³⁵ Se responsabilidade moral não exige o tipo de controle que envolve possibilidades alternativas, então isso abre a possibilidade para uma defesa do semicompatibilismo, mesmo admitindo a solidez do argumento da consequência (FISCHER, 2007, p. 53-82). O agnosticismo é conveniente ao defensor do semicompatibilismo, pois essa posição simplesmente afirma que determinismo causal seria compatível com responsabilidade moral, independentemente se determinismo eliminaria acesso às possibilidades alternativas.

controle que envolveria a capacidade do agente para escolher ou agir de outro modo, mas como compatibilizar determinismo com responsabilidade moral na sequência real dos eventos.

Um dos problemas que o incompatibilista precisaria enfrentar é se no futuro a física vier a descobrir que o determinismo causal é verdadeiro. Sendo o determinismo causal verdadeiro, e o argumento da consequência sólido, então *não* haveria possibilidades alternativas disponíveis ao agente, nem mesmo um mero átimo de liberdade (FISCHER, 2007, p. 61; FISCHER, 2006, p. 45).

Porém, penso que nós não precisaríamos nos comprometer com essa hipótese. Nos casos de tipo Frankfurt, e no caso original de Frankfurt, por exemplo, existem, pelo menos, dois fatores que tornariam plausível afirmar que Jones carece do tipo relevante de acesso metafísico genuíno às possibilidades alternativas: Black e seu dispositivo, por um lado, e a suposição do determinismo causal, por outro. No entanto, esses fatores não teriam desempenhado nenhum papel na sequência real dos eventos. Portanto, não teriam afetado a responsabilidade moral do agente.

A conclusão inicial, como dito anteriormente, de que se Jones não fosse moralmente responsável não seria porque ele careceria de possibilidades alternativas, tem recebido uma série de objeções. Uma dessas objeções, já mencionada anteriormente, foi feita por Goetz, o qual teria imputado a Frankfurt, e aos defensores dos casos de tipo Frankfurt, o uso de mecanismos para nos distrair e nos iludir ao sustentarem que Black e o seu dispositivo neurológico seriam as causas da eliminação das possibilidades alternativas. Para Goetz, não seria nem Black nem o seu dispositivo, mas a relação causalmente determinística entre o estado do universo, em t_1 , e a escolha de Jones, em t_2 , que eliminaria a possibilidade de Jones de começar a escolher diferentemente.

Acredito que Frankfurt talvez possa responder à crítica de Goetz. Em uma interessante nota de rodapé estaria um alerta sobre o papel de Black:

O exemplo é (...) flexível suficiente para permitir a eliminação de Black completamente. Alguém que pensa que a eficácia do exemplo é enfraquecida pela sua dependência de um manipulador humano, que impõe sua vontade sobre Jones, pode substituir Black por uma máquina programada para fazer o que Black faz. Se isso ainda não é bom suficiente, esqueça Black e a máquina, e suponha que seu papel é desempenhado por forças naturais que não envolve nenhuma vontade (FRANKFURT, 1998, p. 7, nota n.4).

A objeção de que seria o determinismo causal e não Black que eliminaria as possibilidades alternativas é controversa. Como argumentado por Fischer, “não é claro que o mero fato de que um estado anterior do universo (em conjunto com as leis da natureza) explica

por que Jones não poderia em t_2 escolher de outro modo implica que *nenhum outro fator* pode desempenhar esse tipo de papel explicativo” (FISCHER, 2010a, p. 332).

Black e o seu dispositivo seriam suficientes para impedir Jones de escolher ou agir diferentemente. Mas intuitivamente esses fatores seriam irrelevantes para a responsabilidade moral de Jones, na medida em que Black não teria desempenhado qualquer papel na sequência real das deliberações ou ações de Jones. Quando algum fator é irrelevante para o que acontece na sequência real em que resultam as escolhas e ações do agente, parece razoável concluir que esse fator é também irrelevante para a responsabilidade moral (FISCHER, 2007, p. 60).

As premissas do argumento para a compatibilidade do determinismo causal com responsabilidade moral poderiam ser formuladas da seguinte maneira: suponha que a tese determinista seja verdadeira, que a presença de Black (e o seu dispositivo) eliminaria a liberdade do agente para escolher ou agir diferentemente. Se Jones não é moralmente responsável por escolher ou agir, não seria em virtude do mero fato de que ele não tinha possibilidades alternativas. Se o determinismo exclui a responsabilidade moral não seria em virtude de eliminar as possibilidades alternativas (FISCHER, 2010a, p. 328).

São dois tipos de argumentos incompatibilistas. Um deles, o argumento indireto, sustentaria que o determinismo é incompatível com a responsabilidade moral, pois o determinismo eliminaria as possibilidades alternativas. O outro, o argumento direto, defenderia que o determinismo é incompatível com a responsabilidade, mesmo supondo que a responsabilidade não exige a capacidade para escolher ou agir diferentemente.

O capítulo a seguir tratará de uma posição intermediária entre o compatibilismo e o incompatibilismo, a saber, o semicompatibilismo defendido por John Fischer, que, partindo da tese de Frankfurt, procura basear as atribuições de responsabilidade moral na sequência real resultante do comportamento.

3 O SEMICOMPATIBILISMO DE FISCHER

Este terceiro capítulo trata de posições compatibilistas que apelam para uma noção de controle que não envolve acesso genuíno metafísico às possibilidades alternativas, e que poderia ser compatível com o determinismo. Também apresentarei o que distingue a teoria de John Fischer da posição de Frankfurt analisada no capítulo anterior.

A posição semicompatibilista, defendida por Fischer, possui algumas vantagens em relação ao compatibilismo e ao incompatibilismo, a saber, ela estaria articulada para aceitar a verdade da tese determinista. Mesmo se futuramente as teorias científicas vierem a divulgar algum estudo que nos mostre que o determinismo é verdadeiro, ainda assim seria razoável e legítimo considerar nós mesmos, e os outros, como pessoas livres e moralmente responsáveis. O semicompatibilista não precisaria revisar as suas crenças. Esse seria um problema que o libertista incompatibilista precisaria enfrentar, na medida em que ele sustenta que o determinismo causal é falso e incompatível com a responsabilidade moral.

A seguir apresentarei a distinção entre o controle de direcionamento e o regulativo.

3.1 Controle de direcionamento e regulativo

A partir de uma distinção entre dois tipos de controle, será possível perceber com mais clareza qual deles estaria associado com atribuições de responsabilidade moral³⁶. Ilustrarei essa diferença por meio de dois casos distintos³⁷. No primeiro caso, suponha que eu esteja dirigindo um carro. O seu sistema elétrico está funcionando bem. Eu desejo dobrar à direita. Como resultado da minha intenção, eu sinalizo com a seta, dobro o volante e guio o carro para a direita. Observe que nesse caso eu era capaz de formar uma diferente intenção: eu poderia ter formado a intenção de dobrar à esquerda. Assim, na medida em que “eu, na realidade, guio o carro em um certo modo, eu tenho controle de direcionamento” (FISCHER, 2010b, p. 232). No caso em que “eu tenho o poder para guiar o carro em um diferente modo, eu tenho controle regulativo” (FISCHER, 2010b, p. 232).

Agora um segundo caso. Eu novamente guio o carro para a direita. O sistema funciona normalmente quando eu conduzo o carro nessa direção. Porém, diferentemente do primeiro

³⁶ A explicação da posição semicompatibilista de responsabilidade de Fischer foi primeiramente introduzida no artigo “Responsibility and Control” (1982). Ela foi posteriormente desenvolvida no livro *The Metaphysics of Free Will* (1994).

³⁷ A apresentação desses dois tipos de casos é baseada no artigo de Fischer (2010). A distinção entre controle de direcionamento e regulativo também pode ser encontrada no artigo de Fischer (1991).

caso, o sistema elétrico foi bloqueado, sem o meu conhecimento. Se eu tentar dobrar à esquerda, o carro dobrará para a direita. Suponha que na realidade eu não tente dobrar o carro para a esquerda, mas para a direita. Assim, o sistema funciona normalmente. Eu controlo o carro no sentido de guiá-lo, em um certo modo, para a direita. Contudo, eu não tenho controle sobre os movimentos do carro. Sendo assim, eu teria o controle de direcionamento, mas não o regulativo: “Este caso (o qual tem alguma característica estrutural de um caso de tipo Frankfurt) ajuda a mostrar que alguém pode ter controle de direcionamento, sem o controle regulativo” (FISCHER, 2010b, p. 233)³⁸.

O tipo de controle exigido para a responsabilidade moral seria o controle de direcionamento e não o regulativo. Desse modo, se preservaria a abordagem tradicional de associar responsabilidade moral com controle de um certo tipo:

Nós acreditamos que existem certas restrições que dão estrutura a *qualquer* especificação plausível das condições da aplicação do conceito de responsabilidade moral. Estas condições são “negativas” no sentido que elas indicam dois modos em que o agente não pode ser moralmente responsável: ignorância e força (...) Estas condições negativas podem ser rastreadas até Aristóteles, e nós devemos chamá-las as condições “Aristotélicas”. A primeira condição, que pode ser denominada a “condição epistêmica”, corresponde à desculpa da ignorância. Ela captura a intuição de que um agente é responsável unicamente se ele conhece os fatos particulares em torno da sua ação e os atos com o tipo correto de crenças e intenções (FISCHER; RAVIZZA, 1998, p.12-13)³⁹.

³⁸ “Esse segundo caso do carro deveria elucidar a intuição de que nós não precisamos de controle regulativo (acesso metafísico genuíno às possibilidades alternativas) para ter o tipo de controle envolvido na responsabilidade moral” (FISCHER, 2010b, p. 233).

³⁹ Para Aristóteles, das pessoas que agem por ignorância, as que se arrependem são consideradas agentes involuntários, e as que não se arrependem, agentes não voluntários, “pois o homem que fez alguma coisa devido à ignorância e não se aflige com o seu ato não agiu voluntariamente, visto que não sabia o que estava fazendo, mas tampouco agiu involuntariamente já que isso não lhe causa dor” (ARISTÓTELES, 1984a, *EN* 1110 a 15-25). Agir *por* ignorância parece diferir também de agir *na* ignorância. O embriagado e o enfurecido agem na ignorância e sem conhecimento do que fazem. Ademais, o termo involuntário tende a ser usado não em relação à ignorância do universal, pela qual os homens são passíveis de censura, mas à ignorância dos particulares, das circunstâncias do ato. Um exemplo que Aristóteles utiliza para ilustrar esse aspecto é o do homem que dá um remédio para curar, mas sem saber exatamente dos efeitos desse remédio, acaba matando o seu paciente. A ignorância pode relacionar-se com qualquer dessas circunstâncias do ato (quais são os efeitos colaterais do remédio, quais são as alergias do paciente etc). A prática de um ato considerado involuntário por causa de uma ignorância desse tipo deve trazer arrependimento. Além da ignorância do particular, há também outra condição que isentaria a responsabilidade moral do agente, qual seja, as ações forçadas ou compulsórias. As ações são forçadas quando a causa se encontra nas circunstâncias exteriores e nada contribui quem é forçado. Por exemplo, alguém que tivesse sido sequestrado (ARISTÓTELES, 1984a, *EN* 1110 b 1-20).

Em síntese, o agente não poderia ser ignorante das circunstâncias da sua ação (condição epistêmica), e também não poderia ter sido forçado a fazer o que ele fez. A origem da ação deveria estar no agente para que se possa considerá-lo apropriadamente como alguém que agiu livremente. Retomando Aristóteles, pode se dizer que nosso caráter tem de estar sob nosso controle voluntário (*EN* 1114 a 20-25). A virtude moral é adquirida pelo hábito, é por praticar ações justas que nos tornamos justos (1103 a 15-1103 b 1); ao escolher o que é o bom ou o mau nós determinamos o nosso caráter (1112 a 1-5). Nesse caso, somos louvados pelas virtudes e censurados pelos vícios desde que esses estejam sob o nosso controle (1106 a)⁴⁰. Por analogia, são os cegos por imprudência e negligência que são censurados e não os cegos por natureza porque admitimos que estava no poder da pessoa tomar as precauções necessárias para ter evitado o acidente que causou a cegueira, como não dirigir embriado, por exemplo.

Levando em consideração que o nosso caráter é determinado pela educação que recebemos dos nossos pais na infância, pela nossa genética ou pelo ambiente em que vivemos, como podemos ser responsabilizados pelo o que somos se o que somos é determinado por fatores além de nosso controle? Adotando a ideia aristotélica inicial de que o nosso caráter é formado pelo hábito como responsabilizar a ação de uma pessoa se ela age determinada pelo seu caráter?

Robert Audi apresenta um exemplo interessante e talvez possa responder essa questão. Jean é desonesta, porém tem a decência de se sentir culpada pelo seu caráter. Ela tem o desejo de reformá-lo. Talvez a desonestidade de Jean foi formada na sua infância, no seu ambiente familiar. Dado essa circunstância, Jean carece de responsabilidade genética pela sua desonestidade. Porém, isso não significa que ela não seja responsável pelos atos desonestos ou que ela não pode mudar o seu caráter (AUDI, 1997, p. 161)⁴¹.

⁴⁰ É importante chamar atenção para o fato de que é possível encontrar em certas passagens da obra de Aristóteles um aspecto determinista quando se trata das disposições de caráter. O caráter é formado a partir da prática de atos repetidos. Ao praticar atos injustos, com conhecimento, o agente se torna injusto voluntariamente. Estava em seu poder, no início, não se tornar injusto, por isso ele é injusto voluntariamente. Porém, depois de se tornar injusto não é mais possível não o ser (ARISTÓTELES, 1984a, *EN* 1114 a 10-30). Apesar desse aspecto estável da disposição, a qual respresentaria a ideia determinista, de que depois de adquirido um certo caráter não seria mais concebível perdê-lo, há outras passagens da *Ética* que apontam para a possibilidade de flexibilidade e de mudança do caráter como por exemplo a seguinte: “os incontinentes por hábito são mais curáveis do que aqueles incontinentes por natureza” (ARISTÓTELES, 1984a, *EN* 1152 a 28-30).

⁴¹ Audi faz uma distinção entre três tipos de responsabilidade pelos traços de caráter: a (a) genética, que é a responsabilidade por produzir o traço em questão, a (b) retencional, que é a responsabilidade por reter o traço de caráter e (c) a prospectiva, que é a responsabilidade por adquirir um determinado tipo de caráter e não outro. Por exemplo, Jean não tem responsabilidade genética pelo seu traço de

O semicompatibilismo seria, ao mesmo tempo, uma doutrina conservadora e radical. Conservadora porque não exigiria uma revisão do conceito de responsabilidade moral ou uma revisão das práticas de responsabilidade, a prática que considera apropriado em certas circunstâncias censurar e punir, e que em outras circunstâncias considera apropriado louvar e recompensar. Também preservaria a ideia tradicional de que a responsabilidade estaria conectada com a liberdade. Todavia, seria radical porque se afastaria significativamente das condições que em geral se postula para a aplicação do conceito de responsabilidade moral, na medida em que não exigiria possibilidades alternativas no que diz respeito à escolha, à ação ou mesmo à formação do caráter (FISCHER, 2007, p. 81-82).

O controle de direcionamento, no qual estariam baseadas as atribuições de responsabilidade moral, e que seria compatível com determinismo causal, é constituído por dois elementos fundamentais: o mecanismo resultante na ação deveria ser do “próprio agente” e deveria ser sensível às razões [*responsive to reasons*] de modo apropriado.

Poderia ser útil começar ilustrando alguns casos em que o agente não seria sensível às razões de modo apropriado, e que, por carecer do tipo de controle relevante sobre o seu comportamento, não poderia ser moralmente responsável. Por exemplo, pessoas que agem sob o efeito da hipnose, lavagem cerebral, drogas potentes, certas formas de manipulação direta do cérebro, certos tipos de distúrbios mentais (como fobias extremas) e determinados tipos de ameaças coercitivas. Considere o caixa de um banco que é ameaçado com uma arma a entregar o dinheiro. Vamos supor que ele tivesse um desejo irresistível para sujeitar-se à ameaça. Se ele tivesse agido a partir de um tal impulso, seria plausível considerá-lo como não tendo o controle de direcionamento da sua ação. De igual modo, quando as pessoas são manipuladas elas passam a se comportar como uma marionete, o que as torna candidatas inaptas para o louvor e a censura. Suponha que uma pessoa é induzida por hipnose a bater em alguém próximo toda a vez que o telefone tocar. Essa pessoa também não seria “sensível às razões de modo apropriado” (FISCHER; RAVIZZA, 1998, p. 35-36), pois dados os efeitos externos sobre o seu comportamento, ela bateria em alguém próximo mesmo se tivesse fortes razões para não bater. Ela não seria sensível porque o que ela faz não seria afetado pelas suas razões para agir.

Em casos de hipnose, ou em casos em que alguém sofre de certos distúrbios mentais, operaria algum fator que enfraquece a responsabilidade na sequência real. Esses casos diferem

desonestidade porque esse traço foi adquirido na convivência com seus pais desonestos. São os pais de Jean que são responsáveis geneticamente pela sua desonestidade. Jean tem responsabilidade prospectiva em adquirir um novo traço de caráter, a honestidade. E ela tem também responsabilidade retencional por reter esse novo traço de caráter (AUDI, 1997, p. 160-161).

daqueles nos quais o mecanismo deliberativo humano opera normalmente. Quando alguém delibera e pesa razões, e decide que deve doar cinco por cento do seu salário para uma instituição de caridade, nós podemos pensar que ela poderia ser legitimamente louvada pela ação praticada em virtude do fato de que a sua decisão não foi induzida por manipulação direta do seu cérebro ou hipnose. Quando o mecanismo deliberativo humano opera normalmente, a pessoa é *sensível* às razões que ela percebe na situação. Por exemplo, se essa pessoa descobrisse que existe alguma fraude na instituição para a qual ele doou, ela provavelmente deixaria de contribuir para essa instituição, sendo sensível às razões envolvidas na situação (FISCHER; RAVIZZA, 1998, p. 36-37).

Há um consenso na literatura sobre o tema da responsabilidade moral de que o agente seria moralmente responsável pelo seu comportamento se ele fosse livre para fazer de outro modo, e que ele não poderia ser apropriadamente responsável moralmente se ele não tivesse essa liberdade. Nesse sentido, “é muito natural e razoável pensar que a diferença entre agentes moralmente responsáveis e àqueles que não são consistiria na “sensibilidade às razões” dos agentes” (FISCHER, 2006, p. 65). Entretanto, existem casos, como os de tipo Frankfurt, em que o agente poderia ser moralmente responsável por realizar uma determinada ação mesmo ele não sendo sensível às razões (quer dizer, embora ele não pudesse ter agido diferentemente). Vejamos a seguinte versão do exemplo:

Imagine que uma pessoa má instalou um dispositivo no cérebro de Jones, que permite monitorar a sua atividade cerebral e também intervir nela, se caso ele desejar. Ela pode eletronicamente manipular a atividade cerebral de Jones por meio de um “controle remoto” para induzir decisões, e vamos imaginar que ele também pode assegurar que Jones aja sobre as decisões assim induzidas. Agora, suponha que Jones está prestes a matar o seu vizinho, e que isto é precisamente o que a pessoa má deseja. Isto é, vamos imaginar que o dispositivo simplesmente monitora a atividade cerebral de Jones, mas não desempenha qualquer papel nas suas decisões e ações reais. Jones delibera e se comporta como se não tivesse sido implantado algum dispositivo no seu cérebro. Mas, podemos também imaginar que tivesse Jones começado a decidir não matar o seu vizinho, o dispositivo teria sido ativado e teria causado ele a escolher assassinar o vizinho (e, assim, realizado a ação) (FISCHER, 2006, p. 65-66)⁴².

Aplicando a distinção entre sequência real e alternativa teríamos: o fator que enfraquece a responsabilidade (manipulação direta do cérebro) operaria *apenas* na sequência alternativa

⁴² Essa é uma versão do caso de tipo Frankfurt que eu adaptei.

(resultando na incapacidade do agente para fazer de outro modo). Em um caso de tipo Frankfurt, o agente possuiria controle de direcionamento e careceria de controle regulativo:

Em um caso de tipo Frankfurt, a sequência real procede em um modo que baseia atribuições de responsabilidade moral, mesmo embora o cenário alternativo (ou talvez um conjunto de cenários alternativos) procede em um modo que elimina a responsabilidade. Em um caso de tipo Frankfurt nenhum fator que enfraquece responsabilidade opera na sequência real, embora tal fator ocorra no cenário alternativo. Tais casos nos levam a adotar uma teoria de responsabilidade moral mais refinada - um modelo de responsabilidade moral da sequência real. Com tal abordagem, nós distinguimos entre dois tipos de mecanismos que operam na sequência real e na(s) sequência(s) alternativa(s) (FISCHER, 2006, p. 66).

Esses exemplos nos fazem pensar que nós teríamos casos nos quais o agente agiria livremente, porém não seria sensível às razões. Além disso, eles nos convidam a desenvolver e adotar uma abordagem de responsabilidade moral baseada sobre a “sequência real”, sobre as características da “sequência real que conduz à ação” (FISCHER; RAVIZZA, 1998, p. 37)⁴³. No curso real dos eventos, o agente escolheria livremente e agiria de acordo com a sua escolha. Nessa abordagem, a responsabilidade moral não envolveria possibilidades alternativas ou a capacidade para agir diferentemente:

Uma implicação é que ser moralmente responsável não exige acesso metafísico genuíno às possibilidades alternativas. (...) Isso não implica que os cenários alternativos ou mundos possíveis não reais são irrelevantes para a especificação das propriedades modais ou disposicionais pertinentes da sequência real (FISCHER, 2006, p. 23).

O que seria significativo é “o que o agente faz, na sequência real, e como suas ações vem a ser realizadas” (FISCHER; RAVIZZA, 1998, p. 37), ou seja, é indispensável analisar a história causal da ação. O critério relevante para atribuição de responsabilidade moral seria procurar identificar como as decisões e as ações se desdobram na sequência real (causal) em nosso mundo. Porém, se o mecanismo deliberativo do agente, diante de diferentes estímulos, o levasse a se comportar do mesmo modo em cenários alternativos (ou mundos possíveis), o agente não seria responsável, pois o mecanismo não teria operado adequadamente. Portanto, cenários alternativos são importantes e relevantes para determinar se o agente é responsável no

⁴³ A abordagem da sequência real implica, segundo Fischer, simplesmente “uma explicação de responsabilidade moral que não exige a disponibilidade de possibilidades alternativas” (1999, p. 125n56).

mundo real. Mas, não significaria que o agente deve ser capaz de agir diferentemente do modo como ele agiu no mundo real.

Mais especificamente, a partir da análise dos casos de tipo Frankfurt, poderia ser desenvolvida uma explicação do controle de direcionamento e uma teoria sobre responsabilidade moral que concentraria no mecanismo que conduz à ação. Essa abordagem distingue-se de outras teorias, as quais baseiam-se unicamente “no agente”. Ademais, é crucial perceber que, nos casos de tipo Frankfurt, o agente não poderia ter feito de outro modo e, sendo assim, ele não seria sensível às razões. Conforme ratificado pela citação: “o mecanismo da sequência real pode ser sensível às razões, embora o *agente* não seja” (FISCHER, 2006, p. 66). Desse modo, o tipo de mecanismo que teria operado no cenário alternativo não seria sensível às razões, sendo esse um diferente tipo de mecanismo (FISCHER; RAVIZZA, 1998, p. 37-38).

Para exemplificar a existência de diferentes tipos de mecanismos, gostaria de retornar um exemplo de tipo Frankfurt, o caso em que Jones vota nos democratas por razões próprias e Black não interfere. Vamos imaginar que a ação de Jones resulte da faculdade normal do seu raciocínio prático. No cenário real, pareceria razoável afirmar que *o mecanismo* operante é sensível às razões⁴⁴. Quer dizer, tivesse existido uma razão suficiente para fazer de outro modo, o mecanismo operante na sequência real teria produzido ou tentado produzir um diferente tipo de ação. No cenário alternativo, se Black tivesse interferido, teria ocorrido um diferente tipo de mecanismo (estimulação eletrônica do cérebro), e Jones não seria sensível às razões.

Após ter esboçado, em linhas gerais, as características de um dos componentes do controle de direcionamento, qual seja, “sensibilidade às razões”, podemos analisar esse conceito mais detalhadamente, para depois explicitar o outro elemento constituinte que diz respeito ao comportamento resultante do mecanismo do próprio agente.

3.2 Sensibilidade às razões: forte, fraca e moderada

A teoria de responsabilidade moral descrita aqui pretende cumprir as seguintes condições: basear a responsabilidade moral do agente nas suas razões para ação e, como implicação dessa condição, fornecer uma teoria que explica como o agente seria moralmente responsável por praticar atos moralmente censuráveis, errados ou imprudentes. A ideia é que existiriam diferentes graus de sensibilidade às razões, sendo os dois extremos o sentido forte (FISCHER; RAVIZZA, 1998, p. 42-43) e o fraco (FISCHER; RAVIZZA, 1998, p. 65-66).

⁴⁴ Por mecanismo da sequência real, Fischer significa “o modo de como a sequência real se desenvolve ou o tipo de processo que ocorre na sequência real” (1999, p. 128n63).

Para a ação resultar a partir de um mecanismo fortemente sensível às razões, três condições deveriam ser satisfeitas: (i) receptividade suficiente às razões, ou seja, o agente deve *reconhecer* as razões que existem e se essas são razões suficientes para agir de outro modo; (ii) reatividade às razões, isto é, o agente deveria *escolher* de acordo com as razões suficientes para agir de outro modo; e (iii) o agente deveria *agir* de acordo com a escolha (de agir diferentemente).

Sensibilidade forte às razões não seria uma condição necessária para o controle de direcionamento e para a responsabilidade moral. Para ver por que, podemos considerar o problema da fraqueza da vontade. Imagine que como resultado da operação do raciocínio prático normal, Pedro rouba um livro da biblioteca, *sabendo* que esse é um ato moralmente errado e que ele será preso. Ele estaria agindo livremente e poderia ser moralmente responsável por esse ato se nenhum fator que enfraquece a responsabilidade estivesse operando na sequência real (como manipulação ou hipnose).

Seria plausível e razoável sustentar que nessa situação o agente é moralmente responsável por roubar o livro, embora o seu mecanismo da sequência real não seja *fortemente* sensível às razões. Nesse caso, existe razão suficiente para fazer de outro modo (razão moral), o agente *reconhece* essa razão, *não escolhe* de acordo com a razão suficiente e *não age* de acordo com tal razão. Apesar de reconhecer a razão de não roubar o livro (seria errado), mesmo assim ele rouba⁴⁵. O problema em aceitar esse tipo de mecanismo no sentido forte como condição necessária para a responsabilidade moral é a implicação de que Pedro *não é* moralmente responsável. Essa seria uma conclusão não razoável visto que a responsabilidade moral deveria abranger ações censuráveis decorrentes de fraqueza da vontade.

Para a ação resultante do mecanismo sensível às razões no sentido fraco, as seguintes condições deveriam ser satisfeitas: (i) a exigência de algum cenário possível (ou mundo possível) no qual existe razão suficiente para fazer de outro modo; (ii) o agente reconhece essa razão; e (iii) e ele faz de outro modo.

Vamos considerar a decisão de Pedro para ir ao jogo. Uma razão suficiente para fazer de outro modo, seria, por exemplo, o fato de que o prazo para a entrega do primeiro capítulo da sua tese de doutorado está acabando e ele está atrasado. Se apesar dessa razão ele fosse ao jogo, nós poderíamos dizer que Pedro sofre de fraqueza da vontade. Entretanto, apesar da

⁴⁵ Há uma distinção entre razão moral e prudencial. O fato de que fumar causa câncer é uma razão prudencial que eu tenho para não fumar, uma vez que eu tenho interesse de preservar a saúde. Por outro lado, o fato de que há outras pessoas na sala é uma razão moral para eu não fumar, uma vez que fumar em um ambiente fechado é um ato de desrespeito com as outras pessoas.

possibilidade da existência de situações como fraqueza da vontade, poderiam existir outras circunstâncias nas quais o agente agiria a partir de um mecanismo apropriadamente sensível às razões. Certamente poderia existir algum cenário, no qual o mecanismo real operaria, em que o agente teria uma razão suficiente para não ir ao jogo e não iria ao jogo. Por exemplo, ele não iria ao jogo se o valor do ingresso fosse duzentos reais. O que quero dizer é que em determinadas situações certas razões seriam mais fortes do que outras. O fato de que o ingresso custa duzentos reais seria uma razão mais forte para Pedro não ir ao jogo, se comparada com a razão mais fraca que Pedro tem para ir ao jogo (o seu time irá jogar). De igual modo, quando alguém pratica uma ação errada, como roubar um livro, existe na realidade alguma razão suficiente (moral e prudencial) para não roubar. O mecanismo real do agente poderia ser sensível a algum incentivo possível para não roubar (digamos que se ele roubasse o livro teria como consequência indesejada uma multa). A princípio, parece razoável responsabilizar moralmente o agente na medida em que a sua ação resulta de um mecanismo fracamente sensível às razões (FISCHER; RAVIZZA, 1998, p. 44-46). Porém, embora esse tipo de mecanismo (sensível às razões no sentido fraco) fosse um elemento constituinte satisfatório e vantajoso para o controle de direcionamento comparado ao outro tipo de mecanismo (sensível às razões no sentido forte), ele seria ainda assim problemático na medida em que não haveria uma conexão adequada entre razões e ações no mecanismo resultante.

Para ver por que o mecanismo sensível às razões no sentido fraco não seria um mecanismo adequado, vejamos o caso de um maníaco. O maníaco do trem não praticaria um ato de crueldade apenas se um dos passageiros do trem estivesse fumando um charuto. Suponhamos que ele esteja em um trem, e que nenhum passageiro esteja fumando um charuto. Como a razão que faria o maníaco não matar todos os passageiros não está presente, ele decide matar todos os passageiros. Nesse exemplo o maníaco estaria agindo a partir do mecanismo sensível às razões no sentido fraco, pois se um dos passageiros estivesse fumando um charuto, ele teria sido sensível a essa razão e teria decidido não matar. Porém, intuitivamente o maníaco não controla as suas ações. Nesse sentido, esse mecanismo seria demasiadamente fraco para descrever o tipo de controle exigido para a responsabilidade moral.

O conceito de sensibilidade moderada às razões poderia ser concebido de modo a evitar os problemas encontrados nos outros dois tipos de mecanismos apresentados nos parágrafos anteriores. Fischer e Ravizza elencam quatro requisitos necessários que esse tipo de mecanismo deveria satisfazer: (i) o mecanismo psicológico sobre o qual o agente age deve ser “*regularmente* receptivo às razões” para agir de outro modo. Uma pessoa que age a partir de um mecanismo regularmente receptivo às razões deve mostrar um certo tipo de padrão de

reconhecimento de razões. Em outras palavras, significa “ver por que uma razão é mais forte do que outra, e entender como a aceitação de uma razão como suficiente implicaria que uma razão mais forte deveria também ser suficiente” (FISCHER; RAVIZZA, 1998, p. 71); (ii) as razões subjetivas do agente devem ser “minimamente baseadas na realidade”, isto é, o agente não pode ser iludido sobre a natureza da realidade; (iii) o mecanismo psicológico sob o qual o agente age deve ser *reativo* às razões reconhecidas subjetivamente em um sentido fraco⁴⁶. Se o mecanismo sob o qual o agente age reage a algum incentivo, isso mostraria que o mecanismo poderia reagir a qualquer incentivo para fazer de outro modo (FISCHER; RAVIZZA, 1998, p. 73). E para ser um sujeito apropriado de julgamento moral, (iv) o agente deve agir a partir de um mecanismo que é moderadamente receptivo às razões *morais* e aos outros tipos de razões. Animais, crianças e psicopatas, embora possam agir sobre um mecanismo que é sensível a razões instrumentais ou prudenciais, não seriam apropriadamente sensíveis às razões morais. Como eles não *reconheceriam* ou *reagiriam* às demandas morais, não seriam agentes morais e moralmente responsáveis⁴⁷.

Agir a partir de um mecanismo apropriadamente sensível às razões não seria suficiente para ter controle de direcionamento. Além desse componente, é exigido que o mecanismo seja *do próprio agente*. Antes de definir esse outro conceito, apresentarei a posição de Frankfurt, pois a análise desse conceito revela divergências entre as posições compatibilistas. Frankfurt, como será visto, defende uma posição a-histórica de responsabilidade moral, enquanto Fischer entende que o conceito de responsabilidade moral deveria ser entendido como um conceito essencialmente histórico.

3.3 Uma noção não histórica de responsabilidade

No capítulo anterior vimos que o tipo de controle necessário e suficiente para atribuições de responsabilidade moral não envolveria o poder para agir diferentemente. Sendo assim, o debate tradicional poderia ser contornado, pois não mais se discutiria a compatibilidade do determinismo com a liberdade metafísica, mas sim a compatibilidade do determinismo com a responsabilidade moral.

⁴⁶ Uma exigência mais forte para *reatividade* tenderia a excluir casos de fraqueza da vontade.

⁴⁷ A psicopatia pode ser definida como um traço de personalidade, em que há ausência de empatia em relação ao outro. O psicopata age levando em consideração unicamente seus desejos individuais. Há graus menores e maiores de psicopatia que vão desde o não reconhecimento de limites até a falta de culpa e remorso.

De acordo com a tradição filosófica clássica, ser livre seria uma questão de fazer o que alguém deseja fazer. Essa definição de liberdade, segundo Frankfurt, não seria completamente clara. Por um lado, essa noção capturaria, pelo menos, parte do que estaria implícito na ideia de um agente que age livremente. Por outro lado, perderia inteiramente o conteúdo peculiar da ideia muito diferente de um agente cuja *vontade* é livre. Seria preciso observar a distinção existente entre liberdade da ação e liberdade da vontade, pois seria a última noção que estaria conectada com a ideia de responsabilidade moral.

Para demonstrar a irredutibilidade da liberdade da vontade à liberdade da ação, Frankfurt apresenta o seguinte argumento:

Nós não supomos que os animais desfrutem de liberdade da vontade, embora reconheçamos que um animal possa ser livre para correr para seja qual for a direção desejada. Assim, ter liberdade para fazer o que alguém deseja fazer não é uma condição suficiente para ter uma vontade livre. Também não é uma condição necessária. Privar alguém da sua liberdade da ação não é necessariamente enfraquecer a liberdade da sua vontade (FRANKFURT, 1971, p. 14).

Em síntese, alguém poderia carecer de liberdade da *ação*, porém sua *vontade* permaneceria livre. A questão da liberdade da vontade não diria respeito à relação entre o que um agente *faz* e o que ele *deseja* fazer, mas sim aos desejos eles mesmos.

Talvez a distinção entre liberdade da ação e liberdade da vontade fique mais clara a partir da definição da estrutura da vontade, pois, para Frankfurt, é esse conceito que distinguiria as pessoas das outras criaturas. Os seres humanos não seriam os únicos a possuir desejos ou os únicos capazes de tomar decisões. Essas capacidades “eles compartilham com membros de outras espécies, alguns parecem mesmo se engajar em deliberação e fazer decisões baseadas em pensamento anterior” (FRANKFURT, 1971, p. 6). A capacidade para formar “desejos de segunda ordem”, no entanto, seria característica unicamente dos seres humanos. Muitos animais parecem ter a capacidade para formar “desejos de primeira ordem”, que “são simplesmente desejos para fazer ou não fazer uma coisa ou outra”, como o desejo de comer isto ou aquilo. Nenhum animal fora o homem, contudo, parece ter a capacidade para autoavaliação reflexiva, que é manifestada na formação de desejos de segunda ordem” (FRANKFURT, 1971, p. 7).

Desejos de primeira ordem podem ser definidos quando alguém deseja fazer ou não fazer isso ou aquilo. Desejos de segunda ordem são desejos de ter ou não ter um certo desejo de primeira ordem. Além disso, alguém pode querer ter um certo desejo, mas querer que esse desejo não seja satisfeito ou efetivo. Por exemplo, um psiquiatra que trabalha com viciados em narcóticos acredita que sua capacidade para ajudar os seus pacientes seria melhorada se ele

entendesse melhor o desejo pela droga na qual eles são viciados. Digamos que o psiquiatra seja conduzido a querer ter um desejo pela droga, que ele seja inclinado ou movido a usar a droga. Seria possível que, embora ele queira ser movido pelo desejo de usar a droga, ele não queira que esse desejo seja efetivo (FRANKFURT, 1971, p. 9).

Alguém também pode ter um desejo de segunda ordem (i) quando simplesmente quer ter um determinado desejo ou (ii) quando quer que determinado desejo seja a sua vontade. Os desejos de segunda ordem do segundo tipo são denominados de “volições de segunda ordem”. O que é essencial para diferenciar as pessoas de outras criaturas, segundo Frankfurt, seriam as “volições de segunda ordem” (FRANKFURT, 1971, p. 10).

O termo *wanton* seria distintivo dos “agentes que têm desejos de primeira ordem, mas que não são pessoas porque, tendo ou não desejos de segunda ordem, eles não têm volições de segunda ordem. A característica essencial do agente impulsivo seria que ele não se importa com a sua vontade” (FRANKFURT, 1971, p. 11). Seus desejos simplesmente o movem a fazer certas coisas. Ele não estaria interessado se ele deseja ser movido por esses desejos ou por outros desejos. Essa classe de indivíduos incluiria todos animais não humanos, crianças jovens e talvez também alguns adultos humanos, em determinadas situações.

O *wanton* poderia possuir e usar faculdades racionais, deliberar sobre como fazer o que ele quer fazer, já que a essência de uma pessoa repousaria na estrutura da sua vontade e não na sua racionalidade. Contudo, não seria possível ser uma pessoa sem racionalidade, pois “é unicamente em virtude das suas capacidades racionais que uma pessoa é capaz de se tornar criticamente consciente da sua própria vontade e de formar volições de segunda ordem” (FRANKFURT, 1971, p. 11-12).

A distinção entre uma pessoa e um *wanton* poderia ser ilustrada pela diferença entre dois drogaditos. Um dos indivíduos manifestaria o desejo de abandonar o vício e lutaria contra o seu desejo pela droga, embora sem êxito. Esse seria o drogadito involuntário [*unwilling addict*], uma pessoa com desejos de primeira ordem conflitantes. Ele quer usar a droga, e também não quer usá-la. Além desses desejos de primeira ordem, ele teria também uma volição de segunda ordem, pois “ele não é neutro em relação ao conflito entre seu desejo de usar a droga e seu desejo de não a usar. É o último desejo e não o primeiro, que ele quer que constitua a sua vontade; é o último desejo e não o primeiro que ele quer que seja efetivo” (FRANKFURT, 1971, p.12). Ademais, ele identificaria a si mesmo, através de uma volição de segunda ordem, com um dos seus desejos conflitantes de primeira ordem, “ele faz um de seus desejos mais verdadeiro do que o outro” (FRANKFURT, 1971, p.13). Seria em virtude dessa identificação:

que o drogadito involuntário pode significativamente fazer declarações analiticamente enigmáticas de que a força que move ele a usar a droga é uma outra força do que a sua própria, e que não é da sua própria vontade livre, mas contra a sua vontade que essa força move ele a usar a droga (FRANKFURT, 1971, p.13).

O outro indivíduo seria o *wanton*. Ele não estaria preocupado se os desejos que o movem a agir são os desejos que ele gostaria que o movessem. A sua falta de preocupação é talvez devida à sua falta de capacidade para reflexão. Ele poderia vir a deliberar como obter a droga ou como administrá-la em si mesmo, “mas jamais ocorre a ele considerar se ele quer que a relação entre seus desejos resulte em ter a vontade que ele tem”. O *wanton* poderia ser um animal, e assim incapaz de se preocupar com sua vontade” (FRANKFURT, 1971, p.12). Além disso, ele poderia sofrer um conflito entre desejos de primeira ordem similar ao conflito sofrido pelo drogadito involuntário. Ele poderia, assim, querer usar a droga e querer abster-se de usá-la. Porém, diferentemente do drogadito involuntário, “ele não prefere que um de seus desejos de primeira ordem constitua a sua vontade” (FRANKFURT, 1971, p.12-13). Ele não teria volições de segunda ordem.

Existiria uma relação muito próxima entre a capacidade para formar volições de segunda ordem e o problema da vontade livre: “é unicamente porque uma pessoa tem volições de segunda ordem que ela é capaz de desfrutar e carecer de liberdade da vontade” (FRANKFURT, 1971, p.14). Esse conceito excluiria todos os *wantons*, humanos ou não humanos, pois esses agentes não satisfazem a condição essencial para o exercício da liberdade da vontade.

A partir da distinção entre liberdade da ação e liberdade da vontade, é possível perceber com mais clareza o ponto em questão. A liberdade da ação, como já apontado anteriormente, seria a liberdade para fazer o que se deseja fazer. Por analogia, uma pessoa desfrutaria de liberdade da vontade se fosse livre para ter a vontade que ela deseja. Nesse sentido, uma pessoa exerceria liberdade da vontade quando houvesse uma conformidade da sua vontade com as suas volições de segunda ordem. A carência dessa conformidade seria um indicativo de que a pessoa não estaria exercendo a sua liberdade da vontade. O drogadito involuntário, por exemplo, por não conseguir determinar a vontade que ele quer, tem uma vontade que não seria livre, pois não é a vontade que ele deseja.

Após feito esses esclarecimentos, o que eu gostaria de chamar atenção é que uma teoria da liberdade da vontade deveria fornecer uma análise sobre as condições para responsabilizar moralmente o agente de forma apropriada. O que é interessante na teoria de Frankfurt é a sua inovação ao pensar a relação entre a responsabilidade moral e a liberdade da vontade. Essa

relação tem sido mal compreendida, pois “não é verdade que uma pessoa é moralmente responsável pelo o que fez unicamente se sua vontade era livre quando ela agiu” (FRANKFURT, 1971, p.18). Uma pessoa poderia ser moralmente responsável mesmo que a sua vontade não fosse livre. Para ilustrar esse ponto, vejamos um terceiro tipo de drogadito. Ele teria as mesmas condições psicológicas e o mesmo impulso irresistível dos dois drogaditos vistos anteriormente, exceto que ele estaria completamente satisfeito com a sua condição. Ele é um drogadito voluntário [*willing addict*], que não quer abandonar a droga e o vício. Cito:

se a força do seu vício de alguma forma enfraquecesse, ele faria qualquer coisa que ele pudesse para restabelecê-la; se o seu desejo pela droga começasse a enfraquecer, ele tomaria as medidas necessárias para renovar a sua intensidade. A vontade do drogadito voluntário não é livre, pois o seu desejo para usar a droga será efetivo independente de se ou não ele quer que este desejo constitua a sua vontade. Mas quando ele usa a droga, ele a usa livremente e da sua própria vontade livre (FRANKFURT, 1971, p.19).

O desejo de usar a droga do drogadito voluntário é efetivo porque ele é psicologicamente viciado, mas também é efetivo porque ele quer que assim seja. A vontade do drogadito estaria fora do seu controle, “mas pelo o seu desejo de segunda ordem que o seu desejo pela droga deveria ser efetivo, ele fez essa vontade a sua própria” (FRANKFURT, 1971, p.20). Sendo assim, ele poderia ser moralmente responsável por usar a droga, pois o seu desejo não é efetivo apenas porque ele é viciado.

A suposição de que uma pessoa é moralmente responsável pelo o que ela fez não implicaria que ela estava em uma posição de ter seja qual for a vontade que ela desejava. Essa suposição “implicaria que a pessoa fez o que ela fez livremente, ou que ela fez da sua própria vontade livre” (FRANKFURT, 1971, p.19). Seria um erro, segundo Frankfurt, acreditar que alguém age livremente unicamente quando é livre para fazer o que deseja ou que age da sua própria vontade livre apenas se a sua vontade fosse livre.

Vejamos a seguinte situação, a qual ilustraria a estrutura da vontade do drogadito voluntário:

Suponhamos que uma pessoa tenha feito o que ela desejou fazer, que ela o fez porque ela desejou fazê-lo, e que a vontade pela a qual ela foi movida quando ela agiu era a sua vontade porque era a vontade que ela desejou. Assim, ela agiu livremente e da sua própria vontade livre. Mesmo supondo que ela poderia ter agido de outro modo, ela não teria feito de outro modo; e mesmo supondo que ela poderia ter tido uma diferente vontade, ela não teria desejado que a sua vontade fosse diferente do que ela foi (FRANKFURT, 1971, p.19).

Visto que a vontade que moveu o agente quando ele agiu era a sua vontade porque ele queria que fosse, não seria correto dizer que sua vontade foi forçada ou que ele é um mero espectador das suas ações. Nessas condições, seria irrelevante para a avaliação da sua responsabilidade moral questionar se as alternativas que ele não optou estavam, na realidade, disponíveis. Para Frankfurt, como vimos no segundo capítulo, a liberdade enquanto poder de fazer escolhas entre alternativas não seria condição para atribuição de responsabilidade moral. Para que o agente seja moralmente responsável pelas suas ações, nessa perspectiva, deveria haver um tipo de harmonia entre os estados mentais do agente e as suas ações. Além disso, as ações deveriam expressar o caráter do agente.

A noção de responsabilidade de Frankfurt tem sido retratada na literatura como não histórica, pois ela identificaria como relevante para a questão da responsabilidade moral unicamente os estados mentais do agente no momento da ação. Apesar dessa explicação hierárquica ser considerada um enriquecimento em relação à interpretação compatibilista clássica de autodeterminação, por ir além dos impedimentos externos da vontade, o seu caráter não histórico tem sido também fonte de críticas. Don Locke foi quem desenvolveu uma objeção à teoria hierárquica, o que me parece revelar uma dificuldade comum a todas as teorias compatibilistas⁴⁸. O ponto poderia ser resumido do seguinte modo: imagine que um neurocientista deletasse a estrutura psicológica (desejos, crenças e valores) do agente e implantasse uma nova no lugar. Isso implicaria afirmar que as volições de ordem superior são meramente produtos de lavagem cerebral, manipulação ou condicionantes que seriam incompatíveis com a autonomia do agente. Desse modo, teorias hierárquicas, como a de Frankfurt, teriam como consequência a responsabilização de agentes que são manipulados, o que, segue a objeção, não é uma conclusão razoável.

O subcapítulo abaixo discutirá o acréscimo de um conceito que pretende evitar essa objeção.

3.4 Mecanismo do próprio agente

A sensibilidade às razões do tipo apropriado estaria entre as condições para legitimamente responsabilizar o agente pelo seu comportamento. Entretanto, esse critério, embora necessário, não parece suficiente uma vez que consideramos que o tipo relevante de sensibilidade às razões poderia ser induzido pela manipulação, hipnose, lavagem cerebral ou outros fatores que tornariam inapropriadas as atribuições de responsabilidade moral. A teoria

⁴⁸ Ver Locke (1974).

de Frankfurt estaria sujeita a essa objeção na medida em que o agente seria moralmente responsável pela sua ação se houvesse uma conformidade entre a volição de segunda ordem e a sua vontade, isto, é, o que importaria para a avaliação moral de responsabilidade seria unicamente a configuração interna dos seus estados mentais no momento da ação.

Acredito que o elemento distintivo da tentativa de Fischer e Ravizza de reafirmar a tese de que há uma compatibilidade entre determinismo e responsabilidade é a apresentação de um segundo elemento constituinte do controle de direcionamento, qual seja, o mecanismo resultante na ação deveria ser do próprio agente⁴⁹.

O seguinte exemplo ilustra a ideia de que não seriam apenas as configurações internas dos estados mentais que importariam para as atribuições de responsabilidade moral. Imagine um caso em que um agente age livremente e exerce o controle de direcionamento no tempo anterior, mas é incapaz de agir livremente no tempo posterior. São comuns histórias nos jornais, na televisão ou na internet, de um indivíduo que dirige embriagado. Recentemente um padre, de 53 anos, embriagou-se com vinho, durante a missa que celebrou, e enquanto voltava para casa, atropelou e matou um funcionário de uma lanchonete no interior de São Paulo. Intuitivamente nós o responsabilizamos porque partimos do pressuposto de que ele poderia ter previsto que ao beber ele poderia causar uma tragédia desse tipo.

Entretanto, nossas intuições concernentes à sua responsabilidade certamente mudariam se nós descobríssemos que o álcool foi injetado nele contra a sua vontade e que ele foi posto atrás do volante nessas condições. Por isso que a história de uma ação seria crucial para saber se o mecanismo resultante na ação seria do próprio agente (FISCHER; RAVIZZA, 1998, p. 194-202). A ação do motorista embriagado não resultaria do seu próprio mecanismo se fosse verdadeira a hipótese de que ele foi embriagado contra a sua vontade⁵⁰.

Michael Slote apresenta um exemplo, que me parece esclarecer que certos tipos de histórias resultam em mecanismos que não seriam, em um sentido apropriado, do próprio agente:

Robert, que está indeciso entre dois desejos de primeira ordem conflitantes X e Y, recebe a visita de um hipnotizador, que decide “resolver” o seu problema ao hipnotizá-lo e induzi-lo a uma volição de segunda ordem em favor de X;

⁴⁹ Esse segundo componente do controle de direcionamento é uma inovação na teoria e é apresentado em *Responsibility and Control: A Theory of Moral Responsibility* (1998).

⁵⁰ De igual modo, dois indivíduos podem se comportar igualmente no que diz respeito ao vício por um determinado tipo de droga. Porém, a distinção no comportamento poderia ser vislumbrada se um deles foi forçado a usar a droga e o outro desenvolveu o mesmo tipo de vício como resultado das suas próprias escolhas e ações, sobre as quais ele tem controle (FISCHER; RAVIZZA, 1998, p. 196).

como resultado de ter essa volição de segunda ordem, Robert, então, age para satisfazer X, nunca suspeitando que a sua decisão foi induzida pelo hipnotizador (SLOTE, 1980, p. 149).

Robert não estaria agindo a partir da sua própria vontade livre. Portanto, o mecanismo resultante na ação não seria propriamente seu. Sendo assim, ele não poderia ser moralmente responsável.

Uma noção histórica de responsabilidade moral estaria conectada à ideia de que o passado deveria conter um processo de aceitação de responsabilidade [*taking responsibility*] para fazer o mecanismo resultante na ação ser do próprio agente. As seguintes condições deveriam ser satisfeitas para que o agente aceite responsabilidade [*take responsibility*]: (i) o agente deveria ver a si mesmo como a origem das suas escolhas e ações (reconhecer as escolhas e as ações como tendo resultado no mundo); (ii) o agente deveria ver a si mesmo como um alvo justo das atitudes reativas, entendidas como expressões morais de louvor ou censura⁵¹. A terceira condição exigiria (iii) que os dois requisitos anteriores, (i) e (ii), sejam adequadamente baseados na evidência disponível ao agente⁵².

Assumir responsabilidade não significaria simplesmente dizer que “alguém assume responsabilidade pelas suas ações” (FISCHER, 2000a, p. 389). Envolve ter certos tipos de crenças sobre si mesmo:

se alguém não vê a si mesmo como responsável, então não é responsável. Agora não é uma questão do que você diz. Você poderia sempre ‘dizer’ que você não é responsável. Mas é uma questão de quais são realmente as suas crenças. Eu diria isso sobre o caso de Clara e Ivan. Clara e Ivan movem os seus corpos no mesmo modo como resultado de processos cerebrais de tipos idênticos. Mas um deles acredita que é responsável e o outro acredita que não

⁵¹ Fischer adota uma abordagem subjetivista para a responsabilidade moral. Para ser moralmente responsável, o agente deveria ver a si mesmo como responsável. Uma notável consequência dessa abordagem subjetivista diz respeito à segunda condição do processo de atribuir responsabilidade a si mesmo. Fischer distingue entre indivíduos não reflexivos e reflexivos. A maior parte dos indivíduos pertence ao primeiro grupo. Seria exigido que eles formem um juízo sobre a justiça das nossas práticas sociais. Para o outro grupo, o juízo social não seria suficiente. Os indivíduos reflexivos, tais como os filósofos que se engajam na reflexão metafísica sobre a relação entre determinismo causal e a justiça da aplicação das atitudes reativas, necessitariam formar “algum tipo de juízo metafísico para o efeito que, todas as coisas consideradas, seria apropriado sujeitar certos indivíduos – incluindo a si mesmo – às atitudes reativas em certos contextos” (FISCHER; RAVIZZA, 1998, p. 226).

⁵² Para motivar essa última condição, é desenvolvida uma metáfora da responsabilidade moral como um tipo de conversa moral. Aplicando essa metáfora, “quando uma pessoa assume responsabilidade, ela está solicitando à comunidade moral para reconhecê-la como um participante legítimo nessa conversa moral. Quando - mas unicamente quando - essa demanda surge em resposta à sua evidência (...) indica que a pessoa é um participante apropriado na conversa que ela está solicitando para envolver-se” (FISCHER; RAVIZZA, 1998, p. 214).

é. Existe uma diferença na responsabilidade moral precisamente porque se você genuinamente acredita que você não é responsável, então você não é responsável (FISCHER, 2000b, p. 406).

O filme “Sob o Domínio do Mal”⁵³ talvez seja um bom exemplo para explicar por que o mecanismo sensível às razões não seria suficiente para atribuições de responsabilidade moral. O filme retrata a história de um grupo de soldados que durante a guerra do Golfo foi sequestrado pelo inimigo. Depois de alguns anos, um dos soldados, já de volta à sua cidade natal, começa a se lembrar do processo de lavagem cerebral pelo qual ele e os outros membros do seu grupo passaram enquanto estavam presos. O processo consistia na aplicação de hipnose fazendo com que ordens fossem obedecidas sem qualquer contestação. Uma dessas ordens consistia em matar (com arma ou asfixia) outro soldado do seu próprio grupo. O ponto interessante do filme é que um dos personagens, o qual foi manipulado, acreditava que as ações que ele praticava não eram “suas”, ou seja, ele não via a si mesmo como a origem das suas ações. Poderia se dizer que ele não assume responsabilidade pelo tipo de mecanismo resultante da sua ação. Dito de outro modo, ele não estaria assumindo responsabilidade pelo raciocínio prático decorrente da hipnose.

Outro exemplo ilustra que a história do comportamento seria importante para avaliar se o agente pode ser legitimamente moralmente responsável. Esse é o caso da Paulina, a paralisada:

Paulina está visitando a Flórida pela primeira vez. Ela leva o seu filho de dois anos de idade, Paul, para caminhar no parque próximo a Tallahassee. Eles eventualmente encontram uma pequena clareira e decidem sentar para almoçar. Depois de um tempo, Paul se levanta e começa a andar pela volta com a comida nas suas mãos, no seu rosto, e nas suas roupas. Paulina, sendo uma mãe responsável, observa Paul e o alerta para não se afastar muito. Perto do fim da clareira, acerca de seis metros, Paul inclina-se para pegar uma pedra que chama a sua atenção. Paulina o observa fazer isso, e nesse momento percebe que há um jacaré encarando Paul a uma distância de três metros. Paulina é paralisada pela surpresa e por algum grau de medo. Ela sente uma crescente sensação de pânico, mas permanece congelada enquanto o jacaré se move em direção a Paul. Infelizmente o jacaré de maneira surpreendentemente faz um movimento rápido e pega o pobre Paul. Só então Paulina finalmente descongela, e então grita. Paulina é responsável por não ter feito mais para salvar Paul? (VARGAS, 2005, p. 272).

Esse caso é apresentado por Manuel Vargas para defender uma tensão entre a condição epistêmica de responsabilidade e a ideia de rastreamento [*tracing*], que é a noção histórica de

⁵³ Um filme norte-americano de 2004, dirigido por Jonathan Demme e distribuído pela Paramount Pictures, cujo título original é “The Manchurian Candidate”.

responsabilidade. Se esse ponto estiver correto, então haveria uma dificuldade concernente às teorias que apelam para a noção histórica de responsabilidade ou para a condição de conhecimento, pois essas seriam condições inadequadas (VARGAS, p. 2005, p. 271).

Vargas pensa que sobre esse caso nós teríamos as seguintes três reações: (i) Paulina seria moralmente responsável por não ter feito mais para salvar o seu filho e (ii) que ela não satisfaria a condição de controle por causa da interação complexa da surpresa e o medo de répteis. Parece razoável que (i) e (ii) nos levariam a invocar a condição de rastreio para explicar a responsabilidade moral de Paulina. Porém, nossa terceira reação seria que (iii) não existe tempo anterior plausível quando Paulina poderia razoavelmente ter previsto que esses fatos interagiriam de tal modo que a deixariam temporariamente paralisada. Ela cresceu sabendo que tinha medo de répteis e sabia que se ela realmente quisesse havia meios disponíveis (como a terapia) para superar o seu medo. Mas antes do acontecimento ela não tinha razão para acreditar que esse medo afetaria sua capacidade para salvar o seu filho (naquela época ela não tinha nem mesmo planos para ter filhos). Diante dessas hipóteses, a condição epistêmica nos impediria de rastrear adequadamente a responsabilidade moral de Paulina a algum momento anterior.

Parece claro que, dado como Vargas descreve o caso, Paulina não é de fato moralmente responsável por não ter feito mais para salvar o seu filho. Nós não a desculparíamos da responsabilidade se descobríssemos que ela estava *literalmente paralisada* pelo seu medo e surpresa? Se o caso é descrito com uma ênfase sobre o papel irresistível desempenhado pelo medo e surpresa, então estaremos inclinados a julgá-la como não responsável (FISCHER; TOGNAZZINI, 2012, p. 221).

Vamos imaginar que Paulina é moralmente responsável pela omissão relevante e que ela falha em satisfazer a condição de controle em relação a essa omissão. Nesse caso, nós precisaríamos invocar a condição de rastreio. Qual seria o momento anterior adequado para rastrear? Houve, sem dúvida, um momento em algum ponto anterior à sua primeira viagem à Florida quando ela poderia ter se instruído sobre os potenciais perigos dos animais selvagens. Conforme Vargas afirma em seu artigo, “era a primeira vez que ela estava frequentando Tallahassee e porque nunca se preocupou a pensar sobre isso, ela não tinha nenhuma expectativa sobre a fauna nativa” (VARGAS, 2005, p. 281). O ponto é que mesmo estando em poder de Paulina procurar informações sobre a fauna, ela simplesmente não poderia ter previsto que a interação de fobias e disposições humanas causaria a tragédia. Dito de outro modo, talvez ela seja consciente dos seus traços de caráter, porém é inconsciente dos efeitos da interação desses traços.

Os defensores da condição histórica de responsabilidade não estão satisfeitos com essa resposta. Acredito que não ter se informado sobre o local que estava visitando certamente foi negligência de Paulina. Tivessem eles viajado para o deserto da Califórnia e seu filho tivesse sido mordido por uma cobra, situação da qual ela não esperava se deparar, não contaria como alguma forma de negligência? Na medida em que Paulina deveria saber sobre os perigos da fauna da Flórida, Tognazzini e Fischer pensam que seria razoável supor que ela é em certo grau moralmente responsável por não ter feito mais para salvar o seu filho. Mesmo que ela não tenha satisfeito a condição de controle no momento do ocorrido, parece que houve algum momento no passado que ela poderia e deveria ter tomado conhecimento dos perigos do local que estava prestes a visitar (FISCHER; TOGNAZZINI, 2012, p. 222). Além disso, para responsabilidade não haveria a necessidade do agente prever um resultado específico. Basta que ele preveja o resultado amplamente, isto é, que as suas ações, omissões ou decisões terão determinadas consequências.

Angela Smith apresenta exemplos em que nós estaríamos inclinados a responsabilizar o agente por alguma atitude sendo essa atitude não dependente de alguma escolha nem rastreável a alguma escolha no passado. Smith afirma que nossas intuições do senso comum não exigem um critério volitivo de responsabilidade, mas um critério racionalista, pois “o tipo de atividade implicado pelas nossas práticas morais não é a atividade da escolha, mas a atividade do juízo avaliativo” (SMITH, 2005, p. 237).

Para a explicação volitiva de responsabilidade, a escolha, a decisão e o controle voluntário são condições necessárias para a responsabilidade pelas atitudes bem como pelas ações. De acordo com essa teoria o que faz uma atitude atribuível a uma pessoa para propósitos de avaliação moral é que essa atitude esteja conectada de algum modo às escolhas do agente. Ele deve ter feito as escolhas no passado que levaram ao desenvolvimento das atitudes atuais; ele necessita ter feito a escolha no presente para se identificar com a atitude em questão; e ainda precisa ter a capacidade para modificar as suas atitudes por meio das escolhas que ele faz no futuro. Embora as teorias compartilhem a ideia de que a escolha é condição para responsabilidade por uma atitude, há uma grande diferença em como essa conexão é estabelecida por diferentes teorias. Por exemplo, na visão da “escolha anterior”, o que seria essencial para atribuições de responsabilidade é que nós sejamos capazes de rastrear o desenvolvimento de uma atitude às próprias escolhas ou decisões anteriores da pessoa. Talvez alguém faça uma decisão deliberada de começar a ler propaganda neonazista e como resultado desenvolva atitudes antissemitas. Nessa perspectiva, uma pessoa seria responsável por essas atitudes na medida em que elas podem ser rastreadas causalmente às suas próprias escolhas

passadas. Já na visão da identificação (ou não histórica), o que seria essencial para atribuições de responsabilidade é que a pessoa mesma se identificasse com a atitude em questão. Nessa visão não importaria como a atitude surgiu (se foi ou não resultado das suas próprias escolhas anteriores). Finalmente, na perspectiva do controle voluntário, o que seria essencial para atribuições de responsabilidade é que a pessoa tenha a capacidade para controlar ou modificar suas atitudes por meio de escolhas ou decisões futuras. Por exemplo, para determinar se uma pessoa é responsável pelas suas atitudes antissemitas é preciso questionar se ela tem a capacidade para eliminá-las por meio dos seus próprios esforços voluntários. Nessa visão também não importa como essa atitude surgiu (SMITH, 2005, p. 238-240).

Vejamos um dos exemplos de Smith sobre isso:

Ano passado eu esqueci o aniversário de uma amiga. Poucos dias depois do fato, eu percebi que essa importante data tinha ocorrido sem que eu tivesse enviado um cartão ou ligado para ela. Eu fiquei atormentada. Que tipo de amiga poderia esquecer tal coisa? Dentro de minutos eu estava no telefone com ela, reconhecendo a minha falha e oferecendo as minhas desculpas. Mas qual era exatamente a natureza da minha falha nesse caso? Afinal de contas, eu não conscientemente *escolhi* esquecer esse dia especial ou deliberadamente *decidi* ignorá-lo. Eu não *pretendi* ferir os sentimentos da minha amiga ou mesmo *previ* que a minha conduta teria esse efeito. Eu apenas esqueci. Não me ocorreu. Eu falhei em observar. E ainda, apesar da aparente involuntariedade dessa falha, não havia dúvida em qualquer uma de nossas mentes de que eu era, de fato, responsável. Embora minha amiga tenha perdoado rapidamente minha negligência e dito se tratar de algo trivial e sem importância, o ato do perdão é simplesmente um modo de rejeitar certas respostas críticas que se reconhece que seriam, em princípio, justificadas (SMITH, 2005, p. 236).

Esse caso representaria uma tensão no nosso pensamento ordinário sobre as condições da responsabilidade moral. Provavelmente muitos de nós diriam que a escolha e o controle voluntário são condições da avaliação moral legítima. Porém, o objetivo de Smith é oferecer “uma explicação alternativa de responsabilidade que faz melhor sentido das nossas intuições ordinárias sobre as condições pelas quais uma pessoa é responsável pelas suas atitudes” (SMITH, 2005, p. 240). Desse modo, o exemplo citado ilustra que nós regularmente responsabilizamos as pessoas por coisas que não parecem envolver uma escolha ou decisão consciente (SMITH, 2005, p. 236). Nós estaríamos inclinados a responsabilizar uma pessoa que esquece o aniversário de um amigo porque parece razoável inferir que essa pessoa não se importa com esse amigo tanto como deveria. Ademais, se o seu esquecimento fosse resultado da hipnose ou manipulação direta do cérebro nós a desculparíamos da responsabilidade não

porque ela careceu de escolha, mas porque seu esquecimento não parece retratar seus juízos avaliativos⁵⁴.

Tognazzini e Fischer concordam com Smith de que parte da explicação da responsabilidade moral pelo comportamento e atitude envolveria a manifestação dos juízos avaliativos. Porém, essa condição não seria suficiente. Com respeito ao caso citado por Smith, por exemplo, é plausível supor que a responsabilidade moral pelo esquecimento remonta à alguma ação ou omissão prévia. Digamos que eu tenha falhado de escolher várias coisas que poderiam ter me ajudado a lembrar do aniversário, como, por exemplo, ter me omitido de escolher colocar a data do aniversário na agenda. Parte do que seria um bom amigo é adotar essa medida para amenizar a probabilidade de eu esquecer uma data importante. Se eu não tomei essa precaução e então esqueci, eu posso ser de maneira legítima moralmente responsável pelo esquecimento, pois estava no meu controle fazer alguma coisa para que esse esquecimento fosse menos provável (FISCHER; TOGNAZZINI, 2012, p. 229).

No capítulo a seguir tratarei do desafio imposto por Robert Kane a toda e qualquer teoria compatibilista, que é o argumento da manipulação. Também procurarei delinear a réplica que Alfred Mele apresenta a esse desafio.

⁵⁴ O critério racionalista que Smith propõe pode ser compreendido da seguinte forma. Quando louvamos ou criticamos alguém por uma atitude, parece que nós estamos respondendo a alguma coisa sobre o conteúdo dessa atitude e não aos fatos sobre sua origem nas escolhas voluntárias anteriores do agente. Nossas atitudes estariam representando o que nós julgamos ser de valor, importância ou significância. Se isso estiver correto, então seria equivocado tentar explicar a responsabilidade pelas atitudes em termos de conexão com as escolhas voluntárias passadas e futuras. Para uma pessoa ser responsável por uma atitude, adotando o critério racionalista, a pessoa deveria ter a capacidade para corrigir ou modificar a atitude através dos próprios processos de reflexão racional (SMITH, 2005, p. 251, 256).

4 TEORIAS ALTERNATIVAS

No primeiro subcapítulo deste capítulo tratarei do desafio imposto pelo libertista Robert Kane a qualquer teoria compatibilista, incluindo a de Fischer. Para Kane, nenhuma explicação compatibilista pode adequadamente lidar com o problema da possibilidade de certos tipos de manipulação. No segundo subcapítulo, eu apresento uma crítica à teoria de Kane, que é a problemática da sorte, e, em seguida, a sua réplica. No terceiro subcapítulo será feita uma distinção entre as visões de Kane e de Fischer. No quarto subcapítulo eu apresento a posição internista de Frankfurt, apontando a sua proposta de solução para o dilema levantado por Kane. No quinto subcapítulo, eu trato da posição de Mele. Sua posição é interessante por apresentar um conjunto de condições compatibilistas na tentativa de responder a Kane. Mele discute alguns casos de manipulação e argumenta que certas condições compatibilistas fornecem uma réplica ao desafio colocado por Kane. Por fim, no sexto e último subcapítulo eu apresento o desacordo entre a posição de Mele e de Fischer, de modo a evidenciar até que ponto suas soluções ao problema são consistentes.

4.1 A posição libertista de Robert Kane

Kane apresenta duas condições para a defesa da tese incompatibilista. A primeira seria a exigência de possibilidades alternativas (PA) e a segunda seria a necessidade da condição que ele denomina de responsabilidade última (RU). A condição de RU não exige que nós poderíamos ter feito de outro modo para *todo* ato feito de nossa própria vontade livre. Porém, é necessário agir diferentemente apenas quando nós formamos o nosso caráter, ou seja, quando realizamos atos passados na nossa história de vida. Essas ações são o que Kane denomina de autoformadoras. Segundo Kane, “nós agimos frequentemente a partir de uma vontade já formada, mas é a “nossa própria vontade livre” em virtude do fato de que nós a formamos por outras escolhas ou ações no passado (ações autoformadoras) pelas quais nós poderíamos ter feito de outro modo” (KANE, 2013, p. 260). Para ser verdadeiramente responsável pela escolha, sendo essa resultante do seu caráter, o agente deve ser responsável pelas escolhas e ações voluntariamente realizadas no passado, as quais formaram o seu presente caráter e motivos.

A condição de PA recebeu várias objeções, como foi visto no capítulo 2, por meio dos contraexemplos de Frankfurt e dos seus seguidores. Porém, incompatibilistas não ficaram convencidos pela réplica de seus oponentes e introduziram mais um passo na argumentação. A condição de RU é a explicação de por que a vontade livre seria, de fato, incompatível com o determinismo. Pelo menos essa é a pretensão de Kane. Supondo a imutabilidade do caráter, e

que as escolhas e as ações são resultantes desse caráter, então para que o agente seja responsável pela escolha ou ação, ele deve ser responsável pelas ações anteriores que formam o seu presente caráter. Há um detalhe aqui, que mais adiante será melhor explicado. A princípio Kane esclarece que *algumas* ações na nossa história de vida (isto é, as ações autoformadoras) não teriam causas suficientes e, portanto, seriam indeterminadas⁵⁵ (KANE, 2013, p. 261).

Obviamente que esse novo caminho para a incompatibilidade da vontade livre com o determinismo levanta algum questionamento por parte dos críticos. Por exemplo, como explicar a possibilidade de ações livres e responsáveis se elas carecem de causas suficientes? Kane oferece a seguinte resposta: indeterminismo não teria de estar envolvido em todos os atos feitos da “nossa própria vontade livre”, pelos quais nós seríamos verdadeiramente responsáveis. Nesse sentido, unicamente as ações autoformadoras deveriam ser indeterminadas (KANE, 2013, p. 262).

Kane reconhece que essas ações autoformadoras ocorrem em momentos difíceis das nossas vidas, quando estamos divididos entre visões conflitantes do que deveríamos fazer ou nos tornar. É provável que estejamos indecisos, por exemplo, entre agir por razões morais ou por motivos autointeressados. Porém, quando nos deparamos com motivações conflitantes devemos fazer um esforço para superar a tentação de fazer alguma coisa diferente que nós também fortemente queremos fazer (KANE, 2013, p. 262). Por exemplo, Kane apresenta um caso sobre uma mulher executiva para explicar em que sentido a vontade estaria dividida por motivos conflitantes. O caso é relatado da seguinte forma. Suponha que a executiva esteja a caminho de uma importante reunião quando se depara com um assalto na rua. Surge um impulso interno. Por um lado, está a sua consciência moral, que lhe diz para parar e chamar por ajuda. Por outro lado, estão as suas ambições de carreira, que lhe dizem que ela não deveria perder a reunião. Ela teria de fazer um esforço para superar a tentação de fazer a escolha contrária à moralidade. O que explicaria a sua escolha de ligar para a delegacia, por exemplo, e se atrasar para a reunião, seria o êxito desse seu esforço. Porém, caso ela escolhesse não ajudar a vítima do assalto, a explicação seria que ela permitiu que seu esforço fracassasse. Poderia ser dito que ela tanto desejou superar a tentação de fazer a escolha contrária à moralidade como também desejou não fazer isso por diferentes razões (KANE, 2013, p. 263). Esse caso representa a ideia de que o agente tenta fazer cada uma das escolhas simultaneamente. Além disso, o agente seria moralmente responsável pela decisão que ele tomasse desde que “ele

⁵⁵ Essa foi uma estratégia para evitar o problema do regresso ao infinito. Eu voltarei a esse ponto no final deste subcapítulo, quando abordar a crítica que Galen Strawson faz à posição de Kane.

endosse o resultado como alguma coisa que ele estava tentando e querendo fazer” (KANE, 1999, p. 233).

É preciso esclarecer qual é o ponto da argumentação de Kane. Kane é um libertista e um incompatibilista. Ele defende a visão libertista de acordo com a qual mesmo que nossas ações sejam *determinadas* pelo nosso caráter, essas ações *podem ser livres* se em algum momento no passado nós livremente criamos o nosso próprio caráter. Além disso, ele assume uma posição incompatibilista, pois vontade livre só faria sentido em um mundo indeterminista. Kane acredita que a condição de PA não é suficiente para estabelecer a tese incompatibilista. É necessário a condição de RU. RU exige que algumas das ações, na nossa história de vida, ou seja, as ações autoformadoras, careçam de causas, portanto, sejam indeterminadas.

Na literatura sobre vontade livre e responsabilidade moral há casos sobre certos tipos de manipulação por outros agentes, como lavagem cerebral, e condicionantes comportamentais, que têm desempenhado um papel significativo não apenas nas críticas às teorias hierárquicas como a de Frankfurt, por exemplo, mas também nas discussões sobre a questão da compatibilidade ou incompatibilidade do determinismo com a liberdade e a responsabilidade moral. Eu desenvolverei esse aspecto mais adiante.

Partindo de exemplos de controle comportamental presentes em obras como *Brave New World*, que retratam uma sociedade futurista, é possível estabelecer uma distinção entre duas espécies de controle que um agente pode exercer sobre o outro agente. Partirei da distinção feita por Kane. O controle pode ser forçado (ou coagido) ou não forçado. Os agentes podem ser forçados a fazer alguma coisa contra a sua vontade. Por exemplo, eles poderiam ser trancados em um quarto e impedidos de fazer o que eles quisessem fazer. Ou eles poderiam ser ameaçados com uma arma ou com uma punição caso não fizessem o que foi ordenado. Os controladores criariam restrições ou impedimentos que ameaçariam a vontade daqueles que estivessem sob os seus comandos, impedindo outros agentes de fazer o que eles quisessem fazer (KANE, 1998b, p. 64). O segundo tipo de controle seria a manipulação secreta. O manipulador controlaria a vontade do agente fazendo com que ele realizasse a ação desejada por ele. Nesse caso o agente não se sentiria ameaçado (como acontece quando ele é forçado ou coagido), nem ele estaria consciente de estar sendo manipulado. Ele agiria de acordo com seus próprios desejos e intenções (KANE, 1998b, p. 65).

Kane conjectura que os compatibilistas estariam sujeitos à uma árdua tarefa. Eles deveriam enfatizar a distinção entre o controle manipulativo por outros agentes e a mera determinação causal por causas naturais, mais especificamente, “eles devem argumentar que enquanto o controle manipulativo elimina a liberdade em um sentido significativo, mera

determinação por causas naturais não” (KANE, 1998b, p. 68). O problema seria localizar a diferença relevante entre os dois. O encargo consistiria em apresentar razões convincentes de que a determinação causal, diferentemente do controle manipulativo, não seria um obstáculo para a liberdade e responsabilidade moral.

Para compreender melhor isso, é útil contrastar a teoria de Frankfurt e a de Kane. Frankfurt, por exemplo, apresentou uma explicação de liberdade da vontade. A liberdade da vontade é exercida quando é assegurada a conformidade da vontade do agente com as suas volições de segunda ordem. Porém, críticos de Frankfurt apontam que ele não trata especificamente das origens da vontade do agente, ou das volições de ordem superior. Apesar da explicação de Frankfurt ser um avanço em relação às concepções clássicas, pois essas estariam mais preocupadas com os impedimentos *externos* à vontade do que propriamente com os impedimentos *internos*, ela ainda estaria sujeita à seguinte crítica: “as preferências de ordem superior podem ser resultado de lavagem cerebral, ou condicionamento do tipo que é incompatível com autonomia” (WATSON, 1987, p. 148).

Kane, ao se concentrar em RU, acredita que o controle manipulativo e a mera determinação teriam mais semelhanças do que diferenças. Em ambos os cenários, “as causas últimas dos fins ou propósitos dos agentes controlados estão fora dos agentes, não neles” (KANE, 1998b, p. 71). Em outras palavras, ou a causa estaria nos manipuladores ou nas leis da natureza (lembrando o argumento da consequência, visto no capítulo 2, de acordo com o qual se o determinismo causal fosse verdadeiro então nossas ações seriam consequências das leis da natureza e do passado). É nesse sentido que a preocupação sobre as origens da nossa vontade conduziria, na perspectiva de Kane, à condição de RU. A origem última da nossa vontade deveria “estar em nós e não fora de nós (KANE, 1998b, p. 71), supondo que nós sejamos verdadeiramente responsáveis pelo o que fazemos e pelo o que somos (KANE, 1998b, p. 70).

Quando compatibilistas e incompatibilistas tratam de casos de manipulação e de determinação causal, eles os analisam a partir dos seus próprios pontos de vistas. Os compatibilistas costumam estabelecer uma distinção entre “controle” e mera determinação causal. Daniel Dennett, por exemplo, afirma que se nós queremos progredir nessa área nós deveríamos evitar confusão a respeito da palavra “controle”. Assim como os compatibilistas clássicos afirmam que uma coisa é impedimento (externo) da vontade e outra coisa é causação, Dennett faz uma diferenciação entre controle e mera determinação. Como ele explica, “a natureza, não sendo um agente, não nos controla” (DENNETT, 1984, p. 61) e “quando um agente tenta controlar outro agente, as coisas ficam complicadas” (DENNETT, 1984, p. 57).

O que explicaria as ações do agente em um mundo determinista? Haveria uma razão para a ocorrência de cada evento ou nos decretos do destino (no caso do fatalismo), ou nas causas antecedentes mais as leis da natureza (no caso do determinismo causal ou científico), ou nos atos predestinados de Deus (no caso do fatalismo teológico). Para Kane, se o agente é verdadeiramente responsável pelo o que ele faz e pelo o que é, então as origens das suas ações estariam “no agente” e não na natureza ou em outros agentes (KANE, 1998b, p. 70-71). Já no caso da mera determinação, se as condições antecedentes e as leis da natureza, ou atos predestinados de Deus, por exemplo, fornecem uma razão ou uma explicação para uma ação, então, para ser verdadeiramente responsável por essa ação, o agente deveria ser responsável por essas condições explicativas, o que seria absurdo. Possuir controle último sobre uma escolha ou ação, na perspectiva de Kane, exigiria que essa escolha ou ação não fosse explicada por fatores externos ao agente como as leis da natureza e as leis do passado (KANE, 1998b, p. 74).

Obviamente que nós não seríamos verdadeiramente responsáveis pela existência dos decretos do destino ou da predestinação divina. Porém, poderíamos ser responsáveis pelos estados de caráter em virtude de atos anteriores (KANE, 1998b, p. 74). Esse ponto já foi mencionado antes. Para ser responsável pela ação decorrente do caráter, o agente deveria ser responsável pelos atos anteriores que formaram o presente caráter. E é justamente aqui que para Kane estaria a necessidade do incompatibilismo. Seria necessário que alguns atos em nossa história de vida fossem indeterminados para evitar o problema do regresso ao infinito. Se vontade livre implica RU, então as ações autoformadoras deveriam ser indeterminadas e o determinismo deveria ser falso. Nesse sentido existiriam dois elementos para o incompatibilismo:

Quando nós consideramos as implicações das PA apenas para o determinismo, nós tendemos a focar sobre noções de “necessidade”, “possibilidade”, “poder”, “capacidade”, “pode”, e “poderia ter feito de outro modo”. RU também envolve essas noções (visto que implica PA para algumas ações livres). Mas, como Marta Klein observou, um interesse na responsabilidade última adiciona um diferente conjunto de preocupações sobre as “origens”, “motivos”, “razões” e “explicações” das ações e eventos, isto é, preocupações de onde eles vieram, o que os produziu, e quem foi responsável por eles. É por se concentrar em tais coisas como as origens e reponsabilidade, eu argumentaria, e não apenas sobre as possibilidades alternativas, que se chega ao incompatibilismo (KANE, 1998b, p. 75).

Tanto a condição de PA como RU seriam necessárias para a responsabilidade moral. Nesse caso, a teoria forneceria elementos tanto para a ação livre como para a vontade livre.

Vontade livre não estaria relacionada apenas com ações livres, mas também com a formação da vontade.

Em resumo, o agente seria responsável moralmente pelo o que faz, mesmo que sua ação tivesse sido determinada pelo seu caráter. Porém, seria necessário que o agente pudesse ter agido diferentemente na formação do seu presente caráter. Acredito que a tese incompatibilista-libertista de Kane estaria sujeita à seguinte objeção feita Galen Strawson:

1. A ação que o agente pratica é resultado do seu caráter.
2. Para ser verdadeiramente moralmente responsável pelo o que o agente fez ele deve ser verdadeiramente responsável pelo o que é.
3. Porém, o agente não pode ser verdadeiramente responsável pelo o que é.
4. Portanto, ele não pode ser verdadeiramente responsável pelo o que fez⁵⁶.

Strawson aponta que como estamos interessados no livre-arbítrio (ação livre), estamos preocupados com as ações que são realizadas por uma razão. Quando um agente age por uma razão, sua ação diz respeito a como ele é mentalmente, ou seja, aos seus valores, desejos e caráter. Para ser verdadeiramente responsável pela ação, ele deveria ser verdadeiramente responsável pelo seu caráter (pelo o que ele é). Mas para ser verdadeiramente responsável pelo o que se é ele deveria ter escolhido, conscientemente, como é. Porém, não se pode dizer que alguém escolheu ser como é, em um modo consciente, a menos que ele estivesse equipado com alguns princípios de escolha, P1 (preferências, valores, ideais e atitudes). Para ser verdadeiramente responsável (no sentido de ter escolhido o modo que é), ele deveria também ter sido responsável por ter estes princípios de escolha, tendo-os escolhido racional e conscientemente. Além disso, para ter escolhido P1 ele deveria ter acesso a outros princípios P2, à luz dos quais ele teria escolhido P1. E assim sucessivamente. O problema é que chegaríamos a um regresso ao infinito. Autodeterminação seria impossível, pois exigiria a realização de uma série infinita de escolhas (STRAWSON, 1994, p. 6-7).

A premissa 3 do argumento é atrativa, pois é inegável que o caráter do agente é, inicialmente, resultado da hereditariedade e experiência anterior, e ele não pode ser responsável por coisas que estão além do seu controle, como características genéticas e hereditárias. Mas o agente não poderia tentar mudar o seu caráter e, portanto, ser responsável por ele e pelas ações decorrentes? A ideia central *não é* que as pessoas não possam mudar o modo que elas são. O

⁵⁶ Esse é o argumento que Galen Strawson apresenta no artigo “The Impossibility of Moral Responsibility”. Para Strawson, o livre-arbítrio é incompatível com o determinismo e com o indeterminismo, o que implicaria que a vontade livre e a responsabilidade seriam incoerentes, pois exigiriam uma concepção de autodeterminação que é logicamente impossível (1994, p.12-14).

problema estaria no fato de que essa mudança de caráter já estaria determinada pelo caráter atual do agente, que é resultado da hereditariedade e experiência prévia. Porém, como observa Strawson, alguém poderia objetar que haveria um ponto negligenciado. Poderia ser argumentado que algumas mudanças no caráter são rastreáveis não apenas à hereditariedade ou à experiência, mas à influência de fatores indeterminísticos e ao acaso. No entanto, Strawson não acredita que essa seja uma boa solução: “é absurdo supor que fatores indeterminísticos ou por acaso, pelos quais alguém não é de forma alguma responsável, possam contribuir de algum modo para ser verdadeiramente responsável pelo caráter” (STRAWSON, 1994, p. 7).

Haveria duas respostas possíveis para essa objeção. A primeira é dada pelos compatibilistas. Divergências específicas à parte, em geral eles acreditam que o agente poderia ser moralmente responsável pelo o que faz desde que a sua ação não seja causada por restrições como impulsos cleptomânicos, neuroses obsessivas, comando pós-hipnótico, ameaças etc. Não haveria aqui exigência de que o agente deveria ser verdadeiramente responsável pelo seu caráter. Portanto, o passo (2) do argumento acima seria falso, na medida em que haveria uma dissociação entre o caráter do agente e a sua ação. Um agente poderia ter responsabilidade compatibilista mesmo se o seu caráter fosse completamente determinado por fatores além do seu controle (STRAWSON, 1994, p. 16). Dessa forma, o compatibilista aceitaria a seguinte proposição: o agente faz o que faz por causa do modo que ele é, e ele não é de nenhum modo verdadeiramente responsável pelo seu caráter (STRAWSON, 1994, p. 17).

A segunda resposta a essa objeção é dada pelo libertista. Por exemplo, Kane argumenta que os agentes possuem vontade livre em um mundo indeterminado: os agentes “têm o poder para fazer escolhas que podem unicamente ser explicadas em termos da sua própria vontade (isto é, caráter, motivos e esforços da vontade), sendo que “ninguém pode ter esse poder em um mundo determinado” (KANE, 1989, p. 254). Kane parece aceitar o passo (2) do argumento acima, pois ele afirma que para sermos moralmente responsáveis pelo o que nós fazemos nós devemos ser verdadeiramente responsáveis por fazer nós mesmos o tipo de pessoas que somos.

Kane apresenta um conceito de liberdade que exige a falsidade do determinismo. O seu incompatibilismo exigiria tanto o poder para agir de outro modo como a noção de responsabilidade última, que envolve o conceito de criação dos próprios fins e propósitos. A hipótese de que as origens dos fins deveriam estar “no próprio agente” e não “nos controladores” é uma tentativa de resposta à crítica que, por exemplo, tem sido feita à teoria de Frankfurt. Ter a vontade que se quer, e ser capaz de agir a partir de volições de segunda ordem, seria apenas parte do que significaria desfrutar de uma vontade livre. Além disso, é preciso

ocupar-se da questão de se nós somos verdadeiramente responsáveis por ter a vontade que temos.

No próximo subcapítulo tratarei das implicações da necessidade do indeterminismo presente na teoria de Kane.

4.2 O problema da sorte para a posição libertista de Kane e a sua resposta às críticas

Em geral, como explorado por Alfred Mele, libertistas corretamente veem o determinismo como sendo uma barreira para a possibilidade de cursos causalmente alternativos de ações, isto é, cursos que sejam consistentes com o passado e as leis da natureza. Em outras palavras, se todo o nosso comportamento é determinado, então há somente um futuro possível. O indeterminismo removeria essa barreira, porém levantaria uma preocupação sobre sorte (MELE, 1998, p. 582).

Conforme Mele, várias posições libertistas, incluindo a de Kane, estariam sujeitas ao problema da sorte. Esse problema é formulado pela imaginação de um mundo possível. Suponha que um agente decide fazer *A* no mundo real, e, ao mesmo tempo, em outro mundo possível com o mesmo passado e as leis da natureza do mundo real, decide fazer *B*. Parece que a diferença entre as decisões do agente nos dois mundos seria apenas uma questão de sorte (MELE, 2006, p. 51). Para compreender melhor qual seria o ponto, talvez o seguinte exemplo possa auxiliar. Aplicando o mesmo exemplo de Kane, pode-se dizer que a executiva (vou chamá-la de Maria₁) acredita, no mundo real, que deve chegar no horário marcado para a reunião. Ela é tentada a se atrasar porque ela quer ajudar a vítima de um assalto. Embora ela tente resistir à tentação, o seu esforço indeterminista falha, e em *t*, um minuto antes da reunião, ela decide chegar atrasada. Em um mundo possível alternativo, com as mesmas leis da natureza e passado até *t*, Maria₂ acredita que deve chegar no horário marcado para a reunião. Ela é tentada a chegar atrasada (pela mesma razão) e tenta resistir à tentação. Diferente do mundo real, aqui o seu esforço indeterminado é exitoso e ela decide, em *t*, chegar no horário e assim o faz. O esforço de Maria₁ no mundo real falha quando ela tenta resistir à tentação, enquanto que o esforço de Maria₂ no mundo possível é bem sucedido (MELE, 1998, p. 583).

O problema é que “se não existe algo sobre os poderes do agente, capacidades, estados da mente, caráter moral que explique essa diferença no resultado, então a diferença é apenas uma questão de sorte” (MELE, 1998, p. 583). Além disso, o fato que os seus esforços são “indeterminados explica porque os esforços podem ter diferentes resultados, mas isso

obviamente não explica (mesmo em um modo indeterminista) por que Maria₁ falhou, enquanto Maria₂ teve êxito” (MELE, 1998, p. 583)⁵⁷.

Se a diferença entre os dois mundos é uma questão de sorte, isso significa que as escolhas, decisões ou ações indeterminadas são arbitrárias, ocorrem por acaso ou são inexplicáveis. A implicação é que escolhas ou ações indeterminadas não seriam livres e, tampouco, responsáveis.

Bruce Waller chama a atenção para o problema da sorte ao criticar a posição de Kane sobre a vontade livre e a responsabilidade moral:

Suponhamos que nós temos duas pessoas idênticas em situações idênticas: cada uma delas tem 10 dólares para gastar, um amigo que está precisando de ajuda, um desejo para ir ao cinema, e suas experiências, desejos, condicionamentos, genes, situações são exatamente muito parecidas. A pessoa A delibera, considera várias razões para cada alternativa, e por acaso, escolhe ir ao cinema. A pessoa B delibera do mesmo modo, e de novo por acaso, exceto que nesse caso ela lembra que um amigo fez um ato de bondade para a comunidade e assim ele merece ser ajudado. Assim B decide fazer o ato de generosidade. Cada um fez uma decisão racional, dado as considerações que ocorrem para cada um; e em cada caso a decisão é feita pela pessoa. Mas existirão bases para distinguir entre A e B, para dizer que um merece censura pela decisão egoísta e o outro merece elogio pela generosidade? Se eles são realmente idênticos e a diferença nos seus atos resulta do acaso, então parece irracional considerar um mais louvável ou (mais censurável) do que o outro (WALLER, 1988, p. 151).

A objeção afirma que se a ação livre e a moral não podem ser compatíveis com o determinismo, tampouco podem ser compatíveis com o indeterminismo. Kane admite que o indeterminismo possa ser um obstáculo para a liberdade⁵⁸. Porém, ele insiste, de modo algum esse obstáculo impossibilitaria a aplicação de noções de responsabilidade moral. Vejamos o caso do assassino. Ele tenta matar o primeiro ministro, mas pode errar o alvo por causa de eventos indeterminados no seu sistema nervoso, como a ocorrência de um tremor no seu braço. O raciocínio de Kane é que, não obstante a presença de eventos indeterminados, se o assassino acerta o alvo e mata o primeiro ministro ele poderia ser moralmente responsável pelo

⁵⁷ Os personagens foram adaptados.

⁵⁸ Kane admite que o indeterminismo poderia diminuir (ou enfraquecer) o controle sobre o que nós estamos tentando fazer, além de ser um obstáculo para a realização de nossos propósitos. Porém, nas ações autoformadoras (como exemplificado pelo caso da executiva), o indeterminismo que estaria diminuindo o seu controle sobre uma certa ação (o ato moral de ajudar uma vítima, por exemplo) origina-se em sua própria vontade, a partir do seu desejo e esforço de fazer o oposto (ir à reunião). Em síntese, o indeterminismo estaria funcionando como uma forma de obstáculo, na forma de resistência no interior da vontade, que tem de ser superado pelo esforço (KANE, 2013, p. 270).

assassinato porque ele voluntariamente e intencionalmente teve êxito em fazer o que ele estava tentando fazer (KANE, 2013, p. 263).

É interessante contrastar o exemplo do assassino e o da executiva. Diferentemente da vontade da executiva, a vontade do assassino não está dividida. Ele quer matar o primeiro ministro, mas também não quer falhar. Se ele falhar, isso será por acaso, e não será voluntário e intencional. A falha seria meramente uma questão de erro ou acidente. Kane acredita que não haveria problema em admitir que o assassino teve sorte ao matar o primeiro ministro. Sorte, nesse sentido, significaria que ele poderia ter errado o alvo. Se o assassino acerta o alvo não faria sentido afirmar que o ato cometido não foi seu, ou que não foi alguma coisa que ele fez. Além disso, tampouco implicaria afirmar que o assassino não foi responsável pelo homicídio.

Para mostrar por que a objeção da sorte também falharia quando se trata da formação da vontade, Kane pensa que nós deveríamos estabelecer uma diferença entre os casos do assassino e o da executiva. Enquanto a vontade do assassino ‘já está determinada’ em um certo modo (matar o primeiro ministro), a vontade da executiva não está formada até que a escolha seja feita (KANE, 1999, p. 230). No caso das ações autoformadoras, Kane afirma:

Imagine que a mulher de negócios está tentando ou fazendo um esforço para resolver dois problemas cognitivos ao mesmo tempo ou para completar duas tarefas deliberativas conflitantes de uma vez só, fazer uma escolha moral e fazer uma escolha autointeressada (...). Com respeito a cada tarefa ela está sendo ameaçada pelo indeterminismo na sua tentativa de fazer o que ela está tentando fazer (...). Mas, no seu caso, o indeterminismo não tem uma origem externa. Vem a partir da sua própria vontade, a partir de seu desejo de fazer o oposto (...). Ela pode falhar de fazer o que ela está tentando fazer, como no caso do assassino. Porém, se ela tem sucesso ela será responsável porque ela terá tido êxito em fazer *o que ela estava tentando fazer*. E isso é verdadeiro seja *qual for a escolha que é feita* porque ela estava tentando fazer ambas as escolhas, e uma delas teve êxito (KANE, 1999, p. 231).

A questão que imediatamente surge é se faria sentido falar do agente tentando fazer duas tarefas conflitantes simultaneamente. Kane acredita que haveria evidências científicas mostrando que o cérebro é um processador paralelo. Por analogia, o cérebro processaria simultaneamente diferentes tipos de informações referentes a tarefas como percepção ou reconhecimento por meio de diferentes caminhos neurais. Essa seria uma capacidade essencial para o exercício da vontade livre e “em casos de autoformação agentes estão, ao mesmo tempo, tentando resolver tarefas cognitivas plurais e conflitantes” (KANE, 1999, p. 231). Embora não se trate de pessoas separadas, é como se os agentes tivessem duas mentes. Por exemplo, a executiva que quer ajudar a vítima do assalto é a mesma mulher ambiciosa que não quer se atrasar para a sua reunião (KANE, 1999, p. 231-232; KANE, 2013, p. 265-266).

O problema da sorte é formulado de diferentes modos quando se trata de casos envolvendo formação da vontade ou ações livres. Porém, a intuição por trás desse problema é a mesma, a saber, que as escolhas ou ações indeterminadas seriam “inexplicáveis”, “arbitrárias”, “não controladas”, pois ocorreriam “por acaso ou por sorte”. É por isso que elas não seriam ações livres e responsáveis⁵⁹. O problema pode ser formulado em termos *de um único agente* em diferentes mundos possíveis. Retomando o caso tratado anteriormente, as hipóteses eram estas: o mundo real e o possível tinham exatamente o mesmo passado até t , o agente acredita que deve chegar no horário marcado para a reunião e ele também deseja ajudar a vítima do assalto. No mundo real o agente se atrasa e, no mundo possível, ele é pontual. Parece que no mundo possível o agente teve sorte ao superar a tentação de se atrasar. Seria justo recompensar um agente e punir o outro? Faria sentido louvar um e censurar o outro se o que diferencia os dois agentes é meramente uma questão de sorte? De modo análogo, um criminoso pode errar o alvo. Caso isso aconteça, ele é responsável unicamente pela tentativa de assassinato. Porém, se ele acerta, não faria sentido, conforme Kane, isentá-lo de responsabilidade moral pelo ato.

A objeção da sorte pode falhar em casos como o do assassino. Mesmo que ele acerte o alvo por sorte, ele ainda seria responsável por esse ato porque foi algo intencional. Porém, essa objeção pode ser reformulada a fim de ser aplicada ao exercício da vontade livre envolvendo as ações autoformadoras, como no caso da executiva. Nesse caso, a falha não é meramente uma questão de erro ou acidente. No mundo real, S_1 (um agente) voluntariamente e intencionalmente faz uma ação A em t . Em um mundo possível, S_2 voluntariamente e intencionalmente faz B em t . Se é indeterminado que se faça A ou B , o agente fazer uma das opções em vez da outra em t seria uma questão de sorte⁶⁰.

Kane faz um alerta sobre como interpretar a questão do esforço indeterminado da seguinte forma:

Não se pode dizer que dois agentes que eles tiveram exatamente o mesmo passado, e fizeram exatamente os mesmos esforços, e um foi sortudo enquanto o outro não. Nem nós podemos imaginar o mesmo agente em dois mundos

⁵⁹ Os termos “acaso” e “sorte” são termos da linguagem ordinária que significam “estar fora de nosso controle”, enquanto que “indeterminismo” é um termo técnico que implica a ausência de “causação determinística” (KANE, 1999, p. 223).

⁶⁰ Para uma discussão acerca do problema da sorte moral ver Williams e Nagel (1993, p. 35-55 e 57-71). Em linhas gerais, vejamos a seguinte situação. Dois motoristas imprudentes voltam para casa dirigindo. Um deles atropela e mata uma criança, ele teve o que se pode dizer má sorte. O outro motorista voltou para casa sem acontecer nada trágico. Não seria arbitrário responsabilizar moralmente apenas o motorista que matou a criança?

possíveis com exatamente o mesmo passado fazendo exatamente o mesmo esforço e tendo sorte em um mundo e não no outro (KANE, 1998, p. 171-172).

Para evitar confusão o ideal seria dizer que dois agentes, em dois mundos, empreenderam os esforços para superar a tentação de fazer a escolha contrária. Em um mundo, um dos esforços resultou em uma escolha, em outro mundo, um diferente esforço resultou em uma diferente escolha. Ambos tiveram sucesso em fazer o que eles estavam tentando fazer e ambos endossaram os resultados dos seus respectivos esforços, de modo que as suas escolhas foram voluntárias e intencionais (KANE, 1999, p. 235).

Kane estaria limitando escolhas livres às escolhas feitas diante de um conflito interno? Quando não existe conflito e nenhum esforço para resistir à tentação, existiria escolha livre (quer dizer, não haveria PA)? Kane endereça essa questão ao explicar que as escolhas livres incompatibilistas “surgem naquelas conjunturas críticas da vida humana quando nós estamos engajados em conflito moral, prudencial e prático, onde existe um conflito interno e o resultado não é óbvio” (KANE, 1989, p. 252).

Parece haver aqui uma assimetria entre a escolha/ação decorrente do caráter (ausência de PA) e aquela escolha/ação que forma o presente caráter (necessidade de PA). Esse aspecto ficará mais claro no próximo subcapítulo, no qual destacarei uma distinção entre a explicação de Kane e a de Fischer.

4.3 Desacordo entre Kane e Fischer

Como visto no capítulo anterior, Fischer aceita que a liberdade para agir diferentemente seja incompatível com o determinismo causal. Porém, ele insiste que a responsabilidade moral seria compatível com o determinismo, pois ela não exigiria a liberdade para agir de outro modo. O primeiro passo do seu argumento em favor dessa conclusão envolve os exemplos de tipo Frankfurt. O desacordo entre Kane e Fischer estaria justamente na definição do conceito de liberdade. Kane acredita que a liberdade necessária para a responsabilidade moral (poder agir de outro modo) seria incompatível com o determinismo causal. Nesse sentido, Kane acredita que um controlador (de tipo Frankfurt) estaria diante de um dilema:

Incompatibilistas e libertistas como eu sobre vontade livre acreditam que se nós somos verdadeiramente responsáveis por ser quem nós somos, então deveria existir algumas escolhas nas nossas vidas que são indeterminadas até o momento quando elas ocorrem. Essas escolhas indeterminadas são (...) ações autoformadoras (...) que são necessárias em alguns pontos de nossas vidas, se nós somos verdadeiramente responsáveis por formar as nossas

próprias vontades. Agora um controlador (de tipo Frankfurt) enfrenta um dilema ao tentar controlar essas (..) escolhas autoformadoras. Visto que elas são indeterminadas até o momento que elas ocorrem, um controlador (...) não pode ter certeza de que modo os agentes vão escolher *antes* deles, realmente, escolherem (KANE, 2007, p. 169).

Dito de outro modo, o indeterminismo exigido para as ações autoformadoras frustraria qualquer controlador potencial. O único modo pela qual o controlador poderia impedir as possibilidades alternativas, no caso das ações autoformadoras, seria interferindo antecipadamente para determinar qual escolha será feita. Nesse caso, quem seria responsável seria o controlador, e não o agente. Se nós somos verdadeiramente responsáveis por formar a nossa própria vontade, e, portanto, para ter vontade livre no sentido libertista, nós devemos praticar *algumas* escolhas autoformadoras em algum momento das nossas vidas. Desse modo, os exemplos de tipo Frankfurt não funcionariam para as escolhas autoformadoras, pois tais exemplos não provariam que a responsabilidade moral seria compatível com *todas* nossas ações se elas fossem determinadas.

O que Fischer acredita ser problemático na argumentação de Kane, e é aqui que a diferença entre os dois autores fica mais clara, é a assimetria na exigência das condições para se ter uma vontade livre. Por exemplo, haveria uma distinção entre duas condições para se ter vontade livre: (a) PA e (b) RU. Para Kane, a condição (b) seria mais importante que a (a). Como já visto anteriormente, RU exige que nós poderíamos ter feito de outro modo com respeito a *alguns* atos na nossa história de vida passada, pelos quais nós formamos o nosso presente caráter.

Para exemplificar como o agente poderia ser moralmente responsável por um determinado ato mesmo sendo incapaz de agir diferentemente, Kane cita o caso de Lutero, discutido por Daniel Dennett em seu livro *Elbow Room* (1984):

Quando Lutero disse “Aqui estou. Eu não posso fazer de outro modo”, na ocasião da sua ruptura com a igreja em Roma, ele quis dizer, segundo Dennett, “que sua consciência lhe faz *impossível* recuar” (p. 133). Lutero pode ter errado sobre isso, Dennett adiciona, mas esse não é o ponto. Pois, mesmo se Lutero tivesse certo e ele não poderia ter feito de outro modo, “nós simplesmente não isentamos alguém de censura ou louvor por um ato porque nós pensamos que ele não poderia ter feito de outro modo” (p. 133). Ao dizer “eu não posso fazer de outro modo”, Lutero não estava evitando responsabilidade pelo seu ato, conforme Dennett, mas assumindo completa responsabilidade por ele. Então, se nós associarmos vontade livre com responsabilidade moral, nós não podemos supor que vontade livre necessariamente implica que o agente “poderia ter feito de outro modo”. PA não seria necessário para vontade livre (KANE, 1998, p. 38).

Obviamente que Kane não concorda com a conclusão de Dennett, pois PA seria necessária pelo menos para *alguns* atos. Nesse sentido, embora Lutero não pudesse fazer de outro modo a não ser romper com a Igreja Católica, a sua responsabilidade por esse ato deveria ser rastreada aos atos passados, nos quais havia PA.

Fischer concorda que nós deveríamos distinguir entre PA e RU. RU implicaria a existência de ações autoformadoras em algum momento na nossa vida. Isso significa que a condição de RU implicaria PA. Porém, Fischer não vê por que a ação autoformadora (formação da vontade) deveria ser diferente da ação ordinária com respeito à exigência das PA. Por que essa assimetria? Por que é exigido PA no nível da vontade mais do que no nível da ação decorrente do caráter? (FISCHER, 2007, p. 185)

Talvez fosse melhor insistir em uma distinção entre a condição de origem (*source*) e a condição de PA: por exemplo, “quando alguém age livremente, alguém deve ser a origem da sua própria ação em um sentido adequado” (FISCHER, 2007, p. 185). Na perspectiva de Fischer:

Se Lutero não livremente rompeu com a igreja em Roma (...), sua responsabilidade moral por ter rompido com a igreja em Roma pode ser rastreada às ações livres anteriores: pode ser rastreada às instâncias passadas de escolher e agir livremente. Mas, escolher e agir livremente não exige liberdade para escolher ou agir diferentemente. Controle de direcionamento não implica controle regulativo. Alguém pode explicar os exemplos como o de “Lutero” via esse tipo de abordagem de rastreio (em que se chega ao passado para encontrar instâncias de agir livremente ou controle de direcionamento), assim como se pode explicar responsabilidade moral pelas consequências de se dirigir embriagado ou comportamento sob influência das drogas. Não existe necessidade de abandonar o modelo de controle de direcionamento; não existe necessidade de colocar possibilidades alternativas em algum ponto na sequência resultante no comportamento (ou na vida dos agentes moralmente responsáveis) (FISCHER, 2007, p. 185).

Enquanto Kane pensa que em algum momento da nossa vida é necessário ter PA para a avaliação moral, para Fischer essa exigência seria desnecessária. Ele acredita que é possível ser moralmente responsável pelo comportamento mesmo carecendo de controle regulativo.

A seguir apresentarei a abordagem internista de Frankfurt. É oportuno conhecer essa interpretação, pois, conforme Kane, ela estaria sujeita à mesma crítica que ele faz a toda e qualquer teoria compatibilista, a saber, de que também não teria os recursos necessários para explicar casos de manipulação.

4.4 A posição internista de Frankfurt

Gostaria de começar com uma citação de Frankfurt que tem sido muito discutida na literatura:

Na medida em que uma pessoa se identifica com as origens das suas ações, ela assume responsabilidade por essas ações e adquire responsabilidade moral por elas; além disso, a questão de como as ações e suas identificações com suas origens são causadas são irrelevantes para as questões de se ela realiza as ações livremente ou é moralmente responsável por realizá-las (FRANKFURT, 1998, p. 54).

É útil compreender o que Frankfurt quer dizer pelo conceito de identificação, conceito esse que não é fácil de definir. No domínio psicológico, particularmente no processo intelectual, nós podemos ser ou ativos ou passivos. Deliberar sistematicamente sobre um problema é uma atividade na qual a pessoa se engaja. Mas nós somos meros espectadores passivos de alguns pensamentos que ocorrem nas nossas mentes, como eventos no nosso corpo⁶¹ e pensamentos obsessivos, eventos cuja procedência pode ser obscura e dos quais não podemos nos livrar (FRANKFURT, 1998, p. 59).

A atividade ou a passividade também pode ser aplicada às paixões. Haveria uma distinção entre as paixões com respeito às quais nós somos ativos e aquelas com respeito às quais nós somos passivos. Entre nossas paixões, existiriam aquelas cujo princípio que move está dentro de nós e outras cujo princípio movente é externa a nós (FRANKFURT, 1998, p. 60). Para exemplificar quais das paixões na história de uma pessoa seriam externas à ela e quais seriam as condições dessa externalidade, vejamos a seguinte situação. Um homem está

⁶¹ Explicitarei o que Frankfurt quer dizer com uma analogia com a ocorrência de eventos corporais. Um dos problemas mais difíceis na teoria da ação é explicar qual é a diferença entre o tipo de coisa que acontece quando uma pessoa levanta o braço (para dar um sinal) e o tipo de coisa que acontece quando o braço da pessoa levanta (por causa de um espasmo muscular). Essa questão envolve um contraste entre eventos que são ações (em que as mais altas faculdades dos seres humanos entram em jogo) e aqueles movimentos do corpo da pessoa (meros acontecimentos corporais) que a pessoa ela mesma não faz. O contraste pode ser generalizado e explicado por meio da distinção entre atividade e passividade. As ações são instâncias de atividade, embora nem sempre. Por exemplo, bater os dedos sobre a mesa certamente não é um caso de passividade. Os movimentos em questão não ocorrem sem alguém fazê-los. Tampouco é uma instância de ação, mas unicamente de *ser ativo*. A ocorrência na vida humana de eventos que são nem ações nem meros acontecimentos, mas são entendidos enquanto ser ativos e passivos, tem sido por algum tempo geralmente negligenciada pelos filósofos. O contraste entre atividade e passividade é perceptível em níveis de existência em que nós não estamos inclinados a supor que existam ações. Por um lado, uma aranha é passiva com respeito aos movimentos das suas pernas quando elas se movem devido a um choque elétrico. Por outro lado, a aranha é ativa com respeito ao movimento das suas pernas (embora não execute nenhuma ação), quando as move a partir do chão. Não é fácil, de acordo com Frankfurt, explicar a diferença entre ser ativo e ser passivo (FRANKFURT, 1998, p. 58-59).

conversando normalmente e, de repente, o seu temperamento sai do seu controle. Embora nada tenha acontecido que torne o seu comportamento compreensível, ele começa a jogar tudo o que ele vê pela frente, cadeiras, copos, e também usa uma linguagem abusiva contra a pessoa com quem ele estava conversando. Ele diz que não sabe o que desencadeou esse “espasmo bizarro de emoção”. Ele não se vê como sendo ele mesmo (FRANKFURT, 1998, p. 63).

Essas desculpas podem ser expressões de arrependimento, mas também podem significar que quando ele estava possuído pela raiva ele não estava na posse de si mesmo, pois “o que ele diz pode apropriadamente expressar seu sentido de que a origem da paixão representava, de alguma forma, uma intromissão sobre ele, que o violava” (FRANKFURT, 1998, p. 63). São as afirmações feitas pelo homem nesse exemplo que expressam, segundo Frankfurt, a experiência da externalidade.

A questão de quais seriam as condições essenciais ou marcas da externalidade tende a ser respondida apelando para o fato de que a pessoa desaprova a paixão que ocorre nela. Porém, desaprovção, embora possa ser uma condição necessária da externalidade, não seria suficiente⁶². E a explicação disso se dá pela seguinte razão: é possível que uma pessoa reconheça uma certa paixão como atribuída a ela, ou sendo sua, mesmo que ela se arrependa disso, desaprova e deseja que a paixão não a mova à ação (FRANKFURT, 1998, p. 63-65).

Seria equivocado pensar que os conceitos de internalidade e externalidade deveriam ser explicados *simplesmente* em termos das atitudes das pessoas, como é explícito na citação: “é equivocado sugerir que a externalidade da paixão é implicada pela desaprovção da paixão pela pessoa ou que a sua internalidade é implicada pela sua aprovção” (FRANKFURT, 1998, p. 65).

Acredito que a caracterização de uma paixão como interna ou externa possa ser elucidada por meio de dois tipos de conflitos de desejos. No primeiro conflito, suponha que uma pessoa queira ir ao concerto e que ela também queira assistir a um filme e que esses dois desejos sejam conflitantes. A pessoa resolveria o problema que esse conflito apresenta decidindo qual desejo ela prefere. É provável que ambos os desejos sejam internos a ela. Ela quer fazer as duas coisas, porém ela não pode fazer ambas ao mesmo tempo. Além disso, conflitos desse primeiro tipo exigem que os desejos sejam ordenados:

⁶² É difícil pensar um exemplo convincente de uma pessoa que aprova a ocorrência de uma paixão que lhe é externa. E essa dificuldade tende a apoiar a conjectura de que a aprovção da pessoa de uma paixão que ocorre na sua história é uma condição suficiente da paixão ser interna a ela (FRANKFURT, 1998, p. 65).

A sua característica essencial é que os desejos conflitantes pertencem ao mesmo ordenamento, ainda que enquanto o conflito não estiver resolvido, a sua posição relativa naquele ordenamento não é definida. O pertencimento ao mesmo ordenamento pode ser explicado como segue. Suponha que uma pessoa decida que ela na realidade prefere ver o filme, mas que ela é incapaz de conseguir o ingresso. Então seria totalmente natural para ela escolher a segunda opção e ir ao concerto. Afinal ela quer fazer isso também. Sua decisão original contrária àquela significou apenas que satisfazer o seu desejo para ir ao concerto era inferior no seu ordenamento do que satisfazer o seu desejo para ver o filme (FRANKFURT, 1998, p. 66).

Já quando uma pessoa enfrenta um conflito do segundo tipo, o problema não seria resolvido atribuindo-se uma posição inferior a um dos desejos, mas sim rejeitando-se completamente um dos desejos. Por exemplo, digamos que, por um lado, Pedro queira fazer um elogio a Paulo por ele ter conseguido um emprego e, por outro, ele também observa dentro de si um desejo ciumento de prejudicar Paulo. Nesse caso, o conflito não seria resolvido ordenando os desejos conflitantes. Digamos que Pedro decidiu fazer o elogio a Paulo, porém não encontrou a oportunidade de fazê-lo. A não satisfação disso não conduziria Pedro a ver se o outro desejo poderia ser satisfeito. Ao rejeitar o desejo de prejudicar Paulo, Pedro estaria tanto se afastando como localizando o desejo rejeitado fora do escopo das suas preferências, como não sendo um candidato para a satisfação (FRANKFURT, 1998, p. 67).

Interessante observar que embora Pedro possa continuar tendo a experiência, na sua história mental, do desejo rejeitado, ele pode fazer com que essa ocorrência seja externa. Embora o desejo rejeitado possa persistir como um elemento da sua experiência, não poderia mais ser atribuído estritamente a Pedro. Em um conflito do primeiro tipo, a pessoa tanto quer assistir a um filme como também quer ir ao concerto. Já no conflito do segundo tipo, a rejeição de Pedro do desejo de prejudicar Paulo significa que Pedro não quer prejudicá-lo, mesmo que esse desejo ocorra na sua experiência mental. É quando Pedro decide que quer fazer um elogio a Paulo e que seu desejo de prejudicá-lo deveria ser excluído dos candidatos para a satisfação, que ele torna o segundo desejo externo a si mesmo e se identifica com o primeiro (FRANKFURT, 1998, p. 67-68).

Antes de finalizar, gostaria de enfatizar uma crítica comum feita a toda e qualquer teoria compatibilista, a saber, que ela não teria justificado apropriadamente a distinção entre o determinismo causal e o caso de manipulação. Embora Frankfurt e Fischer concordem sobre a não necessidade de PA para a responsabilidade moral, ambos diferenciam-se no que diz respeito à explicação dada às origens do comportamento, desejos e preferências e, portanto, distanciam-se na resposta que oferecem a essa crítica.

Fischer acredita que uma abordagem histórica, externista, seria uma solução para o problema. Nesse sentido, olhando os antecedentes causais das ações teríamos condições de distinguir, por um lado, o comportamento que resulta a partir de uma sequência causal e, por outro, o comportamento resultante da hipnose, lavagem cerebral e manipulação. Entretanto, Frankfurt acredita que embora essa abordagem possa ser sugestiva, ela é dispensável e “o que nós necessitamos mais essencialmente é olhar para certos aspectos da estrutura psíquica que é coincidente com o comportamento da pessoa” (FRANKFURT, 2002, p. 27). Isso fica mais claro a partir da seguinte citação:

Resumidamente, parece que se alguém faz algo porque quer, e se ele não tem quaisquer reservas sobre o desejo e está comprometido com ele, então – no que concerne à sua responsabilidade moral pela ação – realmente não importa como ele se tornou o que ele é. Um requisito adicional tem de ser adicionado: os desejos e as atitudes da pessoa têm de ser relativamente bem integrados na sua condição psíquica geral. De outro modo eles não são genuinamente seus, mas apenas intrusos perturbadores na sua verdadeira natureza. Desde que as suas interrelações impliquem que eles são inequivocamente atribuíveis a ele como os seus desejos e as suas atitudes, não faz diferença – no que concerne à avaliação da sua responsabilidade – como ele veio a tê-los (FRANKFURT, 2002, p. 27).

Frankfurt estaria dissociando a responsabilidade pelo caráter da responsabilidade pela ação decorrente desse caráter. A história de como uma pessoa veio a ter um certo caráter não importaria no que diz respeito à responsabilidade moral, pois inevitavelmente o nosso caráter é moldado por circunstâncias sobre as quais nós não temos controle. Então, seria irrelevante se as causas responsáveis por quem nós somos operam por meio de forças naturais que moldam o nosso ambiente ou se elas operam através dos planos manipuladores de outros agentes. O que de fato seria relevante para a avaliação moral de um criminoso, por exemplo, seria analisar se ele é criminoso contrariamente à sua vontade ou se ele está agindo como ele realmente prefere agir (FRANKFURT, 2002, p. 28).

A seguir apresentarei a solução proposta por Alfred Mele ao problema. Embora Mele seja agnóstico em relação à disputa entre compatibilistas e incompatibilistas, ele admite que os compatibilistas possuem os meios necessários para distinguir os casos de manipulação da determinação.

4.5 A proposta de Mele para o compatibilismo

Mele acredita que pode responder à crítica incompatibilista, desenvolvida por Kane, de que os compatibilistas não teriam os recursos necessários para distinguir os casos de

manipulação da mera determinação causal. Sua estratégia consiste em oferecer condições compatibilistas para diferenciar os casos de manipulação de casos de deliberação e ação autônoma.

Mesmo um agente idealmente autocontrolado e mentalmente saudável poderia carecer de autonomia. É muito comum na literatura, segundo Mele, a autonomia envolver o conceito de autocontrole, e, em geral, as duas noções serem tratadas conjuntamente. Essas duas noções são usados na literatura em vários sentidos. Para evitar confusão, Mele explicita como esses termos serão usados no seu livro, *Autonomous Agents: From Self-Control to Autonomy*. Autocontrole é o contrário de acrasia (fraqueza da vontade ou incontinência) e autonomia significa “autogoverno” (MELE, 1995, p. 3). Tradicionalmente, a ação acrática era considerada como aquela em que o agente reconhece a sua ação como sendo contrária ao melhor julgamento. É interessante que para Mele existem vários outros tipos de acrasia. Por exemplo, a acrasia pode ser manifestada em um ato que coincide com o melhor juízo. De igual modo, alguém pode manifestar autocontrole em um ato que vai contra o melhor juízo. Apresento um exemplo, adaptado, para explicar a acrasia e o autocontrole. Digamos que um criminoso decida fazer parte de um grupo que pretende assaltar um banco, mesmo que ele julgue que é melhor não cometer esse ato criminoso. Vamos supor que esse criminoso se recuse a entrar no banco no momento do assalto. O seu ato está, assim, em conformidade com o seu juízo. Mas o que explica a sua recusa a cometer o ato criminoso é o seu medo de ser preso, e não o predomínio da sua crença de que ele não deveria cometer esse ato. Usando o mesmo exemplo, agora para explicar o autocontrole: a ação do criminoso, nesse caso, vai contra o seu melhor juízo. Ele julga que é melhor não participar do crime. Mas no momento do assalto ele domina o seu medo, e entra no banco para participar do assalto. Ele exerceu autocontrole ao dominar o seu medo (MELE, 1995, p. 60).

Considerando o que seria uma pessoa idealmente autocontrolada, Mele afirma ser alguém que pode exibir autocontrole com relação às ações, à aquisição e à retenção de crenças, e às emoções (temos controle, às vezes, sobre como nossas emoções se manifestam em nosso comportamento, sobre quais emoções nós temos ou carecemos) (MELE, 1995, p. 6). Mele questiona se uma pessoa idealmente autocontrolada, mesmo assumindo que ela exerça autocontrole regularmente em todas as áreas da sua vida, é autônoma. A sua resposta é negativa. Mele menciona a definição de Gerald Dworkin, de acordo com a qual autonomia é uma capacidade de segunda ordem para refletir criticamente sobre preferências e desejos de primeira ordem, e a capacidade ou para se identificar com esses ou para mudá-los com base em preferências e valores de ordem superior (DWORKIN, 1988, p. 108). Para Mele, uma pessoa

idealmente autocontrolada teria essa capacidade. O problema seria que mesmo uma tal pessoa pode ter todo o processo de reflexão crítica regulado por princípios ou valores que são produtos de lavagem cerebral ou outras formas de controle da mente (MELE, 1995, p. 122). Mesmo uma pessoa idealmente autocontrolada pode não ser autônoma.

Para Mele, autonomia não seria apenas uma questão das capacidades e disposições que o agente possuiria em um determinado tempo, ou seja, não seria apenas uma questão dos estados psicológicos internos do agente. Mele estaria defendendo assim uma noção externista (histórica) de autonomia ao invés de uma noção internista. As capacidades envolvidas na autonomia seriam de dois tipos: uma direcionada ao ambiente externo e outra ao ambiente interno (psicológico). Por exemplo, alguém acorrentado a uma árvore possuiria uma capacidade limitada para afetar o mundo físico. Outros teriam uma capacidade interna, como a capacidade de tomar decisão e refletir criticamente sobre valores e princípios. Embora seres que careçam de uma vida mental e que são incapazes de decisão e reflexão possam afetar o mundo externo, eles não podem fazê-lo de modo autônomo. É no que consistiria, mais especificamente, a autonomia psicológica, que Mele pretende investigar (MELE, 1995, p. 144). A autonomia psicológica estaria relacionada aos vários aspectos da vida mental ou psicológica do agente, incluindo as proatitudes (isto é, os valores, as preferências e os desejos) (MELE, 1995, p. 138)⁶³.

O que deveria ser adicionado a um tal agente para se obter a autonomia? Essa questão é respondida de formas distintas por compatibilistas e incompatibilistas. Para os compatibilistas, por exemplo, a autonomia não dependeria da falsidade do determinismo. Para os incompatibilistas, a autonomia somente faria sentido em um mundo indeterminista.

O conjunto de condições suficientes para a autonomia psicológica compatibilista proposto por Mele é o seguinte:

1. O agente não tem estados motivacionais compelidos ou coercitivos.
2. As crenças do agente são conduzidas por deliberação informada sobre todas as questões que concernem a ele.
3. O agente é um deliberador confiável (MELE, 1995, p. 187).

⁶³ Para que um agente possua uma pró-atitude de modo autônomo seria necessário que ela fosse possuída autenticamente. A noção de autenticidade para a posse autônoma de uma pró-atitude envolveria a carência de um certo tipo de história. Por exemplo, os valores não devem ser possuídos de modo compulsivo (MELE, 1995, p. 172).

A satisfação dessas três condições bastaria para a ausência de controle manipulativo por outro agente. Um controlador poderia operar a partir de três modos: sobre as atitudes motivacionais da vítima (como será visto, a seguir, no cenário de lavagem cerebral), sobre as atitudes informacionais e sobre as qualidades executivas da vítima (discípulos condicionados). Além disso, essas condições protegeriam mais a autonomia do agente de um “controlador potencial” do que a própria condição de Kane de escolha causalmente indeterminada (MELE, 1995, 187).

Como forma de entender essas três condições compatibilistas que seriam suficientes para autonomia psicológica, eu apresento três cenários nos quais o agente careceria de autonomia. O primeiro é um caso de lavagem cerebral. Podemos começar com a história de dois agentes. Suponha que Ann seja autônoma e uma filósofa aplicada. Ela trabalha várias horas durante o dia nos sete dias da semana. Sua colega Beth atribui valor a muitas outras coisas além da filosofia, e por isso trabalha menos do que Ann, por razões que ela endossou a partir de uma cuidadosa reflexão crítica. Beth se identifica com o seu próprio modo de vida. Vamos imaginar também que o chefe do departamento de filosofia quer que Beth seja como Ann e que ele não conseguiu fazer isso através de formas normais de persuasão. Sem o conhecimento delas, ele contrata uma equipe de psicólogos para investigar o que estaria motivando Ann ao trabalho e uma equipe de manipuladores para fazer com que Beth seja como Ann. Os psicólogos chegaram à conclusão de que seria a hierarquia de valores de Ann que explicaria a sua produtividade. Já a tarefa dos manipuladores consistiu em introduzir a hierarquia de valores de Ann em Beth, e apagar os seus valores conflitantes, por meio de um processo de lavagem cerebral. Eles são bem sucedidos e Beth passa a ser uma filósofa dedicada que valoriza o seu trabalho filosófico acima de tudo (MELE, 1995, p. 145).

Porém, a mudança na hierarquia dos valores surpreende Beth. Ela se questiona o que explicaria o seu repentino entusiasmo pela filosofia e por que as suas atividades sociais são menos gratificantes ou prazerosas do que o trabalho. A sua hipótese é de que ela, talvez, tenha se cansado do seu modo anterior de vida e que finalmente veio a apreciar o valor do trabalho filosófico. Quando ela reflete sobre as suas novas preferências e valores, ela percebe que eles apoiariam uma vida dedicada ao trabalho filosófico (MELE, 1995, p. 145).

Ann é claramente autônoma, mas quanto a Beth? Beth é um clone de Ann. As suas próprias preferências e valores foram apagados e substituídos no processo de lavagem cerebral. Beth não consentiu a esse processo e tampouco foi consciente dele. Pode se dizer que a autonomia de Beth foi violada, pois “ao introduzir novos valores nos estados motivacionais de Beth e eliminar os antigos, os manipuladores deram à sua vida uma nova direção, uma que entra

em conflito com os princípios considerados e valores que ela possuía antes da manipulação” (MELE, 1995, p. 145). Dadas as similaridades psicológicas entre os dois agentes, a diferença na autonomia parece repousar em *como elas vieram* a ter certas características psicológicas, portanto, em alguma coisa externa às suas constituições meramente psicológicas. Nesse sentido, a diferença seria histórica e autonomia seria de algum modo baseada na história.

Uma noção relevante para uma compreensão desses exemplos é a capacidade do agente de controlar a sua vida mental. No caso de lavagem cerebral visto acima e em casos de hipnose, por exemplo, as capacidades dos agentes para controlar as suas vidas mentais estariam sendo contornadas (*bypassed*). Conforme Mele:

Em agentes idealmente autocontrolados, essas capacidades são consideráveis. Tais agentes são capazes de modificar a força de seus desejos à serviço de seus juízos práticos, avaliativos, de fazer com que suas emoções estejam de acordo com os juízos relevantes, e de dominar a motivação que ameaça (às vezes através do desvio do raciocínio prático e teórico) para produzir ou sustentar crenças em modos que violariam seus princípios para aquisição e retenção de crença. Eles são capazes de racionalmente avaliar e revisar seus valores e princípios, de identificar com os seus valores na base de reflexão crítica, informada, e de intencionalmente promover novos valores e pró- atitudes neles mesmos de acordo com os seus juízos avaliativos considerados. Presumivelmente, a maioria dos leitores desse livro tem essa capacidade, em alguma medida. Todas tais capacidades são contornadas nos casos de manipulação de pró-atitude em questão. Em tais casos, novas pró- atitudes não são geradas via um exercício ou ativação das capacidades do agente de controle sobre suas vidas mentais; ao invés, elas são geradas apesar das capacidades do agente para isso (MELE, 1995, p. 166-167).

O papel da história para explicar se o agente realizou livremente uma ação e é responsável por ela pode se tornar mais explícito a partir de uma outra situação. Vamos imaginar que Pat tenha sido um pai medíocre. O seu próprio pai, surpreendentemente, foi muito bom para ele e ocasionalmente Pat se sente culpado sobre o quão pouco ele fez pelos seus dois filhos. Depois de passado um tempo sendo um pai medíocre, ele decide livremente entrar em um programa de automelhoramento. Parte da sua estratégia consistiu em passar mais tempo com seus filhos, preparando atividades que fossem mutuamente agradáveis, e concentrando os seus pensamentos no bem-estar deles. Como resultado da participação nesse programa, ele se tornou um pai admirável, dedicado e adquiriu novos valores parentais. Ele, por exemplo, fez um empréstimo para pagar a mensalidade da faculdade da sua filha, o que foi um grande sacrifício da sua parte (MELE, 2006, p. 168).

Agora compare Pat com Paul. Paul é um pai egoísta que por muitos anos refletidamente se identificou com os seus valores egoístas. Vamos supor que Paul foi dormir como um pai

mediocre de uma filha de dezoito anos. No entanto, durante o sono ele é submetido a um processo de lavagem cerebral. A sua mãe contratou uma equipe de psicólogos e manipuladores, sem o seu conhecimento, para fazer com que a sua hierarquia de valores fosse como a de Pat. Paul acorda um pai extremamente comprometido, tal como Pat (MELE, 2006, p. 168-169).

Ao experienciar subitamente um forte desejo de tornar realidade a vontade da sua filha de estudar na faculdade, Paul se surpreende com a mudança dos seus valores. Ao se questionar sobre o que explicaria essa mudança, a sua hipótese é a de que ele simplesmente se cansou do seu estilo de vida egoísta e que finalmente percebeu a importância das relações entre pai e filha. Quando reflete sobre seus valores, ele percebe que eles sustentam uma vida dedicada ao próximo e ele adota esse estilo novo de vida. Quando ele pensa sobre o desejo de sua filha, é óbvio para ele que ele deveria fazer o empréstimo, e é isso o que ele faz. Dado seus novos valores, ele não poderia ter feito de outro modo (MELE, 2006, p. 169).

O que foi dito sobre Beth pode ser aplicado a Paul. Tanto os valores de Beth quanto os de Paul foram apagados e substituídos no processo de lavagem cerebral. Nenhum dos dois consentiu ou foi consciente do processo. Nenhum dos dois é livre e tampouco merecedores de crédito moral. Embora eles não pudessem ter feito de outro modo, haveria uma distinção na história que explicaria por que eles, e não as suas contrapartes, carecem de responsabilidade moral. Os agentes não mereceriam ser louvados pelas suas respectivas ações porque a origem dos valores está no controle dos manipuladores, e não nos próprios agentes (MELE, 2006, p. 169-170).

O segundo caso, envolvendo manipulação de informação, corresponderia à segunda exigência das condições compatibilistas para a carência de autonomia psicológica⁶⁴. Digamos que um médico delibera sobre o remédio que deveria dar ao seu paciente. Ele pede ao enfermeiro o prontuário do paciente para certificar-se de quais são as suas alergias para poder escolher o melhor medicamento. Vamos supor que o enfermeiro odeia o paciente e fornece a informação falsa de que em seu prontuário não consta nenhuma alergia. O médico prescreve um determinado medicamento ao qual o paciente é alérgico e o paciente morre. Assumindo que o médico não é responsável pela falsa informação, ele não poderia ser censurável pela morte do paciente.

⁶⁴ Eu adaptei o exemplo. Mele apresenta dois exemplos de manipulação informacional, o do investidor enganado e o do rei que toma decisões com base em informações falsas repassadas pelos seus subordinados. Sobre esses exemplos, ver Mele (1995, p. 179-182).

O terceiro caso trata da produção de hábitos deliberativos por meio da manipulação. Aqui o controlador opera sobre as qualidades executivas da vítima. Vejamos o exemplo do líder de um culto religioso. Ele trabalha na hipótese de que os hábitos não conscientemente possuídos são mais resistentes à mudança e, então, faz o procedimento manipulatório secretamente. Ele desenvolve sofisticadas técnicas para produzir certos hábitos deliberativos nos seus seguidores. Além disso, são recrutados discípulos adolescentes dos quais é exigido deliberar várias vezes durante o dia sobre como eles gostariam de passar o tempo livre. As deliberações relacionadas às premissas teológicas são recompensadas, enquanto que as deliberações envolvendo as premissas científicas são punidas, sutilmente (MELE, 1995, p. 182).

A ideia é reconhecer que a satisfação das três condições compatibilistas para a autonomia psicológica elencadas acima bastaria para a ausência de controle manipulativo por outro agente. Desse modo, o problema do controle manipulativo poderia ser, segundo Mele, resolvido pelos compatibilistas (MELE, 1995, p. 187-188). Tal problema poderia ser resolvido de modo satisfatório pela teoria libertista de Kane? Para responder esse questionamento, retomo o caso da manipulação de informação. Digamos que com base na informação repassada pelo enfermeiro, de que o paciente não tem alergia a nenhum medicamento, o médico tem duas ou mais opções de remédios para escolher, todas causando alergia no paciente. Seja qual for a sua escolha, ele não escolherá de modo autônomo, mesmo que a escolha seja causalmente indeterminada. Nenhuma escolha que ele fará irá salvar a vida do paciente, pois ambas as escolhas foram manipuladas⁶⁵.

Ao apresentar essas condições compatibilistas, Mele estaria defendendo uma posição compatibilista? Não. Mele acredita que a afirmação de que existem agentes livres e moralmente responsáveis é mais plausível do que a afirmação de que não existem. Portanto, ele acredita na possibilidade de agentes autônomos. E essa questão é diferente da posição sobre a verdade ou a falsidade do compatibilismo ou do incompatibilismo. Sobre isso, Mele permanece agnóstico. Existe um caso particular que o impede de adotar o compatibilismo. O caso é o seguinte:

Diana cria um zigoto Z em Mary. Ela combina os átomos de Z em uma certa maneira porque ela quer que um certo evento E ocorra trinta anos depois. A partir do seu conhecimento do estado do universo anterior à criação de Z e das leis da natureza do seu universo determinista, ela deduz que o zigoto com a constituição de Z implantado em Mary se desenvolverá em um agente idealmente autocontrolado que, em trinta anos, julgará, com base na

⁶⁵ Essa problemática foi apresentada por Richard Double, mas com um exemplo distinto (1991, p. 214-215).

deliberação racional, que é melhor realizar A, e que fará A com base nesse juízo, portanto produzindo E. Se esse agente, Ernie, tem quaisquer valores inabaláveis naquele momento, eles não desempenham algum papel na motivação de A. Trinta anos depois, Ernie é uma pessoa mentalmente sadia, idealmente autocontrolada, e não tem qualquer atitude relevante compelida ou produzida coercitivamente. Além disso, suas crenças são condutivas pela deliberação informada sobre todas as questões que concernem a ele, e ele é um deliberador confiável. Assim, ele satisfaz uma versão da minha proposta compatibilista das condições suficientes por ter realizado A livremente (MELE, 2006, p. 188).

Mele apresenta o que ele denomina de “argumento do zigoto”:

1. Devido ao modo como o zigoto foi produzido no seu universo determinístico, Ernie não é um agente livre e não é moralmente responsável por nada.
2. Concernente à ação livre e responsabilidade moral dos seres em que o zigoto se desenvolve, não existe uma diferença significativa entre o modo que o zigoto de Ernie vem a existir e o modo que qualquer zigoto humano normal vem a existir em um universo determinístico⁶⁶.
3. Assim, o determinismo impossibilita a ação livre e a responsabilidade moral (MELE, 2006, p. 189).

Por um lado, compatibilistas afirmariam que a premissa 1 é falsa, pois eles acreditam que em alguns mundos determinísticos existem agentes livres e moralmente responsáveis. Por outro lado, incompatibilistas afirmariam que a premissa 1 é verdadeira, pois eles acreditam que Ernie não é livre e moralmente responsável em nenhum mundo determinista.

Por que Mele é agnóstico a respeito do compatibilismo do determinismo causal com a autonomia? Embora ele reconheça que os compatibilistas têm os recursos para responder ao desafio posto por Kane envolvendo a manipulação, os compatibilistas possuiriam a desvantagem de não acomodar uma explicação de autonomia em um mundo indeterminista. Já a partir de uma perspectiva agnóstica, a desvantagem do compatibilismo seria a de que essa posição está comprometida com a tese de que a autonomia é compatível com determinismo que os agnósticos consideram não saber se é verdadeira (MELE, 1995, p. 252).

No subcapítulo a seguir eu trato do desacordo entre as posições de Mele e Fischer, evidenciando, desse modo, até que ponto ambas as posições são consistentes.

⁶⁶ Mele observa que embora a premissa 2 do argumento do zigoto seja parecida com o argumento da consequência, é significativamente diferente. O argumento da consequência é um argumento para o incompatibilismo, ao passo que a premissa 2 do argumento do zigoto é consistente com o compatibilismo (MELE, 2006, p. 190).

4.6 A distinção entre a posição de Mele e Fischer

Apesar de Mele acreditar que o compatibilismo possui os recursos necessários para responder às críticas impostas por Kane, em razão do argumento do zigoto ele permanece agnóstico sobre o compatibilismo.

Para entender como Fischer discorda de Mele, primeiramente precisamos compreender o que está sendo dito nas duas premissas e na conclusão do argumento do zigoto. Na premissa 1, o que significa a sentença “por causa do *modo* que seu zigoto foi produzido no seu universo determinístico”? Significaria que o comportamento resultante (a ocorrência de um certo evento *E* trinta anos depois) foi iniciado por um agente (Diana) que tinha a intenção de produzir esse comportamento. E, sendo assim, o comportamento de Ernie não poderia ser livre, tampouco poderia ser moralmente responsável. Na premissa 2 do argumento é dito que não existe nenhuma diferença significativa entre esse tipo de processo (isto é, em que as intenções de Diana estão presentes) e as sequências ordinárias de determinismo causal. E a conclusão do argumento é que tanto a liberdade como a responsabilidade moral não fariam sentido em um mundo determinista.

Fischer não concorda com a suposição de que as intenções de Diana, ou que a sua existência no início do processo, seria relevante para a sua responsabilidade moral (FISCHER, 2008, p. 199). Para explicitar por que, Fischer imagina, inspirado por esse “cenário Diana” a seguinte situação hipotética. John e Mary tiveram relações sexuais, mas eles não queriam ter um filho. Embora tenham usado contraceptivo com a intenção de evitar a gravidez, infelizmente Mary engravidou. O casal decide ter o filho, que se chama ‘Ernie’. Suponha que o universo seja causalmente determinístico e que trinta anos depois Ernie pratique alguma ação *A*. Também é estipulado que Ernie satisfaz as condições compatibilistas de ação livre quando realiza *A*. Parece muito claro para Fischer que “as circunstâncias da concepção de Ernie trinta anos antes são irrelevantes para a questão de se ela age livremente e é moralmente responsável no contexto em questão (trinta anos depois)” (FISCHER, 2011, p. 1).

Dito de outro modo, o fato de que John e Mary não tinham a intenção de ter um filho, e que eles usaram contraceptivo, é irrelevante para a responsabilidade moral de Ernie. Digamos que a história original fosse alterada. E que, na versão atualizada, o casal tinha a intenção de ter um filho e não usaram o contraceptivo. Mesmo assim as intenções do casal (ou de Diana) seriam irrelevantes para explicar a responsabilidade moral. O mesmo pode ser dito se a história fosse novamente alterada, dessa vez para supor que o casal tivesse desejos específicos com respeito ao filho que eles esperavam criar. De igual modo, não afetaria a nossa avaliação da responsabilidade moral de Ernie “mesmo se nós adicionássemos que John e Mary tiveram

relações sexuais no momento preciso na crença de que, ao fazer isso, eles assegurariam que Ernie se comportaria, como ela faz, no futuro-talvez trinta anos depois” (FISCHER, 2011, p. 2).

A posição de Fischer seria que as circunstâncias particulares da história de John e Mary, em qualquer uma das suas versões, quer dizer, as intenções específicas do casal no momento da relação sexual, parecem ser completamente irrelevantes para a responsabilidade moral de Ernie quando ele se torna uma pessoa adulta (FISCHER, 2011, p. 2). Além disso, esse cenário apontaria para a falsidade da premissa 1 do argumento do zigoto, e a conseqüente falta de solidez desse argumento. O cenário da Diana também não serviria como suporte para essa premissa. O que parece correto no argumento do zigoto é a premissa 2, a noção de que não existiria diferença entre o cenário de Diana e o cenário envolvendo determinismo causal (em que não há a suposição da existência de um criador com intenções específicas). O interessante é que essa premissa poderia ser adotada pelos compatibilistas, não servindo, desse modo, para os propósitos dos seus oponentes. Claro que os incompatibilistas poderiam objetar que ambos os cenários são problemáticos.

Na literatura o argumento do zigoto figura como uma objeção decisiva ao compatibilismo. Fischer acredita que essa objeção não seja decisiva. A existência de ambos os cenários, o da Diana e o do casal, gera um impasse. Uma mesma situação pode ser descrita e avaliada de diferentes modos dependendo da perspectiva de quem a defende. Supondo que não há nenhuma razão independente para começar um argumento com um cenário ou outro, a existência de cada cenário enfraquece a força do outro. Mas o fato é que existem outras razões para aceitar a tese compatibilista. E o argumento do zigoto, mesmo o adaptando para o cenário de John e Mary, não é um argumento para o compatibilismo. Mas dado que existem outras razões independentes para adotar essa posição, esse argumento poderia ajudar a afastar os desafios impostos em relação à sua aceitabilidade (FISCHER, 2011, p. 5-6).

Para finalizar, ainda sobre esse desacordo, quero apresentar duas críticas às quais a teoria de Fischer estaria sujeita. A condição necessária que ele oferece para a responsabilidade moral seria muito forte e a condição suficiente muito fraca (MELE, 2006, p. 146). Por exemplo, a condição necessária seria que o agente “assumisse responsabilidade moral”, o que implicaria que ele aceitasse ser um candidato justo das atitudes reativas. Uma versão dessa exigência, e que é de interesse de Mele, seria aquela de pessoas que “se engajam em significativa reflexão metafísica sobre a relação entre determinismo causal e a justiça das nossas práticas sociais de aplicação das atitudes reativas” (FISCHER; RAVIZZA, 1998, p. 211). Os agentes reflexivos

devem fazer um juízo metafísico sobre se seria apropriado sujeitar certas pessoas, incluindo a si mesmo, às atitudes reativas.

O seguinte contraexemplo pode ser formulado, com o intuito de mostrar que apesar de um sujeito não se ver a si mesmo com um alvo justo para aplicação das atitudes reativas, ele pode ser moralmente responsável pela sua conduta. Digamos que um sujeito relativamente normal, Phil, passa um período de tempo em Londres na companhia de Ted Honderich e Galen Strawson. Vamos imaginar que Honderich persuade Phil de que o determinismo causal seja provavelmente verdadeiro. E Strawson o convence de que o determinismo é incompatível com a responsabilidade moral (MELE, 2006, p. 147). Phil converte-se de compatibilista para “determinista forte”. Ele não mais acredita na existência de candidatos justos de atitudes reativas e está convencido de que atitudes como indignação e gratidão são injustificadas. No entanto, o interessante é que a conversão filosófica de Phil não muda a sua vida na prática. Ele continua acreditando, como o fazia antes da sua conversão filosófica, que as atitudes reativas desempenham um papel importante nas práticas sociais. Ele continua doando dinheiro para as instituições de caridade, tratando bem as pessoas, educando moralmente seus filhos, recompensando a conduta generosa, e punindo a prática egoísta (MELE, 2000, p. 447-448).

Agora sobre a segunda crítica. Um agente é moralmente responsável por uma ação, para Fischer e Ravizza, na medida em que ela resulta do seu próprio mecanismo moderadamente sensível às razões (FISCHER; RAVIZZA, 1998, p. 86). Um agente faz o mecanismo seu próprio “ao assumir responsabilidade por ele” (FISCHER; RAVIZZA, 1998, p. 241). E sensibilidade moderada às razões consiste no mecanismo, na sequência real, ser regularmente receptivo às razões e fracamente reativo às razões. É a condição de reatividade às razões que é alvo da segunda crítica. Reatividade às razões significa ter a “capacidade de *transformar* razões em escolhas (e no comportamento subsequente)” (FISCHER; RAVIZZA, 1998, p. 69). Para o mecanismo do agente ser fracamente reativo às razões, a seguinte exigência precisaria ser satisfeita: no mundo real, o agente faz *A* e, supondo a existência de algum mundo possível com as mesmas leis da natureza em que um mecanismo desse tipo é operativo, “existe uma razão suficiente para fazer de outro modo, o agente reconhece essa razão, e o agente faz de outro modo” (FISCHER; RAVIZZA, 1998, p. 63).

O problema é que essa condição suficiente (o mecanismo ser fracamente reativo às razões) seria muito fraca. Esse ponto pode ser elucidado por meio de um exemplo formulado por Mele:

A agorafobia de Fred é tão forte que ele não tem se aventurado a sair da sua casa nos últimos dez anos, apesar das muitas tentativas da sua família de persuadi-lo e de muitos incentivos. Devido ao seu medo, ele frequentemente decidiu não fazer coisas que ele acreditava que ele tinha boas razões para fazer (e então se comportou de acordo), incluindo algumas coisas que ele acreditava que ele moralmente deveria fazer. Por exemplo, ele recentemente decidiu ficar em casa (e ficou) em vez de comparecer no casamento da sua amada filha, na igreja ao lado. Agora em algum mundo possível com as mesmas leis, houve um incêndio na casa de Fred, no dia do casamento da sua filha. Fred, afinal, tem mais medo do incêndio do que sair de casa. Julgando que ele tem uma boa razão para deixar a sua casa, ele decidiu sair de casa, por essa razão; e então fazendo um esforço heroico, ele caminhou até a igreja, ao lado dela (MELE, 2000, p. 450).

O mecanismo de Fred é fracamente reativo, regularmente receptivo às razões, e é o seu próprio mecanismo. Sendo assim, Fred é moralmente responsável por não ir ao casamento da sua filha, no mundo real, e Phil não é moralmente responsável por tais coisas como ações de caridade e mentira. Essa seria uma consequência inapropriada, conforme Mele, pois Fred tem um problema psicológico. Fred não é moralmente responsável pela sua agorafobia, e por não ter ido ao casamento da sua filha, pois o seu medo é tão debilitante que o incêndio o move a decidir a sair de casa intencionalmente. Já com respeito a Phil não haveria nada que justificasse a não responsabilidade pelas suas ações. A réplica consiste em afirmar que, por um lado, se o indivíduo não é responsável pela fobia e suas escolhas ou ações resultam de impulsos irresistíveis, então, intuitivamente, ele não é moralmente responsável pelo seu comportamento (FISCHER; RAVIZZA, 2000, p. 471). Por outro lado, se o indivíduo é como Fred, então “os seus impulsos são fortes, mas não irresistíveis”. Além disso, dizer que Fred é moralmente responsável não implicaria em considerá-lo censurável por ficar em casa no dia do casamento da sua filha:

Nós estamos inclinados a dizer que ele é diferente de alguém cujos impulsos são genuinamente irresistíveis (ou, por exemplo, alguém cujo cérebro está sendo significativamente manipulado para assegurar que ele se comporte de uma certa forma). O modo que nós marcamos essa distinção é dizer que Fred é moralmente responsável, enquanto que alguém sujeito à manipulação ou alguém com impulsos irresistíveis não é moralmente responsável (FISCHER, RAVIZZA, 2000, p. 471).

Casos como esses são difíceis, pois existem vários tipos de fobias, medos e vícios. Independentemente disso, haveria uma distinção importante aqui. Alguém manipulado e que carece de controle se assemelha a uma marionete (controlado por alguém diferente). E alguém sujeito a impulsos irresistíveis também carece de controle, embora não necessariamente seja controlado por alguém diferente. Fred é diferente do sujeito manipulado e daquele que age por

impulso irresistível. Diferentemente dos outros, Fred é moralmente responsável. Isso não implica que Fred seja censurável, pois esse seria um juízo separado⁶⁷. Então, em um certo sentido, Fischer e Mele estão de acordo sobre esse ponto. Deveria ser enfatizado que existe uma diferença entre Fred e alguém que carece de controle devido à estimulação eletrônica do cérebro (FISCHER; RAVIZZA, 2000, p. 471-472).

Agora sobre Phil. A crítica consistia em afirmar que as atitudes de Phil e seu comportamento não indicam que ele se considera fora de controle. Embora ele acredite que as atitudes reativas sejam metafisicamente injustificáveis, ou não se veja como um candidato apropriado para elas, ele continua aplicando as atitudes reativas, pois ele se importa com o bem-estar dos outros. Portanto, ele poderia ser considerado um alvo apto para essas atitudes. Fischer reconhece que não é fácil apresentar uma réplica a esse importante desafio, especialmente relacionado a uma parte controversa da teoria, que é a natureza subjetiva. Eles oferecem o seguinte esboço de uma resposta, um esboço, eles admitem, que é incompleto. As atitudes reativas envolvidas na responsabilidade moral são reações ou respostas aos agentes cujo comportamento tem uma certa qualidade. As atitudes reativas não parecem ser apropriadas a não ser que essa qualidade esteja presente. A qualidade em questão é um tipo de “autoengajamento” e é subjetiva, tendo a ver com o modo como o agente percebe o seu comportamento. Uma analogia pode esclarecer essa qualidade. Por exemplo, com respeito ao drogadito que tem um impulso irresistível para usar a droga e age para satisfazer o seu desejo porque ele sabe que é inútil resistir, é natural dizer que ele é em certo sentido “passivo” em relação ao seu comportamento⁶⁸. Phil não é como o drogadito em todos os aspectos. Porém, ele é similar na medida em que ambos carecem de um tipo de “autoengajamento” exigido para a responsabilidade moral. Phil, por exemplo, pensa de si mesmo como um alvo inapropriado ou injusto para as atitudes reativas (FISCHER; RAVIZZA, 2000, p. 468-469).

Supondo a existência de uma sociedade hipotética com membros que não possuem atitudes reativas, poderia existir nela agentes moralmente responsáveis? Como tais seres não têm ideia do que sejam as atitudes reativas, eles não consideram a si mesmos como alvos justos

⁶⁷ Ser moralmente responsável pelo comportamento, adotando a concepção de Peter Strawson, é ser um *candidato adequado* para as atitudes reativas. Não se segue disso que a atitude reativa em questão deva ser aplicada. Na medida em que o agente é um *alvo adequado* para a atitude reativa, ele é moralmente responsável (FISCHER; RAVIZZA, 1998, p. 7-8). Nessa abordagem, alguém pode ser moralmente responsável por um ato moralmente neutro.

⁶⁸ É a intuição de Frankfurt e de Fischer que esse drogadito não é moralmente responsável por usar a droga, dado que ele não é moralmente responsável por ter adquirido o vício (FISCHER; RAVIZZA, 2000, p. 469).

de atitudes reativas. Porém, mesmo assim eles têm desejos para o bem-estar dos outros e crenças de que eles devem moralmente fazer certas coisas (MELE, 2006b, p. 293). Essa questão aponta para uma diferença entre várias explicações do conceito de responsabilidade moral. A análise de Strawson, em termos de atitudes reativas, é uma hipótese adotada no livro de Fischer e Ravizza (1998). Eles não estão certos de que seja uma análise correta do conceito. Mais especificamente, eles estão preocupados, nesse livro, em estabelecer as condições de aplicação do conceito de responsabilidade (FISCHER; RAVIZZA, 2000, p. 470).

No próximo e último capítulo dessa tese apresentarei algumas objeções feitas à teoria semicompatibilista. Também procurarei apontar os possíveis equívocos dessas críticas e as sugestões que a própria teoria estaria disposta a aceitar.

5 MECANISMOS E MANIPULAÇÃO

Nesse capítulo, eu começarei tratando da teoria de Fischer com o objetivo de caracterizar os seus conceitos centrais. Depois, apresentarei algumas objeções endereçadas a essa teoria, como o problema da individualização do mecanismo e a possibilidade de manipulação do próprio mecanismo deliberativo real do agente por meio do acréscimo de razões. Após, será vista a resposta de Fischer e um modelo distinto de liberdade baseado na sequência real dos eventos que tenta não invocar o conceito de possibilidades alternativas. Em seguida, tratarei de outra versão do argumento da manipulação que pretende desafiar o critério compatibilista de agência livre e responsável. E, por fim, discutirei a resposta a esse argumento.

5.1 A teoria de Fischer baseada em mecanismos

Responsabilidade moral está, tradicionalmente, conectada com a capacidade para responder às razões. Porém, essa conexão tem sido defendida e desenvolvida em diferentes modos dependendo da teoria adotada.

Apesar de muitas noções presentes no livro *Responsibility and Control: A Theory of Moral Responsibility*, de Fischer e Ravizza, sejam familiares de publicações mais antigas, o livro estende muitos dos seus argumentos anteriores. Por exemplo, a noção fundamental de capacidade moderada para responder às razões, no modo apropriado, não é mais considerada como suficiente, pois as ações podem resultar a partir de um mecanismo responsivo às razões, porém podem não pertencer ao indivíduo de um modo correto⁶⁹.

De acordo com a teoria de Fischer, a responsabilidade moral envolve o controle de direcionamento, e esse controle não exige possibilidades alternativas ou capacidade para fazer de outro modo (isto é, o controle regulativo). Nesse sentido, a responsabilidade moral de um agente por uma ação depende das características da sequência real resultante na ação ao invés da sequência alternativa. Partindo dos exemplos de Frankfurt, eles adotam uma abordagem baseada em mecanismos ao invés de uma abordagem baseada no agente. O mecanismo deveria ser responsivo às razões, não o agente. A tese central e a inovação no livro de 1998 seria que o agente exibe o controle de direcionamento da sua ação se o mecanismo resultante na sequência real é o seu próprio mecanismo moderadamente responsivo às razões. Esse controle seria compatível com o determinismo causal.

⁶⁹ O capítulo 7 e 8 do livro tratam de dois conceitos conectados à responsabilidade: uma noção histórica e de autoria (*ownership*).

Em artigos anteriores à publicação do livro de 1998, Fischer e Ravizza apresentaram um grau fraco e forte da capacidade para responder às razões.

A capacidade para responder fracamente às razões ocorre nas seguintes condições:

Mantendo fixo o tipo de mecanismo real, e então simplesmente é exigido que exista *algum* cenário possível (ou mundo possível) -com as mesmas leis do mundo real-no qual existe uma razão suficiente para fazer de outro modo, o agente reconhece essa razão, e o agente faz de outro modo (FISCHER; RAVIZZA, 1998, p. 63).

Para explicar esse grau de capacidade para responder às razões, vejamos o exemplo do assassino em série. Vamos imaginar que ele entre numa sala de aula e atire com a sua metralhadora contra os alunos e o professor. Digamos que há uma razão suficiente para fazer de outro modo em um mundo possível, qual seja, se um aluno no fundo da sala estivesse manuseando um celular, então essa seria a razão para que ele não atirasse. Esse seu comportamento se enquadraria no grau fraco de capacidade para responder às razões, pois estando presente a razão suficiente para fazer de outro modo ele agiria conforme essa razão.

Então, adotando o critério da capacidade fraca para responder às razões, poderia se dizer que o assassino seria moralmente responsável pela ação exercida? O problema com esse grau de capacidade para responder às razões é que com ele não se pode compreender a motivação do assassino. Como que a presença de alguém com um celular poderia contar como uma razão forte o suficiente para impedir alguém de cometer um crime? (FISCHER; RAVIZZA, 1998, p. 65). É preciso fortalecer esse grau da capacidade para responder às razões, pois uma capacidade fraca acabaria abrangendo agentes incapacitados.

Uma opção para evitar esse problema seria fortalecer a capacidade para responder às razões. Supondo que um mecanismo *K* na realidade resulte em uma ação, a capacidade para responder fortemente às razões ocorreria nos seguintes cenários:

se *K* operasse e existisse razão suficiente para fazer de outro modo, o agente *reconheceria* a razão suficiente para fazer de outro modo, e *escolheria* fazer de outro modo, e *faria* de outro modo. Em outras palavras, nas circunstâncias em que o mecanismo opera, na realidade, se existem razões suficientes para o agente fazer de outro modo, três condições devem ser satisfeitas (...): o agente deve *assumir* as razões como suficientes, *escolher* de acordo com as razões suficientes, e *agir* de acordo com a escolha. Assim, podem haver três tipos de falhas da “sequência alternativa”: falhas na conexão entre que razões existem e que razões o agente reconhece, na conexão entre as razões do agente e sua escolha, e na conexão entre a escolha e a ação (FISCHER; RAVIZZA, 1998, p. 41).

Quando existem razões suficientes para realizar uma ação, mas o agente não reconhece essas razões, pode se dizer que há uma falha em ser receptivo a essas razões. Quando essa falha ocorre, é por causa de alguma incapacidade mental, como psicose. Já quando a falha ocorre na conexão entre as razões do agente e sua escolha, então há uma falha de reatividade, de não ser apropriadamente afetado pelas crenças. O agente, nesse caso, reconhece certas razões como suficientes para a realização de uma ação, porém não escolhe conforme esse reconhecimento⁷⁰. A carência de reatividade envolve tanto o comportamento compulsivo como aquele resultante da fraqueza da vontade e da fobia. E, por último, há a possibilidade de falha na conexão entre a escolha e a ação. Aqui se encaixaria novamente o comportamento resultante da fraqueza da vontade e de incapacidades físicas.

Supondo que nenhuma dessas falhas ocorresse no cenário alternativo, e o mecanismo *K* operasse, então o mecanismo operativo na realidade seria fortemente responsivo às razões. Fischer e Ravizza acreditam que a conexão forte entre as razões e as ações é sem dúvida desejável, porém eles acreditam que essa capacidade para responder fortemente às razões não seja uma condição necessária para o controle de direcionamento e a responsabilidade moral, pois, do contrário, praticamente ninguém agiria livremente ou seria moralmente responsável pelo o que fizesse (FISCHER; RAVIZZA, 1998, p. 42).

Para ter o controle de direcionamento seria necessário que o mecanismo resultante no comportamento seja apropriadamente responsivo às razões. Visto que tanto a capacidade fraca para responder às razões como a forte seriam inapropriadas, então o mecanismo adequado deveria ser moderadamente responsivo às razões (FISCHER, 2006, p. 318).

A capacidade moderada para responder às razões é constituída por dois elementos: (i) moderadamente receptivo às razões e (ii) fracamente reativo às razões. Ser receptivo às razões significa ser capaz de reconhecer quais razões contam como suficientes para a ação. E ser reativo às razões significa ser capaz de reagir às razões de um modo apropriado. Uma das modificações na estratégia dos autores consiste em estipular uma condição de competência moral, pois “o tipo de capacidade para responder às razões que é exigido para a responsabilidade moral deve ser caracterizado não meramente como uma capacidade para responder às razões, mas como uma capacidade para responder a um conjunto de razões que incluam razões *morais*”

⁷⁰ A noção de “razão suficiente” utilizada pelos autores significa que quando um agente reconhece uma razão como suficiente, isso quer dizer que essa razão justifica um certo curso de ação. Nesse sentido, razão suficiente estaria associada à ideia de *justificação*, de que uma razão é mais forte ou melhor para uma determinada ação (FISCHER; RAVIZZA, 1998, p. 41n13).

(FISCHER; RAVIZZA, 1998, p. 81). Dito de outro modo, a competência moral implica que o agente deveria ter a capacidade para reconhecer as razões morais.

A caracterização do conceito de receptividade e de reatividade é indicada não apenas por preocupações sobre a competência moral. Fischer e Ravizza estariam também preocupados sobre a sanidade dos agentes. O exemplo do assassino representa o grau fraco da capacidade para responder às razões de um modo bizarro, na medida em que seu comportamento é “evidência da sua insanidade e conseqüente carência de responsabilidade” (FISCHER; RAVIZZA, 1998, p. 65). Esse caso é enigmático, pois não podemos entender a sua motivação se a presença de alguém manuseando um celular deveria contar como a única razão forte o suficiente para impedi-lo de atirar com a sua arma⁷¹.

Em resumo, uma das condições para que o comportamento do agente possua o controle de direcionamento é que ele resulte a partir de um mecanismo moderadamente responsivo às razões. A capacidade moderada para responder às razões implica uma assimetria entre os elementos de receptividade e de reatividade: “enquanto responsabilidade moral exige que o agente aja a partir de um mecanismo que mostre *regularidade* em reconhecer as razões, essa mesma demanda não pode ser feita com respeito à reatividade” (FISCHER; RAVIZZA, 1998, p. 81). O agente deveria agir a partir de um mecanismo que é regularmente receptivo às razões e fracamente reativo à elas para poder apropriadamente ser moralmente responsável.

O agente deveria exibir um padrão apropriado de reconhecimento de razões. A seguinte citação explica como esse padrão deve ser entendido:

nós queremos saber se (quando agindo a partir de um mecanismo real) ele reconhece como as razões se ajustam, consegue distinguir que uma razão é mais forte do que a outra, e entende como a aceitação de uma razão como suficiente implica que uma razão mais forte deve também ser suficiente (FISCHER; RAVIZZA, 1998, p. 71).

Dito de outro modo, para que o agente seja responsável o seu comportamento deveria ser resultante de um mecanismo regularmente receptivo e fracamente reativo às razões e que exiba um padrão adequado de reconhecimento de razões. Além disso, a reatividade *fraca* às razões “é toda a reatividade exigida para a responsabilidade moral” (FISCHER; RAVIZZA, 1998, p. 73), pois “se o mecanismo de um agente reage a *algum* incentivo para fazer de outro

⁷¹ É necessário também que a capacidade para responder às razões mostre um “padrão compreensível de reconhecimento de razões minimamente baseadas na realidade” (FISCHER; RAVIZZA, 1998, p. 73). Essa exigência serve para indicar que é indispensável que o agente não esteja iludido sobre a natureza da realidade.

modo do que ele na realidade faz, isso mostra que o mecanismo *pode* reagir a *qualquer* incentivo para fazer de outro modo” (FISCHER; RAVIZZA, 1998, p. 73).

Fischer e Ravizza acreditam que os exemplos de Frankfurt ajudam a focar mais cuidadosamente nas características da *sequência real resultante na ação*. Em outras palavras, esses exemplos auxiliam no desenvolvimento de uma explicação da sequência real que leva a “uma abordagem para a responsabilidade moral que não exige possibilidades alternativas” (FISCHER; RAVIZZA, 1998, p. 37). Em contraste com as teorias tradicionais, o modelo da sequência real defende que o importante é o que o agente faz na realidade. Na perspectiva dos autores, seria inútil desenvolver uma interpretação desse modelo deslocando o foco do *agente* e suas propriedades para o *processo* ou o *mecanismo* que na realidade leva à ação. Partindo de um exemplo de Frankfurt, por exemplo, o *agente* não poderia ter feito de outro modo, e portanto, ele não seria responsivo às razões. Porém, o *mecanismo* a partir do qual ele age na realidade seria responsivo às razões, embora o mecanismo que teria operado no cenário alternativo não seja responsivo às razões (FISCHER; RAVIZZA, 1998, p. 38).

Apesar das modificações nos conceitos centrais da teoria de Fischer, algumas críticas permanecem em relação à proposta atualizada. A seguir tratarei da crítica feita por Gary Watson.

5.2 Objeções à teoria baseada em mecanismos

Apesar das revisões feitas por Fischer com respeito à capacidade para responder às razões, alguns problemas ainda permaneceriam. Gary Watson, por exemplo, aponta uma preocupação sobre a tese da assimetria, e oferece uma justificativa para a afirmação de que o componente da reatividade fraca seria insuficiente para contar como uma condição para o controle de direcionamento e responsabilidade moral. A crítica refere-se a uma situação envolvendo um usuário de drogas. A ideia é supor que a única coisa que impediria uma pessoa de usar a droga seria, devido à sua fobia por ratos, se a droga fosse colocada em uma gaiola cheia de ratos. O que a crítica quer mostrar, por meio desse caso, é que a noção de controle não poderia ser identificada com a sujeição de algum contraincentivo (WATSON, 2001, p. 378-379).

Fischer e Ravizza poderiam responder à objeção levantada por Watson afirmando que a presença de certos incentivos alternativos daria origem a diferentes mecanismos. A resposta consistiria em afirmar que o mecanismo operante no cenário alternativo, quando a droga é posta na gaiola e há a presença do medo, difere do mecanismo operante no cenário real, onde há o vício pela droga. Porém, o problema é que não seria suficiente para o controle de

direcionamento que o mecanismo responda diferentemente a uma razão suficiente para fazer de outro modo em um cenário alternativo. É necessário que o mesmo tipo de mecanismo possa reagir diferentemente na presença de razões reais contrárias (FISCHER; RAVIZZA, 1998, p. 73-74).

Para Watson, é difícil de avaliar essa resposta porque ela depende de questões obscuras sobre a identidade de mecanismos (WATSON, 2001, p. 379). Fischer e Ravizza admitem que eles “não podem especificar qual mecanismo é ‘o’ mecanismo relevante para a avaliação da responsabilidade” (FISCHER; RAVIZZA, 1998, p. 47). É uma pressuposição da teoria “que para cada ato existe intuitivamente um mecanismo natural que é apropriadamente selecionado como o mecanismo resultante na ação, para os propósitos de avaliar o controle de direcionamento e a responsabilidade moral” (FISCHER; RAVIZZA, 1998, p. 47).

Watson lança dúvidas sobre o apelo à intuição para a explicação de qual seria o mecanismo relevantemente conectado com o controle de direcionamento e a responsabilidade moral (WATSON, 2001, p. 379). Esse ponto pode ser esclarecido apresentando três exemplos. O primeiro deles eu cito:

Goldie vota em Nader. A ação de Goldie é justificada pela razão de que os outros candidatos são corruptos. Nesse caso, o mecanismo é o processo de deliberação sobre qual candidato apoiar. O teste para saber se o processo é reativo às razões consiste em imaginar um cenário alternativo em que (a) o mecanismo real opera, (b) existe razão suficiente para fazer de outro modo e (c) o agente faz de outro por essa razão. Como que o comportamento de Goldie satisfaz esse teste? Supondo que minutos antes da votação o partido que apoia o candidato oponente fornece informações sobre Nader ao amigo de confiança de Goldie, o qual o influencia a não votar em Nader. Nesse cenário alternativo, é razoável dizer que o mesmo mecanismo, isto é, a deliberação em relação a quem votar, resulta em um diferente comportamento. O processo é responsivo às razões (WATSON, 2001, p. 379-380).

Esse exemplo é o próprio Watson que fornece com o intuito de apontar que o mesmo mecanismo de deliberação resulta em ações diferentes.

Agora sobre o segundo exemplo. Suponha que Brown seja incontinente e que tenha um forte desejo pela droga não viciante Plezu. Embora “ele reconheça que existam fortes razões para *não* usar a droga toda manhã (...) o único cenário em que ele não a usaria seria aquele em que se ele injetasse mais uma vez teria como grave consequência a sua morte” (FISCHER; RAVIZZA, 1998, p. 69-70). Como o seu comportamento é fracamente reativo, ele seria responsável por ter usado a droga.

Segundo Watson, esse exemplo é problemático. Qual é o processo ou o mecanismo que explica por que ele usa a droga? Nesse caso seria o seu desejo de ficar drogado. Em que sentido

esse mecanismo operaria no cenário contrafactual em que ele se abstém de usar a droga por causa do medo da morte repentina? Nesse cenário contrafactual o medo tornaria ineficaz o desejo de usar a droga. Então não se poderia dizer que *o mesmo* mecanismo opera em ambos os cenários, pois no cenário real o mecanismo operante seria o desejo de usar a droga e no cenário alternativo ele não usa a droga por causa do medo. Reiterando, o desejo de ficar drogado não poderia ser o processo ou o mecanismo que explicaria por que o agente, nesse caso, recusa-se a usar a droga (WATSON, 2001, p. 380).

A solução para essa dificuldade consiste para Watson em destacar que o mecanismo que opera em ambos os cenários, real e alternativo, seria a reflexão sobre razões. Aceitando esse modo de interpretar a questão, então esse exemplo não seria diferente daquele da votação. No entanto, talvez Fischer não veja dificuldade nesse ponto, pois o exemplo serviria meramente para explicar o comportamento do incontinente, do agir contrariamente à razão prática.

Agora vejamos o terceiro exemplo tratando do comportamento habitual. Alguém que esteja dirigindo para o trabalho usa sempre a avenida. Ele não delibera sobre qual caminho deveria pegar, pois ir pela avenida faz parte da sua rotina. O exemplo serve para mostrar que o agente pode ser legitimamente moralmente responsável pelo comportamento apesar dele não ser resultante da reflexão prática ou deliberação.

Para avaliar se o comportamento do agente é responsável é preciso fixar o mecanismo da sequência real e questionar se em cenários hipotéticos o agente age diferentemente diante de razões contrárias. Nesse caso, o que é fixado é o fato do “agente não delibera o que fazer”. Aplicando o teste digamos que o caminho pela avenida esteja bloqueado e foi interrompido por causa de obras. É possível que existam cenários em que o motorista automaticamente responde pegando a próxima saída, sem questionar-se ou mesmo deliberar qual caminho seguir. Esse exemplo, segundo Fischer e Ravizza, “ajuda mostrar que a ação resultante na sequência real (a partir de um mecanismo não reflexivo) é moderadamente responsivo às razões” (FISCHER; RAVIZZA, 1998, p. 86).

Watson acredita que é razoável aceitar a conclusão de Fischer e Ravizza desde que seja feita uma correção no processo que explica o comportamento do agente. Para Watson, o ideal seria substituir o mecanismo não reflexivo por um termo mais abrangente, como um certo tipo de capacidade de monitoramento que é dirigida pelo objetivo do agente de chegar em um determinado lugar (WATSON, 2001, p. 381). Seria a capacidade de monitoramento que deveria estar presente em ambos os cenários, real e alternativo. Quando os eventos se desdobram de acordo com a rotina, o comportamento do agente decorreria a partir do hábito. Já quando os eventos fogem da rotina, o agente pode exercer a sua capacidade para fazer os ajustes

necessários. A correção do termo feita por Watson abrangeria tanto comportamentos resultantes de mecanismos reflexivos como de não reflexivos.

Carl Ginet chama a atenção para uma situação em que um agente percebendo que um certo caminho que ele costumava passar está bloqueado sente-se forçado a refletir e tomar uma decisão sobre o que ele deveria fazer, se entrar na próxima saída ou se fazer o retorno e procurar uma entrada anterior. Seria anormal se um agente tivesse razão suficiente para fazer de outro modo que não fosse a ação habitual e fosse incapaz de refletir sobre as alternativas (GINET, 2006, p. 234-235).

Para determinar se um agente seria moralmente responsável pelo seu comportamento, de acordo com Fischer, é preciso avaliar se a sua ação seria resultante de um mecanismo apropriadamente responsivo às razões, dando ao agente uma certa disposição complexa. Ginet acredita que Fischer deveria substituir o conceito de mecanismo pelo conceito de agente para evitar o problema da individualização do mecanismo. O ideal seria afirmar que o agente deve ser responsivo às razões e não o mecanismo. Por que não exigir apenas que o agente tenha essa disposição visto que o mecanismo não parece estar trazendo alguma contribuição para a teoria? (GINET, 2006, p. 235-236)

Watson apontou um desafio que explora a ideia central da teoria de Fischer e Ravizza:

É curioso que Fischer e Ravizza sintam a necessidade de fazer uso dessa afirmação modal. A objeção referente à justiça parece surgir de intuições que apoiam um princípio das possibilidades alternativas (responsabilizar as pessoas é injusto a não ser que elas poderiam ter feito de outro modo). Fischer e Ravizza rejeitam esse princípio por causa dos chamados casos de Frankfurt, em que algum dispositivo permanece à espera para assegurar que um indivíduo se comporte em um certo modo. Por exemplo, supondo que se Goldie fosse mudar de ideia no último momento e estivesse prestes a votar no candidato do partido Verde, o dispositivo teria assegurado que ele escolhesse votar em Nader. Assim, não existe possibilidade de que ele não escolhesse esse candidato. Fischer e Ravizza razoavelmente concluem que esse fato modal não implica que o seu comportamento real de votar não seja responsivo às razões. Isso conduz a eles a rejeitar a ideia que para ser responsável o agente deve ter alternativas ao que ele faz. Fischer e Ravizza afirmam que nos casos de Frankfurt os agentes não poderiam ter respondido diferentemente diante de incentivos contrários, mas o mecanismo operativo na realidade poderia ser responsivo às razões. O que é curioso, então, é que Fischer e Ravizza parecem sentir a necessidade de empregar uma noção de possibilidades alternativas no nível dos mecanismos. Eles parecem estar concedendo, em algum sentido, que a justiça para responsabilizar alguém depende da capacidade do mecanismo em questão responder diferentemente, uma capacidade que deve ser compatível com o determinismo causal. Mas é difícil ver como essa proposta poderia responder às preocupações sobre a justiça, a não ser que em vez de falar sobre as capacidades dos mecanismos nós púdessemos falar sobre o que as pessoas podem fazer. E se nós púdessemos fazer isso, nós deveríamos endossar uma versão compatibilista do princípio das possibilidades

alternativas e não rejeitar a sua relevância completamente (WATSON, 2001, p. 382).

O que Watson quer enfatizar é que embora Fischer e Ravizza rejeitem o princípio das possibilidades alternativas, utilizando-se dos exemplos de Frankfurt, eles argumentam que o mecanismo deveria ser responsivo às razões, isto é, o mecanismo deveria ter uma capacidade para responder diferentemente.

Agora eu retorno à questão da individualização do mecanismo. Mckenna acredita que a teoria de Fischer estaria sujeita à crítica de não conter uma explicação explícita da individualização do mecanismo, fato que é inclusive reconhecido pelos próprios autores como sendo um potencial problema. Para Mckenna, é indispensável que haja uma explicação de como *o mesmo mecanismo* se comporta em resposta às diferentes razões (MCKENNA, 2013, p. 161).

Fischer e Ravizza fazem uso do apelo à intuição para explicar a individualização do mecanismo, o que, segundo o crítico, torna a defesa da teoria como um todo não convincente:

Porque eles (Fischer e Ravizza) não oferecem uma base de *princípios* para individualização do mecanismo, eles devem fazer uso da tese puramente a partir de reações *intuitivas* sobre diferentes casos. Mas, como poderia ser objetado, *quais* elementos a partir do inteiro complexo (de eventos próximos e estados antecedentes à ação) deveriam figurar intuitivamente no mecanismo relevante variará em relação à perspectiva explanatória. (...) Que razão nós teríamos para assumir que a base para individualização de Fischer e Ravizza é a correta? A situação só piora se nós reivindicarmos uma noção hiper-restrita de uniformidade do mecanismo. Na interpretação hiper-restrita, o inteiro complexo de eventos antecedentes e estados funcionaria como o mecanismo pertinente. Se esse fosse o mecanismo relevante, o agente não poderia agir a partir de um mecanismo responsivo às razões em um mundo determinista (MCKENNA, 2001, p. 97).

Vamos imaginar a possibilidade de mundos possíveis. Aplicando a individualização do mecanismo teríamos que assumir que tanto no mundo real como no possível os mesmos fatos deveriam ser compartilhados. O problema surgiria caso se assumisse a verdade da tese determinista no mundo real, pois não poderia existir um mundo possível que compartilhasse os mesmos fatos do mundo real até um certo tempo e que, em um tempo subsequente, diferiria pela existência de um agente que age no mundo possível diferentemente de como ele age no mundo real (MCKENNA, 2001, p. 97).

Já adotando uma noção hiper-restrita, o incompatibilista poderia fazer sentido de um mecanismo responsivo às razões. Apenas em um mundo indeterminista poderia ser o caso que, com respeito a um mesmo mecanismo operando no mundo real e no possível, o agente poderia,

no mundo possível, agir de outro modo na presença de diferentes razões (MCKENNA, 2001, p. 97-98).

Uma teoria baseada em mecanismos seria problemática porque a explicação de por que uma ação ocorreu envolveria um complexo de eventos e estados antecedentes. Esses fatores explanatórios incluiriam desejos, crenças e intenções que estariam entre as razões que explicariam por que um agente agiu de um determinado modo⁷². O problema decorreria do fato de que Fischer e Ravizza estariam adotando a seguinte metodologia para avaliar se uma certa ação resultante de um mecanismo possuiria o controle de direcionamento: eles estariam partindo da hipótese de que teria acontecido se diferentes razões fossem apresentadas ao agente mantendo fixas certas características relevantes ou apropriadas do processo real que levou à ação. A questão é: qual seria o critério que determinaria quais seriam as características apropriadas que deveriam ser fixadas? (GINET, 2006, p. 233)

A teoria estaria pressupondo apenas que “para cada ato existe intuitivamente um mecanismo natural que é apropriadamente selecionado como o mecanismo resultante na ação, para os propósitos de avaliar o controle de direcionamento e a responsabilidade moral” (FISCHER; RAVIZZA, 1998, p. 47). Para Ginet, a intuição não acrescentaria nada para o nosso entendimento para determinar quais seriam as características relevantes. Tudo o que Fischer e Ravizza estariam sugerindo é uma distinção entre mecanismos que são responsivos às razões e mecanismos que não são responsivos às razões, e entre mecanismos responsivos às razões morais de mecanismos que são responsivos apenas às razões prudenciais. No entanto, essa explicação, segundo Ginet, não estaria individualizando os mecanismos (GINET, 2006, p. 234).

Existe uma outra dificuldade: como compatibilizar uma teoria que apela para a capacidade do mecanismo ser responsivo às razões com os exemplos de Frankfurt? Para lembrar, nesses exemplos, da perspectiva de Fischer e Ravizza, a liberdade do agente não poderia ser explicada em termos da capacidade para responder às razões. Devido à presença do interventor contrafactual, o agente não agiria de outro modo mesmo se diferentes razões fossem apresentadas a ele. Nesse caso, a hipótese adotada pelos autores consistiu em elaborar uma teoria baseada no mecanismo. Em um exemplo de Frankfurt o agente não seria responsivo às razões, mas o mecanismo a partir do qual ele age na realidade, que é o seu próprio mecanismo de ação, seria responsivo às razões (FISCHER; RAVIZZA, 1998, p. 38).

⁷² Existem diversas explicações das ações em termos de seus antecedentes: (a) a ação foi causalmente determinada (necessitada) e (b) a ação foi não causada. Tanto (a) como (b) podem ser compatíveis com as possibilidades alternativas no momento da ação (GINET, 2006, p. 233).

Para avaliar essa proposta, Mckenna acredita que precisaríamos responder à seguinte questão: por que o deslocamento para o mecanismo ajudaria a resolver o possível conflito existente entre as duas teses compatibilistas, quais sejam, uma teoria baseada na capacidade para responder às razões e a carência dessa mesma capacidade vinculada ao argumento de Frankfurt? (MCKENNA, 2013, p. 159)

Partindo de um exemplo de Frankfurt, para testar se esse mecanismo resultante na ação na realidade é responsivo às razões seria preciso imaginar a existência de mundos possíveis nos quais a presença do interventor é subtraída. Para explorar esse ponto podemos imaginar o caso de um assassinato cometido por Jones. Suponha que Black não estivesse presente nesse cenário e fosse apresentado a Jones razões pertinentes para não atirar em Smith. A partir do seu mecanismo de ação, Jones reagiria de outro modo e não atiraria em Smith. Aplicando a teoria de Fischer, Jones estaria agindo a partir de um mecanismo apropriadamente responsivo às razões (MCKENNA, 2013, p. 160).

Uma implicação clara e incontroversa da explicação oferecida por essa teoria é que quando Jones age por suas próprias razões na sequência real, o que acontece é muito diferente do mecanismo da sequência alternativa, pois tivesse Black interferido e causado Jones a fazer a ação que ele quisesse, na sequência alternativa, Jones teria carecido de liberdade e responsabilidade moral.

Em um certo momento, Mckenna entende que Fischer e Ravizza deveriam ter oferecido um modo correto de individualizar o mecanismo de ação para que essa posição fosse completa (2001). No entanto, em publicação posterior, ele percebe que existe um problema estrutural com a teoria dos autores que os impede de oferecer qualquer conteúdo específico para as condições de individualização do mecanismo da ação (MCKENNA, 2013, p.163). O autor estaria se referindo aqui ao problema que Watson identificou sobre a fixação do mecanismo de hábito não refletido, visto anteriormente.

Desde a publicação do influente livro de Fischer e Ravizza, *Responsibility and Control* (1998), tem se assumido que o único modo de conciliar uma teoria de liberdade responsiva às razões com a rejeição do controle regulativo fazendo uso do argumento de Frankfurt seria explicando a liberdade em termos de uma teoria baseada em mecanismos. Contudo, devido às dificuldades envolvidas em explicar a individualização do mecanismo, Mckenna propõe uma teoria compatibilista de liberdade responsiva às razões baseada não no mecanismo, mas no agente (MCKENNA, 2013, p. 176).

É preciso enfatizar que para Fischer e Ravizza não seria possível que uma teoria responsiva às razões baseada no agente fosse capaz de se ajustar aos exemplos de Frankfurt e

de explicar o controle de direcionamento, pois esses exemplos excluiriam a capacidade para fazer de outro modo. As teorias responsivas às razões baseadas no agente exigem que os agentes que são adequadamente responsivos às razões sejam capazes de reagir e agir diferentemente.

Também é importante destacar que para Mckenna seria possível conciliar o argumento de Frankfurt com a noção de controle de direcionamento defendido por Fischer e Ravizza, e desenvolver uma teoria responsiva às razões sem a necessidade de contar com o conceito problemático de mecanismos. Sua proposta consiste em apresentar argumentos para mostrar como o agente pode ser responsivo às razões no contexto de um exemplo de Frankfurt.

Vimos que para Fischer e Ravizza essa conciliação não seria possível visto que o agente dentro de um cenário de Frankfurt não seria responsivo às razões. O agente, nesse contexto, seria incapaz de agir ou reagir diferentemente mesmo se diferentes razões fossem apresentadas a ele devido à presença do interventor que estaria preparado para intervir. Já Mckenna acredita que a capacidade para responder às razões poderia ser explorada de outra forma. Em vez de falar que o agente não é responsivo às razões porque ele careceria da liberdade para agir diferentemente, Mckenna propõe que o agente dentro de um contexto de Frankfurt seria adequadamente responsivo às razões, apesar de ser incapaz de reagir diferentemente, quando possui razões suficientes para reagir.

Mckenna estabelece a seguinte distinção que servirá como suporte para a sua proposta. Ser reativo às razões suficientes para agir de outro modo é distinto de ser capaz de reagir de outro modo quando confrontado com razões suficientes. Podemos ilustrar esse ponto através do seguinte exemplo. Vamos imaginar Jones fora de um contexto de Frankfurt. Jones atira e mata Smith por suas próprias razões (vingança). Digamos que a razão pela qual Jones não atiraria em Smith seria se o filho de Smith estivesse com ele. Se Jones tivesse conhecimento dessa razão ele seria reativo e reagiria de acordo com ela, a despeito da sua razão real de vingança, dessa forma não atirando em Smith. Agora imaginando Jones no contexto de Frankfurt, se a presença do filho de Smith se tornasse relevante na tomada de decisão de Jones, então Black interferiria. Nesse caso, Jones seria incapaz de reagir de outro modo por causa de Black. Mas Jones seria reativo a essa razão, pois ser reativo nessa circunstância significaria que quando na presença dessa razão o agente não age por suas próprias razões de vingança (MCKENNA, 2013, p. 171-172).

Agora sobre a resposta que Fischer oferece para algumas das objeções. Ele concorda que uma tese compatibilista deveria ter uma explicação que contasse com um princípio de individualização do mecanismo. Além disso, ele admite que uma teoria carente dessa explicação seria incompleta. Porém, não seria razoável esperar que alguém pudesse apresentar

a defesa de uma tese altamente controversa sobre a vontade livre em que todos os seus componentes fossem discutidos de forma incontestável por meio de princípios sem qualquer apelo a intuições: “como alguém poderia apresentar uma explicação puramente baseada em um princípio de individualização do mecanismo - uma explicação que não apele, em algum nível, às intuições” (FISCHER, 2004, p. 167-168).

Seria importante chamar atenção para o fato de que qualquer teoria que envolva generalidade parece estar sujeita a algum problema em algum nível. A teoria ética de Kant é um exemplo de teoria generalizante. Poderia se questionar se todos os agentes poderiam agir em certos modos, motivados por certos tipos de máximas. Numa teoria como a de Kant não haveria também uma explanação de individualização dos tipos relevantes de máximas e em algum nível alguém aplicando essa teoria deverá fazer uso das suas intuições. Tomando como exemplo o caso de uma pessoa que faz uma doação para uma instituição de caridade. Ela faz essa doação e sua máxima resume-se à promoção do bem-estar das outras pessoas. Entretanto, também poderia se incluir na descrição da máxima a seguinte afirmação “doar bens promove uma vida de bem-estar de acordo com Deus”. Aparentemente, se poderia tanto incluir como excluir elementos na descrição da máxima, o que implicaria que não haveria um princípio ou uma regra que determinasse se a ação exercida foi ou não moralmente correta.

O desafio imposto pela noção da individualização do mecanismo levanta questões metodológicas. Uma característica da metodologia empregada por Fischer e Ravizza talvez ajude a oferecer uma resposta para as preocupações levantadas por Mckenna sobre individualização do mecanismo, além de esclarecer a teoria defendida pelos autores:

Nós objetivamos dar o que Robert Nozick chamou de “explicações filosóficas”, não fazer “filosofia coerciva”. Isto é, nós queremos mostrar que é muito plausível e apelante dizer que (por exemplo) os agentes podem ser moralmente responsáveis pelo seu comportamento independentemente da verdade ou da falsidade do determinismo. E nós tentaremos mostrar *como* esse tipo de visão pode ser desenvolvida e defendida. Mas não estamos supondo que nós púdessemos dar um argumento arrebatador para essa conclusão (ou outra controversa maior do livro). Assim, quando nós afirmamos que nós argumentamos sobre o sucesso da compatibilidade do determinismo causal e responsabilidade moral nós estamos afirmando que temos oferecido um argumento plausivelmente forte para essa conclusão, mas não um argumento que qualquer agente racional é compelido a aceitar (FISCHER; RAVIZZA, 1998. p. 11).

Dito de outro modo, Fischer e Ravizza estariam tentando elaborar uma sistematização com o intuito de chegar a um consenso compartilhado na nossa sociedade, tornando possível distinguir os casos que estariam enfraquecendo a nossa liberdade e responsabilidade moral,

como os casos envolvendo hipnose, lavagem e manipulação cerebral, dos casos em que esses fatores não estão presentes. Eles não estão pressupondo que poderiam provar a verdade da tese compatibilista, ou mesmo convencer um incompatibilista da verdade dos argumentos em favor do compatibilismo. Contudo, seria um trabalho significativo mostrar por que seria desejável o compatibilismo e por que a compatibilidade entre o determinismo causal e a responsabilidade moral teria menos custos ou comprometimentos se comparado com o problema de como compatibilizar o determinismo com a liberdade ou a capacidade para agir diferentemente (FISCHER, 2004, p. 171).

Todd Long, por exemplo, não vê problemas com a abordagem de Fischer e Ravizza da responsabilidade moral baseada em mecanismos, exceto no que diz respeito à aplicação da teoria aos exemplos de tipo Frankfurt. Long elabora um novo exemplo no mesmo estilo usado por Frankfurt com o intuito de tentar compreender como a teoria poderia lidar com certos tipos de manipulação ou indução artificial de estados mentais. Vejamos o seguinte caso:

Schmidt é um oficial de alto escalão no governo alemão antes da ascensão de Hitler ao poder ditatorial. Tendo sido criado por pais com um profundo senso de moralidade e determinação, Schmidt aprendeu desde cedo a diferença entre o que é certo e errado. À medida em que ele cresceu, desenvolveu a sua sabedoria moral. Basta dizer que Schmidt se tornou um homem de impecável juízo moral. Devido ao seu cargo no alto escalão do governo alemão, ele é membro de um grupo secreto que votará para decidir conceder ou não poder supremo a Hitler. Schmidt conhece Hitler há anos. Os dois se tornaram amigos quando estavam iniciando como jovens políticos. Durante aqueles anos eles compartilharam um com o outro os seus sonhos e ideias sobre como construir uma Alemanha melhor. Schmidt admirava tanto a inventividade das ideias de Hitler quanto o profundo senso de moralidade que elas exibiam. Dado o que Hitler disse a Schmidt naqueles anos, Schmidt tinha razões muito boas para pensar que Hitler seria um membro honesto, sincero e moralmente íntegro no futuro da Alemanha. No entanto, alguns anos antes da data da votação, Schmidt passa a ouvir histórias perturbadoras sobre Hitler. Ele ouve de fontes confiáveis que Hitler tem intenções de construir uma raça superior de arianos de sangue puro, expulsando todos os que não são arianos do país, e talvez assassinando aqueles que não saírem. Schmidt acaba tendo profundas dúvidas sobre o caráter moral de Hitler. O papel de Schmidt na decisão é de dar o voto decisivo se os membros votantes chegarem a um impasse. Os membros de fato chegam a um empasse, e Schmidt é assim convocado para dar o voto decisivo. Schmidt delibera. Aguardando a sua decisão está Block, um ser malevolente com poderes causais e preditivos extraordinários. Se ficar claro para ele que Schmidt votará contra Hitler, ele intervirá para garantir que Schmidt vote em favor de Hitler. Do contrário, Block não intervirá. Assim, independentemente das preferências e inclinações iniciais de Schmidt, Block terá o que ele quer (LONG, 2004, p. 157-158).

Apesar das adaptações, esse exemplo não é diferente do já tratado nos capítulos anteriores desta tese. O que a reflexão de Long coloca como novo é a constatação de que haveria duas possibilidades para nós interpretarmos esse exemplo. Primeiro, Schmidt poderia deliberar e votar em Hitler⁷³. Conforme Long, seria plausível que Schmidt fosse moralmente responsável e merecedor de censura pela sua escolha e essa possibilidade poderia ser acomodada pela teoria de Fischer, visto que o agente satisfaria as exigências da capacidade para responder apropriadamente às razões. Segundo, seria imaginar Schmidt tendo uma certa disposição de caráter. Se Block não interferisse, então Schmidt não votaria em Hitler. É pertinente enfatizar que Long divide essa segunda possibilidade de interpretação do exemplo em dois cenários possíveis da sequência alternativa. Desses dois cenários, apenas no primeiro o processo de manipulação estaria eliminando a responsabilidade moral. Eu cito esse primeiro cenário mostrando como que Block pode intervir para que Schmidt fizesse o que ele quer:

Caso 1: Block percebe que Schmidt não vai votar em Hitler um minuto antes dele tomar a decisão. Nesse caso, Block entra em ação manipulando diretamente o sistema nervoso de Schmidt fazendo com que ele causalmente e fisicamente vote em Hitler (LONG, 2004, p. 158-159).

Aqui a manipulação contorna o mecanismo deliberativo normal de Schmidt. Há dois pontos que merecem ser destacados. Primeiro, de acordo com Long, a teoria de Fischer e Ravizza implicaria corretamente que o agente nesse caso não seria moralmente responsável, pois não haveria uma conexão adequada entre o comportamento do agente com as suas próprias razões, de modo que o mecanismo em questão não seria adequadamente responsivo às razões e nem seria de Schmidt. Segundo, a teoria de Fischer e Ravizza explicaria esse caso da seguinte forma: partindo dos exemplos de Frankfurt, o mecanismo deliberativo normal do agente, na sequência real, seria diferente do mecanismo manipulativo, na sequência alternativa. Schmidt não seria moralmente responsável porque sua ação resultou de um processo não responsivo às razões.

Block também poderia desempenhar outro papel. Vejamos a outra hipótese:

Caso 2: assim como no caso 1, Block se torna consciente que Schmidt não fará a ação que ele quer. Block novamente entra em ação: ele adiciona novos inputs no mesmo mecanismo que está operando na sequência real. Digamos que esses novos inputs venham em forma de razões para votar em Hitler. Block age

⁷³ Apesar dessa possibilidade ser inesperada dado o próprio caráter moral de Schmidt e sobre sua desconfiança sobre Hitler. Entretanto, seria um caso possível visto que Fischer e Ravizza pretenderiam que a teoria fosse compatível com os casos de fraqueza da vontade (LONG, 2004, p. 158).

diretamente manipulando o mecanismo deliberativo acrescentando razões suficientes, ou razões fortes suficientes, para certificar-se que Schmidt votará em Hitler (LONG, 2004, p. 159).

A ideia de Long é destacar que simplesmente adicionar novas razões no mecanismo deliberativo normal do agente não deveria mudar o mecanismo. Ele também chama a atenção para o fato de que ninguém havia ainda imaginado a possibilidade de alguém como Block poder acrescentar razões ao próprio mecanismo deliberativo real. Lembrando que, para Fischer e Ravizza, se houve a intervenção de um manipulador o mecanismo já não seria mais o mesmo, pois o agente não seria a origem do seu comportamento.

Para analisarmos esses dois casos teríamos que ter em mente o que está por trás das nossas intuições quando isentamos de responsabilidade moral o comportamento do agente. À primeira vista, parece verdadeiro o seguinte princípio: “Se uma pessoa *S* realiza uma ação *A* em um tempo t_2 unicamente porque outro agente *B* agindo completamente sem o conhecimento de *S* manipulou diretamente o cérebro de *S* no t_1 anterior, então *S* não é moralmente responsável por realizar *A*” (LONG, 2004, p. 155). Seria a manipulação exercida antes do agente agir que suportaria nossas intuições de que ele estaria carecendo de responsabilidade moral.

A reflexão sobre os casos 1 e 2, citados acima, revela intuições conflitantes. Por um lado, em virtude da presença do processo manipulativo, ambos os casos envolveriam o que Fischer e Ravizza chamam de um diferente tipo de mecanismo. Por outro lado, para Long o caso 2 estaria representando uma situação em que a manipulação não comprometeria a responsabilidade moral do agente (LONG, 2004, p. 163).

Embora o princípio que traz a ideia de que a manipulação exercida sem o nosso conhecimento tornaria a prática de responsabilidade moral inapropriada tenha alguma plausibilidade, ao refletir sobre o caso 2 percebe-se a sua falsidade. Dito de outro modo, com base no que Long argumentou, o caso 2 apoiaria a sua tese de que nem toda a manipulação poderia ser considerada como enfraquecedora de responsabilidade (LONG, 2004, p. 163-164).

Esse ponto poderia ser explicado por meio de uma analogia. É facilmente perceptível a possibilidade de manipulação através dos meios externos, como acontece na política, em que não raro os meios de comunicação tentam manipular determinadas informações relevantes seja pela incompletude das matérias seja pela ocultação de fatos, ou nas relações de consumo, em que o trabalho publicitário tem como finalidade nos manipular para comprar o produto que ele quer nos vender. Porém, a possibilidade de haver informações manipuladas não significaria que sejamos meras vítimas da vontade externa sem qualquer capacidade de reflexão. Qual diferença

haveria entre a manipulação de informações seja pelos meios de comunicação, seja pela publicidade, e a manipulação por meio do fornecimento de razões, como acontece no caso 2 citado acima? (LONG, 2004, p. 165).

O caso 2, como já visto, revelaria um problema quando a teoria de Fischer e Ravizza é aplicada aos exemplos de Frankfurt. Se há manipulação o mecanismo resultante na sequência alternativa seria diferente daquele que teria resultado na sequência real. Sendo assim, a manipulação eliminaria a responsabilidade moral. Na perspectiva de Long, esse cenário teria como implicação o fato de que o mecanismo presente na sequência real seria o mesmo da sequência alternativa, isto é, o mecanismo deliberativo do próprio agente.

Vejamos agora o que Fischer teria a dizer sobre o problema levantado por Long. Ele tem consciência de que os casos envolvendo a manipulação ou a indução de estados mentais são complexos e difíceis de avaliar. Sobre o caso 2 tratado por Long, duas interpretações poderiam ser oferecidas. A primeira interpretação envolveria a indução de razões de uma tal forma que não permitiria ao agente uma oportunidade razoável de submetê-las ao escrutínio crítico orientado por seus parâmetros normativos, disposições, valores e preferências. Essa manipulação, conseqüentemente, removeria a responsabilidade moral desse agente. Poderia se dizer que nesse cenário o processo manipulativo mudaria o mecanismo da ação. Não se teria uma conexão apropriada entre as nossas razões e o comportamento (FISCHER, 2012, p. 198-199).

Já a segunda interpretação remeteria à implantação de um novo estado mental como uma nova razão, um novo desejo, ou um desejo com uma nova intensidade, de tal modo que permitiria ao agente uma oportunidade de refletir e filtrar esse novo elemento através do raciocínio prático e disposições de caráter. Nesse contexto, esse tipo de indução poderia ser compatível com a responsabilidade moral (FISCHER, 2012, p. 196-197).

A seguir tratarei da abordagem de Carolina Sartorio que foca exclusivamente nos fatos da sequência real. Ela chama atenção para importantes considerações que têm sido negligenciadas pela literatura, incluindo Fischer.

5.3 Um modelo de liberdade baseado na sequência real

Partindo do argumento de Frankfurt contra o modelo clássico de liberdade, tem sido desenvolvido na literatura recente um modelo de liberdade focado na sequência real dos eventos. Esse modelo é considerado uma alternativa razoável àquele que envolve a capacidade do agente para agir diferentemente.

Sartorio aponta um problema para a teoria que trata da liberdade em termos da sequência real: ao desenvolver a posição em mais detalhes, a explicação tenderia a não respeitar a ideia inicial que a motiva. Por exemplo, Fischer parece estar contando com certos fatores irrelevantes para explicar a liberdade que não fariam parte da sequência real. Avaliar se o agente agiu livremente consistiria em testar se o mecanismo da sequência real foi sensível às razões de um modo apropriado. E para fazer isso seria necessário verificar como que o mecanismo reagiria pressupondo a possibilidade de mundos possíveis. Para Sartorio, seria justamente essa dependência de cenários alternativos que estaria em conflito com a ideia original de exclusividade, isto é, o foco unicamente na sequência real (SARTORIO, 2016, p. 109-115).

Para Sartorio, seria possível desenvolver uma abordagem da liberdade compatível com o pensamento da exclusividade. Para isso, bastaria limitar-se às características centrais da sequência real, sem formulá-la de modo abstrato ou sem conteúdo para que pudéssemos capturar mais claramente o conceito de liberdade (SARTORIO, 2016, p. 114). Seria necessário engajar em uma investigação mais profunda sobre o que significaria dizer que a liberdade é exclusivamente uma questão dos fatos relevantes da sequência real (SARTORIO, 2016, p. 115).

A solução consistiria em dois estágios. O primeiro estágio diria respeito à natureza dos fatos na sequência real e o segundo à interpretação da afirmação de que a liberdade seria exclusivamente uma função desses fatos.

Com respeito ao primeiro estágio, poderia se dizer que as sequências reais incluiriam todos os elementos que desempenham um papel em explicar o comportamento do agente. Nesse caso, além da nossa sensibilidade às razões também teríamos de incluir a nossa sensibilidade às ausências das razões, pois essas também teriam um papel fundamental em explicar por que nós estaríamos agindo de um determinado modo. Por exemplo, digamos que eu decida caminhar para tomar um ar fresco. A sequência real abrangeria mais do que apenas as razões reais da minha decisão, ou seja, o desejo de tomar ar fresco. Também faria parte dos elementos dessa sequência várias ausências de razões para me abster de sair, como a ausência de um impulso repentino que fizesse eu ficar em casa para fazer limpeza, a inexistência de um importante telefonema ou de um pedido de ajuda de um vizinho no momento da minha tomada de decisão. Essas ausências poderiam fazer parte da explicação da minha liberdade (SARTORIO, 2016, p. 115-117).

Conforme Sartorio, essa explicação mais abrangente possuiria vantagens em relação às outras teorias da sequência real. Primeiro, porque seria uma explicação mais simples. Segundo, e talvez o mais importante, ela permaneceria fiel à motivação original, que é manter o foco exclusivamente na sequência real. Ademais, essa forma de interpretar a questão realçaria as

sequências reais como mais substanciais do que elas parecem à primeira vista, na medida em que fariam parte delas tanto as razões reais que motivam o agente a agir como as ausências de muitas razões contrafactuais (SARTORIO, 2016, p. 116).

Já o segundo estágio da explicação estaria relacionado ao conteúdo da ideia da exclusividade. O que significa dizer que a liberdade deveria ser compreendida exclusivamente como sendo uma função da sequência real? Consistiria em considerar a liberdade baseada unicamente sobre fatos da sequência real. Poderiam existir diferentes fatos, alguns externos à sequência real, que seriam relevantes para a composição desse cenário. Para exemplificar esse ponto Sartorio menciona o caso do tubarão, tratado por Fischer e Ravizza. Um homem caminhando pela praia vê uma criança se afogando no mar. Como ele é um nadador experiente, temos razões suficientes para acreditar que ele poderia facilmente salvar a criança. Porém, ele não poderia salvá-la, pois, sem o seu conhecimento, havia muitos tubarões que teriam lhe atacado se ele tivesse entrado na água.

Sartorio esclarece que o nadador não seria moralmente responsável pela morte da criança porque mesmo que ele tivesse decidido entrar no mar, ele não teria conseguido salvá-la. Desse modo, tanto o nadador como o tubarão seriam fatos externos importantes para a explicação da sequência real resultante na morte da criança (SARTORIO, 2016, p. 117-118). Em contraste, poderiam existir outros fatores, também externos, que não seriam relevantes para a composição da sequência real. Os cenários de Frankfurt representariam essa ideia. O papel do neurocientista não seria intuitivamente relevante para a formação da sequência real, na medida em que a explicação de por que o agente seria moralmente responsável consistiria em ele agir por suas próprias razões sem a interferência do manipulador (SARTORIO, 2016, p. 118).

No item a seguir apresentarei o argumento da manipulação proposto por Derk Pereboom contra o compatibilismo. Seu argumento também é direcionado aos agnósticos, aqueles ainda não comprometidos com alguma posição. O argumento de Pereboom consistiria em afirmar que as exigências compatibilistas de agência moralmente responsável seriam insuficientes. Após esclarecido o argumento, o passo seguinte será discutir a réplica.

5.4 O argumento da manipulação

Segundo Pereboom, o argumento da manipulação representa a estratégia anticompatibilista mais eficiente para mostrar a incompatibilidade da responsabilidade moral com o determinismo causal. O argumento consiste na apresentação de quatro situações hipotéticas que têm a pretensão de mostrar que as condições causais compatibilistas são inadequadas, no sentido de que mesmo satisfazendo os critérios compatibilistas o agente estaria

carecendo de responsabilidade moral. O intuito do argumento é converter os compatibilistas e os agnósticos ao incompatibilismo. Ele parte de um caso de manipulação até chegar a um caso de determinação causal, mostrando que não existiriam diferenças entre esses quatro casos. Entretanto, o desafio para o compatibilista consiste em apresentar uma diferença relevante entre um caso de manipulação e um de determinação causal que justificaria afirmar que o agente poderia ser moralmente responsável apesar da sua decisão ou ação ter sido causalmente determinada. Pereboom acredita que não haveria essa diferença entre esses quatro casos. Ele acredita que é possível generalizar do primeiro caso para o último dizendo que intuitivamente o agente não poderia ser moralmente responsável pela sua decisão e ação causalmente determinada.

O argumento começa imaginando um agente, chamado Plum, realizando a mesma ação mental, ou seja, a sua decisão de matar White por vantagens pessoais e egoístas em quatro situações variadas. Em todas as quatro situações, a decisão de Plum seria causalmente determinada por fatores além do seu controle e em todos os cenários o agente estaria satisfazendo todas as exigências compatibilistas para a responsabilidade moral. O primeiro caso caracteriza o tipo mais radical de manipulação:

Um grupo de neurocientistas é capaz de manipular os estados neurais de Plum a qualquer momento por alguma tecnologia de rádio. Nesse caso particular, eles fazem isso pressionando um botão logo antes dele começar a refletir sobre sua situação, o que eles sabem que resultará em um estado neural que produzirá um processo reflexivo fortemente egoísta que deterministicamente resultará em sua decisão para matar White. Plum não mataria White sem a intervenção dos cientistas, uma vez que a sua reflexão não seria suficientemente egoísta para resultar nessa decisão. Mas ao mesmo tempo, o seu desejo de primeira ordem efetivo para matar White está de acordo com os seus desejos de segunda ordem. Além disso, o processo de deliberação do qual a sua decisão resulta é responsivo às razões; em particular, esse tipo de processo teria resultado em Plum não decidindo matar White em situações nas quais as suas razões são diferentes. A sua reflexão é consistente com o seu caráter porque frequentemente ele é egoísta e algumas vezes ele é fortemente egoísta. Ainda assim, ele não é em geral exclusivamente egoísta, pois algumas vezes ele regula o seu comportamento por razões morais, especialmente quando razões egoístas são relativamente fracas. Plum também não é forçado a agir como ele age, pois ele não age em virtude de um desejo irresistível - os neurocientistas não induzem um desejo desse tipo (PEREBOOM, 2014, p. 76-77).

Nesse primeiro contraexemplo às condições compatibilistas, os neurocientistas estariam manipulando diretamente os estados neurais do agente. O que explicaria a carência de responsabilidade moral seria o fato de que a sua decisão estaria além do seu controle por causa

da intervenção causalmente determinística dos neurocientistas. Além disso, Plum não teria decidido matar White se não tivesse ocorrido essa intervenção neurológica.

Vejam os o segundo caso. Similarmente ao primeiro caso ele também envolve manipulação, porém agora a manipulação estaria mais próxima de uma situação comum. A hipótese agora é que Plum tenha sido programado no início da sua vida, de modo que seu processo de raciocínio seja frequentemente, mas não sempre, egoísta. Para Pereboom, seria injustificado afirmar que nesse caso, e em contraste com o primeiro, Plum poderia ser moralmente responsável porque o espaço de tempo entre a manipulação e a decisão teria sido grande o suficiente. Não parece ser uma diferença significativa para a questão da sua responsabilidade moral se a manipulação ocorre pouco tempo antes dele começar a deliberar e decidir ou se ela ocorre quarenta anos antes (PEREBOOM, 2014, p. 77-78).

A decisão estando além do controle de Plum, e sendo causalmente determinada pela intervenção dos neurocientistas, é o que explicaria ele não ser moralmente responsável no primeiro caso. Dada a falta de diferenças entre os dois cenários, seria intuitivo concluir que pela mesma razão ele não seria moralmente responsável no segundo caso (PEREBOOM, 2014, p. 78).

O terceiro caso é ainda mais próximo de uma situação comum. Também é muito similar aos casos anteriores, exceto que agora o papel desempenhado pela manipulação é substituído pelas práticas de convívio da comunidade. Não seria difícil imaginar que Plum poderia ter crescido em um ambiente em que o autointeresse e a violência estivessem fortemente presentes, e que seu processo de raciocínio deliberativo tenha sido causalmente determinado pelos hábitos dessa comunidade, sem que ele tenha tido a oportunidade de modificar essas práticas. Para Pereboom, uma avaliação desses três cenários revelaria que não há diferenças significativas entre eles. De novo, o que explicaria a ausência de responsabilidade seria o fato de que a origem da decisão está além do controle do agente (PEREBOOM, 2014, p. 78).

Já no quarto caso apresentado por Pereboom, o papel dos neurocientistas é substituído pelo determinismo causal. Nesse caso a determinação causal da decisão do agente não é produzida por outros agentes. A questão é se essa diferença seria relevante. Pereboom acredita que não, pois mesmo imaginando casos muito similares aos dois primeiros vistos anteriormente, exceto que os estados em questão fossem produzidos por uma máquina sem nenhuma inteligência humana envolvida, o agente ainda careceria de responsabilidade moral. Poderia se concluir, a partir disso, que a intuição de não responsabilidade não decorreria da determinação causal por outros agentes. O que fundamentaria a intuição de não responsabilidade seria o fato

de que a decisão do agente em discussão é causalmente determinada por fatores além do seu controle (PEREBOOM, 2014, p. 79).

Agora tratarei da réplica ao argumento da manipulação. Duas possíveis réplicas dos compatibilistas se destacam na literatura. Por um lado, há uma réplica forte que acredita que a melhor estratégia de resposta consistiria em aceitar a posição de Pereboom da existência de uma similaridade de um agente manipulado e um determinado. Essa é a posição seguida por Mckenna. Por outro lado, muitos têm adotado uma réplica fraca, propondo uma distinção entre agentes manipulados e agentes determinados, negando que os incompatibilistas tenham verdadeiramente capturado as condições compatibilistas. Esse é o caminho seguido por Fischer.

Para Mckenna, a réplica fraca seria ineficaz. O incompatibilista poderia facilmente objetar propondo a revisão do exemplo de tal modo que as condições compatibilistas fossem atendidas. Algumas posições compatibilistas têm elaborado condições históricas para a responsabilidade moral. E eles têm argumentado que o agente manipulado não seria livre tampouco responsável porque a manipulação impediria que o comportamento do agente tivesse uma história causal apropriada.

Embora Mckenna discorde da réplica fraca, ele acredita que o argumento da manipulação não invalidaria a tese compatibilista. Como visto antes, o argumento consistiria em imaginar situações hipotéticas com o intuito de mostrar que as condições compatibilistas de liberdade e responsabilidade moral são insuficientes. Esse argumento poderia ser representado da seguinte forma:

- P1. Se um agente é manipulado a fazer *A* de um modo *x*, então ele não faz *A* de sua própria vontade livre e, portanto, não é moralmente responsável por *A*.
- P2. Um agente manipulado a fazer *A* não é relevantemente diferente de qualquer agente normal determinado a fazer *A*.
- PC. Portanto, se um agente normal está determinado a fazer *A*, ele não faz *A* de sua própria vontade livre e, portanto, não é moralmente responsável por *A* (MCKENNA, 2008, p. 143).

A ideia básica consiste em afirmar que se esse argumento for sólido, então, a tese compatibilista deveria rever as suas exigências de agência livre e moralmente responsável. Entretanto, para Mckenna esse argumento não seria consistente, pois ele rejeita a premissa 1. Negar a primeira premissa do argumento implica aceitar a tese de que um agente pode ser moralmente responsável apesar de ser manipulado.

A réplica de Mckenna ao argumento da manipulação é composta por quatro passos. O primeiro consistiria em uma interpretação distinta do primeiro caso de Pereboom. Os

neurocientistas estariam restritos no que eles poderiam fazer e o que eles poderiam controlar a respeito de Plum. Por exemplo, Plum “deve ter uma vida mental causalmente integrada e internamente coerente. Suas memórias sobre as considerações do passado deveriam ser capazes de causalmente influenciar suas deliberações atuais” (MCKENNA, 2008, p. 149). O segundo passo envolveria um aperfeiçoamento dos quatro casos de Pereboom. Em todos esses cenários, a ação de Plum seria consistente com a concepção de Strawson de agência livre (STRAWSON, 2008). A ação possuiria significância moral porque Plum seria capaz de sentir culpa ou remorso por ter assassinado Ms. White e apreciar as demandas da comunidade moral (MCKENNA, 2008, p. 151). O terceiro passo consistiria em fixar a atenção nas características de agência moral compatibilista e generalizá-las para os demais casos. Porém, ao invés de partir do primeiro caso, como faz Pereboom, Mckenna propõe partir do último e generalizar dele para o primeiro. Sua estratégia consiste em afirmar que não seria evidente que Plum no quarto caso estaria carecendo de liberdade e responsabilidade moral apesar de agir em um mundo determinista (MCKENNA, 2008, p. 153).

Para Mckenna, essa explicação acabaria em um impasse, visto que ele estaria negando a tese de Pereboom. Se o desacordo entre Mckenna e Pereboom realmente se configurar um impasse, então, para Mckenna, isso seria uma vitória para o compatibilista, pois a sua tarefa seria unicamente invalidar o argumento do incompatibilista e não teria a preocupação de provar a tese compatibilista (MCKENNA, 2008, p. 154).

A ideia é que o custo dialético dos compatibilistas seria mais leve se comparado ao custo dos seus oponentes, pois os incompatibilistas seriam os únicos que estariam afirmando que os agentes apropriadamente manipulados ou determinados não seriam livres e moralmente responsáveis. E também seriam os únicos que pretendem usar essa afirmação em um argumento para o incompatibilismo. Caberia ao compatibilista mostrar que o incompatibilista está equivocado na primeira premissa do argumento da manipulação, ou seja, de que um agente manipulado que cumprisse os critérios compatibilistas não seria livre e responsável. Ademais, o compatibilista poderia adequadamente responder ao argumento meramente lançando dúvidas sobre a premissa incompatibilista, sem a necessidade de oferecer uma defesa da sua tese a respeito da manipulação (MCKENNA, 2008, p. 155).

O quarto passo limitaria-se a tornar os casos de manipulação comuns. Para fazer isso bastaria reformular os exemplos partindo de acontecimentos das próprias experiências sociais. Suponha que na infância de Ana um parente seu tenha tido câncer e tenha feito quimioterapia, e que essa vivência tenha determinado os valores de Ana e o modo como ela vive hoje. Ela passou a ver a vida como algo limitado que não deveria ser desperdiçado. Essa situação

hipotética se assemelharia a um caso de manipulação. Uma questão surgiria imediatamente: Ana seria responsável pela sua conduta resultante desses valores? Ao questioná-la, ela não consideraria suas experiências passadas como um impedimento da sua liberdade e responsabilidade moral (MCKENNA, 2008, p. 156)⁷⁴.

Há outra interpretação na literatura que rejeita o argumento da manipulação, qual seja, a de Matt King⁷⁵. Ele argumenta que a premissa desse argumento dependeria de um conceito particular incompatibilista que seria dialeticamente inapropriado. Para entender a sua crítica ao argumento, eu apresento o argumento da manipulação tal como ele é discutido por King:

- M1. O agente manipulado não é moralmente responsável ao agir com base em estados psicológicos que foram implantados, apesar de satisfazer as condições compatibilistas de responsabilidade moral.
- M2. Não existe uma diferença entre a ação do agente manipulado e a ação do agente ordinário em um mundo determinista.
- MC. Portanto, as condições compatibilistas para a responsabilidade moral são insuficientes. Em resumo, o compatibilismo é falso (KING, 2013, p. 67).

Para King, os incompatibilistas precisam defender a primeira premissa do argumento. Eles precisam explicar por que a manipulação eliminaria a responsabilidade. Sem essa defesa eles estariam sujeitos a uma réplica denominada como o argumento da manipulação paralelo:

- P1⁷⁶. O agente manipulado não é moralmente responsável ao agir com base em seus estados psicológicos criados artificialmente apesar de satisfazer as condições incompatibilistas de responsabilidade moral.
- P2. Não existe uma diferença entre a ação do agente manipulado e a ação do agente ordinário em um mundo indeterminista.
- PC. Portanto, as condições incompatibilistas para a responsabilidade moral são insuficientes. Em resumo, o incompatibilismo é falso (KING, 2013, p. 70).

O incompatibilista poderia responder que esse argumento seria insignificante porque o critério para a agência livre e responsável seria ele mesmo incompatível com a manipulação.

⁷⁴ Segundo Mckenna, o caso de Ana é muito parecido com um caso de manipulação, exceto que nesse caso não há um grupo de neurocientistas (MCKENNA, 2008, p. 156).

⁷⁵ O objetivo de King nesse ensaio não seria defender o compatibilismo, mas tentar convencer o agnóstico que o argumento da manipulação não estaria mostrando o custo do compatibilismo. King não está tomando uma posição no debate sobre o compatibilismo ou o incompatibilismo. Se o argumento que ele apresenta for sólido nenhum lado do debate será favorecido. O objetivo do autor é chamar nossa atenção para o fato de que os argumentos de manipulação não estariam favorecendo os incompatibilistas (KING, 2013, p. 65; p. 81).

⁷⁶ P nesse argumento pode significar tanto premissa como também paralelo, isto é, referindo-se ao argumento da manipulação paralelo.

Duas respostas poderiam ser dadas em apoio à primeira premissa do argumento: (a) fornecer uma condição única à manipulação que explicasse por que o agente manipulado não seria responsável e (b) mostrar que a manipulação estaria eliminando a responsabilidade porque o agente não seria a origem da sua ação.

Para King, tanto a primeira opção como a segunda são problemáticas. A opção (a) porque fornecer uma característica específica à manipulação teria como implicação a falsidade da segunda premissa do argumento, pois ela não abrangeria os casos de ações determinadas ou indeterminadas. A opção (b) porque a explicação dependeria de um critério controverso de responsabilidade. Generalizar para a segunda premissa implicaria em supor a verdade da tese determinista ou indeterminista. Em resumo, King afirma que o argumento da manipulação geraria um impasse dialético ao invés de apoiar um argumento para o incompatibilismo. Os incompatibilistas não poderiam defender a primeira premissa do argumento sem invocar uma condição polêmica de responsabilidade. Porém, sem tal defesa eles estariam na mesma situação dos compatibilistas (KING, 2013, p. 68).

Se os incompatibilistas apelam para uma condição de origem para explicar a não responsabilidade do agente manipulado, então o argumento da manipulação nada mais é do que um argumento de origem:

O1: Se um agente não é a origem apropriada da ação *A*, então esse agente não é moralmente responsável por ter feito *A*.

O2: Se o determinismo é verdadeiro, nenhum agente pode ser a origem apropriada de qualquer ação.

OC: Portanto, se o determinismo é verdadeiro ninguém é responsável por qualquer ação (KING, 2013, p. 79).

O problema com esse argumento é que ele estaria assumindo uma condição incompatibilista de responsabilidade que é dependente da verdade da tese indeterminista. Esse argumento não faria o debate progredir porque o compatibilista também poderia argumentar que um agente poderia ser a origem apropriada da sua ação desde que se aceite a verdade da tese determinista (KING, 2013, p. 80). Seja qual for a defesa que fosse feita da premissa (M1), a explicação deveria ser generalizada para a premissa (M2). Porém, a natureza de tal generalização torna a manipulação supérflua. Se o argumento da manipulação é falho como King tem argumentado, então, seria melhor evitá-lo (KING, 2013, p. 80).

O argumento de Pereboom para a incompatibilidade da responsabilidade moral com o determinismo consistiu na apresentação de quatro exemplos com as seguintes características: manipulação contínua, manipulação inicial pelos neurocientistas, influência da comunidade e

mera determinação causal. Partindo do pressuposto de que o agente, no primeiro caso, não seria moralmente responsável, e que não haveria nenhuma diferença relevante entre os casos, segue que o agente não seria moralmente responsável no último caso. Portanto, determinismo seria incompatível com a responsabilidade moral.

O que Fischer teria a acrescentar sobre o argumento da manipulação de Pereboom? Para Fischer, o agente seria moralmente responsável em todos os quatro casos, mesmo no primeiro o qual envolve manipulação externa. O que na realidade intuimos é o fato de que o agente, no primeiro caso, não é merecedor de censura (FISCHER, TOGNAZZINI, 2011, p. 20).

Muita confusão surgiu no debate por causa da falta de precisão do conceito de responsabilidade moral. Por exemplo, Pereboom não ficou convencido com essa réplica de Fischer, pois ele pensa que a distinção entre a responsabilidade moral e o merecimento de censura ou louvor é insustentável. O termo usado por Pereboom pode ser evidenciado pela seguinte citação: “para um agente ser merecedor de censura por uma ação está de fato implicado em ele ser moralmente responsável no sentido em questão em debate, junto com o seu entendimento que a sua ação foi, de fato, moralmente errada” (PEREBOOM, 2006, p. 211). A questão do merecimento seria básica e não em virtude de considerações consequencialistas, isto é, o agente mereceria censura apenas porque ele entende que a ação foi moralmente errada, e mereceria crédito ou louvor se ele entende que a ação foi moralmente exemplar.

Pode se dizer que Fischer e Pereboom fazem usos distintos do termo responsabilidade moral. Para Pereboom, dizer que um agente é moralmente responsável equivale a dizer que seria justificado censurá-lo (PEREBOOM, 2006). Já Fischer pensa que um agente poderia ser moralmente responsável por uma ação moralmente errada e não merecer censura. Para exemplificar esse ponto vejamos a diferença entre o caso do professor Plum e o caso da Ernie (visto no capítulo anterior). Adotando a terminologia de Fischer, pode se dizer que apenas Ernie mereceria censura. Quando o professor Plum mata o Ms. White ele estaria exercendo as capacidades que compatibilistas (como Fischer) pensam que seriam exigidas para considerá-lo um candidato apropriado para as atitudes reativas. O mesmo pode ser dito sobre Ernie, exceto que apenas Ernie mereceria censura.

Por que a censura estaria justificada apenas para Ernie? A razão para essa explicação estaria nas circunstâncias da manipulação. Apesar de ambos terem sido manipulados, os detalhes da manipulação diferem drasticamente. Plum está sujeito a um tipo de manipulação muito mais invasivo do que o de Ernie. Plum seria como um criminoso que teve uma infância abusiva, já Ernie, que teve apenas as condições iniciais escolhidas (Diana quer que Ernie realize uma ação futura), seria como o criminoso que teve uma infância mais ou menos normal. Assim

como estaríamos inclinados a suspender nossas práticas morais de censura, não necessariamente as práticas institucionais de pena (prisão) em relação ao criminoso se descobríssemos que ele cresceu sob circunstâncias trágicas de formação do seu caráter, assim também estaríamos inclinados a suspender nossas práticas quando nos deparamos com circunstâncias que envolvam a manipulação de valores e desejos, como é o caso de Plum (FISCHER; TOGNAZZINI, 2011, p.23-24).

Como já evidenciado, o termo responsabilidade moral possui vários sentidos. Em um deles, quando se trata de saber se seria legítimo louvar ou censurar um agente, o sentido envolveria questões de imputabilidade. Plum deveria ser desculpado devido às circunstâncias da sua manipulação, portanto censurá-lo seria injustificado. Em outro sentido, a ação moralmente errada pode ser atribuída a Plum, em um sentido reativo, e essa então é uma questão sobre atributabilidade reativa. Quando se trata de censurar o agente, está se falando de imputabilidade, e quando se trata de atribuir uma ação ao agente, está se falando em atributabilidade⁷⁷.

Embora essa discussão acerca da natureza do conceito de responsabilidade moral seja relevante para um exame crítico e completo da posição semicompatibilista de Fischer, em virtude do espaço este trabalho está circunscrito à questão acerca das condições de aplicabilidade desse conceito (isto é, o objeto de investigação é o problema de determinar sob que condições alguém é moralmente responsável pelas suas ações, e não o problema de determinar o que é para alguém ser moralmente responsável por suas ações)⁷⁸. Como resultado,

⁷⁷ Cito um exemplo, de Michael Zimmerman, que explica a distinção entre atributabilidade reativa e imputabilidade: “Vamos supor que Jones se depare com um acidente de carro; o motorista está inconsciente (...) Jones espera que o carro exploda a qualquer momento, e então ele vai até o motorista e o puxa para fora do carro. O resultado: o motorista fica paralisado pelo resto da sua vida (ele não teria ficado paralisado se Jones tivesse o deixado onde ele estava), e o carro não explode. Agora, Jones fez algo errado? (ZIMMERMAN, 1988, p. 41)”. Seria plausível dizer que embora Jones tenha feito algo errado, ele não deveria ser censurado. Embora a ação moralmente errada seja atribuída a Jones em um sentido reativo (ou seja, ele poderia ser moralmente responsável), pois ao exercer a ação em questão ele tinha as capacidades exigidas para torná-lo um candidato apropriado de tais práticas, ele deveria ser desculpado de imputabilidade. Por que Jones não deveria ser censurado? A razão tem a ver com a sua crença quando estava tirando o motorista de dentro do carro. Jones não estava consciente de que a sua ação causaria paralisia permanente e, pressupondo que a sua ignorância não tenha sido culpável, portanto ele não sabia que estava fazendo alguma coisa errada (FISCHER; TOGNAZZINI, 2011, p. 9-10).

⁷⁸ Para uma discussão sobre o conceito de responsabilidade moral, ver Watson (1996). Thomas Scanlon estabelece uma distinção entre responsabilidade no sentido atributivo e substantivo tendo como critério para a sua aplicação a ideia de voluntariedade e do valor da escolha. No sentido atributivo equivale à avaliação de elogio ou censura pelo o que o agente fez ou deixou de fazer. Já a responsabilidade no sentido substantivo está relacionada à submissão de um dever, corresponde afirmar que o agente não

um trabalho como este não está equipado para avaliar de maneira completa o mérito da posição semicompatibilista de Fischer, quando confrontada com o argumento da manipulação de Pereboom. Penso, no entanto, que a discussão até aqui nos permitiu identificar os pressupostos fundamentais envolvidos nesse problema, os custos de se aceitar a objeção de Pereboom ou a réplica de Fischer e o caminho que deve ser seguido por uma discussão mais abrangente sobre a objeção da manipulação.

pode reclamar das obrigações resultantes. É possível que um agente seja responsável em um sentido e não em outro. Por exemplo, podemos censurar alguém por ter usado drogas, portanto usando o sentido atributivo de responsabilidade, porém não necessariamente aplicaríamos o sentido substantivo como a negação do direito à assistência social dado às circunstâncias particulares à qual ele se encontra. Nós censuramos o adicto porque pressupomos que sua ação é consequência da sua escolha não coagida, ou seja, que ele tenha agido voluntariamente. Usando o próprio exemplo de Scanlon para se explicar o sentido substantivo da responsabilidade, pode se dizer que alguém saí para caminhar não levando em consideração a recomendação do setor público sobre a remoção de lixo radioativo. Essa pessoa terá responsabilidade substantiva se sofrer algum dano decorrente da exposição, pois ela poderia ter escolhido não se expor ao material (SCANLON, 1998, p. 256-258).

CONSIDERAÇÕES FINAIS

Para concluir, retomarei os pontos principais discutidos ao longo deste trabalho, procurando destacar os problemas centrais e as possíveis soluções abordadas. A tese defendida foi a de que a teoria de Fischer responde, de modo satisfatório, às críticas feitas pelo incompatibilista. Nesse sentido, a teoria semicompatibilista de Fischer fornece uma solução atrativa para o problema da vontade livre, isto é, para o conflito entre a liberdade e o determinismo. Partindo de uma noção de liberdade que não necessariamente implica uma capacidade para agir ou escolher diferentemente, essa teoria nos permite entender como é possível compatibilizar a responsabilidade moral com o determinismo.

No primeiro capítulo eu caracterizei as principais posições que compõem o debate, a saber, determinismo, indeterminismo, compatibilismo, incompatibilismo e fatalismo. A apresentação das ideias básicas e motivações dessas posições é interessante porque, de uma forma ou outra, os filósofos fazem referência a essas elas seja como crítica, como é o caso do compatibilismo clássico, ou como recurso para apoiar e desenvolver a sua própria posição, como é o caso de Fischer, que concorda com o conceito de responsabilidade moral desenvolvido por Strawson.

No segundo capítulo eu apresentei os experimentos de pensamento e os exemplos hipotéticos de Frankfurt, que introduzem uma metodologia inédita para o debate, com o objetivo de mostrar que uma pessoa pode ser moralmente responsável pelas suas ações mesmo carecendo da capacidade para agir diferentemente. O intuito dessa apresentação foi chamar a atenção para o debate provocado por Frankfurt e os diversos modos pelos quais a sua proposta pode ser explorada, aperfeiçoada e criticada. Como se viu, a grande contribuição de Frankfurt foi desafiar a tradicional exigência de possibilidades alternativas para a responsabilidade moral.

Dado o objetivo mais restrito desta tese de discutir e avaliar a posição de Fischer, a discussão do argumento de Frankfurt também se justifica na medida em que ele é bastante influente na teoria da responsabilidade moral desenvolvida por Fischer. A posição semicompatibilista de Fischer foi discutida no terceiro capítulo. Desenvolvendo o raciocínio de Frankfurt, Fischer defende que a responsabilidade moral exigiria apenas o controle de direcionamento, não havendo a necessidade do controle regulativo, um tipo de poder duplo de ação livre. Dessa forma, para Fischer, a responsabilidade moral seria compatível com o determinismo independentemente se determinismo é compatível ou não com a liberdade.

Em uma obra escrita conjuntamente com Ravizza (1998), Fischer apresenta uma das principais características distintivas da sua teoria, em virtude da qual ela representa inclusive

uma ruptura com a própria versão hierárquica do compatibilismo de Frankfurt. Foi discutido como Fischer, em contraposição a Frankfurt, rejeitaria o internismo e adotaria em seu lugar uma posição externista. Isso significa que para Fischer a responsabilidade moral não seria unicamente uma função das características da configuração interna dos estados mentais do agente. Avaliar se um agente é moralmente responsável por um certo comportamento consistiria em verificar historicamente a origem das suas disposições e estados mentais. Além disso, a conexão do agente com o mundo externo seria fundamental para a avaliação da responsabilidade moral. A ideia é que o mecanismo a partir do qual o agente age deveria ser apropriadamente responsivo às razões apresentadas pelo mundo.

O quarto capítulo trata da posição de Kane, que figura neste trabalho como um representante do incompatibilismo, e de Mele, que representa uma posição agnóstica com respeito ao debate entre compatibilistas e incompatibilistas. Uma discussão da teoria de Kane é pertinente para um trabalho como este porque ela apresenta uma objeção com o intuito de desafiar as posições compatibilistas, como a de Frankfurt, e semicompatibilistas, como de Fischer. A objeção afirma que nenhuma teoria compatibilista poderia adequadamente responder ao assim chamado problema da manipulação. Kane acredita que tanto no controle manipulativo como na mera determinação causal, as origens das ações do agente não estariam nele mesmo, mas nos controladores e no passado ou nas leis naturais. Haveria, ele argumenta, mais semelhanças do que diferenças entre o controle manipulativo e a mera determinação causal.

Como resposta, foi visto que alguns compatibilistas e agnósticos têm argumentado que Kane se equivoca ao supor que o critério para a atribuição de responsabilidade moral dependeria exclusivamente da configuração dos estados mentais do agente. Mele, por exemplo, acredita que a história causal da ação seria importante para diferenciar uma ação responsável de uma não responsável, e analisando essa história seria possível explicar por que um agente manipulado careceria de responsabilidade moral enquanto um agente determinado poderia ser responsabilizado. Quando um agente é manipulado, as origens da sua ação estariam no controlador, mas quando um agente é causalmente determinado, o fato de que sua ação faz parte de uma cadeia causal que remonta ao passado não seria uma razão convincente para isentá-lo de responsabilidade.

Não seria a teoria de Kane que não teria os recursos necessários para lidar com os casos de manipulação? Conforme visto, o critério que ele oferece das possibilidades alternativas ou da escolha sendo casualmente indeterminada poderia não ser suficiente para a atribuição de responsabilidade moral, na medida em que seria possível afirmar que, seja qual fosse a escolha

que o agente fizesse, qualquer uma delas poderia ter sido manipulada sem o seu conhecimento ou consentimento.

No quarto capítulo também apresentei o argumento do zigoto, um influente argumento apresentado pelo incompatibilista. A ideia básica envolve imaginar uma situação fictícia na qual um agente com superpoderes planejou a criação de um zigoto para desenvolver-se em uma pessoa idealmente autocontrolada para realizar uma certa ação no futuro. O criador sabe que essa pessoa fará o planejado, visto que ele possui superpoderes e o cenário é determinista. O argumento supõe que não haveria diferença entre o modo que esse zigoto se origina e qualquer outro zigoto normal em um mundo determinista. Além disso, presume-se que a pessoa que nasceu a partir de um zigoto planejado careceria de responsabilidade moral. E visto que essa pessoa não seria diferente de qualquer outra em um mundo determinista, o defensor do argumento transfere a não responsabilidade do agente planejado para o agente determinado.

No quinto capítulo eu apresentei o argumento da manipulação, uma outra versão do argumento incompatibilista, que possui algumas semelhanças com o argumento do zigoto. O argumento é estruturado de uma tal forma que, se fosse sólido, implicaria que todos os critérios compatibilistas, incluindo o de Fischer, seriam insuficientes. O que estaria por trás do argumento é a ideia de que mesmo satisfazendo as exigências compatibilistas para agência livre e responsável, os estados psicológicos de um agente poderiam ter sido implantados. Supondo que não haveria diferenças entre a ação do agente manipulado e a ação do agente determinado, seguiria que as condições compatibilistas não seriam suficientes. Dito de outro modo, se o argumento do zigoto ou o da manipulação fossem sólidos, isso acatarretaria que os agentes causalmente determinados careceriam de liberdade e responsabilidade moral. A implicação é que somente faria sentido falar em liberdade ou responsabilidade dentro de um contexto indeterminista. Conforme discutido, essa versão do argumento da manipulação revela a importância de estender uma discussão sobre o mérito do semicompatibilismo de Fischer também para o exame do próprio conceito de responsabilidade moral. Na medida em que este estudo está restrito a uma discussão das condições de aplicabilidade desse conceito, ele não está equipado para discutir, de uma maneira completa, o mérito do argumento da manipulação. Este é um aspecto com respeito ao qual esta tese se revela incompleta e em direção ao qual a presente investigação poderia ser naturalmente estendida.

O trabalho apresentou uma problemática existente entre duas posições opostas a respeito dos critérios de aplicação do conceito de responsabilidade moral. Por um lado, a posição compatibilista defende a possibilidade da responsabilidade mesmo em mundo determinado. Por outro lado, a posição incompatibilista argumenta que unicamente em um mundo indeterminado

faria sentido usarmos as noções de liberdade e responsabilidade moral. O objetivo deste trabalho foi mostrar que esse conflito poderia ser resolvido pela posição semicompatibilista. Reivindicando unicamente uma crença na responsabilidade, a teoria semicompatibilista admite que a responsabilidade moral poderia existir tanto em um mundo determinado como em um mundo indeterminado, assumindo uma posição agnóstica sobre a questão da compatibilidade ou incompatibilidade da vontade livre ou liberdade enquanto capacidade para agir diferentemente com o determinismo causal. Se os critérios e argumentos apresentados pela teoria semicompatibilista forem válidos e sólidos, então poderia se dizer que temos boas razões para pensar que um agente poderia ser moralmente responsável mesmo em um contexto causalmente determinado.

Vimos também que a posição semicompatibilista de Fischer procurou diferenciar um mecanismo determinístico de um mecanismo manipulativo, argumentando que se o mecanismo de deliberação do agente sofreu um processo de lavagem cerebral, hipnose ou manipulação, então poderia se dizer que esse mecanismo careceria de autonomia e responsabilidade. E, adotando uma noção histórica de responsabilidade, nós teríamos os meios necessários para diferenciar a ação resultante de um mecanismo responsável de um não responsável, o que revelaria o ganho do semicompatibilismo de Fischer em relação à posição compatibilista de Frankfurt, por exemplo.

Por fim, cabe ressaltar algumas das implicações que a teoria semicompatibilista teria em cenários como de engenharia genética, por exemplo, já que a própria teoria trata de casos de manipulação. Na área da bioética é de conhecimento comum a possibilidade de usar a tecnologia para modificar a composição genética de um ser vivo. Podemos imaginar a possibilidade de manipulação de traços psicológicos. Por exemplo, podemos imaginar um cenário hipotético em que pais vão a um laboratório escolher as seguintes características do seu futuro filho. Eles podem escolher que ele seja uma boa pessoa, que não maltrate os animais e os idosos, e que tenha o hábito de realizar atos de caridade. Podemos também supor que o filho fosse agir a partir de um mecanismo apropriadamente responsivo às razões. A questão que imediatamente surge é esta: já que o filho não escolheu ser a pessoa que ele é, e sim os seus pais, ele teria autonomia e seria merecedor de crédito moral pelas boas ações realizadas, ações que não são decorrentes de um traço de caráter que ele escolheu? Adotando os critérios do semicompatibilismo, poderia se dizer que se o agente não teve a oportunidade de se tornar consciente do processo de manipulação e refletido sobre ele, então ele não estaria assumindo a responsabilidade pelo mecanismo resultante em seu comportamento, no caso, as ações de caridade. Portanto, o mecanismo que foi implantado antes de seu nascimento seria diferente do

mecanismo prático normal de raciocínio e deliberação. Consequentemente, ele seria isento de responsabilidade moral pela ação exercida. Esse exemplo é de relevância porque revela em que medida a qualificação do compatibilismo com um modelo externista e histórico leva a diferentes vereditos de responsabilidade moral. A teoria internista de Frankfurt, por exemplo, implicaria que nesse caso o filho não seria isento de responsabilidade moral, pois as suas ações revelariam os seus valores e a sua identidade, e seriam resultado da sua deliberação.

Casos de manipulação (lavagem cerebral) não são apenas cenários hipotéticos em livros de filosofia. Um desses casos, noticiado amplamente pela imprensa, é o de uma moradora de Lajeado, no Rio Grande do Sul⁷⁹. Enfrentando problemas financeiros, ela alega que participou do evento Fogueira Santa, da Igreja Universal do Reino de Deus, e doou parte dos seus bens à igreja. Segundo o seu relato, os pastores da igreja diziam que quem havia prometido doar tinha que doar, pois quem não doasse seria amaldiçoado. Ela revela ter doado os bens porque ficou com medo da maldição. Tendo se arrependido de ter feito a doação, ela recorreu à justiça exigindo a devolução dos seus bens e uma indenização por danos morais. O tribunal da justiça deferiu o seu pedido, mas a igreja ainda pode recorrer.

Adotando a teoria semicompatibilista de Fischer, podemos explicar que no caso da moradora de Lajeado a posse das capacidades de raciocínio prático normal não estaria em seu poder. E isso se dá pelo fato de que ela foi coagida, ameaçada ou mesmo manipulada por meio de crenças religiosas a decidir doar o dinheiro. Em um momento de desespero, passando por dificuldades financeiras, ela realmente acreditou que se não doasse o dinheiro, ou os bens exigidos, seria amaldiçoada. Sendo assim, ela não poderia ser considerada uma candidata apta para atitudes reativas.

Outro caso que também pode ser mencionado é o do soldado nazista Reinhold Hanning⁸⁰. Ele atuou como guarda no campo de extermínio de Auschwitz e foi declarado culpado como cúmplice na morte de 170 mil presos assassinados nesse local no ano de 1943. Em uma de suas audiências, ele admitiu sentir vergonha por não ter feito algo para impedir as atrocidades que ocorreram em Auschwitz. A justiça admitiu como atenuante o fato dele ter demonstrado arrependimento aos sobreviventes e familiares das vítimas. Como o réu mostrou arrependimento, ele se vê como origem dos seus atos. Ele não nega os atos que lhe foram

⁷⁹ A matéria da BBC está disponível no seguinte endereço: <http://www.bbc.com/portuguese/brasil-43325773>.

⁸⁰ A matéria do G1 pode ser acessada pelo link <https://g1.globo.com/mundo/noticia/morre-guarda-nazista-de-auschwitz-condenado-a-prisao-em-julgamento-em-2016.ghtml>.

atribuídos. Nesse caso, aplicando a teoria de Fischer, poderíamos explicar que ele poderia ser um candidato apropriado para as atitudes reativas, como ressentimento, raiva e perdão.

Como defendido por Strawson, mesmo que aceitássemos que as nossas ações são determinadas, nós ainda assim atribuiríamos importância aos sentimentos e atitudes que são essenciais nas nossas relações interpessoais. Veja o caso do nazista apresentado no parágrafo anterior. Ele foi declarado culpado, por meio da punição institucional e, também, nós reagimos emocionalmente ao seu comportamento, especificamente, dadas as circunstâncias podemos talvez perdoá-lo porque ele se arrependeu, ou podemos permanecer indignados e ressentidos pela crueldade desumana exibida em seus atos. O que quero enfatizar é que nós reagimos ao seu comportamento levando em consideração os seguintes elementos: ele estava no controle da sua ação, isto é, sua ação não foi resultante de um processo de hipnose, lavagem cerebral, ou manipulação. Ele tinha o entendimento de que sua ação era moralmente errada, ou seja, ele tinha capacidade para ser um agente moral, não era uma criança ou um psicopata.

Pode-se perceber que apesar das numerosas pesquisas feitas em áreas como a física, a psicologia, a biologia e a neurociência, que apontam para a ilusão ou diminuição da nossa liberdade, ainda faz sentido falar em responsabilidade moral e ter atitudes reativas, como ressentimento, indignação, gratidão e perdão, por meio das quais nos engajamos na prática de assumir e atribuir responsabilidade moral. Nesse sentido, acredito que a possibilidade de que as nossas ações sejam causalmente determinadas não nos obrigaria a abandonar essas atitudes reativas e os nossos sentimentos morais. Esse parece ser o principal ganho filosófico da proposta de Fischer. Penso que ela oferece uma solução satisfatória para o problema do conflito entre a liberdade e o determinismo, sem precisar contar com uma noção de liberdade metafísica que exigiria, em uma de suas vertentes, uma incompatibilidade com o determinismo, isto é, uma noção de liberdade indeterminista.

REFERÊNCIAS

ARISTOTLE. Nicomachean Ethics. In: BARNES, Jonathan (org.). **The Complete Works of Aristotle**. Volume 1. New Jersey: Princeton University Press, 1984a, p. 1729-1867.

_____. De Interpretatione. In: BARNES, Jonathan (org.). **The Complete Works of Aristotle**. Volume 2. New Jersey: Princeton University Press, 1984b, p. 25-38.

AGOSTINHO. **A Cidade de Deus**. Vol. 1. Lisboa: Fundação Calouste Gulbenkian, 1996.

AUDI, Robert. **Moral Knowledge and Ethical Character**. New York: Oxford University Press, 1997.

_____. Responsible Action and Virtuous Character. **Ethics**, v. 101, n. 2, 1991, p. 304-321.

AYER, Alfred Jules. **The Concept of a Person and Other Essays**. London: The Macmillan Press, 1963.

AZEVEDO, Marco. Responsabilidade moral e Ressentimento. **Ethic@**, v. 8, n. 3, 2009, p. 25-46.

BOBZIEN, Susanne. The Inadvertent Conception and Late Birth of the Free-Will Problem. **Phronesis**, v. 63, n. 2, 1998, p.133-175.

_____. Choice and Moral Responsibility. In: POLANSKY, Ronald. (org.) **The Cambridge Companion to Aristotle's Nicomachean Ethics**. New York: Cambridge University Press, 2014, p. 81-109.

_____. **Determinism and Freedom in Stoic Philosophy**. Oxford: Clarendon Press, 1998.

BOK, Hillary. **Freedom and Responsibility**. Princeton: Princeton University Press, 1998.

BONDENSON, William. Aristotle on Responsibility for One's Character and the Possibility of Character Change. **Phronesis**, v. 19, n. 1, 1974, p. 59-65.

BLUMENFELD, David. The Principle of Alternate Possibilities. **The Journal of Philosophy**, v. 68, n. 11, 1971, p. 339-345.

CLARKE, Randolph. **Libertarian Accounts of Free Will**. Oxford: Oxford University Press, 2003.

_____. Toward a Credible Agent-Causal Account of Free Will. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 285-298.

CLARKE, Randolph; MCKENNA, Michael; SMITH, Angela (org.). **The Nature of Moral Responsibility**. New York: Oxford University Press, 2015.

CHISHOLM, Roderick. Human Freedom and the Self. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 26-37.

COATES, Justin; TOGNAZZINI, Neal. (org.). **Blame: Its Nature and Norms**. New York: Oxford University Press, 2013.

COITINHO, Denis. Virtudes, Caráter e Responsabilidade. **Dissertatio**, n. 39, 2014, p. 121-142.

_____. Contrato, Virtudes e o Problema da Punição. **Dissertatio**, n. 43, 2016, p. 11-40.

COMTE, Jaimir; GELAIN, Itamar (orgs.). **Ensaio sobre a Filosofia de Strawson**. Florianópolis: Editora da UFSC, 2015. p. 245-269.

COSTA, Claudio. Livre Arbítrio: Como ser um Bom Compatibilista? **Princípios**, v. 7, n. 8, 2000, p. 19-33.

DAVIDSON, Donald. Knowing One's Own Mind. **Proceedings and Addresses of the American Philosophical Association**, v. 60, n. 3, 1987, p. 441–458.

DEECKE, Luder; GROZINGER, Berta; KORNHUBER, Hans. Voluntary Finger Movement in Man: Cerebral Potentials and Theory. **Biol Cybern**, v. 23, n. 2, 1976, p. 99-119.

DENNETT, Daniel. **Freedom Evolves**. New York: Viking Penguin, 2003.

_____. **Elbow Room**. Oxford: Oxford University Press, 1984.

DELGADO, Jose. **Physical Control of the Mind: Toward a Psychocivilized Society**. New York: Harper and Row, 1969.

DI MUZIO, Gianluca. Aristotle on Improving One's Character. **Phronesis**, v. 45, n. 3, 2000, p. 205-219.

DOUBLE, Richard. **The Non-Reality of Free Will**. New York: Oxford University Press, 1991.

DUMMETT, Michael. Bringing About the Past. **Philosophical Review**, v. 73, n. 3, 1964, p. 338-359.

DWORKIN, Gerald. **The Theory and Practice of Autonomy**. Cambridge: Cambridge University Press, 1988.

EKSTROM, Laura. Libertarianism and Frankfurt-style Cases. In: KANE, Robert (org.). **The Oxford Handbook of Free Will**. New York: Oxford University Press, 2002, p. 309-322.

ESHLEMAN, Andrew. "Moral Responsibility". *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Edward N. Zalta (ed.). URL: <https://plato.stanford.edu/archives/win2016/entries/moral-responsibility>.

FISCHER, John. **The Metaphysics of Free Will: A Study of Control**. Oxford: Blackwell, 1994.

_____. **My Way: Essays on Moral Responsibility.** New York: Oxford University Press, 2006.

_____. **Deep Control: Essays on Free Will and Value.** New York: Oxford University Press, 2012.

_____. Responsibility and Control. **The Journal of Philosophy**, v. 79, n. 1, 1982, p. 24-40.

_____. Moral Responsibility and the Metaphysics of Free Will: Reply to van Inwagen. **The Philosophical Quarterly**, v. 48, n. 191, 1998, p. 215-220.

_____. Recent Work on Moral Responsibility. **Ethics**, v. 110, n. 1, 1999, p. 93-139.

_____. Semicompatibilism and Its Rivals. **The Journal of Ethics**, v. 16, n. 2, 2012, p. 117-143.

_____. Responsibility and Manipulation. **The Journal of Ethics**, v. 8, n. 2, 2004, p. 145-177.

_____. Responsibility, Control, and Omissions. **The Journal of Ethics**, v. 1, n. 1, 1997, p. 45-64.

_____. Responsibility, History and Manipulation. **The Journal of Ethics**, v. 4, n. 4, 2000a, p. 385-391.

_____. Chicken Soup for the Semi-Compatibilist Soul: Replies to Haji and Kane. **The Journal of Ethics**, v. 4, n. 4, 2000b, p. 404-407.

_____. The Frankfurt Cases: The Moral of the Stories. **Philosophical Review**, v. 119, n. 3, 2010a, p. 315-336.

_____. Precis of My Way: Essays on Moral Responsibility. **Philosophy and Phenomenological Research**, v. 80, n. 1, 2010b, p. 229-241.

_____. Frankfurt-type Examples and Semi-Compatibilism. In: KANE, Robert. (org.). **The Oxford Handbook of Free Will.** Oxford: Oxford University Press, 2002, p. 281-308.

_____. The Zygote Argument Remixed. **Analysis**, v. 71, n. 2, 2011, p. 1-6.

_____. Autonomous Agents: From Self-Control to Autonomy by Alfred Mele. **Noûs**, v. 33, n. 1, 1999, p. 133-143.

_____. Free Will and Luck by Alfred Mele. **Mind**, v. 117, n. 465, 2008, p. 195-201.

_____. Reply: The Free Will Revolution. **Philosophical Explorations**, v. 8, n. 2, 2005, p. 145-156.

_____. The Free Will Revolution (Continued). **The Journal of Ethics**, v. 10, n. 3, 2006, p. 315-345.

FISCHER, John; RAVIZZA, Mark. Replies. **Philosophy and Phenomenological Research**, v. 61, n. 2, 2000, p. 467-480.

_____. Responsibility and Inevitability. **Ethics**, v. 101, n. 2, 1991, p. 258-278.

_____. **Responsibility and Control: A Theory of Moral Responsibility**. New York: Cambridge University Press, 1998.

FISCHER, John; KANE, Robert; PEREBOOM, Derk; VARGAS, Manuel. **Four Views on Free Will**. Oxford: Blackwell Publishing, 2007.

FISCHER, John; TOGNAZZINI, Neal. The Triumph of Tracing. In: FISCHER, John. (org.). **Deep Control: Essays on Free Will and Value**. New York: Oxford University Press, 2012, p. 206-233.

_____. The Physiognomy of Responsibility. **Philosophy and Phenomenological Research**, v. 82, n. 2, 2011, p. 381-417.

FRANKFURT, Harry. **The Importance of What We Care About**. Cambridge: University Press, 1998.

_____. Alternate Possibilities and Moral Responsibility. In: FRANKFURT, Harry. **The Importance of What We Care About**. New York: Cambridge University Press, 1998, p. 1-10.

_____. Reply to John Martin Fischer. In: BUSS, Sarah; Overton, Lee (eds.). **Contours of Agency: Essays on Themes from Harry Frankfurt**. Cambridge, 2002, p. 27-31.

_____. What We Are Morally Responsible For. In: FISCHER, John; RAVIZZA, Mark (org.). **Perspectives on Moral Responsibility**. London: Cornell University Press, 1993, p. 286-295.

_____. An Alleged Asymmetry Between Actions and Omissions. **Ethics**, v. 104, n. 3, 1994, p. 620-623.

_____. Freedom of the Will and the Concept of a Person. **The Journal of Philosophy**, v. 68, n. 1, 1971, p. 5-20.

_____. Three Concepts of Free Action. **Proceedings of the Aristotelian Society**, v. 49, n. 1, 1975, p. 113-125.

FREUD, Sigmund. Civilization and Its Discontents. In: STRACHEY, James (ed.) **Complete Psychological Works of Sigmund Freud**. London: W. W Norton, 1961.

GARRETT, Brian. **What is this Thing Called Metaphysics?** New York: Routledge, 2006.

GINET, Carl. 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.

_____. Working With Fischer and Ravizza's Account of Moral Responsibility. **The Journal of Ethics**, v. 10, n.3, 2006, p. 229-253.

GOETZ, Stewart. Frankfurt-Style Counterexamples and Begging the Question. **Midwest Studies in Philosophy**, v. 29, n. 1, 2005, p. 83-105.

HAJI, Ishtiyaque. Historicism, Non-Historicism, or a Mix? **The Journal of Ethics**, v. 17, n. 3, 2013, p. 185-204.

HAJI, Ishtiyaque; CUPERS, Stefaan. Hard-and Soft-Line Responses to Pereboom's Four-Case Manipulation Argument. **Acta Analytica**, v. 21, n. 4, 2006, p. 19-35.

HOBBS, Thomas. **The English Works of Thomas Hobbes of Malmesbury, Vol. 3: Leviathan**. London: John Bohn, 1839.

HOBUSS, João. A Responsabilidade Moral e a Possibilidade de Agir de Outro Modo. **Veritas**, v. 57, n. 1, 2012, p. 9-25.

HOBUSS, João. Aristóteles e a possibilidade de mudança de caráter. **Dois Pontos**, v. 10, n. 2, 2013, p. 291-313.

HORGAN, Terry. Injecting the Phenomenology of Agency into the Free Will Debate. In: Shoemaker, David (ed.). **Oxford Studies in Agency and Responsibility**. V.3. Oxford: Oxford University Press, 2015, p. 34-60.

HORGAN, Terence. 'Could', Possible Worlds, and Moral Responsibility. **Southern Journal of Philosophy**, v. 17, n. 3, 1979, p. 345-358.

_____. Compatibilism and the Consequence Argument. **Philosophical Studies**, v. 47, n. 3, 1985, p. 339-356.

HUBY, Pamela. The First Discovery of the Freewill Problem. **Philosophy**, v. 42, n. 162, 1967, p.353-362.

HUME, David. **Investigações sobre o Entendimento Humano e sobre os Princípios da Moral**. São Paulo: UNESP, 2003.

_____. **Tratado da Natureza Humana: Uma Tentativa de Introduzir o Método Experimental de Raciocínio nos Assuntos Morais**. São Paulo: UNESP, 2009.

HUXLEY, Aldous. **Brave New World: Revisited**. New York: Rosetta Books, 2000.

IRWIN, Terence. Reason and Responsibility in Aristotle. In: RORTY, Amélie (Org.). **Essays on Aristotle's Ethics**. Los Angeles: University of California Press, 1980, p. 117-155.

KANE, Robert. **The Significance of Free Will**. New York: Oxford University Press, 1998.

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

_____. **Free Will and Values**. New York: Albany State University of New York Press, 1985.

_____. (ed.). **The Oxford Handbook of Free Will**. New York: Oxford University Press, 2002.

_____. Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 299-321.

_____. Responsibility, Indeterminism and Frankfurt-style Cases: A Reply to Mele and Robb. In: WIDERKER, David; MCKENNA, Michael. (org.). **Moral Responsibility and Alternate Possibilities: Essays on the Importance of Alternative Possibilities**. Aldershot: Ashgate Publishing, 2006, p. 91-105.

_____. Can a Traditional Libertarian or Incompatibilist Free Will Be Reconciled with Modern Science? Steps Toward a Positive Answer. In: SUAREZ, Antoine; ADAMS, Peter. (org.). **Is Science Compatible with Free Will? Exploring Free Will and Consciousness in the Light of Quantum Physics and Neuroscience**. New York: London, 2013, p. 255-272.

_____. Responsibility, Luck and Chance: Reflections on Free Will and Indeterminism. **Journal of Philosophy**, v. 96, n. 5, 1999, p. 217-240.

_____. Two kinds of Incompatibilism. **Philosophy and Phenomenological Research**, v. 50, n. 2, 1989, p. 219-254.

KANT, Immanuel. **Fundamentação da Metafísica dos Costumes**. Lisboa: Edições 70, 2007.

_____. **Crítica da Razão Prática**. São Paulo: Martins Fontes, 2016.

KING, Matt. The Problem with Manipulation. **Ethics**, v. 124, n. 1, 2013, p. 65-83.

LAPLACE, Pierre-Simon. Determinism, Ignorance, and Probability. In: MCGREW, Timothy; ALSPECTOR-KELLY, Marc; ALLHOF, Fritz (eds.). **The Philosophy of Science: An Historical Anthology**. Oxford: Wiley-Blackwell, 2009, p. 250-254.

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

LEVY, Neil; MCKENNA, Michael. Recent Work on Free Will and Moral Responsibility. **Philosophy Compass**, v. 4, n. 1, 2009, p. 96-133.

LEWIS, David. Are We Free to Break the Laws? **Theoria**, v. 47, 1981, p. 113-121.

LOCKE, Don. Three Concepts of Free Action I. **Proceedings of the Aristotelian Society**, v. 49, 1974, p. 104-106.

LONG, Todd. Moderate Reasons-Responsiveness, Moral Responsibility, and Manipulation. In: CAMPBELL, Joseph; O'ROURKE, Michael; SHIER, David (eds.). **Freedom and Determinism**. Cambridge: The MIT Press, 2004, p. 151-172.

MARMODORO, Anna. Moral Character *versus* Situations: an Aristotelian Contribution to the Debate. **Journal of Ancient Philosophy**, v. 5, n. 2, 2011, p. 1-24.

MELE, Alfred. Manipulation, Compatibilism, and Moral Responsibility. **The Journal of Ethics**, v. 12, n. 3, 2008, p. 263-286.

_____. Moral Responsibility and Agents' Histories. **Philosophical Studies**, v. 142, n. 2, 2009, p. 161-181.

_____. Moral Responsibility and History Revisited. **Ethical Theory and Moral Practice**, v. 12, n. 5, 2009, p. 463-475.

_____. Moral Responsibility, Manipulation, and Minutelings. **The Journal of Ethics**, v. 17, n. 3, 2013, p. 153-166.

_____. **Free Will and Luck**. New York: Oxford University Press, 2006a.

_____. Fischer and Ravizza on Moral Responsibility. **The Journal of Ethics**, v. 10, n. 3, 2006b, p. 283-294.

_____. **Autonomous Agents. From Self-Control to Autonomy**. New York: Oxford University Press, 1995.

_____. The Significance of Free Will by Robert Kane. **The Journal of Philosophy**, v. 95, n. 11, 1998, p. 581-584.

_____. Reactive Attitudes, Reactivity, and Omissions. **Philosophy and Phenomenological Research**, v. 61, n. 2, 2000, p. 447-452.

MELE, Alfred; ROBB, David. Bbs, Magnets and Seesaws: The Metaphysics of Frankfurt-style Cases. In: WIDERKER, David; MCKENNA, Michael. (org.). **Moral Responsibility and Alternative Possibilities. Essays on the Importance of Alternative Possibilities**. Aldershot: Ashgate Publishing, 2006, p. 127-138.

_____. Rescuing Frankfurt-Style Cases. **The Philosophical Review**, v. 107, n. 1, 1998, p. 97-112.

MCKENNA, Michael. Alternative Possibilities and the Failure of the Counter-example Strategy. **Journal of Social Philosophy**, v. 28, n. 3, 1997, p. 71-85.

_____. Defending Nonhistorical Compatibilism. **Philosophical Issues**, v. 22, n. 1, 2012, p. 264-280.

_____. Moral Responsibility, Manipulation Arguments, and History: Assessing the Resilience of Nonhistorical Compatibilism. **The Journal of Ethics**, v. 16, n. 2, 2012, p. 145-174.

_____. Responsibility and Control: A Theory of Moral Responsibility by John Martin Fischer and Mark Ravizza. **The Journal of Philosophy**, v. 98, n. 2, 2001, p. 93-100.

_____. A hard-line Reply to Pereboom's Four-Case Manipulation Argument. **Philosophy and Phenomenological Research**, v. 77, n. 1, 2008, p. 142-159.

_____. Reasons-Responsiveness, Agents, and Mechanisms. In: Shoemaker, David (ed.). **Oxford Studies in Agency and Responsibility**. Volume 1. Oxford: Oxford University Press, 2013, p. 151-183.

MEYER, Susan. Aristotle on the Voluntary. In: KRAUT, Richard (org.). **The Blackwell Guide to Nicomachean Ethics**. Oxford: Blackwell, 2006, p. 137-157.

_____. **Aristotle on Moral Responsibility: Character and Cause**. New York: Oxford University Press, 1993.

MILL, John Stuart. **Utilitarianism and On Liberty**. Oxford: Blackwell Publishing, 2003.

NAGEL, Thomas. Moral Luck. In: STATMAN, Daniel (org.). **Moral Luck**. New York: State University of New York Press, 1993, p. 57-71.

NAGEL, Thomas. **The View From Nowhere**. New York: Oxford University Press, 1986.

NATALI, Carlo. Por que Aristóteles escreveu o Livro III da EN? *Analytica*, v. 8, n. 2, 2004, p. 47-74.

NAYLOR, Margery. Frankfurt on the Principle of Alternative Possibilities. **Philosophical Studies**, v. 46, n. 2, 1984, p. 249-258.

NUSSBAUM, Martha. **The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy**. New York: Cambridge University Press, 2001.

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

_____. Reasons-responsiveness, Alternative Possibilities, and Manipulation Arguments against Compatibilism: Reflections on John Martin Fischer's My Way. **Philosophical Books**, v. 47, n. 3, 2006, p. 198-212.

RACHELS, James. **Problemas da Filosofia**. Londres: Gradiva, 2009.

RIBEIRO, Leonardo. Semicompatibilismo, Responsabilidade e Manipulação. **Ethic@**, v. 10, n. 2, 2011, p. 255-279.

ROSS, David. **Aristotle**. London: Routledge, 1995.

RUSSELL, Paul. "Hume on Free Will". *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition). Edward N. Zalta (ed.). URL: plato.stanford.edu/archives/win2016/entries/hume-freewill.

SARTORIO, Carolina. A Partial Defense of the Actual-Sequence Model of Freedom. **The Journal of Ethics**, v. 20, n. 1-3, 2016, p. 107-120.

- SLOTE, Michael. Understanding Free Will. **Journal of Philosophy**, v. 77, 1980, p. 136-151.
- SMITH, Nowell. Freewill and Moral Responsibility. **Mind**, v. 57, n. 225, 1948, p. 45-61.
- SMITH, Angela. Responsibility for Attitudes: Activity and Passivity in Mental Life. **Ethics**, v. 115, n. 2, 2005, p. 236-271.
- SCANLON, Thomas. **Moral Dimensions: Permissibility, Meaning, Blame**. Cambridge: Harvard University Press, 2008.
- _____. **What We Owe to Each Other**. Cambridge: Harvard University Press, 1998.
- _____. The Significance of Choice. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 38-57.
- SHARPLES, Robert. **Stoics, Epicureans and Sceptics: An Introduction to Hellenistic Philosophy**. London: Routledge, 1996.
- SKINNER, Burrhus. **Walden Two**. Cambridge: Hackett Publishing Company, 2005.
- _____. **Beyond Freedom and Dignity**. Middlesex: Penguin Books, 1971.
- STRAWSON, Peter. **Freedom and Resentment and Other Essays**. New York: Routledge, 2008.
- _____. Freedom and Resentment. **Proceedings of the British Academy**, v. 48, 1962, p. 1-25.
- STRAWSON, Galen. **Freedom and Belief**. New York: Oxford University Press, 1986.
- _____. The Impossibility of Moral Responsibility. **Philosophical Studies**, v. 75, n. 1/2, 1994, p. 5-24.
- VAN INWAGEN, Peter. **An Essay on Free Will**. Oxford: Clarendon Press, 1983.
- _____. Ability and Responsibility. **The Philosophical Review**, v. 87, n. 2, 1978, p. 201-224.
- _____. Moral Responsibility, Determinism, and the Ability to do Otherwise. **The Journal of Ethics**, v. 3, n. 4, 1999, p. 343-351.
- _____. An Argument for Incompatibilism. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 38-57.
- VARGAS, Manuel. On the Importance of History for Responsible Agency. **Philosophical Studies**, v. 127, n. 3, 2006, p. 351-382.
- _____. The Trouble with Tracing. **Midwest Studies in Philosophy**, v. 29, n. 1, 2005, p. 269-291.

WALLACE, Jay. **Responsibility and Moral Sentiments**. Cambridge: Harvard University Press, 1994.

_____. Addiction as Defect of the Will: Some Philosophical Reflections. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 424-452.

WALLER, Bruce. Free Will Gone Out Of Control: A Review of Kane's Free Will and Values. **Behaviorism**, v. 16, n. 2, 1988, p. 149-157.

WATSON, Gary. Soft Libertarianism and Hard Compatibilism. **The Journal of Ethics**, v. 3, n. 4, 1999, p. 353-368.

_____. Free Action and Free Will. **Mind**, v. 96, n. 382, 1987, p. 145-172.

_____. Reasons and Responsibility. **Ethics**, v. 111, n. 2, 2001, p. 374-394.

_____. Introduction. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 1-25.

_____. Two Faces of Responsibility. **Philosophical Topics**, v. 24, n. 2, 1996, p. 227-248.

WIDERKER, David. Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. **The Philosophical Review**, v. 104, n. 2, 1995, p. 247-261.

WIDERKER, David.; MCKENNA, Michael. (org.). **Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities**. Aldershot: Ashgate Publishing, 2006.

WIGGINS, David. Towards a Reasonable Libertarianism. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 94-121.

WILLIAMS, Bernard. **Moral Luck**. Cambridge: Cambridge University Press, 1981.

_____. Moral Luck. In: STATMAN, Daniel (org.). **Moral Luck**. New York: State University of New York Press, 1993, p. 35-56.

_____. **Shame and Necessity**. Los Angeles: University of California Press, 1993.

WYMA, Keith. Moral Responsibility and Leeway for Action. **American Philosophical Quarterly**, v. 34, n. 1, 1997, p. 57-70.

WOLF, Susan. **Freedom within Reason**. New York: Oxford University Press, 1990.

_____. Asymmetrical Freedom. **Journal of Philosophy**, v. 77, n. 3, 1980, p. 151-166.

_____. Sanity and the Metaphysics of Responsibility. In: WATSON, Gary (org.). **Free Will**. 2nd Edition. Oxford: Oxford University Press, 2013, p. 372-387.

ZIMMERMAN, Michael. **An Essay on Moral Responsibility**. Totowa: Rowman & Littlefield, 1988.

ZINGANO, Marco. **Aristóteles: Ethica Nicomachea I 13-III 8**. Tratado de Filosofia Moral. São Paulo: Odysseus, 2008.

_____. Aristóteles, Alexandre e o que está em nosso poder: libertarianismo e responsabilidade moral. In: HOBBUS, João (org.). **Ética das Virtudes**. Florianópolis: Editora da UFSC, 2011, p. 85-108.