

RAWLS E A JUSTIFICAÇÃO DA PUNIÇÃO

Denis Coitinho¹

Resumo: O objetivo central deste artigo é procurar compreender a concepção normativa híbrida usada por John Rawls para justificar moralmente a punição legal, fazendo uso de um recurso consequencialista, a fim de justificar a instituição da punição, isto é, a eficácia social e um recurso retributivista para justificar os atos punitivos particulares, ou seja, a culpa do agente. Analisaremos, especialmente, o artigo “Two Concepts of Rules” e o livro *A Theory of Justice*. Defenderemos que Rawls faz uso de uma visão liberal de punição, bem como utiliza o método de equilíbrio reflexivo, estipulando uma regra para aprovação da punição legal, a partir da coerência com nossos juízos morais ponderados, o que parece reconciliar a justiça retributiva com a justiça distributiva.²

Palavras-Chave: Punição. Responsabilidade moral. Equilíbrio reflexivo. John Rawls.

1 INTRODUÇÃO

A questão de como podemos justificar a instituição da punição e suas práticas punitivas é um dos problemas centrais da filosofia política e da filosofia do direito. Se olharmos a história da filosofia, veremos que os mais importantes filósofos, tais como Aristóteles, Kant, Hegel, Bentham e Mill, se preocuparam com esse problema, a saber, o de como é possível justificar uma instituição estatal, em fazendo uso do seu aparato legal para infligir sanções que causam dano intencional a uma parte significativa de seus cidadãos? A dificuldade da questão está em justificar o dano intencional, uma vez que causar dano intencional a alguém é algo errado, em situações normais. É a complexidade da questão está na sua inter-relação com um conjunto adicional de problemas morais, tais como as questões de erro, autoridade, liberdade, censura e responsabilidade, por exemplo.

John Rawls não será exceção nesse cenário, já que o problema da punição receberá sua atenção, ao longo de seu percurso filosófico, desde os

¹ Professor do Programa de Pós-Graduação da Universidade do Vale do Rio dos Sinos - UNISINOS e bolsista de produtividade do CNPq. E-mail: deniscoitinhosilveira@gmail.com

² Usaremos as seguintes abreviações para os textos de Rawls: “Punishment” (PU), “Two-Fold Basis of Justice” (TB), “Two Concepts of Rules (CR), *A Theory of Justice* - Original Edition (TJ), *A Theory of Justice* - Revised Edition (TJ rev.) “The Sense of Justice” (SJ), *Justice as Fairness (JF)*, *Political Liberalism (PL)*, *The Law of Peoples (LP)*, “The Idea of Public Reason Revisited” (PR).

<http://dx.doi.org/10.1590/S0101-31732017000300005>

primeiros artigos, em especial CR, até a sua obra central, *TJ*. Acontece que sua concepção de responsabilidade jurídica apresenta um sério problema, para os intérpretes de Rawls: ela parece ser contraditória com as premissas centrais da justiça como equidade, porque Rawls defenderá uma posição híbrida entre consequencialismo e retributivismo, sustentando a justificação da instituição da punição pela eficácia social, isto é, pela estabilidade social e a justificação dos atos particulares de punição pela culpa do agente em agir erroneamente, ou seja, pelo seu mérito. Isso talvez explique, ao menos parcialmente, o desinteresse dos comentadores em abordar a concepção de justiça retributiva defendida por Rawls.

O problema é como acomodar a concepção utilitarista de Rawls sobre a justificação da instituição da punição com a justiça como equidade como um todo, em razão de ela ser uma clara contraposição ao modelo utilitarista usado em filosofia moral e política, por não respeitar adequadamente os direitos das pessoas enquanto livres e iguais. Também, há um forte aspecto retributivista em sua forma de justificar os atos punitivos, conectando intrinsecamente mérito e punição, o que distingue a sua compreensão de justiça retributiva da justiça distributiva, a qual não considera o mérito como critério central para a distribuição dos bens primários. Seria trágico para uma teoria que toma a coerência como critério decisivo para a justificação apresentar uma incoerência interna dessa proporção entre as suas partes. Dessa maneira, como compreender a concepção de justiça retributiva de Rawls?

Para tentar mostrar que não há uma contradição entre as partes da justiça como equidade, gostaria de apontar duas importantes questões iniciais, a saber, que a concepção rawlsiana de punição defende uma visão liberal e usa o mesmo método empregado nas questões de justiça distributiva, quer dizer, o método de equilíbrio reflexivo.

Deixem-me iniciar com a visão liberal. De forma geral, uma visão liberal defende que as instituições políticas e jurídicas (i) devem respeitar a todos como sendo livres e iguais e (ii) só podem ser justificadas em termos políticos. Dito isso, podemos afirmar que a concepção de punição de Rawls é claramente liberal, por defender uma neutralidade ética, bem como os direitos individuais dos agentes. De acordo com Rawls, as razões para punir um ofensor não podem ser parte de uma doutrina moral abrangente, o que implicará uma concepção contraposta ao paternalismo e ao moralismo jurídico. Esse aspecto é claramente compartilhado pelos utilitaristas, em especial Mill, ao estipular o “dano aos outros” como único critério para a restrição da liberdade individual.

Ademais, Rawls defende os direitos de liberdade e igualdade das pessoas, o que revela uma grande proximidade com o retributivismo, em especial o kantiano, uma vez que este também parece valorizar a liberdade e autonomia humana com o critério de mérito. Por isso, parece que a concepção de punição defendida por Rawls não é contraditória com a justiça como equidade, mas revela uma teoria liberal que incluirá, também, uma argumentação contratualista para a justificação das normas em sociedades que são marcadas pelo pluralismo moral razoável, através do uso público da razão.

Em acréscimo, a respeito do equilíbrio reflexivo, é importante identificar que tanto a justiça distributiva como a retributiva farão uso do mesmo método coerentista, o qual tem a característica de procurar identificar os juízos morais ponderados e ver que tipo de regra seria estipulada a partir desses valores morais compartilhados. A estratégia geral usada é procurar ver quais seriam os valores morais dos envolvidos e tentar especificar que regras seriam aceitas por sua coerência com esses valores. A liberdade ou autonomia é claramente um valor socialmente compartilhado, uma vez que nos vemos como responsáveis, porque censuramos a nós mesmos e aos outros. De modo similar, valoramos socialmente a dignidade humana, o que parece trazer por consequência a garantia dos direitos individuais, bem como o desejo pela estabilidade social. Esses valores são centrais para quando pensamos sobre as razões para punir. Por outro lado, parece que a igualdade apenas se mostra como um valor compartilhado, quando pensamos sobre os nossos direitos sociais e econômicos, não sendo tão relevante para quando pensamos sobre a responsabilização. Com isso, não seria difícil ver que a regra de justiça que permitiria a punição teria que incluir a intenção de prevenir futuros crimes com a garantia das liberdades individuais, o que significaria tomar o sujeito individual como responsável. Voltaremos a isso posteriormente.

2 O PROBLEMA DA PUNIÇÃO

Entretanto, qual é mesmo o problema da punição de que estamos tratando? Rawls especifica o problema da instituição da punição como um problema ético, isto é, o de como podemos justificar normativamente a punição, visto que ela implica um dano intencional infligido pelo Estado aos ofensores:

O problema, então, é um problema ético: como a instituição da punição pode ser justificada. Isso não é uma questão de como ela se desenvolveu

historicamente, ou quais motivos as pessoas têm para mantê-la, ou que benefícios alcançamos com seu funcionamento, ou que funções sociais ela desempenha. A questão é apenas: qual é a sua justificação ética – como são os fundamentos éticos para a instituição [da punição] ser estabelecida. (PU, p. 1).

Em um texto não publicado, intitulado PU, de 1952-1953, que parece ser uma primeira versão do artigo CR, Rawls se pergunta a respeito da justificação normativa da punição legal. A questão é a de como poderíamos justificar moralmente o dano ou sofrimento intencional retributivo reprobatório infligido pelo Estado ao ofensor? Isso é um problema, uma vez que querer causar sofrimento a alguém não é algo correto. Mas, quando o Estado aplica medidas punitivas, ele causa sofrimento a alguns dos seus membros que descumpriram a lei. Então, como justificar essas ações punitivas legais?

A estratégia que Rawls usará será conciliar as tradicionais respostas dadas a esse problema, a saber, as respostas consequencialista e retributivista, as quais fazem uso dos critérios normativos de eficácia social e mérito, respectivamente, procurando superar os seus problemas. O utilitarismo sustenta que podemos justificar a instituição da punição por seus bons efeitos sociais, porque os atos punitivos teriam a função de inibir futuros crimes. O problema do utilitarismo é que justificar a instituição da punição pelos efeitos sociais positivos poderia trazer por consequência punir um inocente, ou mesmo punir um agente culpado em demasia. Por sua vez, o retributivismo defende que a justificação da punição se encontra no mérito do agente em sofrer em razão do ato errado de que é culpado. O erro do retributivismo parece estar na sua compreensão equivocada de livre-arbítrio, de forma a ver o agente que comete um erro como merecedor de sofrimento, porque teria escolhido livremente. A particularidade dessa conciliação se encontra especificamente na distinção feita por Rawls entre a (i) justificação da instituição da punição (J1) e a (ii) justificação de um ato punitivo particular (J2). E, assim, nesse modelo híbrido, J1 seria alcançada pela prevenção de futuros crimes, enquanto J2 seria dada pela condenação apenas do culpado.

Essa posição híbrida hoje é classificada como um retributivismo negativo ou fraco, pois estabelece uma cláusula para só se punir alguém, se o mesmo for culpado do crime. Assim, seria errado punir o inocente, porque, não havendo cometido nenhum crime, ele não mereceria ser punido e, também,

seria errado punir os ofensores mais duramente do que eles merecem. Veja-se a diferença com o retributivismo positivo, que oferece uma justificação positiva para a punição: devemos punir o culpado, porque ele merece sofrer. O retributivismo negativo, por sua vez, não nos diz por que devemos punir. Ao contrário, ele nos diz que, seja lá qual for a justificação positiva para a instituição da punição, nós não devemos punir o inocente ou puni-lo mais drasticamente do que ele merece.³

Isso já parece nos mostrar que Rawls não está defendendo uma teoria moral abrangente retributivista para a justificação da punição, uma vez que apenas está usando um critério retributivista de mérito, para reafirmar os direitos individuais como a liberdade e a dignidade humana. De forma semelhante, parece que não está defendendo uma teoria utilitarista, mas usando apenas o critério de estabilidade social para a justificação da punição, de forma a possibilitar a coerência entre os nossos valores morais compartilhados e as regras. No restante do artigo, analisaremos os textos CR e *TJ*, a fim de ver como Rawls trata o problema da justificação da punição. Posteriormente, procuraremos identificar as principais virtudes e os principais limites dessa concepção híbrida.

3 CR E A RECUSA DO *TELISHMENT*

Em CR, Rawls defende uma concepção de punição vinculada a um consequentialismo de regras e não de atos, fazendo uma importante distinção entre justificar uma ação e justificar uma prática. O argumento central do texto é enfatizar que a regra que aprovaria a punição do inocente visando ao bem social, isto é, à maximização do bem-estar, não seria aprovada por nossos juízos morais ponderados que valorizam a dignidade humana e a liberdade. E, então, a solução será encontrada em um modelo híbrido, o qual garante J1 pela eficácia social com a cláusula de só punir o culpado (J2).

³ Thom Brooks esclarece que, para o retributivismo negativo, o mérito é uma condição necessária, porém, não suficiente para a punição. Ele salienta: “O mérito é necessário, mas não é suficiente para o retributivismo negativo. O mérito é necessário pelas mesmas razões que ele é necessário para o retributivismo positivo. A punição é apenas aplicada ao agente que merece [...]. Entretanto, o mérito não é suficiente para determinar se devemos punir uma pessoa que merece.” (BROOKS, 2012, p. 97). Observa ainda que o retributivismo negativo é um tipo de consequentialismo de regras. Ver Brooks (2012, p. 96-99).

3.1 O PROBLEMA MORAL DA PUNIÇÃO

Rawls inicia o seu texto apontando para o problema moral da punição, ou seja, como se pode justificar moralmente a instituição da punição, já que ela causa um dano intencional aos agentes, e isso é um erro em situações ordinárias. Parte das duas respostas mais conhecidas para esse problema, a saber, a resposta retributivista e a utilitarista. Para o retributivismo, a punição é justificada pelo ato errado cometido. A pessoa que age erroneamente mereceria sofrer em proporção ao seu erro, em razão de sua culpa em ter cometido um crime. Essa concepção tem por foco central o passado, uma vez que o que justificaria a prática punitiva seria a culpa do agente em ter agido erroneamente, de forma livre. Por outro lado, para o utilitarismo, a punição seria justificada pelas futuras consequências que traria, a saber, a eficácia social, e isso por ter um efeito inibidor de futuros crimes. Parece implicar que o seu foco seria o futuro, porque pretende assegurar a estabilidade social, não estando centrada exclusivamente na culpa do agente.

A estratégia argumentativa de Rawls é estabelecer uma conciliação entre essas duas posições, fazendo uma distinção entre (i) justificar uma prática como um sistema de regras e (ii) justificar uma ação particular que recai sobre essas regras. Ressalta que o argumento utilitarista parece mais apropriado com respeito à prática da punição, enquanto o argumento retributivista parece se encaixar melhor na aplicação dos atos punitivos.⁴ Para esclarecer essa importante diferença, imagina um diálogo entre um pai e seu filho. Imagine que um filho pergunte para o pai “Por que J foi preso ontem?” e o pai responda: “Porque ele roubou o banco B. Ele foi devidamente julgado e considerado culpado”. Agora, imagine que a pergunta do filho seja: “Por que as pessoas colocam as outras na cadeia?” E, então, a resposta do pai seria: “Para proteger as pessoas boas das más” ou “para possibilitar que as pessoas possam dormir em paz.” (CR, p. 5).

⁴ Importante ver que o argumento utilizado por Hart, em “Prolegomenon to the Principles of Punishment”, parece similar, uma vez que a justificação geral da punição se daria pela minimização dos erros futuros com a maximização da liberdade. Hart concorda com Bentham que a prática da punição pode ser justificada pelas boas consequências de evitar ações erradas no futuro e, também, faz uso da distinção rawlsiana entre culpado e inocente, a fim de valorizar a liberdade. Diferentemente de Rawls, dirá que a regra da justiça que restringe a punição ao inocente é inteiramente instrumental. Hart igualmente não compartilhará da perspectiva rawlsiana de que as regras de distribuição (regras como práticas), com ênfase na distinção entre culpado e inocente, possam ser defendidas adequadamente com o apelo indireto para a justificação utilitarista da prática como um todo, que é para reduzir o futuro ato errado (regras como sumários). Ver Hart (1959-1960, p. 12).

O que esse exemplo parece mostrar é que existem duas questões diferenciadas a respeito da punição. Uma questão é por que alguém é punido, ao invés de outra pessoa, enquanto a outra questiona a razão de termos a instituição mesma da punição, isto é, por que punimos as pessoas, ao invés de perdoá-las, por exemplo? Veja-se que a resposta do pai nos revela que uma pessoa é punida porque é culpada, e ela é culpada porque descumpriu a lei. Nesse caso, tanto a lei, o juiz, bem como o júri olham para trás e a pena é estabelecida pelo erro que o agente cometeu. Por outro lado, o exemplo também evidencia que a instituição da punição possui um olhar diferenciado, ou seja, ela olha para o futuro, de forma similar ao olhar do legislador (ideal), que deveria procurar ver qual é a eficácia social de uma instituição política. Nesse caso, a instituição da punição teria por função evitar futuros crimes e possibilitar uma sensação de segurança na comunidade política (CR, p. 6).

O ponto de Rawls aqui é procurar integrar em uma mesma concepção normativa um olhar dado tanto pelo legislador como pelo juiz, de forma a melhor responder ao complexo problema da justificação da punição. Em suas palavras:

Se pode dizer, então, que o juiz e o legislador estão em diferentes posições e olham em diferentes direções: um olha o passado e o outro olha o futuro. A justificação do que os juízes fazem, enquanto juízes, parece ter proximidade com a visão retributivista; a justificação do que um legislador (ideal) faz, enquanto legislador, se aproxima da visão utilitarista. Então, ambas concepções têm um ponto importante [...], e a confusão inicial desaparece quando se vê que estas concepções se aplicam a pessoas que possuem diferentes funções e diferentes deveres, e estão situadas diferentemente com respeito ao sistema de regras que constitui o direito penal. (CR, p. 6).

A posição assumida por Rawls é claramente a de conciliar essas concepções vistas tradicionalmente como antagônicas. Ele concorda que a punição deva ser aplicada apenas no caso de haver uma violação da lei, quer dizer, havendo um agente culpado. Mas, sua preocupação recai sobre a instituição da punição como um sistema de regras. E, assim, sua justificação moral seria dada pelo efetivo bem social criado. Importante ressaltar que Rawls está aludindo à punição legal, isto é, dos atos de privação de certos direitos dos agentes que violaram as regras legalmente instituídas, violação reconhecida por um tribunal, através do devido processo legal. Também, que essa privação de direitos é estabelecida por uma autoridade legítima, isto é, o Estado, que

cria o sistema penal que estabelece o que é a ofensa e qual é a penalidade, sendo esses estatutos jurídicos anteriores às ofensas (CR, p. 10).

Essa tentativa de conciliação do retributivismo com o utilitarismo pode acarretar em dois problemas, segundo Rawls. O primeiro é que poderia ser inaceitável para o retributivista, em razão da diferença de opiniões a respeito do critério adequado para a punição, uma vez que vê a punição como uma retribuição de uma ação errada, enquanto o utilitarista não vê propriamente a existência de ações erradas (CR, p. 8-9). O segundo problema é que a posição utilitarista poderia justificar a existência de instituições bem diferentes da descrita acima, instituições essas que fossem cruéis e arbitrárias para a garantia da eficácia social. É exatamente para resolver esse segundo problema que Rawls criará um procedimento hipotético de *telishment*, de sorte a assegurar a culpa do agente como um critério normativo da instituição da punição, em adição à estabilidade social.⁵

3.2 EXPERIMENTO MENTAL DE *TELISHMENT*

A forma que Rawls encontra para solucionar esse segundo problema, formulado especificamente por Carrit, a saber, que o modelo utilitarista poderia justificar em demasia nos casos de punição e, assim, poderia punir uma pessoa inocente ou mesmo punir demasiadamente o culpado, visando à segurança da sociedade, é fazendo uma distinção entre a justificação de uma instituição e a justificação de uma ação particular que recaí sobre essa prática. Além disso, procurará esclarecer as diferentes funções dos agentes em instituir essa prática e decidir os casos particulares que recaem sobre ela. Essa distinção é realizada através de uma experimento mental que tem a função de destacar os nossos juízos morais ponderados para a justificação da punição. Vejamos.

Imagine uma instituição chamada *telishment*. Nessa instituição, os oficiais estatais teriam a autoridade de punir uma pessoa inocente, objetivando

⁵ Rawls faz referência a uma crítica de E. F. Carrit ao utilitarismo, em *Ethical and Political Thinking* (Oxford, 1947), a saber, a de que o utilitarismo poderia justificar a punição de um inocente em razão da expectativa de um bem-estar social. Nas palavras de Carrit: “O utilitarista pode assegurar que estamos justificados em infligir dor sempre e apenas para prevenir uma dor pior ou trazer mais felicidade. Isto, então, é tudo que precisamos considerar na punição, que deve ser puramente preventiva, Mas, se algum tipo de crime muito cruel vier a se tornar comum e nenhum dos criminosos puderem ser apanhados, pode ser uma alternativa, como exemplo, enforcar um homem inocente, se a acusação contra ele poder ser tramada de forma a ele ser considerado culpado de forma universal.” (CARRIT, 1947 apud CR, p. 10).

o melhor interesse da sociedade. E, mais, o poder discricionário dos oficiais seria limitado pela regra de não condenar um inocente, a menos que exista uma onda muito grande de crimes que desestabilizam a sociedade e não se sabe quem é o criminoso. E os oficiais que teriam esse poder discricionário seriam os juízes de alta corte, em consulta com o chefe de polícia, o ministro da justiça e o comitê dos legisladores (CR, p. 11).⁶ A questão é: nós aprovaríamos uma instituição desse tipo, se fôssemos consultados a esse respeito?

A resposta intuitiva é que, se imaginarmos uma instituição que pode punir o inocente, visando ao bem-estar social, ela teria a nossa desaprovação, porque haveria um risco muito alto de que os juízes que punem um inocente (*telishing*), sob condições de segredo, poderiam abusar de seu poder discricionário e a punição ao inocente (*telishment*) invalidaria a legitimidade da punição aos olhos dos cidadãos, já que todos se sentiriam inseguros por essa aleatoriedade permitida. Nas palavras de Rawls:

Uma vez que se identifica o que está envolvido no estabelecimento de uma *instituição*, se vê que os perigos [do *telishment*] são muito grandes. Por exemplo, quem fiscaliza os oficiais? Como se pode dizer se suas ações são autorizadas ou não? Como se pode limitar os riscos envolvidos em permitir tais fraudes sistemáticas? Como evitar dar um poder discricionário total para as autoridades *telish* quem eles quiserem? Além dessas considerações, é óbvio que as pessoas viriam a ter diferentes atitudes em relação ao seu sistema penal quando o *telishment* fosse adicionado a ele. Elas teriam dúvida se um homem condenado estaria sendo punido ou *telished*. (CR, p. 11-12).

A conclusão do raciocínio é que a instituição de *telishment* é intuitivamente errada, pois ela não é coerente com as “nossas opiniões morais ponderadas”. E, assim, as considerações utilitaristas deveriam ser entendidas como aplicadas à instituição da punição em primeiro lugar e não às ações particulares que recaem sobre ela, a menos que essa prática admita essas ações particulares (CR, p. 12-13).

Uma outra forma de ver esse experimento mental é ressaltando os diferentes personagens e funções que estariam envolvidos na criação de uma instituição política tal como a punição legal. Nesse cenário, teríamos,

⁶ *Telishment* é uma palavra cunhada por Rawls, em CR, e significa a prática ou instituição de punir um inocente para garantir a segurança da sociedade. É uma junção da palavra grega *télos*, que significa “finalidade”, com a palavra inglesa *punishment*, que significa punição: tel = *télos* (fim) + ishment = (*pun*)ishment (punição).

por um lado, os legisladores, os quais criariam uma dada instituição como a punição e, por outro, teríamos os juízes, que julgariam as pessoas a partir das regras contidas no sistema penal instituído pelo legislativo. Dessa maneira, o argumento de Rawls poderia ser expresso da seguinte maneira:

(i) A melhor escolha para um legislador seria aprovar uma regra que proíba os juízes de punirem os inocentes, deliberadamente. Essa seria a regra da punição ideal que apenas autoriza a punição aos ofensores. A aprovação da regra da punição ideal teria melhores consequências em longo prazo do que a aprovação da regra que autorizaria a punição do inocente pela boa finalidade social.

(ii) O fato de que a melhor regra para os legisladores escolherem proíba os juízes de punirem um inocente deliberadamente significa que os juízes deveriam sempre seguir essa regra. Quer dizer, um juiz ideal seguiria a regra da punição ideal, mesmo quando, em um caso particular, seriam previstas melhores consequências, no caso de sua violação.

Veja-se que a afirmação (i) é fundamentada com o experimento de *telishment*. Porém, e a afirmação (ii)? Aqui parece que se terá que lidar com a questão de como superar o problema da adoração da regra (*rule worship problem*), que é um problema recorrente para qualquer utilitarismo de regras. Mais claramente, a questão é: se alguém acha que se poderia produzir mais bem, descumprindo a regra, faria sentido para o utilitarista adotar esse ponto de vista anterior no tempo e, assim, um utilitarista diria para o agente violar a regra, mesmo aceitando a regra que foi inicialmente justificada sobre uma adequada base utilitarista. Conforme Rawls, isso é apenas um problema para um utilitarismo de regras que faz uso de uma concepção de regras como sumários e não como práticas, uma vez que o que está sendo proposto é que as pessoas deveriam agir de acordo com regras que promovem o melhor resultado, mesmo que, em um caso específico, a alternativa da maximização do bem-estar exigisse violar essa regra.⁷

3.3 REGRAS COMO PRÁTICAS

Mas qual é a diferença de tomar as regras como práticas, ao invés de interpretá-las como sumários? Seguindo o estilo de Wittgenstein, Rawls con-

⁷ Para mais detalhes do problema da adoração da regra (*rule worship problem*) e a sua relação com a concepção rawlsiana de regras como práticas e não como sumários, ver Boonin (2008, p. 70-75).

cebe as regras como constituintes de uma prática. E uma prática “[...] significa qualquer forma de atividade especificada por um sistema de regras que define os ofícios, papéis, movimentos, penalidade, defesas etc., e que dá à atividade a sua estrutura.” (CR, p. 3). Como exemplos de práticas, temos os jogos, rituais, tribunais e parlamentos. Por outro lado, quando as regras são tomadas como sumários, elas são vistas como relatórios, os quais informam que casos de um certo tipo foram fundados sobre outras bases para ser propriamente decididas de uma certa maneira. Quer dizer, as regras seriam sumários de decisões passadas a que se chega pela aplicação direta de um princípio aos casos particulares. Nessa perspectiva, as decisões tomadas nos casos particulares seriam logicamente anteriores às regras (CR, p. 19-22).

O ponto central de Rawls é defender um tipo específico de utilitarismo, a saber, o utilitarismo de regras, a fim de evitar a objeção de adoração da regra que é comumente feita aos modelos utilitaristas. Para o utilitarismo de regras, o que é requerido é o estabelecimento de uma prática, quer dizer, a especificação de uma nova forma de atividade. Assim, pode-se reconhecer que uma prática envolveria necessariamente abdicar da liberdade integral de agir sobre as bases utilitaristas e prudenciais. Isso parece implicar que aqueles que se engajam em uma prática reconheceriam as regras que a definem (CR, p. 24). Nessa concepção, as regras não são generalizações de decisões individuais, aplicando o princípio utilitarista diretamente aos casos particulares recorrentes. Ao contrário, as regras definem uma prática e seriam elas mesmas o objeto do princípio utilitarista (CR, p. 24).

É importante identificar que as regras seriam logicamente anteriores aos casos particulares. E, assim, não poderia haver um caso particular de uma ação que recaia sobre a regra de uma prática sem a existência da prática mesma. Veja-se um jogo de beisebol, por exemplo. Muitas ações realizadas em um jogo de beisebol podem ser feitas fora do contexto do jogo. Alguém pode bater em uma bola com um bastão, correr, ou mesmo apanhar uma bola com um certo tipo de luva. Todavia, não poderá marcar um ponto, roubar uma base ou rebater, a menos que essas ações sejam efetivadas no contexto da prática do jogo de beisebol, que é constituída por suas regras. O argumento formulado por Rawls parece muito claro: alguém engajado em uma prática não poderia realizar o ato definido pela prática, sem seguir as regras que constituem essa prática.

Vejamus isso aplicado à prática da punição. Sobre a influência da concepção de regras como sumários, seria natural supor que os oficiais de um

sistema penal decidissem o que fazer em casos particulares sobre uma base utilitarista de maximização de bem-estar. A tradicional objeção feita ao utilitarismo é que se atribuiria aos juízes uma autoridade plena para decidirem os casos particulares sobre uma base utilitarista. Nesse caso, o juiz poderia punir o inocente, visando à eficácia social, porque a base utilitarista de raciocínio se sobreporia à regra que proíbe o *telishment*. Mas, uma vez que o utilitarismo é conectado a essa noção de prática, a decisão de um juiz não poderia contrariar a regra da punição ideal, já que ela é logicamente anterior aos casos particulares julgados (CR, p. 29-30).

Com isso, temos que uma sentença de um juiz não seria uma decisão particular tomada caso a caso, em afastamento às próprias regras que constituem o sistema penal. O juiz deveria punir seguindo as regras da punição, porque uma ação que não estiver de acordo com as regras da punição não será punição. Nas palavras de Rawls:

É impossível dizer o que a punição é, ou descrever uma instância particular dela, sem referências às autoridades, ações e ofensas especificadas pelas práticas. A punição é um lance em um elaborado jogo legal e pressupõe a complexidade das práticas que constituem a ordem legal. O mesmo é verdadeiro em tipos menos formais de punição: um pai, responsável, ou alguém com autoridade apropriada pode punir uma criança, mas ninguém mais pode. (CR, p. 42).

Essa concepção de regras como práticas parece resolver o problema de adoração da regra e, assim, possibilitar uma concepção de punição mais atraente, em razão de sua cláusula retributivista, a qual garante a liberdade e a dignidade humana. No entanto, ainda restaria um problema a ser investigado, a saber: essa concepção não implicaria conservadorismo? Parece que não, pois Rawls não diz que as práticas sociais de uma dada sociedade providenciariam um padrão de justificação para as ações dos agentes. O ponto é apenas lógico: quando uma ação é justificada por uma prática, não se pode justificar uma ação particular de um agente sem referência às regras que constituem essas práticas (CR, p. 32).

4 *TJ* E A JUSTIFICAÇÃO DA PUNIÇÃO

Há poucas passagens em *TJ* em que o tema da justificação da punição é discutido, mas a posição defendida por Rawls é muito clara e não é distinta

da já apresentada em CR, a saber, que se alcançaria J1 pela estabilidade social e J2 pela condenação do culpado pelo seu demérito. Podemos localizar essas passagens na seção 38, que discute questão do estado de direito (*The Rule of Law*), na seção 42, que faz observações sobre os sistemas econômicos (*Some Remarks about Economic Systems*), na seção 48, que trata do problema do mérito moral (*Legitimation Expectation and Moral Desert*) e, por fim, na seção 86, que aborda o problema da estabilidade (*The Good of the Sense of Justice*). Vejamos isso detalhadamente.

4.1 PUNIÇÃO COMO RECURSO ESTABILIZADOR E QUE GARANTE A LIBERDADE

Rawls inicia abordando a questão da punição como um recurso estabilizador e que garante as liberdades individuais. Sustenta que é razoável assumir que, mesmo em uma sociedade bem ordenada, o poder coercitivo do governo é necessário em algum grau para a estabilidade social. Com isso, o sistema penal parece relevante para a segurança dos agentes que acordaram seguir as regras. Veja-se que isso está de acordo com uma abordagem hobbesiana, uma vez que agentes, a partir de seu autointeresse, verão a punição como uma forma de garantir o cumprimento do acordo. Aqui estamos aludindo a uma estabilidade exterior, dada pelo medo da punição (*TJ*, p. 240, 211 rev.).

A questão é que as partes, no segundo estágio, isto é, em uma convenção constitucional, poderiam questionar sobre a correção de um sistema punitivo. Por exemplo, as partes poderiam achar que essa instituição seria muito onerosa e, mais, poderiam achar que um sistema de sanções seria uma ameaça a sua liberdade. O ponto de Rawls aqui é importante. Ele destaca que “[o] estabelecimento de uma instituição coercitiva é racional apenas se estas desvantagens forem menores que a perda da liberdade pela instabilidade” e, assim, o melhor arranjo é aquele que pode minimizar esses riscos (*TJ*, p. 241, 211 rev.). De forma similar, a questão das penalidades parece que teria que ser estabelecida no terceiro estágio, ou seja, no estágio legislativo, levando em consideração que qualquer argumento para restringir a liberdade deve proceder do princípio mesmo da liberdade (*TJ*, p. 242, 211-212 rev.).

A ideia aqui é ver o sistema punitivo como necessário, dadas as condições da vida humana, mas que deve ser justificado por um princípio fornecido pela concepção ideal – e esse será o princípio da liberdade. A estratégia rawlsiana parece ser a de conectar um esquema não ideal, a punição legal, a partir da concepção ideal de justiça. Para Rawls:

Fica claro a partir das observações anteriores que precisamos de uma concepção de sanções penais, entretanto, limitada por uma teoria ideal. Dadas as condições normais da vida humana, alguns arranjos são necessários. Eu tenho defendido que os princípios que justificariam essas sanções podem ser derivados do princípio da liberdade. A concepção ideal mostra nesse caso como o esquema não ideal deve ser ajustado; e isto confirma a conjectura que é a teoria ideal que é fundamental. (*TJ*, p. 241, 212 rev.).

Com isso, a instituição da punição (i) deve prevenir futuros crimes, tendo um papel estabilizador, mas (ii) deve garantir as liberdades dos agentes, o que integra um elemento consequencialista de punição com um elemento retributivista. Por sua vez, qual é concepção de responsabilidade moral de Rawls? O agente seria tomado como responsável em razão de seu livre-arbítrio, isto é, como tendo uma capacidade de escolher livremente e, assim, quando erra, ele teria uma possibilidade alternativa de não errar? Rawls estaria fazendo uso de uma concepção indeterminista de responsabilidade moral, como aparenta ser para o retributivismo? Não parece ser o caso, visto que ele ressaltará a importância de os cidadãos terem a habilidade de conhecer o que a lei é e terem oportunidade equitativa para tomar a direção que a lei aponta para se considerar adequada a aplicação das sanções penais (*TJ*, p. 241, 212 rev.). O princípio da liberdade é, então, apenas a consequência de tomar um sistema legal como uma ordem de regras públicas endereçadas a agentes racionais, de forma a regular sua cooperação, dando o apropriado peso à liberdade.

No entanto, que concepção de responsabilidade moral Rawls estaria usando? Creio que seja uma concepção compatibilista de responsabilidade, a qual vê o agente como tendo capacidade de escolha a partir dos diversos condicionamentos que vão muito além de sua vontade. Essa capacidade de escolha parece estar mais ligada aos elementos cognitivos da responsabilidade de saber o que é o certo e errado e menos à capacidade do agente em agir diferentemente, isto é, no livre-arbítrio ou PAP.⁸ Isso parece ser confirmado

⁸ O princípio das possibilidades alternativas – PAP – diz que uma pessoa tem responsabilidade moral pelo que fez apenas se ela tivesse podido ter agido de outra forma. Frankfurt argumenta que esse princípio é falso, pois mesmo não podendo agir diferentemente, o agente seria moralmente responsável, se os seus desejos estiverem em sintonia com uma presumida coerção. Vejamos um exemplo desse tipo. Black quer que Jones mate o prefeito, todavia, tem dúvidas se Jones fará a ação. Então, implanta um dispositivo cerebral em Jones, possibilitando que as ações de Jones sejam programadas e operadas por um computador. Se, na hora marcada, Jones não matar o prefeito, Black acionaria o dispositivo e passaria a controlar as sinapses cerebrais de Jones. A questão é que Jones quer matar o prefeito. E, na

pela referência problematizadora que Rawls faz a *mens rea*, assim como ela é compreendida no direito penal. Rawls assinala:

Acredito que essa concepção de responsabilidade nos permite explicar a maioria das desculpas e defesas reconhecidas pelo direito penal sob o nome de *mens rea* e que isso pode servir de guia para uma reforma legal. Entretanto, essas questões não podem ser perseguidas aqui. Mas, é suficiente notar que a teoria ideal requer uma concepção de sanções penais como um mecanismo estabilizador [...]. (TJ, p. 241, 212 rev).⁹

Rawls, ao concordar com Hart nesse ponto, parece estar dizendo que deveríamos reformar o sistema penal que toma por base o princípio do “dever” implica “poder” e, assim, a pessoa que não poderia ter deixado de fazer o que fez não seria moralmente culpada e, logo, não poderia ser punida, devendo passar a ter por foco a proteção da sociedade. Isso parece nos apontar para uma concepção de sistema punitivo mais eficiente e justo, em razão de estar baseado em uma compreensão de responsabilidade moral e legal compatibilista, a qual não tem por centro a capacidade de agir diferentemente. Mas ele não desenvolve essa teoria aqui e em nenhum outro lugar de sua teoria da justiça como equidade. Contudo, já é suficiente para notarmos que a teoria ideal requer uma concepção de punição tomada como um instrumento estabilizador que deve respeitar o princípio da liberdade e da responsabilidade.

4.2 PUNIÇÃO E CONFIANÇA

A questão central na seção 42, que faz algumas observações sobre os sistemas econômicos, é identificar que a punição serviria para garantir a

hora marcada, Jones mata o prefeito e Black não precisou acionar o dispositivo. O que isso mostra? Que Jones não poderia ter agido de outro modo, mas tem responsabilidade moral por sua decisão, por seus desejos. Ver Frankfurt (1969, p. 834-836).

⁹ Aqui Rawls faz referência ao texto de Hart, *Punishment and Responsibility*, (2008, p. 173-183), afirmando que seguirá Hart sobre a questão da necessidade da reforma legal. O ponto central de Hart é que a punição deve ter por foco a proteção da sociedade antes que estar preocupada com questões metafísicas sobre a responsabilidade que, em geral, recaem nas questões de determinismo e livre-arbítrio. O problema é o que o princípio fundamental da moralidade diz que uma pessoa não poderia ser censurada e punida pelo o que fez, se ela não pudesse ter deixado de fazer o que fez. Esse é o princípio do “dever” implica “poder”. Entretanto, como determinar o que é essa capacidade de agir de outra forma (livre-arbítrio)? Uma observação importante de Hart é que o sistema penal deveria se concentrar mais nos elementos cognitivos da responsabilidade e menos na voluntariedade do ato. Também, que o princípio da responsabilidade recairia sobre a capacidade e oportunidade equitativa do agente para ajustar o seu comportamento à lei. Ver Hart (2008, p. 158-185).

confiança dos agentes que teriam acordado seguir determinadas regras comuns para o seu convívio. Aqui Rawls faz o contraste entre o problema do isolamento e o problema da confiança. O primeiro tem relação com a identificação de certas situações em que a escolha racional isolada dos indivíduos traria um pior resultado, defendendo que a escolha de um ponto de vista coletivo seria melhor. Isso é facilmente explicado com o caso do dilema dos prisioneiros. Por outro lado, o problema da confiança é diferente. A questão aqui é a de como assegurar para as partes cooperantes que o acordo comum será seguido.

Veja-se que, no modelo contratualista defendido por Rawls, as partes contratantes especificam as regras para o convívio público e precisam ter uma disposição para seguir essas regras. Essa é a primeira característica da razoabilidade, isto é, é a capacidade de senso de justiça. Mas isso pressupõe uma situação contingente de todos seguirem essas regras. Quer dizer, a disposição de um agente em seguir as regras acordadas depende de sua confiança de que os outros agentes também seguirão essas regras. Entretanto, pode ser o caso de nem todos agirem a partir de sua capacidade de senso de justiça. E, assim, para garantir a confiança pública nesse esquema como superior a qualquer outro, “[...] algum mecanismo para administrar multas e penalidades deve ser estabelecido” (TJ, p. 270, 238 rev.).¹⁰

Com isso, podemos facilmente identificar que aqui também o ponto essencial da punição é ser tomada como um recurso estabilizador externo, não acrescentando muita coisa ao que já foi exposto na seção 38. Por outro lado, parece reforçar a importante constatação de que a instituição da punição legal recai em uma parte não ideal da teoria da justiça, o que parece implicar tomar a punição como uma instituição contingente, a qual deve ser justificada à luz da razão pública e não por uma teoria moral abrangente.¹¹

¹⁰ O ponto crucial, por conseguinte, não seria o estabelecimento da instituição da punição *per se*, entendida na forma de uma instituição que necessariamente deveria causar sofrimento intencional ao ofensor, para a garantia da confiança. Como expresso por Rawls, para garantir a confiança, deve-se estabelecer algum sistema de multas e penalidades. A definição de quais multas e que penalidades e para que casos específicos seria feita apenas no terceiro estágio, isto é, no estágio legislativo. E isso poderia estar de acordo com certas práticas coercitivas que não ocasionariam o sofrimento intencional ao agente, como no caso das multas, por exemplo. E isso parece responder à objeção feita por Hanna, ao identificar um raciocínio deficiente para a justificação da punição na concepção rawlsiana. Para ele, mesmo que a coerção fosse necessária para a garantia da confiança, a punição em si mesma não o seria, uma vez que há formas não punitivas de coerção. Ver Hanna (2009, p. 332).

¹¹ Esse é um ponto importante na concepção rawlsiana de punição: ela não é justificada com o uso de uma teoria moral abrangente, tal como o utilitarismo ou retributivismo, ou mesmo uma teoria mista, mas deve ser justificada pelos cidadãos que vivem em sociedades marcadas pelo pluralismo moral razoável, através de sua razão pública. Melissaris faz uma observação importante a esse respeito. Ele

4.3 PUNIÇÃO E MÉRITO MORAL

Vimos até agora a forma que Rawls responde à questão de como nós podemos justificar moralmente a instituição da punição, isto é, pelo seu efeito estabilizador social. Porém, o que fariam os atos particulares punitivos justificados? Não poderia ser apenas a tentativa de prevenção de futuros crimes, uma vez que isso poderia significar condenar um inocente, ou mesmo, condenando um culpado, tratar o indivíduo como um mero meio e não como um portador de direitos. Na seção 48, Rawls responderá que é o (de)mérito do agente o que justifica esses atos punitivos, quer dizer, os atos punitivos podem ser normativamente justificados pela culpa do agente em ter agido erroneamente.

O ponto de Rawls é mostrar que, em uma sociedade razoavelmente bem ordenada, aqueles que são punidos por violarem leis justas normalmente fizeram algo errado, ou seja, geralmente cometeram um crime. É isso é assim, porque o propósito do direito penal é assegurar os deveres naturais básicos, aqueles que nos proíbem de causar danos às outras pessoas, tais como tirar a vida, a liberdade ou a propriedade dos outros. A punição, então, serve para esse fim. Além disso, Rawls acrescentará que a propensão de cometer tais atos é uma marca do mau caráter do agente e que, em uma sociedade justa, a punição deve recair apenas sobre esses agentes que cometem essas faltas. Nas palavras de Rawls:

É verdade que em uma sociedade razoavelmente bem ordenada aqueles que são punidos por violarem leis justas normalmente fizeram algo errado. Isto é porque o propósito do direito penal é assegurar os deveres naturais [...]. Ele não é simplesmente um esquema de taxas e ônus destinado a colocar um preço em certas formas de conduta e, dessa forma, guiar a conduta dos homens para a vantagem mútua [...]. Então, a propensão para cometer tais atos é uma marca de mau caráter, e em uma sociedade justa a punição legal apenas recairá sobre aqueles que desempenham essas faltas. (TJ, p. 314-315, 276-277 rev.).¹²

sublinha que a punição, na concepção de Rawls, deve ser vista como uma instituição contingente que proporciona uma resposta apropriada às violações dos deveres políticos sob a base da responsabilidade dos ofensores, e que não faz uso de uma fundamentação moral *a priori* ao problema. Ver Melissaris (2012, p. 124).

¹² Importante ressaltar que os “direitos naturais” aqui não são vistos como critérios normativos *a priori*. Ao contrário, os direitos só surgem em razão dos deveres que são assumidos no modelo contratualista. Quer dizer, são os deveres estabelecidos na posição original o que garante os direitos dos agentes, sendo um ato errado apenas uma extensão de uma violação de um dever político estabelecido pelos termos do contrato. Creio que uma maneira bastante esclarecedora para entender a relação entre direitos e deveres, na justiça como equidade, é compreender os deveres como estabelecidos em uma perspectiva de segunda

O problema aqui é tentar explicar por que Rawls não faz uso do critério do mérito moral para as questões de justiça distributiva, mas utiliza esse critério para as questões de justiça retributiva. Importante ressaltar que é exatamente por essa aparente incoerência que Rawls é criticado, em sua concepção de punição, por vários intérpretes.¹³ Creio que o ponto central seja identificar que as questões de como nós devemos distribuir os bens primários na sociedade recai sobre o esquema ideal da teoria da justiça como equidade, enquanto a questão de quem punir e o quanto punir recai sobre o esquema não ideal da teoria, isto é, recai sobre uma perspectiva de funcionamento de uma sociedade real.

Vejamos a diferença. Se imaginarmos pessoas hipotéticas que estariam em uma situação de simetria para escolher os princípios de justiça, a fim de reger as principais instituições econômicas, políticas e sociais, é razoável pressupor que elas recusariam a ideia de tomar os bens primários como relacionados ao mérito dos indivíduos, isto é, como tendo relação com o valor moral dos agentes, vendo-se todos como igualmente merecedores desses bens sociais. Aqui temos um ponto de vista coletivo para a escolha, uma vez que as partes não saberiam sobre suas concepções profundas do bem. E, assim, todos mereceriam igualmente os bens primários, independentemente de seu caráter moral (*TJ*, p. 315, 277 rev.).

Porém, por que, na justiça retributiva, Rawls não faz uso do mesmo raciocínio? A questão parece ser a de que o crime e a punição aconteceriam em uma situação real, pois, em uma sociedade bem ordenada, em princípio, as pessoas não teriam razões para cometer crimes, porque as suas instituições seriam justas e as pessoas teriam senso de justiça exatamente por terem sido criadas em uma sociedade com instituições justas, isto é, regradas pelos

peessoa, isto é, como deveres que se devem aos outros que podem exigir de nós o seu cumprimento. O próprio Darwall parece confirmar essa aproximação. Em suas palavras: “Substancialmente o mesmo ponto pode ser colocado em termos rawlsianos, dizendo que é a perspectiva de segunda pessoa que nos coloca no espaço de ‘razão pública’ e do ‘razoável’ preferencialmente ao (meramente) individualmente ‘racional’.” (DARWALL, 2006, p. 23).

¹³ Essa é a crítica feita por Honing, em seu artigo “Rawls on Politics and Punishment”. Seu ponto central é demonstrar que a teoria de Rawls busca a estabilidade social com uma estratégia de reconciliação que progressivamente aproximaria os agentes e estreitaria seus laços, na esfera política. Mas, ao tratar da punição, ele abandona esses pressupostos coletivos da justiça distributiva, reintroduzindo um discurso de mérito moral na justiça retributiva, de sorte a demonizar os criminosos como portadores de um “mau caráter”. E isso revelaria uma inconsistência na teoria rawlsiana. Ver Honing (1993, p. 110-120). Ainda sobre essa pretensa incoerência entre a justiça distributiva e a retributiva em Rawls, ver Sandel (1998, p. 89-90).

princípios de igual liberdade, igualdade equitativa de oportunidade e pelo princípio da diferença. Veja-se que o senso de justiça, conforme Rawls, nada mais é do que a disposição para aceitar os princípios de justiça que é possibilitada pela sensibilidade moral compreendida pelos sentimentos morais de culpa, ressentimento e indignação, em conexão com as atitudes de confiança mútua e afeição que seriam adquiridos no convívio social, quer dizer, na família e na escola, por exemplo.¹⁴ E, assim, em cometendo crimes, essas pessoas demonstrariam uma marca de um caráter vicioso, sendo essa marca não uma questão coletiva, mas individual. A questão parece ser a de que, em uma sociedade justa, o crime seria a expressão de um traço comportamental não desejável do agente, exatamente por colocar em risco a segurança dos cidadãos. Isso seria censurável e passível de punição, explicando por que a justiça distributiva e retributiva não seriam naturalmente intercambiáveis (*TJ*, p. 315, 277 rev.).¹⁵

4.4 PUNIÇÃO E ESTABILIDADE

Por fim, na seção 86 de *TJ*, Rawls aborda um problema central em sua teoria, que é o problema da estabilidade. A questão é a de saber se cidadãos reais fariam uso dos princípios de justiça em sua vida pública, uma vez que eles poderiam agir oportunisticamente, isto é, acordando sobre as regras e não as cumprindo. A resposta que ele dá em *TJ* é a da congruência entre o justo e o bem, resposta que será posteriormente alterada para a do consenso sobreposto em *PL*. Contudo, é interessante notar que não haverá uma diferença no tratamento dado ao papel da punição, pois, em ambas as respostas, a instituição da punição se justificaria pelo seu papel de garantir a estabilidade social. Vejamos isso detalhadamente.

Mesmo considerando uma congruência do justo e do bem, é possível que certas pessoas ajam erroneamente, isto é, ajam de forma a manifestar que seu senso de justiça não seria um bem para ela e, assim, a punição legal

¹⁴ No artigo "Sense of Justice", Rawls afirma que o senso de justiça não apenas está conectado com os sentimentos morais de ressentimento e indignação, mas também com as atitudes naturais de confiança mútua e afeição. Ver Rawls (1968, p. 281-282). Para a distinção dos três tipos de sentimento de culpa (*authority, association, principle*) e o seu processo de aquisição, ver Rawls (1968, p. 286-293).

¹⁵ Scheffer usa um argumento semelhante, para responder ao problema da assimetria entre justiça distributiva e retributiva. Ele enfatiza que o mérito moral é tomado por seu caráter holístico na justiça distributiva, porém, é tomado por seu caráter individualístico na justiça retributiva e, assim, o mérito seria uma condição necessária mas não suficiente para a punição. Ver Scheffer (2000, p. 983-987).

teria um papel estabilizador. O ponto é que, mesmo em uma sociedade bem ordenada, poderiam existir pessoas que não agiriam a partir de um senso de justiça, embora todas fossem portadoras dessa capacidade, e isso parece implicar que a justiça não seria uma virtude para esses agentes.¹⁶ Considerando essa possibilidade, como se garantirá a estabilidade social? A resposta tem um certo tom de lamento, mas é muito clara: estabelecendo um sistema penal que punirá aqueles que agirem erroneamente, pois esse problema “pressupõe uma teoria da punição.” (*TJ*, p. 575, 504 rev.).

Concordando com as sanções penais para estabilizar o esquema cooperativo, as partes aceitariam o mesmo tipo de restrição ao autointeresse, o qual elas já aceitaram, ao escolher os princípios de justiça, inicialmente. Quer dizer, tendo acordado com os princípios de justiça como razões para reger as principais instituições políticas, sociais e econômicas, é racional autorizar medidas necessárias para manter as instituições justas, assumindo que as restrições de igual liberdade e estado de direito sejam reconhecidas. Nas palavras de Rawls:

Além do mais, os princípios de direito e justiça são coletivamente racionais; e é no interesse de cada um que todos devem cumprir com os acordos justos. Este também é o caso de que a afirmação geral do senso de justiça é um grande bem social, estabelecendo as bases para a confiança mútua da qual todos são beneficiários. Então, acordando sobre as penalidades que estabilizam um esquema de cooperação, as partes aceitam o mesmo tipo de restrição ao autointeresse que elas reconheceram ao escolher os princípios de justiça inicialmente. Tendo consentido a estes princípios vistos como razões, é racional autorizar medidas necessárias para manter as instituições justas [...]. (*TJ*, p. 576, 504 rev.).

O ponto é que, para justificar uma concepção de justiça, não é necessário que todos ajam a partir de sua capacidade moral, isto é, de seu senso de justiça. E, tendo essa consideração em mente, instrumentos penais desempenharão um importante papel no sistema social (*TJ*, p. 576-577, 504-

¹⁶ É importante observar que, de forma similar a Chomsky, Rawls faz uma distinção entre competência e performance, através da analogia entre o senso de gramaticalidade e o senso de justiça, estabelecendo uma diferença entre a competência e a performance moral propriamente dita na justiça como equidade. Assim, a competência moral ou senso de justiça estaria associada ao conhecimento moral individual, sendo as condições mentais do agente para ação, enquanto a performance estaria associada a como esse conhecimento moral é utilizado, isto é, estaria associada ao comportamento moral dos indivíduos. Isso parece significar que ter capacidade moral não significa agir moralmente de forma necessária, o que parece justificar a necessidade da instituição da punição e as sanções legais. Ver *TJ*, p. 41 rev.

505 rev.). No entanto, veja-se que a ideia básica é a de que agentes criados sob instituições justas adquiririam um senso de justiça num certo grau que seria suficiente para manter estáveis as instituições. Mesmo em *JF* e *PL*, ao tratar do problema da estabilidade, Rawls assinala que, dadas certas assunções de uma psicologia humana razoável e condições normais de vida, aqueles que crescem sob instituições básicas justas adquiririam um senso de justiça e uma lealdade a essas instituições, que seriam normalmente suficientes para mantê-las estáveis (*JF*, p. 185; *PL*, p. 141-143).¹⁷

Embora *JF* e *PL* não façam menção à instituição da punição, não parece muito problemático afirmar que o sistema penal teria a mesma função estabilizadora que foi ressaltada em *TJ*, uma vez que, inclusive, não se contará mais com uma pressuposição de congruência entre o justo e o bem, contando-se apenas com a expectativa de consenso sobreposto para a garantia da estabilidade, o que parece resultar em um certo deflacionamento moral com relação ao senso de justiça. E isso parece nos apontar que o tipo de estabilidade requerida na justiça como equidade é tanto interna como externa, ou seja, tanto tem por base a sensibilidade moral dos agentes em assumirem seus deveres quanto usa o sistema penal para punir aos que descumprirem com o que foi livremente consentido. No entanto, é importante frisar que a justificação de uma instituição política contingente como a da punição deve ser pública, quer dizer, deve ser estabelecida nos termos de uma razão pública.

5 CONCLUSÃO: FUNÇÃO CORRETIVA DA PUNIÇÃO

Após a análise realizada em *CR* e *TJ*, podemos perceber uma vantagem inicial dessa concepção híbrida de punição, a saber, que ela parece incluir os acertos dos modelos retributivista e consequencialista, visto que tanto garantirá os direitos individuais dos agentes como terá uma preocupação apropriada com a estabilidade social. Mas, diferentemente do retributivismo, não exigirá o sofrimento do ofensor em razão de não fundamentar a responsabilidade moral do agente no livre-arbítrio. Também, diferentemente do utilitarismo,

¹⁷ Rawls escreve, em *JF*, que “[a] estabilidade é assegurada por uma motivação suficiente de tipo apropriado adquirida sob instituições justas” (*JF*, p. 185). Em *PL*, Rawls acrescentará: “Agora todos reconhecem que a forma institucional de sociedade afeta seus membros e determina em grande parte o tipo de pessoa que elas querem ser, bem como o tipo de pessoa que elas são.” (*PL*, p. 269). E, em *LP*, salientará: “A estabilidade pelas razões corretas significa uma estabilidade causada pelos cidadãos em agindo corretamente de acordo com os princípios apropriados de seu senso de justiça, princípios que eles adquiriram por terem crescidos sob instituições justas e participado dessas instituições.” (*LP*, p. 13).

não cairá no problema de adoração da regra. Uma vantagem adicional dessa posição é que os critérios normativos de “estabilidade social” e “liberdade”, por exemplo, não serão considerados como valores prescritivos de uma teoria moral abrangente, devendo ser justificados com o uso público da razão.¹⁸

Isso parece revelar uma teoria liberal de fundo, nesse modelo híbrido, a qual mostra uma proximidade com a teoria expressivista da punição, de forma que medidas punitivas serviriam para comunicar ao ofensor a censura que ele merece por seus crimes, sendo uma expressão de desaprovação pública ao ato ilícito cometido. A punição diria a todos que o ofensor não tem o direito de fazer o que fez. Veja-se que há uma dimensão pública importante aqui, uma vez que o crime é um erro público, isto é, aquilo que a comunidade política vê como um erro moral e que deve ser prevenido. Com isso, tanto é necessária a esfera legislativa como a esfera judiciária para a especificação desse erro, distinguindo claramente as concepções privadas de bem com a esfera dos direitos dos cidadãos, nos marcos de uma sociedade caracterizada pelo pluralismo moral razoável.¹⁹

Essa distinção entre os valores morais privados dos valores morais públicos é uma importante marca liberal da concepção rawlsiana. E isso parece trazer por consequência um tipo de neutralidade ética do Estado, já que não seria adequado dizer qual é o bem a ser protegido e que tipo de valores morais privados as pessoas deveriam professar. Quer dizer, parece haver uma recusa ao paternalismo e ao moralismo jurídico, o que se demonstra coerente com o respeito à autonomia dos cidadãos e a busca pela estabilidade social.

Todavia, uma limitação na forma de justificar a punição é que essa concepção rawlsiana não parece apontar para a função de correção do caráter do agente, a qual é claramente percebida no contexto interpessoal da punição, como na punição dos pais dadas a um filho que comete um erro. Uma explicação que se poderia dar é que, como esse modelo híbrido quer evitar o paternalismo e defender o liberalismo, uma instituição política não deveria legislar sobre a esfera íntima de seus membros, recomendando quais valores

¹⁸ Creio que isso já responda por que seria infundada a preocupação de alguns intérpretes, por exemplo, Thomas Pogge, de que a justiça como equidade regressaria a um modelo consequencialista, em razão de as condições de realização institucional serem invisíveis na posição original. A explicação para tal é que as razões utilitaristas identificadas na justiça como equidade estariam conectadas com as exigências de publicidade e de responsabilidade. A esse respeito, ver Melissaris (2012, p. 142).

¹⁹ Feinberg afirma que a punição “[...] é um instrumento convencional para expressar atitudes de ressentimento e indignação, e juízos de desaprovação e reprovação [...]”, tendo um forte significado simbólico. Ver Feinberg (1970, p. 98).

morais deveriam ser perseguidos. Nesse sentido, a especificidade “pública” na punição legal poderia ser uma razão suficiente para explicar a ausência desse aspecto corretivo. Mas, não penso ser esse o caso, uma vez que Estados liberais, na concepção de Rawls, determinam, para além dos direitos, quais são os deveres públicos de seus membros, por exemplo, o dever de ser civilizado ou razoável, o dever de cidadania, bem como o dever de imparcialidade, conforme é exigido para agentes públicos, por exemplo, a um juiz.²⁰

Como Estados liberais exigem o respeito aos deveres públicos, não haveria um motivo específico para eles não assumirem o compromisso, tanto com a formação do caráter de seus membros em relação a essas normas políticas quanto com a correção do caráter, quando do caso de alguns de seus cidadãos não cumprirem com os deveres que foram aceitos por todos. Asseverar que as instituições teriam um importante papel na formação e correção moral-política de seus membros não parece ser contraditório com as premissas liberais que exigem que as instituições políticas e jurídicas sejam justificadas politicamente, isto é, com neutralidade ética, e defendam a liberdade e a igualdade de seus membros.²¹

Creio que, se aplicarmos o mesmo método empregado por Rawls, poderíamos superar esse limite apontado acima. Imaginemos uma instituição chamada *ortishment*, que é uma instituição de punição cujo objetivo é evitar futuros crimes, com a condição de só punir o culpado e pretendendo a correção do caráter do ofensor. Nessa instituição, os oficiais estatais teriam a autoridade de somente punir o culpado por um crime, com a expectativa de que os ofensores reconhecessem seu erro, se arrependessem do ato cometido que

²⁰ O papel da razão pública, para Rawls, é possibilitar um forte compromisso de todos os envolvidos com os ideais e valores morais e políticos de uma sociedade democrática, tomando como ponto de partida (i) o critério de reciprocidade (*reciprocity*), que exige o (ii) dever de civilidade (*duty of civility*), o que implica a defesa da virtude da (iii) amizade cívica (*civic friendship*) e de um (iv) ideal de cidadania democrática (*democratic citizenship*), o qual toma por base a (v) lei legítima (*legitimate law*), o que significa defender os (vi) princípios de tolerância e liberdade de consciência (*principles of toleration and liberty of conscience*). Ver PR, p. 588-591.

²¹ No texto “Two-Fold Basis of Justice”, Rawls compreende que o aprendizado moral se dá pelo ensinamento de uma percepção. Essa percepção moral seria desenvolvida quando os agentes se arrependem do erro cometido e se desculpam diante dos outros pela injúria causada. E essas desculpas têm o papel de restabelecer a conexão entre as pessoas, a qual seria rompida, quando não se leva em consideração os sentimentos e aspirações dos outros. Assim, a responsabilidade do agente estaria fundada exatamente nesse reconhecimento das pessoas como entidades legítimas de concernimento, e isso parece consistente com uma concepção de punição que teria uma função formativa, e não parece contraditório com uma posição liberal, uma vez que a autonomia do agente é respeitada. Ver TB, p. 3-4.

prejudica toda a sociedade e, por fim, modificassem o seu caráter. A cláusula seria que os únicos atos passíveis de *ortishment* seriam os erros tomados como públicos. Ademais, que os oficiais que teriam esse poder de determinar qual a pena que melhor alcançaria esse objetivo corretivo seriam os juízes, em conjunto com um certo tipo de representação dos cidadãos. Se consultados, será que nós aprovaríamos uma instituição desse tipo?

A resposta intuitiva parece ser positiva, uma vez que ela se mostra coerente com os nossos juízos morais ponderados, os quais valorizam, para além da estabilidade social e liberdade individual, o bem-estar dos membros da comunidade política da qual fazemos parte. Veja-se que essa expectativa no progresso moral não parece ameaçar a autonomia do agente, visto que não implicaria necessariamente coação, sendo apenas um apelo à consciência moral da pessoa, da mesma forma que parece não ameaçar a sua privacidade, pois estaria restrita somente aos deveres públicos estipulados em um ponto de vista de segunda pessoa. A vantagem desse acréscimo corretivo é que ele parece revelar uma clara perspectiva de reconciliação entre os membros de uma comunidade política, perspectiva essa que se mostra fortemente harmonizada com uma teoria liberal e coerentista da punição.

COITINHO, Denis. Rawls and the justification of punishment. *Transformação*, Marília, v. 40, n. 3, p. 67-92, Jul./Set., 2017.

Abstract: The principal aim of this paper is to understand the hybrid normative concept used by John Rawls to morally justify legal punishment. Rawls uses a consequentialist argument – that is, social efficacy – to justify the institution of punishment, and a retributivist argument – that is, the guilt of the agent – to justify particular acts of punishment. The article “Two Concepts of Rules” and the book *A Theory of Justice* are analyzed. We argue that Rawls uses a liberal view of punishment, as well as the method of reflective equilibrium, stipulating a rule for approval of legal punishment based on its coherence with our considered moral judgments; this seems to reconcile retributive justice with distributive justice.

Keywords: Punishment. Moral responsibility. Reflective equilibrium. John Rawls.

REFERÊNCIAS

BOONIN, D. *The problem of punishment*. Cambridge: Cambridge University Press, 2008.

BROOKS, T. *Punishment*. London: Routledge, 2012.

DARWALL, S. *The second-person standpoint: morality, respect, and accountability*. Cambridge, MA: Harvard University Press, 2006.

FEINBERG, J. *Doing and deserving: essays in the theory of responsibility*. Princeton, N.J: Princeton University Press, 1970.

FRANKFURT, H. Alternate possibilities and moral responsibility. *The Journal of Philosophy*, v. 66, n. 23, p. 829-839, 1969.

HANNA, N. Liberalism and the general justifiability of punishment. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, v. 145, n. 3, p. 325-349, 2009.

HART, H. L. A. Prolegomenon to the principles of punishment. *Proceedings of the Aristotelian Society, New Series*, v. 60, p. 1-26, 1959-1960.

_____. *Punishment and responsibility: essays in the philosophy of law*. 2. ed. Oxford: Oxford University Press, 2008.

HONIG, B. Rawls on political and punishment. *Political Research Quarterly*, v. 46, n. 1, p. 99-125, 1993.

MELISSARIS, E. Toward a political theory of criminal law: a critical Rawlsian account. *New Criminal Law Review: An International and Interdisciplinary Journal*, v. 15, n. 1, p. 122-155, 2012.

RAWLS, J. Two-fold basis of justice. *The papers of John Rawls*, Harvard University Archives, HUM 48 Box 9 Folder 1, 1950s.

_____. Punishment. *The papers of John Rawls*. Harvard University Archives, HUM 48 Box 7 Folder 2, 1952-1953.

_____. Two concepts of rules. *Philosophical Review*, v. 64, n. 1, p. 3-32, 1955.

_____. The Sense of justice. *Philosophical Review*, v. 72, n. 3, p. 281-305, 1963.

_____. *A Theory of justice*. Original edition. Cambridge, MA: Harvard University Press, 1971.

_____. *Political liberalism*. New York: Columbia University Press, 1993.

_____. *The Law of peoples*. Cambridge, MA: Harvard University Press, 1999a.

_____. *A Theory of justice*. Revised edition. Cambridge, MA: Harvard University Press, 1999b.

_____. *Justice as fairness: a restatement*. Edited by Erin Kelly. Cambridge, MA: Harvard University Press, 2001.

_____. The idea of public reason revisited. In: S. FREEMAN (Ed.). *Collected papers*. Cambridge, MA: Harvard University Press, 1999, p. 573-615.

COITINHO, D.

SANDEL, M. *Liberalism and the limits of justice*. 2. ed. Cambridge: Cambridge University Press, 1998.

SCHEFFLER, S. Justice and desert in liberal theory. *California Law Review*, v. 88, n. 3, p. 965-990, 2000.

Recebido em 26/02/2016

Accito em 18/09/2016