

LIBERDADE, DETERMINISMO E NATURALISMO

Roberto Horácio de Sá Pereira

UFRJ

O presente artigo possui dois objetivos centrais. O primeiro consiste na realização de uma crítica às tentativas contemporâneas de compatibilizar a imputabilidade com a tese do determinismo causal a partir das chamadas análises condicionais dos conceitos de liberdade da ação e liberdade do arbítrio. O segundo se identifica à busca de uma nova formulação para a concepção incompatibilista de liberdade entendida como espontaneidade do arbítrio, nos termos da qual a imputação é inteiramente incompatível com qualquer forma de determinismo causal. Fornecerei uma explicação naturalista, mas não determinista, para o conceito de liberdade da vontade enquanto espontaneidade.

Para atingir o primeiro objetivo, apresentarei o compatibilismo nos termos da análise condicional do conceito de liberdade proposta por Tugendhat. Essa escolha se justifica, em primeiro lugar, pela clareza conceitual com que Tugendhat descreve e desenvolve a chamada análise condicional do conceito de liberdade proposta inicialmente por Moore. Em segundo lugar, essa opção se justifica na medida em que a versão proposta por Tugendhat é, sem sombra de dúvidas, a mais refinada, buscando responder às principais dificuldades apresentadas nas versões anteriores.

Contra a versão do compatibilismo proposto por Tugendhat, apresento duas considerações que me parecem decisivas. Em primeiro lugar, busco assinalar que Tugendhat na sua análise condicional para a noção de “querer refletido”, realizada com base em uma análise igualmente condicional do conceito de “capacidade de reflexão prática”, incorre em petição de princípio. Em segundo lugar,

apresento um diagnóstico sobre o equívoco que motiva o condicionalismo proposto pelo autor.

Para atingir o segundo objetivo, retomarei as principais teses kantianas acerca da deliberação racional (tal como elas foram interpretadas por Allison). Rejeitarei, contudo, o caráter inteligível que Kant atribui à espontaneidade do arbítrio como também sua tentativa de compatibilizar o conceito (incompatibilista) de liberdade como espontaneidade com o determinismo causal à luz do seu Idealismo Transcendental. Neste particular, pretendo mostrar que a espontaneidade do arbítrio não precisa ser concebida metafisicamente como um substrato inteligível e incognoscível, produtor das ações observadas no plano fenomênico, podendo ser pensada como um dos elementos essenciais da nossa teoria explicativa das ações de terceiros.

Minha tese central pode ser resumida nos termos de que, segundo o enfoque performativo daquele que age, livres e responsáveis são aqueles indivíduos que escolhem suas ações com base em desejos primários, ao incorporá-los em máximas ou regras de preferência espontaneamente. E segundo o enfoque cognitivo daquele que interpreta a conduta alheia, livres e responsáveis por suas ações são aqueles indivíduos cujo comportamento em geral pode ser compreendido à luz de razões primárias e secundárias. Nestes termos, um indivíduo só será responsável pelas suas ações quando seus intérpretes potenciais se virem forçados a postular, no esquema interpretativo do seu agir, desejos e convicções espontâneos de segunda ordem (acerca dos seus desejos primários e dos fatores externos à sua vontade) para tornarem seu comportamento em geral inteligível.

Na filosofia contemporânea o compatibilismo é entendido como a tentativa de conciliação entre a tese da liberdade da ação e da vontade e a tese do determinismo causal. Embora livres e imputáveis, as ações intencionais seriam determinadas pelas leis da causalidade natural como tudo o mais na natureza que tem lugar no espaço e no tempo. A principal motivação (negativa) do compatibilismo reside na rejeição de um suposto caráter incondicional atribuído ao arbítrio humano. Não faria sentido concebermos a liberdade da vontade como

uma capacidade sobrenatural, pertencente a um “reino à parte”, distinto do reino natural sujeito ao determinismo causal.

Inúmeros são os filósofos que se filiam a esta tradição. Mas, segundo meu juízo, a contribuição contemporânea que melhor exemplifica tal empreendimento é de autoria de Tugendhat. Paradigmaticamente esse autor afirma, por exemplo:

Para Kant a própria moral era algo metafísico, parecendo-lhe portanto razoável ligar o fato metafísico da moral ao fato sobrenatural da liberdade. Contudo, aquela imputabilidade de que fazemos uso na reflexão e à qual nos referimos quando fazemos repreensões é um fato empírico da nossa vida cotidiana. (...) Esse fato não tem que ser provado, tem de ser apenas elucidado.¹

Na sua análise dos conceitos de liberdade e de imputação, Tugendhat nomeia três domínios distintos. O primeiro seria o domínio da moral, o segundo seria o domínio do direito penal e o terceiro seria o domínio da reflexão prática em geral. Considerando que os dois primeiros domínios pressupõem o terceiro (o qual apenas acrescenta fatores adicionais), Tugendhat circunscreve sua contribuição teórica ao plano da reflexão prática em geral entendido como o domínio básico ou fundamental.

Sua análise encontra o ponto de partida na formulação do problema proposta por Moore em “Free Will”². Nesse célebre artigo, após caracterizar como *boas* apenas aquelas ações que produziriam as melhores conseqüências possíveis frente às demais ações que o agente *poderia* realizar, Moore conclui que a avaliação de uma ação singular sempre supõe que seu agente *poderia* tê-la evitado ou *poderia* ter atuado de forma distinta (em um determinado sentido que ele busca ao longo do artigo elucidar)³. Retomando a análise de Moore, Tugendhat reafirma que o

(1) Tugendhat, E.: “Der Begriff der Willensfreiheit”, em: E. Tugendhat: *Philosophische Aufsätze*, Frankfurt a. M. 1992. P. 340.

(2) Moore, G. E.: “Free Will”, em: G.E. Moore: *Ethics*, London 1912. P. 299-311.

(3) Comparar, Moore, G. E.: “Free Will”. P. 300.

problema da liberdade enquanto imputabilidade se resume à elucidação da seguinte proposição: *ele poderia ter agido diferentemente*⁴.

Segundo Tugendhat devemos ter claro, em primeiro lugar, que o conceito de “possibilidade” em questão não se opõe ao conceito de causalidade natural, mas ao conceito de coação no sentido de causa alheia à vontade do agente. Com base em Aristóteles (*Ética a Nicômaco*), ele afirma que a imputabilidade deveria satisfazer duas condições. (1) Antes de mais nada, o agente deveria saber o que está fazendo ao agir.(2) Além disso, a ação não poderia ser produto de uma coação externa, i. e., de uma causalidade externa contrária à vontade do agente. Se a estas duas condições acrescentássemos uma terceira, qual seja, (3) a de que uma ação imputável tampouco pode ser o resultado de uma coação interna, i.e., de uma causalidade interna mas igualmente contrária à vontade refletida do agente, não haveria mais nenhum impedimento aparente para considerarmos a ação imputável como um acontecimento natural como outro qualquer. O relevante seria encontrarmos critérios empíricos para a sua identificação e distinção frente aos demais acontecimentos naturais⁵.

Sabemos até aqui que o conceito de possibilidade se define negativamente como ausência de coação, queremos saber agora como devemos entendê-lo positivamente. Neste contexto, Tugendhat retoma a principal tese de Moore. Para Moore, a palavra “poderia” se caracteriza essencialmente por um duplo sentido. Em uma primeira acepção, o termo “poderia” é empregado negativamente como um operador modal de proposições que exprimem a ocorrência de acontecimentos passados à luz do princípio de causalidade: (i) se houve efetivamente uma segunda guerra mundial e (ii) se tudo o que ocorre, ocorre necessariamente, então a segunda guerra não *poderia* deixar de ter ocorrido. Segundo esta acepção, nada do que ocorreu segundo o princípio da causalidade - nem mesmo as ações humanas - *poderia* deixar de ter ocorrido.

(4) “Sie hätte auch anders können”. Comparar, “Der Begriff der Willensfreiheit”. *Philosophische Aufsätze*. P. 334-5.

(5) Comparar, Tugendhat, E. “Der Begriff der Willensfreiheit”. *Philosophische Aufsätze* . P. 342.

Há, no entanto, uma segunda acepção para a expressão “poderia” nos termos da qual ela seria compreendida como expressão de uma capacidade ou habilidade. Tal acepção seria aquela presente em afirmações, como por exemplo, de que Moore *poderia* ter corrido uma milha em vinte minutos esta manhã, mas *não poderia* de forma alguma ter corrido vinte milhas em cinco minutos. De acordo com esta segunda acepção, a expressão *poderia* assumiria a forma de um predicado disposicional, usualmente analisada nos termos de *condicionais comportamentais*. Assim, seria possível definirmos a propriedade “poderia se quebrar” (ou simplesmente “quebrável”), traduzindo enunciados que a exprimem (por exemplo, “um determinado vaso *poderia* se quebrar”) nos termos de condicionais comportamentais (por exemplo, “o vaso se quebraria, *se* caísse de uma certa altura, ou *se* fosse atingido por outro objeto com determinada massa, velocidade etc”).

Sendo assim, parece haver pelo menos um sentido para a expressão “poderia” que não se oporia à necessidade ou ao “ter que” incondicional da causalidade natural. Pelo contrário: assumindo a forma de um condicional irreal, cujo antecedente descreve as condições causais subjuntivas para a ocorrência dos comportamentos descritos pelo seu conseqüente, este “poderia” estaria indissociavelmente ligado ao conceito de causalidade. Ora, mas esse sentido disposicional de “poderia” seria justamente aquele relevante para qualquer imputação ou atribuição de responsabilidade. Ao afirmarmos que um agente é responsável pelo que fez, supomos então que ele *poderia* ter atuado de forma distinta, no sentido disposicional em que possuía a habilidade ou a capacidade para tal.

Mas se a liberdade enquanto imputabilidade é uma capacidade expressa por um predicado disposicional, é indispensável distingui-la de outras capacidades que por ventura também possam ser representadas da mesma forma. O próximo passo na análise de Tugendhat consiste então na distinção da liberdade entendida como uma capacidade humana frente às capacidades próprias aos objetos inanimados.

Mais uma vez a análise de Moore é uma referência obrigatória. A liberdade enquanto capacidade humana se distinguiria das demais propriedades

disposicionais na medida em que nela *o querer* desempenha um papel decisivo. Ao afirmarmos, por exemplo, que “um determinado indivíduo poderia tocar violino”, temos em mente, além dos fatores externos que condicionam sua ação (como a presença do referido instrumento) e a existência de uma habilidade específica, a vontade como o fator determinante. Tugendhat propõe então uma primeira análise condicional para a noção de liberdade nos termos:

(1) *S* pode fazer *A* \equiv Def. *S* realiza *A*, se ele quer.

Entretanto, como o próprio Moore já assinalara em seu artigo, qualquer definição similar a essa estaria circunscrita à esfera do agir, deixando portanto em aberto a pergunta sobre se tal indivíduo *poderia escolher* de forma distinta. Além disso, essa definição também é insatisfatória pois em seus termos seria possível atribuímos responsabilidade a animais, crianças pequenas e adultos senis, tendo em vista que suas ações, como as nossas, são voluntárias ou produzidas por um querer.⁶

Tugendhat pensa poder contornar tal dificuldade introduzindo na definição proposta por Moore o conceito de “querer refletido”. Seria justamente este conceito que, negligenciado por toda a tradição compatibilista (de Hume a Moore), teria levado Kant a postular sua concepção incompatibilista de liberdade, a liberdade transcendental. O desafio que Tugendhat tem pela frente consiste, portanto, em apresentar uma análise condicional para a própria noção de “querer refletido” que torne a liberdade e a atribuição de responsabilidade compatíveis com o determinismo causal.

(6) A resposta de Moore consiste simplesmente em reintroduzir no plano da vontade a distinção entre duas acepções da expressão “poder”. No sentido em que “poder” exprime uma necessidade natural, um indivíduo tendo escolhido uma forma determinada de ação não poderia deixar de tê-la escolhido. No sentido, entretanto, em que “poder” exprime uma capacidade, não haveria mais uma vez nenhum problema ao afirmarmos que um determinado indivíduo poderia ter escolhido diferente, *desde que ele tivesse escolhido fazer tal escolha*. Neste termos, Moore parece pretender analisar condicionalmente o próprio conceito de liberdade da vontade. Comparar, Moore, p. 310.

Antes de mais nada, o “querer refletido” deve ser distinguido do “querer sensível”: enquanto este último seria definido como aquelas formas de desejos sensíveis primários determinados pelas sensações de prazer ou do agradável, o querer refletido seria aquela forma de querer determinada por razões e argumentos. Isso parece ser tudo o que Tugendhat precisa para propor (ainda de forma provisória) uma nova definição condicional para a noção de responsabilidade entendida como uma capacidade humana:

(2) S é responsável pela ação A ≡ Def. S faz A, quando ele quer A de forma refletida. E S pode querer A de forma refletida.

O problema de Tugendhat a partir deste ponto consiste em apresentar um critério empírico para verificarmos quando um determinado sujeito S pode querer uma certa forma de ação (A) refletidamente. Considerando, contudo, que a escolha refletida de uma forma de ação supõe a capacidade de refletir, para tornar (2) empiricamente operacionalizável, Tugendhat apresenta uma nova análise condicional, desta vez para a noção de capacidade de reflexão. Sua proposta consiste em defini-la nos seguinte termos:

(3) S pode refletir ≡ Def. S reflete, quando ele quer.

Para evitar qualquer acusação de circularidade, Tugendhat se apressa em esclarecer que o querer que aparece no antecedente causal do condicional não é o querer refletido - que justamente se busca definir - mas o querer sensível, definido com base nas sensações do agradável e do prazeroso. Nos termos desta definição, *pode* refletir quem reflete quando de alguma forma possui um desejo sensível. Entretanto, esta definição ainda não nos fornece um critério operacional. Considerando provavelmente que a sensação do agradável tem lugar normalmente quando o indivíduo consegue otimizar as vantagens e minimizar as desvantagens que resultam das ações que ele

escolhe refletidamente, Tugendhat acredita que através das noções de vantagem e desvantagem seria possível formularmos a seguinte análise condicional para a noção de capacidade de reflexão:

(4) *S pode refletir* \equiv Def. em situações nas quais é necessário que se reflita, ou seja, em situações nas quais S se encontra diante de uma escolha entre vantagens e desvantagens, S normalmente reflete.

Nestes termos, saberíamos quando um indivíduo seria capaz de refletir. Bastaria isolarmos situações paradigmáticas nas quais as vantagens e desvantagens fossem inequívocas. Se em tais situações o sujeito escolhe como forma de ação a opção claramente mais vantajosa, maximizando assim a satisfação dos seus desejos sensíveis, atribuir-lhe-emos capacidade de refletir sobre suas ações. Podendo saber se um determinado indivíduo *pode refletir*, poderíamos igualmente saber se o mesmo *pode querer agir refletidamente* e, assim, se ele é *responsável* pelas suas ações.

Nos termos de tal análise condicional do querer refletido, o determinismo causal não traria mais nenhum prejuízo para a atribuição de responsabilidade e mesmo para a atribuição da capacidade de autodeterminação. Mesmo que pudessemos mostrar que a existência de desejos sensíveis e a presença de circunstâncias que maximizam sua satisfação conduzem os agentes a optar *inexoravelmente* por aquelas formas de ação que maximizam a satisfação de tais desejos, os agentes ainda assim deveriam ser considerados responsáveis e autodeterminados. Estes agentes seriam responsáveis por suas escolhas e autodeterminados (pela sua própria vontade), uma vez que as suas escolhas seriam determinadas por razões e não por causas alheias à sua vontade.

Como procurei destacar, a originalidade da proposta de Tugendhat consiste na introdução do conceito de “querer refletido” (negligenciado pela tradição compatibilista) na tentativa de fornecer uma nova análise condicional para a noção de imputabilidade ou de atribuição de responsabilidade. Em primeiro

lugar, gostaria de salientar que a série (1-4) das formulações propostas por Tugendhat não consegue analisar efetivamente a capacidade de “querer agir de forma refletida” em termos de condicionais comportamentais sem de algum modo supor outras formas de “querer refletido”. Recordando, Tugendhat pretendia definir a ação imputável como aquela realizada com base em uma vontade refletida [passo (2)]. Mas para isso seria necessário fornecer uma análise condicional para a própria capacidade de reflexão prática sem lançar mão da noção de “querer refletido”, apelando apenas à noção de “querer sensível” [passos (3) e (4)]. Nestes últimas tentativas de definição, atribuiríamos ao agente capacidade de reflexão prática quando, e apenas quando, ele fizesse a escolha da opção mais vantajosa diante de si. Entretanto, a escolha concreta da opção que melhor satisfaça seus desejos sensíveis diante de opções claramente mais e menos favoráveis não depende apenas da capacidade de reflexão prática por parte do seu agente, mas também do concurso de uma série de atos voluntários irredutíveis à sensação do agradável.

Antes de mais nada, para que um indivíduo escolha a opção que melhor satisfaça suas inclinações não basta que ele possua a capacidade de refletir praticamente e se encontre diante de opções claramente favoráveis e desfavoráveis à satisfação das suas inclinações. É igualmente indispensável que ele *queira* adotar espontaneamente a seguinte regra de preferência (prudencial): *confrontado com alternativas, sempre buscarei escolher de forma refletida ao invés de seguir irrefletidamente meus impulsos mais imediatos*. Se o indivíduo tiver adotado por máxima a regra de preferência contrária, digamos, de *sempre buscar satisfazer seus desejos mais imediatos*, mesmo possuindo a capacidade de reflexão, em muitos casos sua opção recairá sobre a alternativa menos favorável à maximização dos seus desejos, bastando para isso que a opção mais favorável não se apresente como aquela imediatamente mais agradável.

Mas para que um indivíduo escolha a opção claramente mais favorável é igualmente necessário que ele *queira* adotar uma segunda regra básica de preferência como sua máxima subjetiva: *confrontado com alternativas exclusivas de ação, sempre escolherei as opções que melhor satisfizerem minhas inclinações*. Se o indivíduo tiver adotado por máxima a *renúncia dos prazeres sensíveis*, ele naturalmente não escolherá a

opção que melhor satisfaz seus desejos sensíveis, ainda que (i) se veja diante de alternativas de ação que claramente satisfaçam seus desejos sensíveis, (ii) possua a capacidade de refletir e (iii) tenha se predisposto a agir de forma refletida.

Por conseguinte, não conseguimos definir a capacidade de reflexão prática em termos comportamentais senão sob a suposição de uma série de atos voluntários que exprimem diferentes formas de “querer refletido”. Mas como Tugendhat pretende definir a noção de “querer refletido” justamente através de uma definição da noção da “capacidade de reflexão prática”, acaba por incorrer em um círculo vicioso.

Talvez se pudesse alegar, contudo, que o objetivo central de Tugendhat não consiste em fornecer uma análise condicional para o conceito de liberdade em termos estritamente comportamentais. Embora suas quatro formulações decisivas sejam apresentadas inequivocamente como definições sob a forma de bicondicionais comportamentais, sua principal proposta consistiria em formular apenas um “critério” empírico e, portanto, falível para a noção de vontade livre que fosse ao mesmo tempo empiricamente operacionalizável e compatível com a idéia de determinação causal.

Mesmo nestes termos, a contribuição de Tugendhat me parece condenada ao fracasso. Antes de mais nada, continua havendo uma circularidade viciosa na formulação dos critérios empíricos para a atribuição de responsabilidade. Pois, para sabermos empiricamente quando o agente quer algo de forma refletida, temos que saber previamente se o agente é capaz de reflexão prática. Mas para sabermos se um determinado agente é capaz de refletir praticamente, não nos basta observá-lo escolhendo as opções que melhor satisfaçam suas inclinações ou desejos sensíveis. A atribuição de capacidade de refletir com base na observação da escolha de formas de ação que melhor satisfazem desejos sensíveis é indissociável da atribuição da vontade superior de satisfazer seus desejos sensíveis, sempre que possível.

Mas como a atribuição de racionalidade prática é indissociável da atribuição de toda uma série de estados intencionais aos agentes, não faz sentido formularmos um critério empírico comportamental para identificá-la isoladamente dos demais estados intencionais. Não temos como correlacionar inequivocamente cada estado intencional ou capacidade mental com um tipo padrão de comportamento

observável. Dito em outros termos, a base epistêmica de atribuições de estados e capacidades é holista: o que fazemos é correlacionar os comportamentos observados com toda uma série de estados e capacidades. Por esta razão, sob a atribuição de diferentes estados intencionais, podemos em princípio acomodar a atribuição de racionalidade prática com a observação de qualquer forma de comportamento. Ao observarmos (no exemplo de Tugendhat) um determinado agente escolher uma opção claramente menos vantajosa (diante de outras claramente mais vantajosas), nem por isso precisamos concluir que ele não possui a capacidade de refletir praticamente. Seu comportamento continua sendo perfeitamente racional se supusermos simplesmente que o agente acolheu como máxima (ou a regra de preferência) renunciar aos prazeres sensíveis sempre que possível.

Este caráter holista dos estados e capacidades mentais torna igualmente problemático o compatibilismo proposto por Tugendhat, pelo menos se compreendermos o determinismo como a tese da existência de leis causais estritas entre a vontade sensível e a reflexão prática. Podemos resumir a análise de Tugendhat em duas teses causais. Como em Moore, tratava-se em primeiro lugar de definir (ou pelo menos identificar) as ações livres e imputáveis como aquelas produzidas ou causadas por um querer refletido (o agente poderia ter agido de forma diversa, se assim tivesse escolhido agir). Posteriormente tratava-se de definir este querer refletido, identificando a reflexão prática como a forma de atividade produzida ou causada por uma forma de querer sensível na busca daquelas opções que melhor pudessem satisfazê-lo.

Ora, mas se o querer sensível só nos conduz à escolha refletida da opção que melhor o satisfaça, sob a suposição de que tenhamos adotado por máxima agir de forma a satisfazer os nossos desejos, parece óbvio que *nem sempre* que nos encontrarmos diante da opção mais agradável iremos escolher de forma refletida aquelas formas de ação que melhor satisfaçam os nossos desejos sensíveis. Não há, portanto, nenhuma lei causal estrita entre o querer sensível (causa) e a atividade de reflexão prática (efeito).

No meu entender, o erro clássico na tentativa contemporânea de compatibilizar o determinismo causal com a imputação de responsabilidade tem

origem em uma teoria equivocada sobre a motivação do agir racional. Esta teoria, que remonta a Leibniz e Hume, se caracteriza basicamente por duas teses centrais. A primeira se exprime pela afirmação de que apenas sentimentos ou inclinações são capazes de motivar o arbítrio humano a agir, de modo que a razão não desempenharia qualquer papel motivador. A segunda, como consequência direta da primeira, consiste na afirmação de que os sentimentos atuam sobre a vontade humana como forças (cegas) de atração e repulsão.

Nestes termos, todas as ações humanas deveriam ser consideradas como efeitos causais inteiramente determinados pelas forças que atuariam sobre a nossa vontade. Deste modo, as pretensas definições disposicionais nada mais fariam senão exprimir tal relação de determinação causal, buscando analisar os enunciados que descrevem o arbítrio humano (no exemplo de Tugendhat, “S quer uma ação A de forma refletida”) em enunciados condicionais comportamentais (“Quando S se encontra em situações nas quais há opções claramente mais e menos vantajosas, S escolhe a mais vantajosa”). Tugendhat assume essa teoria da motivação explicitamente:

Se nos interessarmos pela tese de Hume de que apenas sentimentos (como inclinações) podem ser determinantes de ações, então a suposição de Kant de que o mandamento, enquanto livre de afeto, poderia ser determinante da ação, revelar-se-ia como ficção e dever-se-ia afirmar que de qualquer modo só um afeto nos dá consciência do bem determinante da ação.⁷

Naturalmente, a satisfação das inclinações através do sentimento de prazer é a causa cega de uma série de comportamentos *não intencionais*, atuando como uma força de atração sobre o arbítrio humano. Mas quando consideramos o caráter intencional da conduta humana, não indagamos por qualquer relação causal, mas apenas por aquela capaz de justificar epistemicamente a escolha de ações a partir de razões, motivos ou argumentos. Por essa razão, mesmo a justificação da

escolha de uma forma de ação motivada *exclusivamente* pela sensação de prazer não se reduz de modo algum à relação causal cega entre impulso ou sensação de prazer e ação escolhida.

Em primeiro lugar, para que alguém possa justificar sua preferência por uma forma de ação motivada exclusivamente pela sensação de prazer, além da simples presença da sensação de prazer é absolutamente indispensável que o sujeito *saiba* se referir à tal sensação como *sua* sensação de prazer, através de diferentes predicados de sensação. Além disso, é igualmente necessário que o sujeito *saiba que* a ação por ele escolhida possui a propriedade de produzir a sensação de prazer. Sem o concurso de atos judicativos espontâneos que tipifiquem sensações e diferentes formas de ação, a ação perde seu caráter intencional. Ora, é desnecessário acrescentar que tais atos de classificação das sensações e das ações são irreduzíveis a qualquer sensação.

Mas isso não é tudo. Se a premissa da justificação encerra uma caracterização de determinadas formas de ação como aquelas que produzem a sensação de prazer e a conclusão enuncia a *preferência* justamente por alguma destas ações (que produzem prazer), para inferirmos tal conclusão a partir da referida premissa necessitamos de uma premissa adicional que enuncie justamente a preferência em geral pelas ações que produzem prazer. Caso contrário, haveria um hiato na justificação. Mais uma vez é desnecessário acrescentarmos que tal regra de preferência por formas de ação que produzem a sensação de prazer é irreduzível à própria sensação de prazer.

Assim, mesmo como explicação daquelas ações motivadas exclusivamente pela sensação de prazer, a teoria humeana da motivação do arbítrio humano se mostra inteiramente improcedente. Sem o concurso de juízos e regras de preferências (irreduzíveis à sensação), o afeto não nos dá nenhuma consciência do bem determinante de uma ação, contrariamente ao que afirma Tugendhat na passagem acima citada. Não é por outra razão que Kant afirma em franca oposição a Hume:

que as ações humanas são induzidas [veranlasst] mas não inteiramente determinadas pela sensibilidade, pois a Razão tem que fornecer um complemento de suficiência [ein Complement der zulanlichkeit].⁸

Na *Crítica da razão pura* a mesma tese se exprime através da célebre caracterização do arbítrio humano. Segundo Kant, somos seres portadores de um *arbitrium* ao mesmo tempo *sensitivum* e *liberum*, pois embora sejamos afetados pelos objetos que satisfazem nossas inclinações, nossas ações não são por eles exigidas patologicamente.⁹ Como Allison assinala,¹⁰ o cerne da teoria da motivação kantiana está contido na tese por ele denominada “tese da incorporação”. Embora tal tese já se encontre presente em trabalhos anteriores, ela só é apresentada com o devido destaque na *Religião nos limites da simples razão*, onde Kant a formula nos seguintes termos:

A liberdade do arbítrio [Willkür] possui uma constituição de tal forma peculiar que não pode ser determinada a uma ação por nenhum incentivo [Triebfeder], a não ser na medida em que a pessoa o tenha acolhido [aufgenommen] na sua máxima (tenha feito dele uma regra geral para si, em conformidade com a qual ela quer se comportar).¹¹

Ora, mas na medida em que a adoção de máximas por parte do agente exprime um *posicionamento* frente às suas inclinações, ela será irredutível às mesmas inclinações. Ao contrário do que sugere Tugendhat, creio que na afirmação de tal espontaneidade do arbítrio não há necessariamente nada de “metafísico” ou “sobrenatural”, mas um fato empírico e relativamente trivial da nossa vida ordinária. A grande maioria das nossas interações humanas parte da suposição que somos,

(8) Comparar, Kant, I.: “Reflexion” 5611 18, *Gesammelte Werke*, Berlin 1902.

(9) Comparar, Kant, I.: *Kritik der reinen Vernunft*, *Gesammelte Werke*, publicado pela Preuss. Akad. Berlin 1902 . P. A534=B562.

(10) Comparar, Allison, H. E.: *Kant’s Theory of Freedom*. P. 35-41.

(11) Comparar, Kant, I.: *Werke*, VI. P. 23.

com efeito, senhores das nossas ações no sentido indicado em que *somos capazes de um modo geral de nos posicionamos espontaneamente* frente aos nossos desejos primários e inclinações, rejeitando alguns e acolhendo outros sob a forma de máximas ou regras de preferências.

A partir deste ponto minha análise começa a divergir inteiramente da concepção kantiana. Apesar do conceito incompatibilista de liberdade enquanto espontaneidade que emerge da sua análise do agir racional, Kant parece endossar sem nenhum problema a tese do determinismo irrestrito, mesmo para o plano das descrições psicológicas da vontade e do agir, quando realizadas sob o enfoque da terceira pessoa. Para Kant, sob o enfoque cognitivo mesmo o arbítrio humano e as ações por ele causadas estariam em conformidade com uma concepção inteiramente determinista da natureza:

Se pudéssemos investigar exaustivamente todos os fenômenos da vontade humana, não encontraríamos um única ação humana que não pudesse ser predita com certeza e ser reconhecida como tendo resultado necessariamente das suas condições antecedentes.¹²

Como assinalam seus melhores intérpretes, o que Kant pretende com a sua *terceira antinomia* é curiosamente compatibilizar o conceito incompatibilista de liberdade (tese) com a doutrina do determinismo (antítese). Tal conciliação seria feita com base no seu Idealismo Transcendental. Nos termos desta doutrina, todos os eventos enquanto fenômenos que têm lugar no espaço e no tempo (incluindo aí as ações humanas), possuiriam uma realidade ou caráter empírico e seriam cognoscíveis segundo o princípio da causalidade natural. Mesmo o agir racional, quando considerado a partir desta perspectiva cognitiva, deveria ser considerado como determinado.

Mas além da perspectiva cognitiva das coisas como fenômenos sujeitos à causalidade natural, poderíamos também enfocá-las a partir de uma suposta

(12) Kant, I. *Kritik der reinen Vernunft*. Werke. P. A550=B578.

forma não empírica como entidades *inteligíveis* ou como coisas em geral = x. Relativamente às nossas ações isso significaria supor que elas têm origem em um substrato noumênico e incognoscível. Assim, Kant tornaria o conceito incompatibilista de liberdade enquanto espontaneidade compatível com a tese do determinismo simplesmente assinalando dois pontos de vista possíveis frente às ações. Sob o enfoque cognitivo do espectador, poderíamos conhecer o agir racional como um acontecimento temporal dentre outros como um fenômeno inteiramente determinado pela causalidade natural. Mas sob o enfoque performativo do agente, deveríamos *supor* que a nossa ação resultaria da nossa espontaneidade concebida como um substrato *inteligível* livre de quaisquer determinações.

As dificuldades inerentes ao suposto caráter *inteligível* da espontaneidade bem como a referida tentativa de conciliação idealista deste conceito incompatibilista com a tese do determinismo causal já foram exaustivamente discutidas pela literatura e sua implausibilidade me parece mais que evidente. Só gostaria de destacar – o que já foi afirmado também por inúmeros autores – que nos termos propostos pelo Idealismo kantiano as nossas práticas mais usuais de imputação ou atribuição de responsabilidade não fariam mais qualquer sentido. Sob um enfoque cognitivo, jamais poderíamos saber ou afirmar legitimamente que as ações humanas (mesmo as nossas próprias ações) são livres e espontâneas. Assim, para aqueles que consideram as práticas de imputação bem mais plausíveis do que a atribuição de um caráter inteligível à espontaneidade da nossa deliberação racional e do que a doutrina do Idealismo Transcendental, a alternativa que resta consiste em buscar uma formulação para a noção de liberdade da vontade, compreendida como espontaneidade, sob o enfoque cognitivo da terceira pessoa. Seguindo os passos de Kant, buscaremos uma definição para a espontaneidade prática a partir de um esclarecimento prévio da espontaneidade epistêmica.

Como os compatibilistas contemporâneos, Kant entende por liberdade da vontade uma capacidade ou faculdade peculiar à condição humana. Compreendida, entretanto, como espontaneidade, esta capacidade não poderia jamais assumir a forma de um predicado disposicional a ser analisado nos termos de condicionais comportamentais. Enquanto espontaneidade, a liberdade da vontade é

expressa antes de mais nada pelos atos cognitivos do entendimento: “O Eu penso que deve poder acompanhar minhas representações”. Restrita assim ao domínio epistêmico, a liberdade do arbítrio é expressa pelo emprego dos operadores epistêmicos conhecidos: eu *penso* que, eu *considero* que, eu *julgo* que, eu *tenho boas razões* para acreditar que, eu *sei* que etc.¹³ Se a letra *A* representar um dos acima mencionados operadores epistêmicos, a letra *S* uma terceira pessoa qualquer, a letra *P* uma sentença qualquer e a letra *T* um instante no tempo, podemos definir a espontaneidade epistêmica nos seguintes termos:

$A_{S,T}$ P (no tempo *t* o sujeito *S* *pensa* ou *julga* que *P* é o caso).

Com base nesta definição podemos agora buscar uma formulação para a espontaneidade prática. Como é sabido, Kant a define como a capacidade de agir segundo a representação de leis, ou seja, como a capacidade de agir segundo o que julgamos como sendo necessário ou bom. Podemos exprimir tal espontaneidade prática nos termos de mais um operador proposicional. Se a letra *C* simbolizar tal operador e, como antes, *S* representar o sujeito, *P* a sentença regida pelo operador e *Q* a convicção (normativa ou axiológica) com base na qual o sujeito age, teremos então:

$C_{S,T,Q}$ P (no tempo *t* o sujeito *S* é capaz de tornar a sentença *P* verdadeira com base na sua convicção *Q*).

Nestes termos, livres e responsáveis serão aqueles sujeitos capazes de tornar frases verdadeiras com as suas ações a partir daquilo que consideram bom ou necessário. Ora, como intérpretes do comportamento de terceiros, não observamos

(13) Distintamente dos operadores usuais de asserção (“é um fato que”, “é o caso que” etc.), estes operadores sentenciais se caracterizam por não penetrarem todas as conseqüências lógicas da sentença sobre a qual operam, mas apenas aquelas epistemicamente relevantes.

obviamente nem o agente se posicionando frente a conteúdos proposicionais nem tampouco tornando frases verdadeiras ou falsas. Tudo o que podemos observar são as suas ações concretas. Por esta razão, sob o enfoque cognitivo de um espectador externo, só poderemos saber se terceiros agem de forma espontânea se pudermos tornar inteligíveis as ações que observamos, atribuindo ao seu agente o desejo de tornar determinadas frases verdadeiras a partir de suas convicções normativas e axiológicas. Assim, a naturalização da liberdade do arbítrio como espontaneidade que proponho só poderá ser realizada nos termos de uma teoria da interpretação seguindo o chamado modelo da racionalização.

Este modelo de naturalização do agir intencional foi adotado por alguns intérpretes contemporâneos de Kant¹⁴. Minha abordagem se distingue, contudo, da desses autores em uma série de aspectos fundamentais. Em primeiro lugar, ao buscar uma naturalização para a espontaneidade do arbítrio à luz do modelo da racionalização, não pretendo fornecer, de modo algum, uma interpretação para a doutrina kantiana do estatuto inteligível da espontaneidade, mas um modelo alternativo. Abandonando por completo a doutrina do estatuto inteligível da espontaneidade, distintamente dos intérpretes de Kant, não tenho tampouco a intenção de reconsiderar o Idealismo Transcendental à luz do “monismo anômalo” de Davidson, como propõem Meerbote e outros¹⁵. Recusando todas as diferentes formas contemporâneas de materialismo e monismo, freqüentemente associadas ao modelo da racionalização, a naturalização que proponho não objetiva, por último, compatibilizar a espontaneidade com determinismo causal, mas pelo contrário, pretende assinalar acima de tudo o seu caráter essencialmente incompatibilista.

(14) Conferir Meerbote, Ralf. “Kant on Freedom and the Rational and Morally Good Will”. Em *Self and Nature in Kant's Philosophy*, ed. Allen W. Wood. Ithaca and London: Cornell University Press, 1984. P. 57-72.

(15) Na medida em que supõe que a causalidade entre eventos descritos mentalmente só pode ser implementada por leis causais que emergem da descrição física dos mesmos eventos, o monismo anômalo acaba por reduzir os estados intencionais de uma forma geral à condição de epifenômenos e a espontaneidade do arbítrio a uma quimera.

O problema consiste em saber, então, sob que condições podemos interpretar o comportamento observado de terceiros como ações que, a partir de convicções normativas e axiológicas, tornam determinadas frases verdadeiras. As primeiras condições são aquelas pressupostas na interpretação de condutas como ações intencionais. Em primeiro lugar, a interpretação de uma conduta como uma ação intencional é indissociável da atribuição de um conjunto coerente de estados sob a forma de *atitudes favoráveis* à ação que se busca explicar, ou seja, intenções, desejos, volições, preferências, ambições, etc. Aquilo que sob o enfoque imanente do agente é normalmente descrito como um desejo favorável à realização de uma determinada ação, sob o enfoque cognitivo daquele que interpreta a conduta observada, assume a forma da atribuição de uma “pró-atitude”, como a denomina Davidson.

Da mesma forma que, sob o enfoque performativo do agente, os desejos primários favoráveis à realização de ações são indissociáveis das convicções acerca das propriedades das ações em questão, sob o enfoque cognitivo do intérprete, a atribuição de pró-attitudes será igualmente indissociável da atribuição de convicções acerca das propriedades da ação. Pró-attitudes e convicções, quando tomados conjuntamente, constituem aquilo que Davidson denomina “razão primária”:

*C1. G é uma razão primária do porquê uma pessoa realizou a ação A sob a descrição b se G se compõe de uma pró-atitude do agente frente às ações com uma propriedade determinada e da convicção por parte do agente de que A possui tal propriedade sob a descrição b.*¹⁶

Com base nessa definição, podemos fornecer uma explicação naturalista para a liberdade da ação. Se, sob o enfoque performativo, uma ação é livre quando for produto de uma escolha baseada em máximas ou regras de preferências,

(16) Davidson, Donald. “Actions, Reasons and Causes”, em: *D. Davidson: Essays on Actions and Events*, p. 22, Oxford 1980. É importante salientar que Davidson é mais um dentre os muitos adeptos da análise condicional do conceito de liberdade.

sob o enfoque cognitivo, a mesma ação será livre e intencional quando pudermos interpretá-la à luz de razões primárias, atribuindo ao seu agente pró-attitudes e convicções. Inversamente, uma ação não será livre quando (sob o enfoque performativo) não resultar de uma escolha baseada em máximas e (sob o enfoque cognitivo) quando não for possível interpretá-la ou explicá-la intencionalmente segundo razões primárias.

Mas como bem salienta Tugendhat na sua crítica a Moore, a imputabilidade exige mais do que a simples liberdade da ação. É possível interpretarmos o comportamento de crianças pequenas, de adultos senis, de doentes mentais e até, quem sabe, de alguns mamíferos superiores à luz de razões primárias, mas nem por isso os responsabilizamos pelas suas ações. A imputabilidade exige também uma explicação sob o mesmo enfoque cognitivo exterior para a própria liberdade de escolha. A minha sugestão aqui consiste na introdução de uma segunda ordem de razões interpretativas que, por analogia às primeiras, denominarei “razões secundárias”. De modo similar à definição de razão primária, podemos definir as razões secundárias nos seguintes termos:

C2. F é uma razão secundária do porquê uma pessoa realizou uma determinada ação se F se compõe de (i) uma pró-atitude (desejos ou volições de segunda ordem) do agente frente ao desejo primário f (favorável à realização da referida ação) que sob uma descrição c possuiria uma determinada propriedade e da (ii) convicção (igualmente de segunda ordem) por parte do agente de que seu desejo f possui efetivamente tal propriedade sob a descrição c.

Assim, se sob o enfoque performativo a liberdade de escolha se manifesta no posicionamento do agente frente aos seus desejos primários, sob o enfoque cognitivo esta mesma liberdade será conhecida pelos seus intérpretes quando as ações puderem ser interpretadas à luz de razões secundárias, ou seja, quando puderem ser atribuídos aos agentes não apenas desejos primários e convicções acerca das propriedades da ação, mas também atitudes frente aos próprios desejos primários e convicções acerca das suas propriedades. Os exemplos mais paradigmáticos são as ações contrárias a vícios estabelecidos. Para bem explicarmos um

comportamento de um agente qualquer como a ação de renunciar ao tabagismo, por exemplo, não nos basta atribuir ao mesmo razões primárias, i. é, atitudes contrárias à prática descrita como hábito de fumar (juntamente com a convicção de que tal prática possui determinadas propriedades sob a descrição em questão). Como o agente na condição de viciado possui, com efeito, desejos primários favoráveis à ação de fumar, só tornaremos tal ação plenamente compreensível ao atribuir-lhe razões secundárias, ou seja, uma atitude (desejo ou volição de segunda ordem) contrária ao próprio desejo primário de fumar, juntamente com a convicção de que tal desejo possui determinadas propriedades sob a descrição em questão.¹⁷

Entretanto, responsabilizaremos agentes por suas ações mesmo em situações nas quais estas ações não poderão ser interpretadas a partir de “razões secundárias”. Em muitos casos responsabilizaremos agentes por suas ações e omissões, mesmo sabendo que eles não refletiram sobre seus desejos primários, tendo agido de forma impulsiva. Nesses casos, basta supormos que o agente não refletiu por negligência, mas tinha de algum modo a capacidade para fazê-lo. Mesmo nestes contextos, o conceito de “razão secundária” permanece decisivo. Pois um agente responsável será aquele cujo comportamento *em geral* pode ser interpretado com base em razões secundárias.

O conceito de “razões secundárias” (desejos e convicções de segunda ordem acerca dos desejos sensíveis de primeira ordem) no modelo da racionalização vem cumprir uma função análoga àquela desempenhada pelo conceito “querer refletido” no condicionalismo proposto por Tugendhat. Em ambos os casos, busca-se para a imputação uma explicação naturalista (sob o enfoque cognitivo da terceira pessoa) que permita identificar as ações intencionais daqueles seres que

(17) De modo similar Frankfurt fala também de preferências e volições de segunda ordem, sem mencionar, contudo, as noções de convicção e de razão. Conferir Frankfurt, Harry. “Freedom of the Will and the Concept of Person”. *Journal of Philosophy* 68 (1971): 5-20.

Contudo, contrariamente tanto ao incompatibilismo que venho defendendo, quanto ao condicionalismo de Moore e outros, Frankfurt é de opinião de que a imputação ou atribuição de responsabilidade nem mesmo supõe a existência de possibilidades de alternativas de ação.

possuem liberdade de escolha e autodeterminação. Contudo, é importante salientarmos as distinções cruciais.

Como vimos, a análise condicional define o “querer refletido” como uma propriedade disposicional a ser descrita nos termos de um condicional causal. Na posição de conseqüente deste condicional causal figuraria a descrição de uma determinada ação, enquanto na de antecedente, a descrição de desejos sensíveis de primeira ordem e de circunstâncias que melhor os satisfariam. O “querer refletido” é identificado através da realização de ações maximizadoras dos desejos de primeira ordem que seriam produzidas ou causadas por estes mesmos desejos sensíveis e pela presença das circunstâncias que o maximizam. Nestes termos, o “querer refletido” seria sempre compatível com um possível determinismo psicológico. Fosse esta análise correta, não haveria nenhum prejuízo para os conceitos de liberdade da vontade e de autodeterminação, se pudessemos mostrar que diante de um determinado “querer sensível” e das circunstâncias que melhor o satisfazem, a opção refletida dos agentes *sempre* recairá sobre estas formas de ação mais adequadas à satisfação ou maximização de tais desejos sensíveis.

Em contrapartida, no modelo da racionalização proposto, as razões secundárias são antes *postuladas* no esquema interpretativo como um fator causal adicional que teria origem de forma espontânea no próprio agente. A característica fundamental dos sistemas autodeterminados consiste justamente no fato de que sua conduta (*output*) excede de forma avassaladora o conjunto das diferentes ações causais exercidas sobre o sistema sob a forma de estímulos (*inputs*). Assim, ao buscar compreender a conduta daqueles agentes autodeterminados, o intérprete se vê forçado a atribuir-lhes uma ordem de razões que seria, negativamente falando, irreduzível aos fatores causais externos e, positivamente, produto da atividade espontânea do próprio agente. Nesses termos, sob o enfoque cognitivo da terceira pessoa, um agente só será responsável por suas ações quando seus intérpretes potenciais puderem inserir no esquema interpretativo de seu comportamento posicionamentos espontâneos frente aos fatores externos à sua vontade refletida que introduziriam algo de “novo” na série causal.

Assim, no quadro de uma teoria empírica da interpretação (que tem sempre de se adequar ao comportamento observável), a espontaneidade do arbítrio perde seu suposto estatuto *inteligível*. Ao invés de ser definida como o substrato noumênico incognoscível que causaria, de forma atemporal, ações observadas no plano fenomênico, ela passa a ser localizada em uma teoria empírica da interpretação como aquele fator causal adicional cuja postulação interpretativa se faz necessária para tornar o comportamento observado compreensível. Esta explicação naturalista da espontaneidade do arbítrio tem naturalmente como pressuposto uma concepção não determinista e não fisicalista de natureza. Para atribuímos razões secundárias aos agentes temos que supor que as relações causais existentes entre os estímulos externos (inputs), seu posicionamento e as ações observadas (outputs) não sejam definidas por leis estritas. E para estipularmos interpretativamente que o posicionamento frente aos fatores externos à sua vontade introduz um novo fator na série causal, temos que romper com o pressuposto fisicalista do fechamento causal do universo, i. é, com o postulado de que todas as ocorrências físicas possam ser explicadas exclusivamente em termos físicos. Mas esta é uma discussão que pretendo deixar para um outro trabalho.