

## ENTRE O ÚTIL E O CERTO: HUME, MOTIVAÇÃO E RAZÕES PARA SER JUSTO

### BETWEEN THE USEFUL AND THE RIGHT: HUME, MOTIVATION AND REASONS FOR BEING JUST

**PEDRO FIOR MOTA DE ANDRADE<sup>1</sup>**  
(UFAC/PPGD-UFRJ - BRASIL)

#### RESUMO

Neste artigo, pretendo revisitar uma das mais antigas e importantes questões da história da filosofia moral, a saber, que razões temos para ser justos. Eu a abordarei a partir de uma controvérsia interpretativa existente no âmbito da teoria moral de David Hume. Duas vertentes interpretativas se consolidaram neste debate. A primeira sustenta que, para Hume, é uma condição suficiente para sermos justos que sejamos motivados pelos motivos não morais apropriados. A segunda sustenta que, para Hume, é uma condição necessária para sermos justos que sejamos motivados pelos motivos morais apropriados. Pretendo argumentar em favor da primeira vertente interpretativa oferecendo uma leitura que possa responder às possíveis objeções que são frequentemente levantadas contra esta interpretação.

**Palavras-chave:** Hume; Justiça; Ética; Motivação; Autointeresse.

#### ABSTRACT

In this paper, I intend to revisit one of the oldest and most important questions in the history of moral philosophy, namely, what reasons we have to be just. I will address it from the perspective of an interpretive controversy within the framework of David Hume's moral theory. Two interpretative lines set this debate. The first holds that, for Hume, it is a sufficient condition for us to be just that we are motivated by the appropriate non-moral motives. The second holds that, for Hume, it is a necessary condition for us to be just that we are motivated by the appropriate moral motives. I intend to argue in favor of the first interpretative line by offering a reading that can answer the possible objections that are often raised against this interpretation.

**Keywords:** Hume; Justice; Ethics; Motivation; Self-interest.

#### Introdução

Uma das mais antigas e importantes questões da história da filosofia moral é que razões temos para ser justos. Ela é discutida no Livro II da *República*, onde Sócrates é confrontado por Glauco acerca da origem e natureza da justiça. Glauco apresenta a tese sofística de que a justiça não

é um bem desejado em si mesmo, mas somente em função de suas consequências vantajosas, tais como assegurar recompensas, evitar punições e preservar a boa reputação na sociedade (357a - 358b), ao passo que Sócrates sustenta a tese oposta de que a justiça é um bem desejado em função de suas consequências vantajosas, mas muito mais em si mesma (367c-d). O argumento que norteia esta controvérsia ao longo de toda a *República* é aquele que pretende definir conclusivamente qual é a melhor vida, a do indivíduo justo ou a do injusto. Sócrates se propõe a defender que a vida do justo é a melhor a partir de uma investigação da origem e da natureza da justiça que se inicia no contexto da organização política da cidade e posteriormente se redireciona para o exame da alma do indivíduo justo e do injusto. Uma importante consequência desta discussão socrática é que ela fixou paradigmaticamente a controvérsia entre Sócrates e Glauco em termos de duas posições filosoficamente opostas: i) é uma condição necessária que as razões que os indivíduos possuem para serem justos devam ser exclusivamente morais ou ii) é uma condição suficiente que as razões que os indivíduos possuem para serem justos possam ser exclusivamente autointeressadas. Filósofos tais como Platão, Aristóteles, Immanuel Kant, John Rawls, Thomas Scanlon e outros sustentam uma versão ou outra da primeira tese, ao passo que filósofos como Thomas Hobbes, David Gauthier, James Buchanan, Michael Moehler e outros sustentam uma versão ou outra da segunda tese.

Entre todos os filósofos que trataram desta questão talvez seja David Hume o mais controverso deles. Em uma *Investigação sobre os Princípios da Moral* (1751), Hume famosamente apresenta a tese de que a "utilidade pública é *única* origem da justiça e que as reflexões sobre as consequências benéficas dessa virtude são a *única* fundação de seu mérito" (HUME, 2003, 241). Talvez jamais tenha sido o caso de um filósofo apresentar tão claramente sua tese relativa à justiça enquanto uma prática estabelecida por estritas razões de utilidade e interesse e, ao mesmo tempo, os demais elementos conceituais de sua teoria moral pareçam fornecer tão ampla evidência textual para se concluir o contrário. De fato, a natureza dos estados psicológicos que nos motivam a ser justos é um dos pontos mais controversos na teoria moral de Hume, suscitando diversas leituras e interpretações, seja para esclarecer a sua exata posição, seja para oferecer críticas ou seja para reiterar a validade de seus argumentos.

Neste artigo, eu pretendo apresentar e contrastar as principais leituras referentes à natureza da motivação para sermos justos na teoria da justiça de Hume e defender uma solução interpretativa para esta questão. Na primeira seção, eu exponho a concepção humeana de virtude e a máxima geral que a orienta na definição de qualquer prática como

virtuosa, dentre elas a própria justiça. Também apresento as circunstâncias que tornam possível a emergência, consolidação e expansão da justiça enquanto uma prática estabelecida por uma convenção social na genealogia imaginada por Hume. Assim, é possível descrever a interpretação ortodoxa da motivação da justiça em Hume, a interpretação do “autointeresse esclarecido” ou “autointeresse redirecionado”, cuja tese sustenta que a razão para sermos justos é um motivo de natureza estritamente não moral.

Na segunda seção, eu contrasto a interpretação ortodoxa com as interpretações heterodoxas de Rachel Cohon, Stephen Darwall e Marcia Baron, cuja tese sustenta que um motivo de natureza não moral não seria suficientemente confiável para poder adequadamente motivar os indivíduos a serem justos em todos os casos relevantes. Portanto, a teoria da justiça de Hume só poderia ser consistente se admitisse que a razão para sermos justos deve, em última instância, incluir um motivo genuinamente moral.

Na terceira seção, eu apresento mais um desafio trazido pelas interpretações heterodoxas à interpretação ortodoxa na figura hipotética do chamado “espertalhão ardiloso”. Esta figura imaginada por Hume é caracterizada de modo bastante similar a “Giges” de Glauco, ao “tolo” de Hobbes e ao “maximizador irrestrito” de Gauthier e, portanto, levanta problemas similares a uma tal concepção puramente autointeressada de justiça. Afinal, um tal indivíduo como o espertalhão, motivado por razões puramente autointeressadas, pode se perguntar se de fato existem boas razões para ser justo sob todas as circunstâncias possíveis.

Finalmente, na última seção, eu ofereço uma solução interpretativa que se alinha à leitura ortodoxa baseada em: i) uma leitura sobre como a prática da justiça poderia se constituir, consolidar e preservar sem necessariamente envolver motivos morais dos indivíduos e ii) uma interpretação particular da forma como Hume pode encarar o desafio colocado pelo espertalhão ardiloso.

### **Virtude, justiça e autointeresse esclarecido**

Hume é um tipo de ético da virtude. Para ele, discernir valores morais é uma questão de definir o virtuosismo ou o vício de objetos particulares que podem ser avaliados moralmente. Isto significa que os objetos que constituem o centro da moralidade na filosofia de Hume são virtudes e vícios, a saber, certas “qualidades mentais” que nos induzem a comportamentos que “despertam” certos sentimentos de “prazer” ou “dor” em um dado “espectador” que os “contempla” a partir de ponto de vista “geral” (HUME, 2000, 510-12, 614; 2003, 372). Tendo em vista que a própria justiça é definida por Hume como um tipo de virtude, devemos

iniciar com a análise do que constitui o nosso sentido de virtude.

Hume defende que as ações em si não possuem qualquer qualidade intrinsecamente moral, mas servem meramente como um “signo” que nos informa sobre as qualidades mentais “duradouras” e “constantes” do “caráter”, as quais são verdadeiramente o objeto de nossa estima ou desprezo moral (HUME, 2000, 447, 614). Desta tese pode-se deduzir duas conclusões importantes, a saber, que todas as nossas ações derivam de certos estados mentais e que ações virtuosas somente são assim consideradas se são originadas a partir de estados mentais avaliados moralmente como virtuosos. Portanto, o ajuizamento moral tem como objeto não as ações em si, mas os motivos que causam as ações. Este raciocínio se aplica a qualquer forma de virtude.

Entretanto, Hume ressalva que o motivo de uma ação virtuosa nunca pode ser a consideração pela virtude da mesma ação. Um raciocínio desta natureza implica em circularidade. É necessária a existência de um motivo distinto da consideração da própria moralidade da conduta. É esta a primeira máxima que Hume conclui de sua explicação sobre a virtude em geral no Livro 3, Parte 2, Seção 1 do *Tratado da Natureza Humana* (1739-40):

Em resumo, podemos estabelecer como uma máxima indubitável que *nenhuma ação pode ser virtuosa ou moralmente boa, a menos que haja na natureza humana algum motivo que a produza, distinto do sentido de sua moralidade* (HUME, 2000, 519).

A circularidade que Hume aponta parece clara. Deve preexistir algum motivo natural que possa motivar uma ação virtuosa, de modo que o conteúdo deste motivo preexistente não seja a mera consideração pelo mérito moral desta ação. Em termos gerais, o que Hume argumenta é que todos os indivíduos são possuidores de um certo conjunto de estados mentais naturais desprovidos de qualquer qualidade moral, tais como motivos, desejos, paixões, emoções e sentimentos, os quais são exteriorizados através de ações e comportamentos de cada um. Ao contemplarem tais ações, que desempenham a função de signos dos motivos que as causaram, os indivíduos aprovam indiretamente o motivo natural subjacente a estas ações, o que implica que este motivo é um tipo de estado mental não moral que é posteriormente objeto de um ajuizamento moral apropriado. Esta máxima tem como propósito ressaltar a distinção existente entre o motivo que causa uma ação e a posterior aprovação moral deste motivo, de maneira que afirmar que o motivo que

origina uma ação seja um senso de aprovação da própria ação é simplesmente assumir o que se quer demonstrar, a saber, que a virtuosidade de uma ação consiste em sua aprovação moral. A compreensão sobre o que é virtude a partir da máxima que Hume desenvolve em seu raciocínio pode ser resumido pela frase de Rachel Cohon: "cada virtude consiste, ao menos em parte, em um motivo aprovável de característica não moral." (COHON, 1997, 93).

Para entender a razão de Hume apresentar um tal raciocínio para caracterizar o nosso sentido de virtude, é importante atentar-se a sua proposta filosófica como um todo. Hume está engajado com um tipo de investigação geral da natureza humana. Este é o tema norteador de toda a sua filosofia. Esta investigação é conduzida por Hume a partir de uma abordagem fortemente naturalista. Neste sentido, a proposta filosófica é descobrir as causas e efeitos gerais que explicam e preveem as tendências e regularidades comportamentais humanas com base nos métodos da filosofia experimental, ou seja, por meio de um tipo específico de investigação empírica. No domínio moral, Hume pretende especificar a estrutura psicológica natural que subjaz às disposições comportamentais humanas. Seu propósito é identificar que elementos motivacionais específicos estão associados a cada uma de nossas ações, além de rejeitar descrições que ou pretendam explicar nossos padrões comportamentais na ausência dos elementos motivacionais relevantes ou que flagrantemente se contraponham aos elementos motivacionais relevantes constituintes de nossa estrutura psicológica natural.<sup>2</sup>

Neste sentido, que razões então temos para sermos justos ou, no contexto definido pela máxima da virtude de Hume, qual seria então o motivo natural que nos levaria a praticar atos justos?<sup>3</sup> Hume elenca três alternativas que poderiam se constituir como o motivo natural de atos justos: i) o puro autointeresse individual, ii) a benevolência pública, isto é, a consideração pelo interesse público, ou iii) a benevolência privada, isto é, a generosidade individual. No entanto, Hume descarta estes três estados psicológicos como estados motivacionais apropriados de atos justos argumentando que cada um deles, por si só, está apenas contingentemente e não necessariamente ligado à motivação da justiça. Por esta razão, ele conclui em favor da inexistência de qualquer motivo natural não moral, distinto do próprio senso de aprovação moral da justiça, que possa se constituir como elemento motivacional adequado de um padrão comportamental de atos justos. A conclusão que forçosamente se impõe a partir destas considerações é que a disposição comportamental justa não é um traço comportamental originário da psicologia natural da espécie humana, resultando, necessariamente, que a prática da justiça só pode se

originar artificialmente, através de acordos e convenções entre os indivíduos. Portanto, para Hume, a justiça é uma virtude artificial. A questão então é como um padrão comportamental de atos justos pode ser criado na sociedade e que estados motivacionais estão em jogo aqui.

Na genealogia delineada por Hume, ele especula que o ser humano, em seu "estado selvagem e inculto", é definido por ter "inúmeras carências e necessidades" ao mesmo tempo em que é dotado de "escassos meios" para satisfazê-las. A única maneira de satisfazer suas necessidades básicas é unindo-se e estabelecendo uma sociedade que aumente nossa "força", diversifique e especialize nossa "capacidade" e promova nossa "segurança" (HUME, 2000, 525-6). No entanto, indivíduos precisam enfrentar os obstáculos postos por sua condição psicológica natural e as circunstâncias do ambiente. Em particular, eles devem procurar evitar conflitos pela aquisição de "objetos externos escassos" e superar os obstáculos relacionados à disposição moderadamente "egoísta" de todos os seres humanos (HUME, 2000, 534; 2003, 247). Hume acredita que a concorrência destes fatores, quando envolvendo indivíduos em escala suficiente, é capaz de envolvê-los em conflitos que podem ser "diretamente destrutivos para a sociedade", pois são impelidos por "avidez" a "obter bens e posses" escassos em quantidades cada vez maiores para "nós e nossos amigos mais íntimos" (HUME, 2000, 532). Tal é a natureza do maior impedimento à formação da sociedade.

Deste modo, indivíduos devem se esforçar por unir forças e coexistir em uma ordem social pacífica pautada por certas regras coletivamente aceitas para guiar a sua interação social recíproca. Eles concordam em estabelecer um conjunto específico de regras de justiça por acreditarem que, ao se comportarem de acordo com elas, todos podem esperar resultados socialmente vantajosos. As regras de justiça são assim estabelecidas com o propósito de estruturar direitos de propriedade, garantir a segurança na posse de bens, permitir a divisão social do trabalho, facilitar a troca de bens e serviços e promover o desenvolvimento material da sociedade.

Hume descreve a justiça como um conjunto de regras de comportamento estabelecidas por meio de uma série de acordos ou convenções acordadas por indivíduos por razões de interesse comum. Essas convenções são definidas como "um sentido geral do interesse comum, que todos os membros da sociedade expressam mutuamente, e que os leva a regular sua conduta segundo certas regras" (HUME, 2000, 530). Nesse esquema, o conjunto mais básico de regras de justiça compreende a regra da estabilidade da posse, a regra da transferência da posse por consentimento e a regra do cumprimento de promessas (HUME, 2000, 565).



A regra para estabilizar as posses de objetos externos entre os indivíduos é a primeira assim introduzida quando se observa que será de meu “interesse deixar que outra pessoa conserve a posse de seus bens, *contanto que* ela aja da mesma maneira em relação a mim”, o que a leva a ter “consciência de um interesse semelhante em regular sua conduta” (HUME, 2000, 530). É assim que essa convergência de interesses “produz uma resolução e um comportamento adequado” e esse tipo de “ato se torna um exemplo para os demais”, o que gradualmente consolida a disposição comportamental de respeito à regra de propriedade privada (HUME, 2000, 530, 538).

Ao longo de sua genealogia, Hume também detalha a expansão deste sistema básico de regras de justiça e a consolidação do padrão comportamental de atos justos entre os indivíduos em sociedades mais complexas ao longo do tempo. À medida que eles se conscientizam das vantagens que uma ordem social pautada por regras de justiça lhes proporciona, eles passam a desenvolver um senso de aprovação moral do comportamento justo e desaprovação moral do comportamento injusto com base em laços de simpatia, o qual se baseia, inicialmente, na reflexão sobre as vantagens proporcionadas por essa prática, mas que se estende além dela para se tornar um genuíno senso de apreciação moral pela própria prática justa em si. Esse aprimoramento natural da psicologia coletiva humana é o que Hume chama de “progresso dos sentimentos”. É o que garante que a sociedade cresça de sua primeira “formação” para “uma tribo ou uma nação” de forma pacífica e ordenada e as “fronteiras da justiça se ampliem ainda mais” junto com ela ao longo do tempo (HUME, 2000, 539-40; 2003, 252-53). Posteriormente, esse senso de aprovação moral do comportamento justo, que inicialmente se desenvolve de maneira natural e orgânica, é reforçado artificialmente por meio de alguns mecanismos sociais, tal como a atuação educativa dos “políticos”, “pais” e “instrutores” que procuram treinar as próximas gerações neste padrão comportamental justo (HUME, 2000, 540; 2003, 330).

Portanto, é quando os indivíduos observam um interesse comum em regular suas condutas por um conjunto de regras que introduzem e regulam as relações envolvendo a propriedade privada que a justiça é criada, e a disposição comportamental de respeitar essas regras se consolida enquanto uma virtude artificial ao longo do tempo. Para os propósitos metodológicos presentes, convém dividir esta genealogia em dois momentos, a da criação e a da expansão da justiça, pois, em cada um desses estágios, existe um estado psicológico distinto que parece contribuir na estrutura motivacional constituinte da virtude da justiça entre os indivíduos, como Hume resume no *Tratado*:

Assim, o *interesse próprio* é o motivo original para o *estabelecimento* da justiça, mas uma simpatia com o *interesse público* é a fonte da *aprovação moral* que acompanha essa virtude (HUME, 2000, 540).

À luz dessa genealogia, é necessário examinar as possíveis interpretações relativas à motivação da justiça em Hume. A interpretação ortodoxa da motivação da justiça é a do "autointeresse esclarecido" (enlightened self-interest) ou do "autointeresse redirecionado" (redirected self-interest), a qual é endossada pela maioria dos estudiosos de Hume incluindo Anette Baier (1991, 233; 2010, 41, 49), Don Garret (2015, 268), Jaqueline Taylor (1998, 9, 11, 17), Margareth Tate (2005, 101), David Gauthier (1992, 410), John L. Mackie (1980, 80), Barry Stroud (1977, 206), entre outros.

Em linhas gerais, a interpretação do autointeresse esclarecido se apresenta da seguinte forma. Ela sustenta como tese central que o motivo original e não moral que induz os indivíduos a criar regras de justiça por meio de uma convenção e a manter uma disposição comportamental de respeito a essas regras é justamente a paixão da autointeresse. Entretanto, esta paixão, em seu "movimento" natural "cego e impetuoso", não tende naturalmente à cooperação mútua, mas contribui para fortalecer a nossa própria parcialidade (HUME, 2000, 530). Portanto, é necessário a criação de um artifício, isto é, a criação, através de um exercício do "juízo" e do "entendimento", de um sistema de regras sociais que possa refrear, redirecionar e coordenar esta paixão mais básica e comum de todos os indivíduos visando resultados socialmente mais vantajosos. O estabelecimento das regras de justiça através de um acordo ou convenção tem como único propósito a satisfação de nossas necessidades mais básicas por meio da regulação da propriedade privada, de modo que é de nosso maior interesse respeitar essas regras. Não é o tipo de interesse que poderia ser alcançado pelas "paixões naturais e não artificiais dos homens", mas é o interesse "controlado" e cuja "direção" foi "alterada" por ele próprio ao ser guiado por uma reflexão mais ampla (HUME, 2000, 532, 537). Esse autointeresse esclarecido, informado ou redirecionado, enquanto associado ao respeito às regras de justiça, na medida em que há a expectativa que este sistema de regras possa produzir resultados socialmente vantajosos para todos, é a única fonte motivacional da justiça para Hume. Talvez a passagem que ofereça o suporte textual mais robusto a esta interpretação seja encontrada no *Tratado*:



Desse modo, as leis da justiça surgem de princípios naturais de um modo oblíquo e artificial. Sua verdadeira origem é o amor por si mesmo; e como o amor que uma pessoa tem por si mesma é naturalmente contrário ao das outras pessoas, essas diversas paixões interessadas são obrigadas a se ajustar umas às outras de maneira a concorrer para algum sistema de conduta e comportamento (HUME, 2000, 568).

Semelhantes passagens que reforçam esta interpretação são encontradas em outros momentos também na *Investigação*, o que sugere que Hume não parece ter alterado substancialmente sua posição filosófica ao longo de sua vida (HUME, 2003, 337). Entretanto, é importante lembrar que a interpretação ortodoxa não é unânime entre os estudiosos de Hume. Alguns intérpretes julgam que a atuação motivacional de um motivo de natureza não moral, tal como é o autointeresse, é insuficiente para que se qualifique como a adequada motivação da justiça em Hume. A disposição justa necessitaria de suplementação por meio de um outro componente motivacional, a saber, um motivo genuinamente moral. Esta linha interpretativa alternativa se baseia em três elementos textuais, como procurarei mostrar: uma interpretação particular de certos conceitos da teoria moral de Hume, a descrição genealógica da justiça e a figura hipotética do espertalhão ardiloso.

### **Aprovação moral, autorregulação e educação**

Em linhas gerais, a interpretação heterodoxa da motivação da justiça sustenta que atos justos devem ou ser exclusivamente motivados por ou incluir, como componente motivacional, um motivo de natureza moral, o qual seria então o apropriado motivo da justiça em Hume. Diferentes versões desta tese interpretativa são endossadas por estudiosos de Hume incluindo Rachel Cohon (1997, 95), Sophie Botros (2015, 268), Stephen Darwall (1995, 296), Marcia Baron (1981, 541), entre outros. Neste texto, pretendo me restringir à análise das versões de Cohon, Darwall e Baron para ilustrar a tese geral da interpretação heterodoxa.

Rachel Cohon critica a interpretação ortodoxa com base no que, em sua visão, se apresenta como em uma discrepância conceitual entre a definição de virtude e a máxima da virtude em relação à justiça em Hume. Em primeiro lugar, Cohon observa que Hume define a justiça em termos de uma virtude, a saber, um tipo de disposição de caráter que nos comunica sentimentos de prazer ao ser contemplada por um observador a partir de um ponto de vista geral. Em segundo lugar, a justiça, enquanto uma virtude, se insere no escopo da máxima que enuncia que toda virtude deve

ser motivada por um motivo natural distinto de um senso de aprovação moral. Em terceiro lugar, como não existe um motivo natural na psicologia humana que nos induza à prática de atos de justiça, Hume é levado a concluir que a virtude da justiça não é natural, mas criada artificialmente pela atuação redirecionada do autointeresse. Não obstante, na opinião de Cohon, não é claro como uma disposição de caráter motivada por autointeresse pode vir a ser aprovada moralmente como uma disposição comportamental de atos justos. Não parecemos retratar pessoas de disposição justa como pessoas que atuam meramente por motivos de autointeresse. Ser justo e ser autointeressado parecem significar coisas diferentes. Justiça e autointeresse parecem categorialmente distintos enquanto tipos de disposição de caráter. Essa é a raiz da controvérsia.

A solução que Cohon oferece para esclarecer esta discrepância conceitual consiste em relativizar a exigência da máxima da virtude de Hume concedendo que existam um conjunto de virtudes, as virtudes artificiais, que possam ocasionalmente ser motivadas exclusivamente por um senso de aprovação moral, descartando o motivo natural não moral anterior. Desta forma, Hume conseguiria sustentar consistentemente três teses em aparente conflito em sua teoria moral: i) a justiça é uma virtude, ii) a justiça é, tal como qualquer virtude, constituída por um motivo que deve ser aprovado moralmente e iii) inexistente um motivo na psicologia natural dos seres humanos que os induza a agir justamente (COHON, 1997, 95).

Ao sugerir esta mudança na afirmação de Hume, Cohon deseja mostrar que a justiça não pode ser considerada uma virtude se entendida no mesmo sentido que as virtudes naturais, as quais caem rigidamente sob o escopo da máxima da virtude de Hume. Entretanto, ela é uma autêntica virtude se entendida conforme sua definição geral, a saber, uma disposição de caráter que desperta sentimentos agradáveis de aprovação moral em um observador que a contempla a partir um ponto de vista geral. Esta caracterização da justiça enquanto uma virtude exige, na opinião de Cohon, que este próprio sentimento de aprovação moral seja o motivo apropriado do comportamento justo. Assim, em princípio, estaríamos na posição de poder aprovar moralmente a conduta justa através de um ajuizamento moral apropriado sob qualquer circunstância. Em resumo, a justiça poderia ser sempre reconhecida como uma espécie do gênero das virtudes, porém não se situaria no subgênero das virtudes naturais. Porém, para que isto seja possível, esta virtude não pode ser motivada pelo autointeresse, pois se trata de um motivo que não pode ser moralmente aprovado sob qualquer circunstância.

A interpretação que Cohon mantém, em oposição à interpretação ortodoxa, se baseia em sua crença de que o motivo que impulsiona a virtude da justiça é originado através de um longo processo descrito genealogicamente por Hume, do qual o autointeresse é somente um dos componentes motivacionais em jogo, mas não o exclusivo. É necessário que a disposição justiça seja suplementada por um estado motivacional adicional, isto é, um senso de aprovação moral de atos justos, para que o padrão comportamental de atos justos seja considerado uma virtude (COHON, 1997, 105). É ilustrativo notar como Cohon traça uma distinção entre a prudência e a honestidade aqui. A disposição de agir em mera conformidade às regras de justiça é definida como a prudência, cuja fonte motivacional é a paixão moralmente neutra do autointeresse. Em contrapartida, a disposição de agir por respeito e consideração pelas regras da justiça é definida como a honestidade, que é um traço comportamental associado à virtude da justiça, e cuja fonte motivacional deve ser algum senso de aprovação moral.<sup>4</sup> Na visão de Cohon, Hume possui o cuidado de não identificar a conduta honesta com a conduta prudente, o que os defensores da interpretação ortodoxa parecem confundir (COHON, 1997, 104).

Em resumo, a interpretação de Cohon reforça que o papel motivacional do autointeresse se restringe ao estabelecimento inicial e preservação temporária da justiça enquanto um sistema de regras sociais e uma disposição comportamental nascente de segui-las em vista de resultados socialmente vantajosos para seus participantes, tal como apresentado na primeira parte da genealogia de Hume. No entanto, o autointeresse não pode continuar a se manter como o motivo apropriado da justiça à medida que a sociedade se expande e se complexifica ao longo do tempo, tal como imaginado na segunda parte da genealogia de Hume. Isto porque uma disposição motivacional de natureza não moral não é capaz de fazer com que justiça seja concebida como uma genuína virtude nos termos em que Hume a define. O motivo em que o comportamento justo se baseia deve ser um senso de aprovação moral que viemos a desenvolver habitualmente ao contemplar outros atos de natureza semelhante:

O interesse redirecionado é o motivo que inicia a convenção e que a preserva e a estabiliza em uma pequena comunidade em nível de subsistência. Mas não é o motivo que faz da honestidade uma virtude. O motivo que faz da honestidade uma virtude é um motivo moral, a aprovação de atos que se conformem à regra e a desaprovação de violações, o que foi fortalecido e reforçado pelo hábito, de modo a se tornar um sentimento motivador (COHON, 1997, 107).

Além da leitura de Cohon, uma outra versão da interpretação heterodoxa é desenvolvida por Stephen Darwall. Essa leitura se baseia em uma interpretação particular do conceito de obrigação, o que, em Hume, é associado à sua teoria da motivação. No que concerne à justiça, Hume apresenta estes conceitos correlatos do seguinte modo no *Tratado*:

A obrigação *natural* da justiça, ou seja, o interesse, já foi explicada em sua totalidade; quanto à obrigação *moral*, ou sentimento do que é certo ou errado, será preciso primeiramente examinar as virtudes naturais, antes que possamos fornecer uma explicação completa e satisfatória (HUME, 2000, 539).

Darwall argumenta que Hume parece endossar a tese sustentada pela interpretação ortodoxa, a saber, que as regras de justiça e a disposição comportamental para respeitá-las são originadas e mantidas pelo motivo não moral do autointeresse, uma vez que agir de acordo com esse conjunto de regras sociais tende a produzir resultados socialmente vantajosos para seus participantes. No contexto desta interpretação, a obrigação natural da justiça, como Hume afirma na passagem acima, seria o autointeresse, ao passo que a obrigação moral da justiça seria o sentimento de aprovação moral posterior desta disposição comportamental. É importante ressaltar que estes dois tipos de obrigação não são relacionados à emergência do sistema de regras de justiça, ocorrendo onde quer que subsistam motivos morais ou não morais constituindo virtudes e vícios naturais, por exemplo. No entanto, Darwall baseia-se neste mesmo conceito de obrigação para sustentar a existência de uma terceira forma obrigacional nova, específica e totalmente distinta das outras duas. Esta nova obrigação deve emergir somente depois do estabelecimento da convenção de justiça. Darwall traz como suporte textual para esta leitura a afirmação de Hume de que é “uma vez firmada essa convenção sobre a abstinência dos bens alheios” que “as ideias de justiça e injustiça” e as de “propriedade, direito e obrigação” surgem na humanidade (HUME, 2000, 531). O argumento de Darwall é apresentado nestes termos:

Estamos agora em uma posição para apreciar o sentido que Hume deve ter em mente quando ele diz que o conceito de obrigação surge somente *após* as convenções que estabelecem as regras de justiça terem sido “feitas” (p. 531). Como eu mencionei no início, nem a obrigação natural (o motivo do autointeresse) nem a obrigação moral (o

sentimento de aprovação) requerem convenções anteriores para serem inteligíveis (DARWALL, 1995, 296)

Esta terceira forma obrigacional a que Darwall se refere é criada a partir da convenção de justiça e seu componente motivacional tem como conteúdo a atitude de se autorregular de acordo com quaisquer regras que forem estabelecidas por este acordo. Para ilustrar a natureza desta nova obrigação que comanda a aderência a regras, Darwall lembra a distinção realizada por H. L. Hart entre a perspectiva externa e interna das regras. O aspecto externo caracteriza a atitude de cumprimento das regras sem necessariamente envolver qualquer preocupação com o endosso pessoal em relação ao seu conteúdo. O aspecto interno revela a atitude de aceitação pessoal de tomar a regra como base para regulação da própria conduta. Neste sentido, o aspecto interno não se refere propriamente ao conteúdo mesmo das regras, mas à atitude de autorregulação por regras em geral (DARWALL, 1995, 297; HART, 2007, 98-99). É justamente esta última concepção que Darwall evoca para caracterizar a obrigação que nasce com a convenção de justiça e que se constitui como fonte motivacional do padrão comportamental de atos justos, a saber, a disposição para autorregular a própria conduta conforme as regras de justiça previamente acordadas.

Assim como Cohon, Darwall também sustenta que seja necessário um estado motivacional adicional de natureza moral que suplemente a motivação da justiça em Hume, a saber, uma atitude de se autorregular de acordo com regras de justiça previamente acordadas. O autointeresse é necessário para motivar os indivíduos a inicialmente estabelecer um conjunto de regras de justiça por meio de uma convenção, porém é insuficiente para manter o padrão comportamental de atos justos ao longo do tempo.

Em particular, Darwall objetivava, através de sua versão da interpretação heterodoxa, solucionar algumas das dificuldades centrais da teoria moral de Hume, tais como, por exemplo: i) evitar incorrer na circularidade na motivação da justiça tal como prescrita pela máxima da virtude de Hume, ii) esclarecer a relação entre as teorias da vontade e obrigação de Hume e iii) explicar o surgimento do sentimento de aprovação moral da justiça em Hume.<sup>5</sup> Entretanto, o mais importante objetivo é esclarecer consistentemente a motivação da justiça em Hume sem recorrer à interpretação ortodoxa.<sup>6</sup>

Finalmente, gostaria de brevemente apresentar a versão da interpretação heterodoxa de Marcia Baron. A leitura de Baron também se baseia em uma interpretação específica da genealogia de Hume, em particular, de sua segunda parte. Baron ressalta que o autointeresse, por mais esclarecido, informado ou redirecionado que possa ser, não pode

servir para outro propósito que não seja desempenhar um papel motivacional que leve os indivíduos a inicialmente estabelecerem as regras de justiça por meio de uma convenção. Entretanto, este motivo não é capaz de, por si só, manter o padrão comportamental de respeitar estas regras ao longo do tempo. Em outras palavras, o autointeresse é o motivo que nos leva a criar as regras de justiça, mas não pode se constituir como o motivo apropriado da virtude da justiça. Isto porque o autointeresse pode apenas nos motivar a agir de modo a maximizar nosso benefício individual, o que pode frequentemente nos levar a agir em detrimento das regras de justiça em busca do maior benefício possível. Tal como Cohon, que pareceu oferecer um argumento de natureza intensional contra a identificação de atos justos com atos benéficos, Baron, por sua vez, parece assumir que o conjunto de atos justos e o conjunto de atos benéficos não são correferenciais.

A hipótese interpretativa que Baron propõe visando solucionar esta controvérsia consiste em admitir a necessidade de adicionar um componente motivacional de natureza moral e não menos artificial que o autointeresse esclarecido à motivação da justiça. Este motivo deve ser um sentimento de puro dever com respeito às regras da justiça. Entretanto, diferentemente de Cohon, Baron sustenta que este componente motivacional deve ser associado a atos justos enquanto motivo apropriado da justiça pela atuação social de políticos e educadores morais. Aqui vemos uma distinção na ênfase interpretativa de Cohon e Baron com relação a segunda parte da genealogia de Hume. Ao passo que Cohon compreende o motivo da justiça como um senso de aprovação moral de atos justos que emerge, se desenvolve e se consolida natural e habitualmente como sua fonte motivacional apropriada, Baron mantém que o sentimento de puro dever com respeito às regras de justiça só pode vir a ser associado a atos justos enquanto sua motivação apropriada ao sermos constantemente ensinados pela atuação de um conjunto de atores políticos e sociais relevantes.

Este sentimento de puro dever, que comanda o respeito às regras de justiça e que nos é artificialmente inculcado por políticos e educadores morais, é, na visão de Baron, o componente motivacional da virtude da justiça em Hume. Sua versão da interpretação heterodoxa se resume no seguinte parágrafo:

Minha interpretação não é diferente da interpretação padrão da explicação de Hume sobre nosso interesse em ter regras de justiça. A divergência ocorre em um segundo grau: enquanto outros sustentam que o autointeresse esclarecido



deve nos motivar a adequar nossa conduta às regras de justiça, eu afirmo que, na visão de Hume, o autointeresse, não importa o quão esclarecido, não vai sempre nos dirigir a agir justamente. Neste caso, nem a benevolência ou qualquer outro motivo natural. O autointeresse esclarecido precisa ser complementado, assim como a benevolência, por um artifício de políticos e educadores morais. Nossos educadores incutem em nós um motivo artificial – o sentido de dever – que nos leva agir justamente (BARON, 1981, 541).

A leitura de Marcia Baron possui implicações muito originais na forma como se pode interpretar a genealogia da justiça, a motivação da justiça, e a aprovação moral da justiça na teoria moral de Hume. No entanto, é importante ressaltar que sua leitura permanece como uma versão da interpretação heterodoxa. Esta versão interpretativa sustenta a tese de que o motivo apropriado da justiça em Hume deve ser ou incluir um componente motivacional de natureza moral se este conceito de justiça pretende apresentar-se como inteligível nesta teoria moral. Desta forma, apresentamos três alternativas motivacionais para a virtude da justiça apresentadas por defensores da interpretação heterodoxa: i) o sentimento de aprovação moral de atos justos, ii) a autorregulação por regras e iii) o sentimento de dever com respeito às regras da justiça inculcado por meio de educação pública e privada.

### **Giges, tolos, maximizadores irrestritos e espertalhões ardilosos**

Até este momento, a controvérsia sobre da motivação da justiça se pautou pelo debate acerca da correta interpretação de alguns elementos conceituais que constituem o arcabouço teórico da filosofia moral de Hume, tais como sua definição geral da virtude, a máxima da virtude, a definição geral de obrigação e sua descrição genealógica da justiça. Entretanto, tal como notado por Cohon e Darwall, existe um outro elemento conceitual, não sujeito a grandes controvérsias interpretativas, que, por si só, se apresenta como um obstáculo ainda mais desafiador à interpretação ortodoxa. Se trata da figura hipotética do “espertalhão ardiloso” (*sensible knave*) postulada por Hume na Seção 9 da *Investigação* (COHON, 1997, 105; DARWALL, 1995, 301).

O espertalhão ardiloso pode ser caracterizado como uma espécie de cético moral que, como regra geral, está disposto a seguir as regras de justiça nos casos em que o beneficia, porém, nos casos em que elas se contrapõem aos seus interesses, constata que é mais vantajoso violá-las para conseguir plena satisfação de seus desejos. Assim Hume o descreve:

E embora se admita que nenhuma sociedade pode subsistir sem um respeito pela propriedade, ocorre que, em razão da maneira imperfeita pela qual são conduzidos os assuntos humanos, um espertalhão ardiloso (*sensible knave*) poderá julgar, em certas circunstâncias, que um ato de injustiça ou deslealdade trará um importante acréscimo à sua fortuna sem causar nenhuma ruptura considerável na confederação ou união social. “A honestidade é a melhor política” pode ser uma boa regra geral, mas é suscetível de muitas exceções, e pode-se pensar que o homem que se conduz com a máxima sabedoria é aquele que observa a regra geral e tira proveito de todas as exceções (HUME, 2003, 364).

É importante identificar as similaridades entre esta figura e outros cétricos morais hipotetizados ao longo da tradição, tais como “Giges” de Glauco no Livro II da *República* na antiguidade, o “tolo” de Hobbes no capítulo XV do *Leviatã* (1651) na modernidade, e o “maximizador irrestrito” de Gauthier no capítulo VI de *Morals by Agreement* (1986) na contemporaneidade. Em primeiro lugar, todos são caracterizados como pautando seus comportamentos não por razões morais, mas por razões de puro autointeresse. Tal é o caso de Giges que procurava “assenhorar-se do poder” (360b), do tolo cuja “razão” dita perseguir o “seu próprio bem” (HOBBS, 1983, 87) e do maximizador irrestrito que procura “maximizar sua utilidade dadas as estratégias daqueles com quem ele interage” (GAUTHIER, 1986, 167). Em segundo lugar, todos são cétricos morais no sentido de se perguntarem se há razões convincentes para serem justos a partir do paradigma da motivação puramente autointeressada para, ao fim, concluírem que é irracional serem justos. Tal é o caso daquele que seja presenteado com o anel de Giges e ainda assim refreie-se de cometer injustiças, julgando-se desta forma “desgraçado e insensato” (360d), do tolo que conclui que não seria “contra a razão” agir por “violência injusta” nos casos em que “é impossível que daí resulte qualquer dano” (HOBBS, 1983, 87) e do maximizador irrestrito que “escolhe não cooperar” com outros em cenários semelhantes ao do “dilema dos prisioneiros” (GAUTHIER, 1986, 170). Em terceiro lugar, todos são derrotáveis. Em outras palavras, eles são hipotetizados como instrumento metodológico para se defender que, de fato, existem razões justificáveis para ser justo. E é neste momento que é igualmente importante enfatizar no que estas figuras se diferem.

Fundamentalmente, é importante notar como as estratégias para se responder a cada cétrico moral são bem distintas. Tanto o tolo de Hobbes

como o maximizador irrestrito de Gauthier são caracterizados como, em princípio, suscetíveis à punição por parte seus semelhantes caso sejam flagrados agindo injustamente. Hobbes derrota o tolo mostrando que é irracional ser injusto pois o tolo, ao agir assim, corre o risco de ser “deixado fora” ou “expulso de uma sociedade”, de modo que estará agindo “contra a razão de sua preservação” ao ser injusto (HOBBS, 1983, 87-8). Gauthier derrota o maximizador irrestrito mostrando que é irracional ser injusto pois eles são retratados como “translúcidos”, ou seja, nem “transparentes” ou “opacos” com relação à certeza pela qual outros podem identificar sua disposição para cooperação, de modo que eles investirão em desenvolver uma disposição para identificar e punir maximizadores irrestritos injustos (GAUTHIER, 1986, 174).

No entanto, a hipótese de Giges de Glauco é bem mais radical. Este tipo de cético moral é caracterizado como, em princípio, insuscetível à punição por parte de seus semelhantes, pois estaria agindo em total segredo por meio de seu anel. Em princípio, ele nunca poderia ser flagrado agindo injustamente. E Hume admite esta mesma possibilidade no caso do espertalhão ardiloso (HUME, 2003, 365). Tendo em vista que Giges só é derrotado não pelo paradigma do autointeresse, mas por meio de um argumento socrático que assume um paradigma eudaimonista para concluir que justo e o injusto vivem vidas qualitativamente bem distintas, é neste ponto fundamental que se revela a fraqueza da conexão entre virtude da justiça e a motivação autointeressada sustentada pela interpretação ortodoxa de Hume. À luz da hipótese de um espertalhão ardiloso capaz de agir em segredo, apenas um componente motivacional de natureza moral seria capaz de adequadamente motivá-lo a ser justo e, assim, torná-lo apto a preservar um padrão comportamental de atos justos longo do tempo. De acordo com defensores da interpretação heterodoxa, é impossível que a virtude da justiça tenha como único motivo o autointeresse, pois assim esta virtude seria facilmente enfraquecida pela existência de espertalhões ardilosos já imaginados por Hume.

### **Forma geral da estrutura cooperativa e o agir em segredo**

Em face do que foi exposto, parece que defensor da interpretação ortodoxa se encontra sem saída. Porém, pretendo agora oferecer minha contribuição interpretativa em sua defesa. Em razão da natureza das objeções, acredito que meu argumento deva ter duas partes: i) ele deve mostrar que indivíduos motivados por autointeresse possam ser capazes de mais do que apenas criar, consolidar e preservar temporariamente as regras de justiça em pequenas comunidades, o que exige pacificar a interpretação

da posição da virtude da justiça no contexto da genealogia de Hume e ii) ele deve responder ao seu maior desafio na figura hipotética do espertalhão ardiloso capaz de agir em segredo a partir do paradigma do autointeresse. Na primeira parte, o argumento deve ser de natureza contrafactual no contexto da genealogia imaginada por Hume. É preciso estabelecer se é, em princípio, possível que o padrão comportamental de atos justos pudesse se constituir, consolidar e preservar indefinidamente sem necessariamente envolver a existência de motivos morais adicionais que os indivíduos possam vir a ter, tal como especificado na segunda parte da genealogia de Hume.

Isto pode ser demonstrado a partir da análise das propriedades que constituem a forma geral da estrutura cooperativa em Hume. Assume-se que os indivíduos são moderadamente autointeressados. Assume-se que participar da interação social definida pelas regras de justiça produza, em certa medida, resultados socialmente vantajosos para todos os participantes. E assume-se que, após as regras de justiça terem sido acordadas por estritas razões de autointeresse como os termos básicos da interação social entre os participantes, a opção do indivíduo de seguir as regras de justiça possa ser interpretada como uma escolha cooperativa e a opção de violar as regras de justiça, como uma escolha competitiva. Se os membros da convenção de justiça são caracterizados como interagindo apenas uma vez, a descrição dessa estrutura de interação social assemelha-se muito à do Dilema dos Prisioneiros. Neste cenário, a desconfiança entre indivíduos autointeressados se reforça. Eles são incentivados a escolher competir violando as regras de justiça, pois isso garante que nenhum deles acabe em uma pior situação em uma estrutura de interação social em que todos têm fortes incentivos para não cooperar ao seguir as regras de justiça. Aqueles que cooperarem pode acabar arcando com a maior perda possível frente a possíveis competidores que buscam maximizar seus benefícios. É trivial concluir que, sob estas circunstâncias, a consolidação de um padrão comportamental de atos justos é absolutamente impossível.

Entretanto, Hume admite que a convenção de justiça se consolida ao longo do tempo, o que pressupõe uma dinâmica de reiteradas interações sociais. Se os membros da convenção de justiça são caracterizados como interagindo infinitas vezes, a descrição dessa estrutura de interação social assemelha-se muito à do Dilema dos Prisioneiros infinitamente repetido. Argumento que este tipo de estrutura de interação social é a que especifica a forma geral da estrutura cooperativa em Hume. Neste cenário, o comportamento dos indivíduos é muito diferente. Indivíduos autointeressados são incentivados a considerar a escolha que os beneficia no longo e não no curto prazo, pois serão capazes de aprender como se

comportar com base em expectativas de que interações norteadas por regras de justiça continuarão a ocorrer no futuro. Eles serão capazes de prever as consequências negativas caso não cooperem seguindo as regras de justiça.

Disto resulta que eles devem adotar, como escolha racional, um tipo de comportamento cooperativo condicional que prescreve que cada indivíduo autointeressado deva seguir as regras de justiça, desde que os outros também se comportem dessa maneira. Se, no entanto, os outros optarem por não seguir as regras de justiça, então cada um também não deverá cooperar.<sup>7</sup> Como resultado, os cooperadores procurarão outros cooperadores com quem possam se envolver em interações sociais norteadas por regras de justiça, enquanto os competidores serão punidos com o ostracismo de interações cooperativas com base em sua má reputação construída ao longo de interações reiteradas. Para competidores, a escolha de não pautar suas ações em um comportamento cooperativo condicional pode acabar se mostrando tão arriscada e custosa à luz das possíveis punições que pode se caracterizar como irracional agir de forma diferente. Desta forma, os competidores terão fortes incentivos para se tornarem cooperadores a fim de usufruir das vantagens sociais proporcionadas por um esquema de justiça socialmente vantajoso para todos. Portanto, é totalmente possível que um padrão comportamental de atos justos entre indivíduos autointeressados possa se tornar autoimposto (*self-enforcing*) na sociedade.<sup>8</sup>

Ao contrário do que sustentam os defensores da interpretação heterodoxa, uma correta interpretação da primeira parte da genealogia de Hume revela que ele não precisa se comprometer com componentes motivacionais de natureza moral para que um padrão comportamental de atos justos possa se constituir, consolidar e preservar indefinidamente ao longo do tempo. Tudo é o que é necessário é que indivíduos autointeressados se atenham a uma adequada disposição de comportamento cooperativo condicional que os permita manter um sistema social de punições suficientemente eficaz de possíveis competidores em um cenário interativo que se assemelhe à forma geral da estrutura cooperativa em Hume. Assim, a preservação do padrão comportamental de atos justos depende somente da eficácia geral do sistema social de punições e não necessariamente do tamanho da comunidade ou da existência de motivos morais, como os defensores da interpretação heterodoxa sustentam.<sup>9</sup> Isto não significa que o surgimento de componentes motivacionais de natureza moral atrelados à justiça, tal como retratado na segunda parte da genealogia, não seja importante na teoria moral de Hume. Tais motivos desempenham um papel importante na transformação moral do ser

humano e na forma como ele passa a encarar a justiça como uma virtude ao longo do tempo. Entretanto, é um erro supor que a própria disposição justa dependa de quaisquer um deles.

No que concerne à segunda parte do argumento, eu sugeri que deveria existir uma interpretação convincente da resposta de Hume ao espertalhão ardiloso capaz de agir em segredo a partir do paradigma do autointeresse. Para que isso seja possível, é necessário compreender como Hume apresenta sua solução a este cético moral em sua obra:

Mas ainda que fossem completamente secretos e bem-sucedidos, o homem honesto, se tem algumas noções de filosofia, ou mesmo de observação e reflexão ordinárias, descobrirá que esses espertalhões são, no fim das contas, os maiores simplórios, pois trocam a inestimável satisfação associada a um caráter, pelo menos perante si mesmos, pela aquisição de bagatelas e quinquilharias sem valor. Quão pouco se requer para suprir as *necessidades* da natureza! E, relativamente ao *prazer*, que comparação poderia haver entre a satisfação gratuita que obtemos da conversação, do convívio social, do estudo, e até mesmo da saúde e das belezas comuns da natureza, mas, acima de tudo, da tranquila reflexão sobre a própria conduta. [...] Esses prazeres naturais, na verdade, não têm preço, pois qualquer preço é superior ao que custa para obtê-los e, ao mesmo tempo, inferior ao prazer que proporcionam (HUME, 2003, 365).

Com base nesta resposta, pode parecer que voltamos ao ponto inicial deste texto em que resumimos a resposta socrática a Glauco sobre a origem e a natureza da justiça e seus efeitos na alma do homem justo, pois é convidativo interpretar esta passagem de Hume a partir de um paradigma eudaimonista. Argumento que este não precisa ser o caso aqui. É possível interpretar esta passagem a partir do paradigma do autointeresse.

Mais uma vez, assume-se que os indivíduos são moderadamente autointeressados, o que, neste contexto, implica que pretendem satisfazer tantos desejos e desfrutar de tantos prazeres quanto possível. O que Hume parece argumentar aqui é que o espertalhão ardiloso, mesmo quando é capaz de realizar atos “completamente secretos e bem-sucedidos”, deve fazer escolhas entre os tipos de prazeres que poderá desfrutar ao longo de sua vida. Escolher desfrutar alguns prazeres implica em renunciar a outros. Desfrutar o ócio pode implicar em renunciar ao prazer dos “estudos”, desfrutar uma dieta específica pode implicar em renunciar ao prazer de uma “vida saudável”, desfrutar dos benefícios da injustiça pode implicar em



renunciar ao prazer de uma “satisfação associada a um caráter”, e desfrutar dos benefícios da injustiça ao extremo, em detrimento da própria subsistência da sociedade, pode implicar em renunciar até aos prazeres da “conversação” e do “convívio social”. Em suma, o simples desfrutar de certos prazeres torna outros relativamente inacessíveis. E o espertalhão, na medida em que pretende desfrutar tantos prazeres quanto possível, terá que considerar esta “troca” de prazeres ao longo de sua vida. Sua escolha de vida só será racional a partir do paradigma do autointeresse se conseguir manter o tipo de existência mais coerente possível em termos do máximo de prazeres que puder desfrutar.<sup>10</sup>

Além disso, podemos conceber não apenas que um conjunto relevante de prazeres esteja indisponível ao espertalhão por meio da constante troca resultante de seu estilo de vida, mas também que ele possa incorrer em certo sofrimento decorrente de seus atos. Por exemplo, assumamos que o espertalhão, tal como Gíges, nasça e cresça em um contexto plenamente social e adquira sua capacidade de agir em segredo apenas tardiamente em sua vida. Pode ser o caso que, como resultado de seu comportamento injusto, ele, ainda que colha benefícios, possa simultaneamente internalizar sentimentos de culpa, arrependimento e remorso habitualmente associados a tais violações durante sua vida, que operam como verdadeiras punições ou sanções internas na consciência do espertalhão.<sup>11</sup> Nestas circunstâncias, o espertalhão capaz de agir em segredo pode, em princípio, efetivamente incorrer em sofrimento ao praticar atos injustos, o que revela, neste caso específico, o descompasso entre o comportamento injusto e sua pretensão de desfrutar tantos prazeres quanto possível.

Pode-se com razão objetar se esse argumento é realmente convincente. Eu argumento que sua força não depende de um ponto exclusivamente filosófico, mas também das contribuições da psicologia empírica. Afinal, estamos agora nos referindo à incidência psicológica de diferentes tipos de prazeres e dores na psiquê de cada indivíduo. Entretanto, à parte desta discussão, a conclusão filosófica é mais radical aqui. Em ambos os casos apresentados acima, o que o argumento de Hume parece provar é que, se for possível conceber um homem justo capaz de manter uma existência mais coerente que a do espertalhão ardiloso capaz de agir em segredo em termos do máximo de prazeres que puder desfrutar, conclui-se que o espertalhão terá sido derrotado a partir do próprio paradigma do autointeresse. Tendo em vista que Hume não considera esta hipótese como conceitualmente impossível, o espertalhão ardiloso terá razões viver uma vida justa tomando como inspiração este homem justo se desejar desfrutar o máximo de prazeres que puder.

Concluo afirmando que a teoria da moral de Hume, quando

adequadamente interpretada, possui recursos conceituais suficientes para sustentar que o motivo apropriado da justiça é o autointeresse esclarecido, tal como reza a interpretação ortodoxa. Não é necessário recorrer nenhum outro componente motivacional, seja um sentimento de aprovação moral de atos justos, seja uma disposição de autorregulação por regras, ou seja um sentimento de dever com respeito às regras da justiça inculcado por atores políticos relevantes.

## **Conclusão**

Neste artigo, eu pretendi revisitar uma das mais antigas e importantes questões da história da filosofia moral, a saber, que razões temos para ser justos. Procurei abordá-la a partir de uma controvérsia interpretativa no contexto da teoria moral de David Hume. Acredito ser Hume um autor muito frutífero neste aspecto. Ele não apenas se propõe a responder os difíceis desafios colocados por algumas variantes do ceticismo moral, como o próprio arcabouço conceitual que compõe sua teoria moral parece ser muito amplo, o que abre margem interpretativa para compreendê-lo a partir das duas concepções diametralmente opostas neste debate: as razões para termos para ser justos são de natureza moral ou de natureza não moral. Aqui procurei oferecer uma interpretação da resposta de Hume que sustenta a tese de que é possível sermos justos motivados exclusivamente por razões autointeressadas, o que pode servir de inspiração para futuras tentativas de responder a esta questão ética.

## **Notas**

<sup>1</sup> Professor substituto em filosofia da Universidade Federal do Acre (UFAC) e pós-doutorando no Programa de Pós-Graduação em Direito da Universidade Federal do Rio de Janeiro (PPGD-UFRJ).

<sup>2</sup> Neste contexto, uma forma útil de compreender a lógica da máxima da virtude de Hume é interpretá-la como nada mais que uma instância muito geral do princípio "ought implies can" em relação à dependência entre a estrutura psicológica humana como um todo e nossos padrões comportamentais relevantes para a moralidade. Os racionalistas ingleses do século XVII, tais como Ralph Cudworth, Samuel Clarke e William Wollaston, cuja proposta filosófica era fundamentar a moralidade na razão a priori, foram criticados por Hume exatamente por violarem este princípio.

<sup>3</sup> É importante ressaltar que a filosofia de Hume não parece oferecer recursos conceituais para distinguir claramente entre duas posições metaéticas contemporâneas: o internalismo motivacional e o internalismo de razões. Ao passo que a primeira é uma tese motivacional que sustenta que subsiste uma relação interna ou necessária entre juízos morais e a motivação, a segunda é uma tese normativa que sustenta que a natureza das razões que possuímos é necessariamente interna ao conjunto de desejos e interesses que o agente possui. Na filosofia de Hume, ambas parecem se sobrepôr de modo que o internalismo motivacional parece implicar o internalismo de razões, como sua discussão sobre as obrigações naturais ou morais evidencia (2000, 539). Para uma versão de cada uma das duas posições, ver Michael Smith (1994, 92-3) e Bernard Williams (1982, 101-2) respectivamente.

<sup>4</sup> A distinção de Cohon entre a prudência e a honestidade ecoa a distinção de Kant entre “agir conforme o dever” e “agir por respeito ao dever”. Enquanto a primeira forma de motivação é destituída de valor moral, a última forma de motivação é a única que possui valor moral genuíno. Para mais sobre a distinção, ver Kant (2009, 115-19). Admito que não é surpreendente que Cohon trabalhe com categorias semelhantes às de Kant na interpretação de Hume, tendo em vista seu estudo de longa data sobre esses autores.

<sup>5</sup> Cohon, por exemplo, parece violar a primeira exigência estabelecida pela máxima da virtude de Hume.

<sup>6</sup> Esta interpretação proposta por Darwall é alvo de críticas tanto de Margaret Tate quanto de Rachel Cohon. A hipótese interpretativa que postula uma terceira forma obrigacional motivada pela mera atitude de se autorregular por regras de justiça previamente acordadas não é sustentada pelo arcabouço conceitual e teórico que caracteriza a filosofia de Hume. Hume nunca poderia concordar que somos motivados a seguir as regras de justiça por serem dotadas de uma qualidade intrinsecamente normativa, isto é, que impõem obediência independentemente da finalidade a que elas sejam projetadas a favorecer. Esta concepção seria contrária à teoria da vontade de Hume. Ver Cohon (1997, 112) e Tate (2005, 105).

<sup>7</sup> O comportamento de cooperar condicionalmente é retratado por Hume como a disposição adotada por indivíduos autointeressados ao acordar as regras de justiça por meio da convenção. Por exemplo, Hume afirma que devemos respeitar a regra de propriedade privada “contanto que ela (outra pessoa) aja da mesma maneira em relação a mim” e “com a condição de que os outros façam o mesmo” (HUME, 2000, 530, 538).

<sup>8</sup> Para uma interpretação paradigmática da teoria da justiça de Hume em termos de uma análise de teoria dos jogos na qual eu me baseei, ver Vanderschraaf (1998, 224-6).

<sup>9</sup> Defensores da interpretação heterodoxa devem aceitar esta caracterização da forma geral da estrutura cooperativa em Hume, já que as premissas que fundamentam esta caracterização são assumidas por ele já na primeira parte de

sua genealogia. Em particular, é assumido que: i) os indivíduos sejam moderadamente autointeressados, ii) eles concordem em estabelecer regras de justiça por meio de uma convenção motivados por estrito autointeresse, iii) eles se disponham a cooperar condicionalmente, iv) ao participar do esquema estabelecido pelas regras de justiça, eles possam, em certa medida, esperar resultados socialmente vantajosos para todos, e v) o padrão comportamental de atos justos se consolide gradualmente ao longo do tempo entre eles.

<sup>10</sup> É importante ressaltar que, para que este argumento seja bem-sucedido, não é necessário que Hume caracterize o espertalhão como se comprometendo com qualquer distinção qualitativa de prazeres. Exige-se apenas que ele se preocupe em manter uma existência em que o máximo de prazeres possa ser desfrutado coerentemente ao longo do tempo. Para uma distinção entre prazeres superiores e inferiores, ver Mill (2005, 49-51).

<sup>11</sup> A estratégia de responder a cétricos morais capazes de agir em segredo por meio de sanções internas ou pela internalização de sentimentos pode ser observada em Mill (2005, 68-9) e Stemmer (2000, 171-75).

## Referências bibliográficas

BAIER, A. *A Progress of Sentiments. Reflections on Hume's Treatise*. Cambridge: Harvard University Press, 1991.

BAIER, A. *The Cautious Jealous Virtue. Hume on Justice*. Harvard: Harvard University Press, 2010.

BARON, M. "Hume's Noble Lie: An Account of His Artificial Virtues". *Canadian Journal of Philosophy*, v. 12, n. 3, p. 539-555, 1981.

BOTROS, S. *Hume, Reason and Morality. A Legacy of Contradiction*. London and New York: Routledge, 2006.

COHON, R. "Hume's Difficulty with the Virtue of Justice". *Hume Studies*, v. 23, n. 1, p. 91-112, 1997.

DARWALL, S. *The British moralists and the internal 'ought': 1640-1740*. Cambridge: Cambridge University Press, 1995.

GARRETT, D. *Hume*. London and New York: Routledge, 2015.

GAUTHIER, D. "Artificial Virtues and the Sensible Knave". *Hume Studies*, v.

18, n. 2, p. 401-427, 1992.

GAUTHIER, D. *Morals By Agreements*. Oxford: Oxford University Press, 1986.

HART, H. L. A. *O Conceito de Direito*. Tradução: A. Ribeiro Mendes. Lisboa: Fundação Calouste Gulbenkian, 2007.

HOBBS, T. *Leviatã ou Matéria, Palavra e Poder de um Governo Eclesiástico e Civil*. Tradução: João Paulo Monteiro e Maria Beatriz Nizza da Silva. São Paulo: Editora Nova Cultural, 1983.

HUME, D. *Tratado da Natureza Humana*. Tradução: Déborah Danowski. 2 ed. São Paulo: Editora Unesp, 2000.

HUME, D. *Uma Investigação sobre o Entendimento Humano*. Tradução: José Oscar de Almeida. São Paulo: Editora Unesp, 2003.

KANT, I. *Fundamentação da Metafísica dos Costumes*. Tradução: Guido Antônio de Almeida. São Paulo: Barcarolla, 2009.

MACKIE, J. L. *Hume's Moral Theory*. London and New York: Routledge, 1980.

MILL, J. S. *Utilitarismo*. Tradução: Pedro Galvão. Porto: Porto Editora, 2005.

SMITH, M. *The Moral Problem*. Oxford: Basil Blackwell, 1994.

STEMMER, P. *Handeln zugunsten anderer: Eine moralphilosophische Untersuchung*. Berlin: De Gruyter, 2000.

STROUD, B. *Hume*. London and New York: Routledge, 1977.

PLATÃO. *A República*. Tradução: Maria Helena da Rocha Pereira. 9 ed. Lisboa: Fundação Calouste Gulbenkian, 2001.

TATE, M. W. "Obligation, Justice and the Will in Hume's Moral Philosophy". *Hume Studies*, v. 31, n. 1, p. 93-122, 2005.

TAYLOR, J. "Justice and the Foundations of Social Morality in Hume's *Treatise*". *Hume Studies*, v. 25, n. 1, p. 5-30, 1998.

VANDERSCHRAAF, P. "The Informal Game Theory in Hume's Account of Convention". *Economics and Philosophy*, v. 14. p. 215-247, 1998.

WILLIAMS, B. "Internal and External Reasons". In: *Moral Luck*. Cambridge: Cambridge University Press, 1981. cap. 8, p. 101-13.

*Received/Recebido: 27/09/2022*  
*Approved/Aprovado: 27/08/2023*