

Cooperação e conflito nos processos deliberativos em John Rawls e Jonathan Haidt

8

*Cooperation and conflict in the deliberative processes
in John Rawls and Jonathan Haidt*

*Cooperación y conflicto en los procesos deliberativos
en John Rawls y Jonathan Haidt*

DOI: 10.18226/21784612.v22.n2.08

Marcos Fanton*

Walter Valdevino Oliveira Silva**

Resumo: Nosso objetivo, neste texto, tomando como autor representativo o filósofo norte-americano John Rawls (1921-2002), é mostrar que os fundamentos da teoria política contemporânea, em sua vertente contratualista e de inspiração kantiana, estão situados historicamente e foram influenciados de modo determinante pelas pesquisas sobre psicologia e desenvolvimento moral, realizadas nas décadas de 1970 a 1990, nos EUA, particularmente por Lawrence Kohlberg (1927-1987). Atualmente, o embasamento científico dessas pesquisas “caiu por terra”, derrubando os pressupostos adotados por Rawls e por outros teóricos contratualistas. Iniciaremos detalhando como Rawls pressupõe um tipo de psicologia moral como elemento fundamental para garantir a estabilidade de sua teoria, para, em um segundo momento, referir alguns problemas desse modelo.

Palavras-chave: John Rawls. Lawrence Kohlberg. Justiça como equidade. Psicologia moral.

Abstract: Our objective, in this text, following the works of the North-American philosopher John Rawls (1921-2002), is to show that the grounds

* Doutor em Filosofia. Professor no Departamento de Filosofia da Universidade Federal de Pernambuco (UFPE). *E-mail:* fanton.marcos@gmail.com

** Doutor em Ética e Filosofia pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Professor-Adjunto no Departamento de Filosofia da Universidade Federal Rural do Rio de Janeiro (UFRRJ). *E-mail:* waltervaldevino@gmail.com

of contemporary political philosophy, in its contractualist, Kantian-inspired, trend are heavily influenced by the research on moral psychology and moral development of the period, which incidentally shows how Rawls' own moral psychology is historically situated by this research, particularly by the works of Lawrence Kohlberg (1927-1987). The scientific grounds of such research, however, have not aged well, forcing us to revisit and review the presuppositions adopted by Rawls and other contractualist theories of this period. In this article we will detail Rawls' own approach to moral psychology as a paramount element for the soundness of his doctrine, proceeding with an analysis of the problems within this approach and the consequences of the unviability of his moral psychology for the political model proposed within Rawls' works.

Keywords: John Rawls. Lawrence Kohlberg. Justice as fairness. Moral psychology.

Resumen: Nuestro objetivo, en este texto, tomando como autor representativo el filósofo norteamericano John Rawls (1921-2002), es mostrar que los fundamentos de la teoría política contemporánea en su vertiente contractual y de inspiración kantiana, están situados históricamente, y fueron influenciados de modo determinante por las investigaciones sobre psicología y desenvolvimiento moral, realizadas en las décadas de 1970 hasta 1990, en los Estados Unidos, particularmente por Lawrence Kohlberg (1927-1987). Actualmente, el fundamento científico de esas investigaciones “cayó por tierra”, derrumbando los supuestos conocimientos adoptados por Rawls y por otros teóricos contractuales. Iniciaremos detallando como Rawls presupone un tipo de psicología moral como elemento fundamental para garantizar la estabilidad de su teoría, para, en un segundo momento, señalar algunos problemas de ese modelo.

Palabras clave: John Rawls. Lawrence Kohlberg. Justicia como equidad. Psicología moral.

1 O argumento da estabilidade na teoria da justiça como equidade

É famosa a frase de abertura de *Uma teoria da justiça* (1971): “A justiça é a primeira virtude das instituições sociais, como a verdade é dos sistemas de pensamento”. (RAWLS, 2016, p. 4). Rawls estabelece, assim, a absoluta prioridade do papel distributivo de uma concepção política. Contudo, algumas páginas adiante, ele se vê às voltas com outros “problemas sociais fundamentais”, como os da coordenação de planos

de vida, da eficiência de arranjos sociais e econômicos e da estabilidade e unidade sociais (RAWLS, 2016, p. 7). Ainda que tais problemas estejam subordinados a questões de justiça, cumprem um papel importante na escolha entre teorias de justiça, pois, segundo Rawls, “permanecendo constantes as demais condições, uma concepção de justiça é preferível a outra quando suas consequências mais amplas são mais desejáveis”. (RAWLS, 2016, p. 8). Na terceira parte de *Uma teoria da justiça*, Rawls torna essa restrição ainda mais forte: “[P]or mais atraente que uma concepção de justiça possa ser em outros aspectos, ela estará seriamente comprometida se os princípios da psicologia moral foram tais que deixem de gerar nos seres humanos o necessário desejo de agir com base nela”. (RAWLS, 2016, p. 561).

Tais problemas incluem a preocupação com as consequências da aplicação de princípios de justiça em determinada sociedade e a própria viabilidade prática de uma concepção política. Isso foi previsto por Rawls na famosa função de utopia realista da *teoria da justiça como equidade*: “[E]ntendemos a filosofia política como realisticamente utópica; ou seja, como exame dos limites da possibilidade política praticável”. (RAWLS, 2003, p. 5-6).

Nesse sentido, o problema da estabilidade é introduzido exatamente para testar os limites de adesão a uma concepção política de acordo com as circunstâncias sociais de uma sociedade democrática. Sua função, dito de modo mais específico, é mostrar ao leitor de *Justiça como equidade...* que essa concepção mantém-se em “equilíbrio estável”. (RAWLS, 2016, p. 564). Esse estado ocorre na medida em que suas normas básicas são regularmente cumpridas pelos seus cidadãos e, além disso, distúrbios e perturbações (como crises econômicas, escândalos políticos ou debates intensos sobre leis que influenciam *nas* ou restringem *as* visões de mundo, como a regulamentação do aborto, da eutanásia ou do uso de drogas, por exemplo) invocam “forças internas” para o retorno da estabilidade. No caso de Rawls, tais *forças* podem ser identificadas através das leis psicológicas que são “cristalizadas” no senso de justiça dos cidadãos e de um suposto consenso sobreposto entre doutrinas abrangentes e uma concepção política. Todos os cidadãos adotariam o mesmo ponto de vista político, que é público e compartilhado mutuamente.

Um dos grandes problemas na leitura de uma *Teoria da justiça* é estabelecer claramente a posição do argumento da estabilidade no seu interior. Em obras posteriores a *Uma teoria da justiça*, Rawls começa a

falar desse argumento como sendo o *segundo* da posição original, a fim de deixar mais explícito que a escolha de uma concepção política, por um lado, também leva em consideração o problema da estabilidade e, por outro, ainda ocorre no modelo da posição original. Isso significaria, portanto, que o argumento é hipotético e essencialmente comparativo.

O trecho decisivo para compreender a função do segundo argumento encontra-se em *Justiça como equidade: uma reformulação*:

[A]s *partes* devem indagar se as pessoas que crescem numa sociedade bem-ordenada pelos dois princípios de justiça – os princípios adotados na primeira parte do argumento – adquirem um senso de justiça suficientemente forte e eficaz para que possam normalmente concordar com dispositivos justos e não sejam levadas a agir por outros motivos, por exemplo, pela inveja e pelo desprezo sociais, por uma vontade de dominar ou por uma tendência a se submeter. (RAWLS, 2003, p. 258).

Aqui, podemos compreender que a preferência de uma concepção de justiça com base na sua estabilidade é decidida pelas *partes* da posição original. Contudo, seu ponto de vista é outro: não mais o das partes da posição original, que buscam definir o problema da justiça distributiva, mas o dos cidadãos de uma sociedade bem-ordenada, que buscam definir se ser justo é bom (racional) para si próprio. As partes avaliam, como se verá mais adiante, “a probabilidade de os cidadãos que crescem nesse “pano de fundo” [institucional] se deixarem dominar por atitudes especiais desestabilizadoras”. (RAWLS, 2003, p. 258, grifo nosso).

Além disso, a abordagem a respeito de emoções e sentimentos é adiada para garantir a autonomia na decisão de princípios de justiça, que devem ser escolhidos “com base nos interesses fundamentais das pessoas consideradas cidadãos livres e iguais”. (RAWLS, 2003, p. 258). Caso se colocasse esse argumento em uma etapa apenas, além de o procedimento ser mais oneroso e complexo, do ponto de vista argumentativo, correr-se-ia o risco de escolher princípios com base em tais “fatores contingentes”. (RAWLS, 2016, §40).

Poderíamos dizer que Rawls está endossando uma “divisão de trabalho” entre concepções políticas normativas e princípios psicológicos básicos, algo realizado em suas aulas sobre J. S. Mill. (RAWLS, 2012, p. 293). Portanto, Rawls pretende mostrar, nessa segunda etapa do argumento, que seu procedimento hipotético dá conta de

aspectos e fatos cotidianos dos seres humanos. É nesse momento que Rawls realiza uma “restrição humeana” a concepções políticas: ou sua justificação está de acordo com as condições da nossa vida, conforme as conhecemos, ou não são possíveis. Podemos acrescentar a isso, conforme comentado, que uma teoria da justiça deve ser “psicologicamente adequada às inclinações humanas”. (RAWLS, 2016, p. 562).

Weithman, em seu *Why political liberalism?...* (2010) é, de longe, o autor que melhor esmiuçou esse argumento tão negligenciado pela literatura. De acordo com ele, o argumento acerca da estabilidade depara-se com dois problemas estabelecidos ainda na segunda parte de *Uma teoria da justiça*. (§42). Ali, Rawls vê-se às voltas com o problema (ou a tragédia) dos bens públicos e o secciona, entre os conhecidos problemas do isolamento (que envolve, tradicionalmente, a figura do carona [*free rider*]) e o problema da garantia mútua. (RAWLS, 2016, p. 332 ss). Rawls quer evitar, por um lado, que uma sociedade bem-ordenada recaia em uma “generalização do dilema do prisioneiro”, na medida em que cidadãos prefiram adotar pontos de vista alternativos aos da justiça como equidade. Não apenas a *irracionalidade* pode colocar em risco uma concepção política, mas também a adoção contínua de atitudes racionais autointeressadas, que se acomodam mais facilmente aos planos de vida de cada um. Ao mesmo tempo, Rawls pretende encontrar uma solução para a dificuldade de se estabelecer um reconhecimento público e compartilhado entre os cidadãos sobre a manutenção do acordo estabelecido. A disposição para sermos justos pode se ver fragilizada, caso não tenhamos garantias de que não sofreremos ou não nos arruinaremos ao adotar determinada atitude. O senso de justiça, para Rawls, faz com que busquemos promover esquemas justos, cumprindo nossa parte, quando percebemos que os outros também farão a parte deles. Um dos motivos de se considerar a justiça como equidade uma utopia realista tem este objetivo: mostrar que seus ideais são possíveis sem assumir motivações heroicas, pois não exigem uma posição completamente altruísta nem completamente egoísta.

Há diferentes modos de se resolver esses problemas, e a literatura cresce exponencialmente em diferentes campos. Para Rawls, o soberano hobbesiano é um dos exemplos mais emblemáticos de solução para o dilema do prisioneiro. Esse poder político, em sua opinião, é um elemento exterior, um “mecanismo acrescentado a um sistema de cooperação”, na medida em que sua função, como *terceiro imparcial*, é

modificar a tabela de ganhos dos jogadores mediante a coerção e a ameaça e, conseqüentemente, modificar a ordem de preferências desses. O soberano é responsável pela alteração das “condições fundamentais que possibilitam que os indivíduos pensem racionalmente”. (RAWLS, 2012, p. 86).

Essa solução radical afastou Rawls de Hobbes, sendo esse colocado ao lado de Santo Agostinho e Fiódor Dostoiévski como as “mentes sombrias do pensamento ocidental”. (RAWLS, 2012, p. 329). O modo como Hobbes acreditava que sentimentos morais deveriam ser desenvolvidos e a visão humana de um “sujeito sem autocontrole” tornaram-no um autor intratável para o problema da estabilidade de uma concepção político-liberal. Acreditamos que essas razões podem explicar a tão citada (e pouco interpretada) observação de Rawls a respeito de seu vínculo com o contratualismo moderno e a conseqüente exclusão de Hobbes.

Portanto, Rawls quer evitar soluções que compreendam a estabilidade como um assunto meramente prático, que pode se manter por meio da persuasão e da retórica ou do cerceamento de cidadãos. Um dos fatos gerais pressupostos por Rawls, desde o início de sua teoria, é o de que um regime democrático duradouro e estável deve ser apoiado de modo livre e voluntário pela maioria dos seus cidadãos. Por isso, do mesmo modo que a solução hobbesiana, a Mentira Nobre de Platão e a religião como instrumento para a coesão social, descrita em *Os irmãos Karamazov*, de Fiódor Dostoiévski, são excluídas do rol de recursos filosóficos disponíveis. (RAWLS, 2016, p. 560).

Contra Hobbes, Rawls quer deixar claro que os problemas que afetam a estabilidade de uma sociedade podem ser resolvidos *da mesma forma* com o desenvolvimento de um senso público de justiça apoiado em atitudes naturais, como amizade, amor ao próximo e confiança mútua:

Quando esses sentimentos são fortes o bastante para anular as tentações de transgredir as normas, os sistemas justos são estáveis. Cada pessoa passa a considerar que cumprir com o próprio dever e obrigação é a reação correta às ações dos outros. O plano racional de vida de cada pessoa, regido pelo senso de justiça, leva a essa conclusão. (RAWLS, 2016, p. 613).

2 Atitudes naturais e atitudes morais: o desenvolvimento do senso de justiça

Como podemos perceber, a plausibilidade do desenvolvimento de um senso de justiça é fundamental para o desfecho positivo do argumento. Gostaríamos de defender, aqui, que Rawls parte de duas premissas muito importantes: a primeira é metodológica e afirma que a “interpretação do desenvolvimento moral está totalmente vinculada à concepção de justiça”. (2016, p. 569-570). Isso significa dizer que, ao menos na formulação da teoria da justiça, a psicologia moral e o papel dos sentimentos e das emoções são sempre introduzidos após a escolha de princípios de justiça. Essa divisão de trabalho, como foi afirmado, serve para manter a autonomia da política em relação a assuntos e visões de mundo controversos e determinar a direção das exigências recíprocas.

Isso leva Rawls a uma segunda premissa, que exige comprovações empíricas (por mais que ele se negue a fornecê-las): a moralidade de princípios, apesar de ser definida através da teoria dos estágios de desenvolvimento, pode conter conteúdos diversos, de acordo com o contexto dos indivíduos. Rawls afirma, constantemente, que as instituições jurídicas e políticas e a compreensão que possuímos do mundo social e do político influenciam enormemente na nossa autocompreensão como pessoas: “[A] natureza humana e sua psicologia natural são elásticas: podem limitar as concepções de pessoa e os ideais políticos praticáveis, além das psicologias morais que podem servir-lhes de base, mas não determinam quais concepções e ideais deveríamos adotar. (2011, p. 104).

De um lado, Rawls está afirmando uma visão da natureza humana *plástica* muito próxima da de J.-J. Rousseau, como fica evidente em seus comentários sobre a psicologia moral do referido autor: “[O]s princípios da natureza humana são como uma função matemática: a depender das condições sociais e históricas, eles determinam os tipos de caráter que se desenvolverão e serão adquiridos na vida em sociedade”. (2012, p. 225).

De outro, ele está negando a possibilidade de explicação de conteúdos morais com base na referência direta à natureza humana, ainda que esta seja fundamental para impor limites a nossas expectativas de realização de um mundo (perfeitamente) justo: “[C]omo qualquer outra concepção política, para que seja praticável, suas exigências e seu ideal de cidadania devem ser aqueles que as pessoas possam entender, aplicar e se sentir suficientemente motivadas a respeitar”. (2011, p. 104).

É importante mencionar, ainda, que o argumento de estabilidade, no que se refere ao senso de justiça, permanece o mesmo após a reformulação da *teoria da justiça como equidade*, salvo alterações ligadas à restrição do domínio do político. Isso adquire evidência textual segundo a nota de rodapé de Rawls em *Justiça como equidade: uma reformulação*: “[A] psicologia moral que está por trás das premissas descritas nesse texto é exposta muito mais detalhadamente em *Teoria*, Cap. VIII, §§70-2, 75-6. Apenas faço referência a essas seções já que não as mudaria substancialmente”. (2003, p. 280, n. 17). Se lermos o final do prefácio de *O liberalismo político*, Rawls insiste na necessidade de uma “lealdade razoável e informada” por parte de seus cidadãos para a estabilidade de sociedades democráticas e a manutenção de um fórum público de deliberação política.

O desenvolvimento do senso de justiça, tal como exposto por Rawls, passa, então, pelas seguintes etapas: 1) a explicação da relação entre sentimentos naturais e sentimentos morais como dois tipos de disposições sobrepostas e, até certo ponto, semelhantes; 2) a elucidação das leis psicológicas envolvidas no processo de desenvolvimento moral, que descreve a modificação da compreensão e da relação afetiva que estabelecemos com nós mesmos e com nossas relações com o mundo a partir de princípios; e 3) a explicação de como o senso de justiça pode tornar-se um ponto de vista que regula nosso plano de vida e restringe tanto nossas concepções de mundo quanto nossas ações com os outros. Ao final, Rawls pretende explicar como alcançamos determinadas disposições para a cooperação e a justiça e como a valoração de diferentes objetos, pessoas e ações é *transformada* com a formação de uma consciência moral.

Rawls quer mostrar que sentimentos morais fazem parte de nossa vida humana comum e, com sua eliminação, estaríamos desfigurando a nós mesmos. Para tanto, Rawls parte da constatação, realizada no §74 de *Uma teoria da justiça*, de que há uma “relação psicológica íntima” entre atitudes naturais e atitudes morais. Negar sentimentos morais significa negar sentimentos e vínculos naturais; estaríamos destruindo ou reprimindo disposições como amor, amizade, confiança e diversos outros afetos. O egoísta, para Rawls, não é apenas incapaz de sentir remorso e indignação (sentimentos normalmente vinculados à moral), mas é incapaz de estabelecer laços de amizade e confiança mútua com seus próximos, ainda que possa manifestar emoções características, como:

raiva, tristeza e felicidade (2016, p. 602-603). A força do argumento atinge colorações contrafactuais quando Rawls afirma o seguinte: “[A]o entendermos como seria não ter um senso de justiça – que seria não ter parte de nossa humanidade também –, somos levados a aceitar o fato de termos tal sentimento”. (2016, p. 604).

Do mesmo modo, e esse pode ser visto como um ponto importante para a estabilidade entre concepções políticas, como se verá mais tarde, ao levarmos em consideração determinadas exigências de justiça e o teor de princípios de justiça, somos levados a compará-los com o tipo de vínculo afetivo que irá permear nossas relações cotidianas. Esse modo de colocar a questão envolve, ao menos implicitamente, o pressuposto de uma *figura* de nossa humanidade, isto é, uma comparação com os “recursos naturais” (RAWLS, 2016, p. 540) encontrados na ciência e no senso comum da cultura pública. Rawls tem em mente ao menos três critérios psicológicos básicos de comparação que cada psicologia moral deve fortalecer: 1) a afirmação de uma preocupação incondicional com nosso próprio bem, desenvolvido pelo senso de nosso próprio valor; 2) a consideração de que princípios e valores políticos devem possuir uma percepção clara de seus ideais morais; e 3) a consideração de que pessoas justas, que fazem sua parte no esquema de cooperação, receberão admiração e reconhecimento dos demais. (2016, p. 687-688). Poderíamos dizer que tais critérios são retirados de convicções e conhecimentos gerais a respeito da natureza humana e dos princípios psicológicos estabelecidos pela justiça como equidade. Um dos principais problemas desse modo de abordagem reside, como iremos argumentar no final, que esse é apenas *um* modo de articular as diferentes fundações de nossa *matriz moral*.

O modo como Rawls estabelece a relação entre sentimentos naturais e morais leva à confirmação de sua adesão ao que ele chama de “teoria racionalista” do desenvolvimento moral, baseada em Rousseau, Kant, Mill, Piaget e Kohlberg:

A outra tradição da aprendizagem moral deriva do pensamento racionalista e é ilustrada por Rousseau e Kant, algumas vezes por J. S. Mill, e mais recentemente pela teoria de Piaget. A aprendizagem moral não é tanto uma questão de fornecer motivos que faltam, mas sim do livre desenvolvimento de nossas capacidades intelectuais e emocionais inatas, de acordo com sua tendência natural. Quando as capacidades

de entendimento amadurecem e as pessoas vêm a reconhecer o seu lugar na sociedade, sendo capazes de adotar o ponto de vista dos outros, elas apreciam os benefícios mútuos do estabelecimento de termos equitativos de cooperação social. Temos uma simpatia natural com as outras pessoas, e uma suscetibilidade inata para os prazeres proporcionados pelo sentimento de companheirismo e pelo autodomínio, que fornecem uma base afetiva para os sentimentos morais, uma vez que sejamos capazes de ter um entendimento claro de nossas relações com nossos consócios, de uma perspectiva adequadamente geral. Assim, essa tradição considera os sentimentos morais como uma consequência natural de uma plena valorização de nossa natureza social. (RAWLS, 2016, p. 510).

Ou seja, essa teoria racionalista do desenvolvimento moral parte de duas premissas acerca do aprendizado moral: 1) este é o livre-desenvolvimento de nossas capacidades intelectuais e emocionais inatas; e 2) ele depende do reconhecimento de determinadas regras. A inspiração de Rawls para a construção de seus três estágios da moralidade (*moralidade de autoridade*, vinculada à fase infantil; *moralidade de grupo*, vinculada a padrões morais de associações, e *moralidade de princípios*), vem imediata e diretamente do psicólogo americano Kohlberg (1927-1987), criador da “Teoria do Desenvolvimento Moral”:

Embora a visão de desenvolvimento moral a ser seguida nos §§ 70-72 se destine a adequar-se à teoria da justiça, eu a tomei emprestada de várias fontes. A ideia dos três estágios cujo conteúdo é dado pelos preceitos, ideais de funções e princípios se assemelha a William McDougall, *An Introduction to Social Psychology* (Londres, Methuen, 1908), caps. VII-VIII. A obra de Piaget, *The Moral Judgment of the Child*, me sugeriu a contraposição entre a moralidade de autoridade e as moralidades de grupo e de princípios, e grande parte da descrição desses estágios. Ver também a elaboração que Kohlberg faz desse tipo de teoria, nas referências na nota 6 acima, esp. pp. 369-389, sobre os seus seis estágios. Na última parte do § 75, observo algumas diferenças entre a visão que apresento e a desses autores. A respeito da teoria de Kohlberg, eu gostaria de acrescentar aqui que acredito que a moralidade de grupo equivale aos seus estágios de número três a cinco. O desenvolvimento nesse estágio consiste em tornar capaz de assumir papéis mais complexos, exigentes e abrangentes. Mas o mais importante é que suponho que o estágio final, a moralidade de princípios, pode ter

conteúdos diferentes, apresentados por qualquer uma das doutrinas filosóficas que discutimos. É verdade que defendo a superioridade da teoria da justiça, e elaboro a teoria psicológica a partir dessa suposição; mas essa superioridade é uma questão filosófica e não pode, julgo eu, ser estabelecida apenas pela teoria psicológica do desenvolvimento. (RAWLS, 2016, p. 692, nota 8).

A relação entre Kohlberg e Rawls, ambos professores em Harvard, é mútua, sendo que o psicólogo educacional também vê a obra do filósofo como uma excelente aplicação de sua teoria. Em seus *Ensaio sobre o desenvolvimento moral*, Kohlberg afirma que “alguns, como Rawls e eu mesmo, identificam a moralidade com a justiça”. (1992, p. 196).

Além de Rawls, Kohlberg influenciou fortemente o filósofo alemão Habermas, que chegou a formular, em uma conferência clássica de 1974, intitulada “Desenvolvimento da Moral e Identidade do Eu”, um sétimo estágio de desenvolvimento moral, vinculado à “liberdade moral e política” e cuja esfera de validade seria “todos enquanto membros de uma fictícia sociedade mundial”. (HABERMAS, 1983, p. 68).

A partir disso, gostaríamos de defender o seguinte: 1) a moralidade de princípios, determinante para a formação do senso de justiça, é toda construída por analogia à moralidade de associação e, dessa forma, parece realizar um *salto mortale* na solução do problema da estabilidade; e 2) Rawls distingue-se de Kohlberg, entretanto, na medida em que, como afirmamos, pressupõe que nossa consciência moral possa ser modificada de acordo com a cultura e as instituições políticas. Essa é a interpretação psicológica da famosa frase de Rawls a respeito da influência profunda das instituições no autorrespeito e nas expectativas legítimas dos cidadãos:

A justiça ou injustiça dos arranjos institucionais da sociedade e as convicções humanas sobre essas questões influenciam de maneira profunda os sentimentos sociais; em grande parte, definem como vemos a aceitação ou a rejeição de uma instituição por outra pessoa, ou seu empenho em reformá-la ou defendê-la. (RAWLS, 2016, p. 607).

Vejamos como essa interpretação pode adquirir plausibilidade.

O senso de justiça é formado em uma sequência de três estágios de desenvolvimento moral, todos com a pressuposição do que Rawls chama de “leis psicológicas da reciprocidade”, pois derivam da interpretação

dos princípios de justiça da justiça como equidade e de seu arranjo institucional justo. Elas evidenciam um “fato psicológico profundo” dos seres humanos, cujas condições de sociabilidade seriam muito diferentes, se não tivessem essa tendência de retribuir, na mesma moeda e de responder de maneira recíproca, a comportamentos e manifestações afetivas.

É assim que a moralidade de autoridade, o primeiro estágio, é alcançada: o bebê cria sentimentos de amor, admiração e confiança em relação aos seus pais, na medida em que esses o amam “de forma manifesta” ou com “intenção evidente”, isto é, através de atos explícitos de cuidado, encorajamento, preocupação com suas necessidades e em enunciar regras claras e inteligíveis de acordo com o nível de compreensão da criança. Se tais afetos recíprocos não ocorrerem, as relações de amor e confiança ou a compreensão de que as normas e limites impostos pelos pais indicam modos de ação tornam-se ameaças e repressões.

A moralidade de associação, por outro lado, ocorre com a expansão do círculo afetivo familiar e é caracterizada pela capacidade do indivíduo de assumir diferentes papéis nas associações a que ele pertence. Cada papel possui padrões ou ideais estabelecidos de acordo com o contexto dos objetivos ou fins da própria associação – lazer, esporte, vínculos afetivos, crenças religiosas, etc. Necessárias para essa etapa são as habilidades de perceber que se é distinto das outras pessoas e, assim, discernir nossas próprias aspirações, convicções e desejos de outros e conseguir colocar-se no lugar desses outros (a habilidade de empatia). (RAWLS, 2016, p. 579).

O reconhecimento desses papéis e exigências normativas expande nossa sensibilidade moral com a criação de vínculos afetivos como laços de amizade, admiração e confiança mútua da mesma forma que sentimentos de culpa, vergonha e indignação. Cria-se, com isso, diferentes ideais e virtudes de caráter, que vinculam os indivíduos reciprocamente.

Contudo, o ponto fundamental, aqui, é que Rawls estabelece uma relação muito próxima entre moralidade de associação e moralidade de princípios:

A moralidade de associação assume muitas formas, dependendo da associação e do papel em questão, e essas formas representam muitos níveis de complexidade. Porém, se analisarmos as posições mais exigentes que são definidas pelas principais instituições da sociedade, os princípios

da justiça serão reconhecidos como reguladores da estrutura básica e como parte do conteúdo de inúmeros ideais importantes. De fato, esses princípios se aplicam ao papel do cidadão que pertence a todos, pois supõe-se que todos, e não só aqueles que participam da vida pública, tenham opiniões políticas acerca do bem comum. Assim, podemos supor que existe uma moralidade de associação na qual os membros da sociedade se veem como iguais, amigos e associados, juntos em um sistema de cooperação que se sabe destinar-se ao benefício de todos e regido por uma concepção de justiça comum. (2016, p. 582-583).

Portanto, os estágios de desenvolvimento moral de associação e de princípios encontram-se, da mesma forma que as atitudes naturais e morais, superpostos. Assim como uma pessoa passa a desejar ser um bom esportista ou um bom professor, com vínculos afetivos como admiração e vergonha, indignação e culpa, ela também passa a desejar ser uma pessoa justa: “[A] moralidade de associação leva, *naturalmente*, a um conhecimento dos padrões de justiça”. (RAWLS, 2016, p. 584). O senso de justiça, portanto, diferencia-se dos vínculos afetivos de associação, na medida em que, por um lado, liga-se a instituições sociais justas, que possibilitam a cooperação justa (a união social de uniões sociais), e, por outro, por tornar os sentimentos morais “independentes de contingência”, pois toda relação justa é uma relação entre cidadãos livres e iguais (em sentido político).

De acordo com essa “teoria dos sentimentos” de Rawls, as leis psicológicas que regem o desenvolvimento moral não apenas modificam os laços afetivos, mas, dada a modificação desses, modificam-se *também* os interesses e objetivos últimos das pessoas. Isso não significa dizer, por exemplo, que, à luz de alguma afirmação adicional, encontramos um melhor meio de satisfazer nossos desejos (como a afirmação de rotas alternativas de viagem ou restaurantes melhores com preços mais atrativos). Isso seria, diz Rawls, um juízo racional. A explicação da psicologia moral fornece um esclarecimento de como transformamos o *padrão* de objetivos últimos, isto é, uma transformação do modo como hierarquizamos nossas preferências em um plano de vida específico. Esse é o *mesmo resultado* que Rawls buscava ao substituir o Leviatã, de Hobbes, e fugir de um mecanismo externo, desvinculado do próprio sistema de justiça. Ao formar novos vínculos afetivos, como o senso de justiça, o próprio modo de compreender nossas relações com os outros indivíduos

e os diferentes ideais vinculados a elas são modificados e ganham um ponto de vista moral como norte.

A busca pela resolução do problema da cooperação social, com isso, recebe uma resposta baseada na capacidade dos seres humanos de criarem laços afetivos fortes o suficiente para restringirem suas atitudes autointeressadas ou seus desejos de hierarquização e dominação. A figura do *free rider* é substituída pela figura do *reciprocity rider*. (WEITHMAN, 2010, p. 148).

Essa tradição racionalista da psicologia moral é escolhida por Rawls porque abre caminho para uma separação – que será mantida e reforçada ao longo de sua obra – entre o justo e o bem. Em outras palavras, isso significa que a ideia de justiça seria algo proveniente de nossa natureza e que saberíamos separá-la de nossas concepções particulares de bem. Nesse sentido, é indispensável, para a construção dos estágios da moralidade, segundo Rawls, que

a explicação psicológica da aprendizagem moral seja verdadeira e esteja de acordo com o conhecimento concreto. [...] A não ser que a explicação psicológica tenha defeitos que coloquem em dúvida o reconhecimento dos princípios da justiça e não o padrão de utilidade [...], dela não decorre nenhuma dificuldade irreparável. (2016, p. 512).

Qual é o problema, então? Bem, a teoria psicológica da explicação moral adotada por Rawls não é verdadeira, não está de acordo com o conhecimento concreto e tem sérios defeitos que colocam em dúvida o reconhecimento dos princípios de justiça – ao menos na forma pela qual Rawls entende que eles devam se manter estáveis.

Passamos, então, à segunda parte deste trabalho, para indicar quais são esses defeitos e em que sentido eles podem contribuir para repensarmos as teorias contemporâneas do liberalismo político e seus impasses.

3 Objeções à psicologia moral rawlsiana

Uma das exposições mais claras sobre as implicações decorrentes das teorias das emoções para a psicologia moral é feita pelo filósofo e psicólogo americano Haidt, professor na New York University's Stern School of Business.

Em seu *The righteous mind: why good people are divided by politics and religion* (2012, p. 817), Haidt adota uma visão humeana da moralidade, ou seja, sistemas morais são conjuntos de virtudes, práticas, regras, instituições e mecanismos psicológicos cujo propósito principal é restringir o autointeresse e possibilitar a cooperação social. Para ele, todo esse sistema está assentado em intuições morais que surgem, de modo automático e instantâneo, muito antes de ter sido iniciado qualquer processo de raciocínio moral. Essa visão sobre a moralidade seria mais adequada para explicar, portanto, como as pessoas, ao se engajarem em debates morais ou sobre princípios de justiça, independentemente de seu nível de instrução e mesmo de conhecimento específico sobre moralidade e política, podem ser tão tendenciosas, ilógicas e reprodutoras de mentiras e teorias da conspiração. A ideia fundamental da ética kantiana – retomada pelos psicólogos racionalistas, por Rawls, Habermas, etc., como vimos – de se colocar no lugar do outro simplesmente não ocorre no debate político, no qual os participantes estão imensos em suas próprias narrativas juntamente com os membros de seu grupo.

Haidt está fazendo, portanto, uma crítica à psicologia racionalista de Piaget e Kohlberg, que estariam justificando a ordem moral-liberal e secular – voltada contra a autoridade, a hierarquia e a tradição – através do princípio moral de colocar-se no lugar do outro.

Após estudar as obras de Piaget, Kohlberg e Turiel, justamente os teóricos dos quais Rawls e Habermas retiram a fundamentação de suas psicologias morais, Haidt se dá conta de que há algo errado nesses modelos.

Além da crítica a essa “ilusão racionalista”, fruto do exagerado otimismo com a nossa capacidade reflexiva, Haidt descobre, a partir de diferentes pesquisas, que a amplitude do domínio da moral varia de acordo com o contexto de aprendizado cultural. Isso explicaria, portanto, por que até mesmo elementos que o individualismo liberal contemporâneo, em tese, considera como não portadores de caráter moral, mas simples convenções sociais, têm – sim – peso moral: vestimenta, alimentação, escolhas sexuais, etc. Pessoas da mesma sociedade terão diferentes visões do domínio moral, a depender do grupo social e cultural em que nasceram, cresceram e se desenvolveram. É aqui que Haidt extrai uma conclusão surpreendente ao comentar os resultados finais de suas pesquisas realizadas na Filadélfia, em Porto Alegre e em Recife:

“[E]u viajei cinco mil milhas ao sul em busca de variações morais quando, na verdade, há mais para se encontrar a alguns quarteirões a oeste do campus, no bairro pobre das cercanias de minha universidade”. (2012, p. 91).

Mas não é só isso: suas pesquisas também indicaram uma nova interpretação ao funcionamento da moralidade: as pessoas que condenavam certos atos, mesmo que eles não oferecessem danos a terceiros, o faziam porque já os consideravam ofensivos, e não porque utilizaram processos de racionalização a respeito deles. Quando questionadas a respeito de suas decisões morais, várias pessoas simplesmente não conseguiam apresentar argumentos para suas decisões intuitivas. O processo de racionalização, portanto, é posterior à intuição e tem como função tentar justificar as reações emocionais, ou seja, não só a moralidade varia em diferentes meios, mas ela própria seria produto de sentimentos anteriores de desgosto e desrespeito relacionados ao meio em que se vive.

Haidt, com isso, faz uma revisão das principais conquistas de três áreas distintas que se voltaram ao estudo da formação de concepções morais: a psicologia, a antropologia e a teoria da evolução, inspirando-se nas obras do cientista cognitivo Pinker, do entomologista e biólogo Edward O. Wilson, do primatólogo Frans de Waal, do neurocientista Antonio Damasio, entre outros.

Seu ponto central é mostrar que é simplesmente inútil reproduzir a clássica dicotomia entre cognição e emoção, já que a emoção é parte inseparável do próprio processamento de informação. Haidt usa a metáfora de um elefante (emoção, afetos e intuições, que ocorrem primeiro), sendo conduzido por alguém que está montado nele (cognição e raciocínio estratégico, que vem depois). Com essa imagem, é possível entender – literalmente – o peso das emoções em nosso processo moral-decisório, mas também visualizar que é possível que o guia altere o curso do elefante, embora isso não seja tão simples e fácil como modelos racionalistas/cognitivos pressupõem. Segundo Haidt,

o julgamento moral não é uma tarefa puramente cerebral na qual ponderamos questões a respeito de danos, direitos e justiça. Ele é um processo de tipo rápido e automático semelhante aos julgamentos que os animais fazem quando se movem no mundo, sentindo-se impelidos por ou contra várias coisas. (2012, p. 220).

Isso ocorre porque a razão, do ponto de vista evolutivo, surgiu muito depois nos seres humanos. É essa característica que faz a razão ser uma ferramenta que busca justificativas, e não a verdade, o que estaria ligado a processos de garantia de reputação perante o grupo social em que se vive, elemento fundamental à sobrevivência. Haidt (2012) compara nossos processos de decisão moral a políticos querendo votos: ao contrário de cientistas, que buscam a verdade, os políticos não estão preocupados com *estarem* certos, mas com *parecerem* certos. Isso basta para garantir-lhes apoio e reputação.

Com isso, Haidt estabelece sua “Teoria das Fundações Morais”: diferentes concepções morais e visões da sociedade estimulam e desenvolvem diferentes fundações ou sistemas psicológicos. Dependendo do modo como são desenvolvidas e estimuladas, tais fundações formam determinada matriz moral, responsável por modelar nossas intuições morais. Assim, Haidt tira conclusões radicalmente contrárias as de Rawls: nossas deliberações políticas são essencialmente intuitivas. A própria ideia de separação entre o justo e o bem, entre o domínio do político e o domínio das doutrinas abrangentes, articula *uma* fundação moral específica (liberdade/opressão), ou seja, nem todas as concepções morais, surgidas no *interior* da *mesma* sociedade, desenvolvem uma disposição a dar primazia à dimensão política sobre outras. (HAIDT, 2012, p. 737). Segundo Haidt, “nós somos múltiplos desde o início”, isto é, temos a possibilidade de nos tornar justos sobre diferentes causas e preocupações, mas apenas algumas são ativas ao longo do nosso aprendizado moral. Por isso, a tentativa de estabelecer um “módulo cognitivo universal” (um “monismo moral”) é *desumano*. (p. 357). Como nossas visões políticas são essencialmente intuitivas, não há como depositar todas as nossas “fichas” em uma deliberação racional para a estabilidade política de princípios de justiça que priorizam valores específicos. Isso simplesmente não iria funcionar. Somos muito mais glauconianos e dependentes da ocasião, muito menos coerentes e pacientes com justificativas alheias do que estamos dispostos a admitir.

4 Conclusão

Como pensar, então, nos processos de deliberação política a partir dessa perspectiva? A explicação de Haidt tem o mérito de descrever muito melhor esses processos. Opiniões políticas – seja a respeito de princípios, seja a respeito de questões triviais – funcionam na esfera

pública como ferramentas de criação de laços sociais. As pessoas aderem a certas ideologias, partidos políticos, etc. da mesma forma que aderem a times de futebol, religiões, associações, tradições, etc. Não é o egoísmo que caracteriza a política, mas o comportamento de busca de aceitação no grupo ao qual se pertence. Esse processo independe do grau de instrução: militantes políticos, especialmente os radicalizados, tornam-se facilmente alienados em relação a fatos e adeptos a crenças paranoicas e teorias da conspiração.

Se isso aponta às limitações para o debate a respeito de questões políticas triviais, o mesmo é válido para a discussão de princípios políticos e de justiça. Rawls aposta, como vimos, que a estabilidade do sistema político democrático-liberal ocorrerá através da adesão ao que ele chama de “consenso sobreposto” a respeito de princípios políticos fundamentais, ou seja, liberdades básicas e promoção de uma sociedade equitativa. A pergunta é: A estabilidade (ou instabilidade) dos regimes democráticos realmente ocorre a partir dessa ideia de cidadãos que teriam chegado a um suposto estágio mais avançado de desenvolvimento moral?

Não parece ser esse o caso, como mostram tanto pesquisas contemporâneas sobre desenvolvimento moral quanto a própria prática política.

Referências

- GREENE, Joshua. *Moral tribes*. New York: The Penguin Press, 2013.
- HABERMAS, Jürgen. *Para a reconstrução do materialismo histórico*. São Paulo: Brasiliense, 1983.
- HAIDT, Jonathan. *The righteous mind: why good people are divided by politics and religion*. New York; Toronto: Pantheon, 2012.
- KOHLBERG, Lawrence. *Psicologia del desarrollo moral*. Bilbao: Desclée de Brower, 1992.
- RAWLS, John. *Uma teoria da justiça*. São Paulo: M. Fontes, 2016.
- _____. *Justiça como equidade: uma reformulação*. São Paulo: M. Fontes, 2003.
- _____. *O liberalismo político*. São Paulo: M. Fontes, 2011.
- _____. *Conferências sobre a história da filosofia política*. São Paulo: M. Fontes, 2012.
- WEITHMAN, Paul J. *Why political liberalism?: on John Rawls's political turn*. New York: Oxford University Press, 2010.

Submetido em 28 de outubro de 2016.
Aprovado em 14 de março de 2017.