

A TEORIA DA JUSTIÇA DE JOHN RAWLS: VIABILIDADE, ESTABILIDADE E APLICABILIDADE*

JOHN RAWLS' THEORY OF JUSTICE: FEASIBILITY, STABILITY, AND APPLICABILITY

Ulysses Ferraz de Camargo Filho

<https://orcid.org/0009-0004-9723-6518>

ferraz.ulysses@gmail.com

*Instituto de Estudos Sociais e Políticos (IESP-UERJ),
Rio de Janeiro, Brasil.*

RESUMO *O artigo analisa as condições de viabilidade, estabilidade e aplicabilidade da justiça como equidade de John Rawls. Com base em Pablo Gilabert, defende que a viabilidade deve ser entendida de modo dinâmico, incorporando a possibilidade de transformação institucional e mobilização coletiva, sem reduzir a teoria normativa às limitações práticas imediatas. Em seguida, examina a estabilidade como critério de validade normativa, reconstruindo a transição de Rawls de uma abordagem fundada na congruência entre justo e bem para o modelo do consenso sobreposto em contextos de pluralismo razoável. A estabilidade é interpretada como sustentação da justiça “pelas razões certas”, ou seja, pelo reconhecimento de sua legitimidade por cidadãos livres e iguais. Por fim, discute-se a aplicabilidade da teoria ideal, com destaque para o papel do princípio de diferença como critério regulativo tanto*

* Artigo submetido em: 26/08/2025. Aprovado em: 27/07/2026.

em sociedades bem-ordenadas quanto em contextos de injustiça. A crítica de Charles Mills à inaplicabilidade da teoria rawlsiana em sociedades racialmente estruturadas é contraposta pela defesa de Samuel Freeman, que interpreta a teoria ideal como um recurso metodológico indispensável para avaliar injustiças e orientar reformas. O artigo conclui que a justiça como equidade pode ser concebida como viável, estável e aplicável, oferecendo diretrizes normativas mesmo em sociedades marcadas por desigualdades estruturais.

Palavras-chave: *Aplicabilidade. Estabilidade. John Rawls. Justiça como equidade. Princípio de diferença. Viabilidade.*

ABSTRACT *The article examines the conditions of feasibility, stability, and applicability of John Rawls' justice as fairness. Drawing on Pablo Gilabert, it argues that feasibility should be conceived dynamically, incorporating the potential for institutional transformation and collective mobilization, without reducing normative theory to immediate practical constraints. It then analyzes stability as a criterion of normative validity, reconstructing Rawls' shift from an approach grounded in the congruence between the right and the good to the model of overlapping consensus in contexts of reasonable pluralism. Stability is interpreted as the support of justice "for the right reasons," namely, through the recognition of its legitimacy by free and equal citizens. Finally, the article discusses the applicability of ideal theory, emphasizing the role of the difference principle as a regulative standard both in well-ordered societies and in contexts of injustice. Charles Mills's critique of Rawlsian theory as inapplicable to racially structured societies is contrasted with Samuel Freeman's defense, which interprets ideal theory as a methodological device indispensable for evaluating injustices and guiding reforms. The article concludes that justice as fairness can be conceived as feasible, stable, and applicable, providing normative guidance even in societies marked by structural inequalities.*

Keywords: *Applicability. Difference principle. Feasibility. John Rawls. Justice as fairness. Stability.*

1 Introdução: os problemas de viabilidade, estabilidade e aplicabilidade

De acordo com Pablo Gilabert (2023), a expressão "...é viável" é frequentemente utilizada quando se consideram processos políticos que envolvem agentes individuais ou coletivos que procuram produzir determinados resultados ou estados de coisas em determinadas circunstâncias. Assim,

recorremos à noção de viabilidade para abordar se, e em que medida, os agentes, em dadas circunstâncias, têm a capacidade ou o poder de produzir os resultados que procuram. Para Gilabert, a viabilidade é um conceito relacional de poder ou capacidade que conecta três elementos básicos: um agente, certos resultados e certas circunstâncias. A estabilidade de uma concepção de justiça e de uma sociedade bem-ordenada que a incorpora parece ser de grande importância para sua viabilidade (Freeman, 2007a). Se uma concepção de justiça não é estável, então, para Rawls, isso levanta sérias questões e dúvidas sobre sua justificação (Freeman, 2007a). Nesse sentido, embora afirme que a justiça é a “virtude primeira das instituições sociais”, reconhece que há problemas sociais fundamentais que vão além da justiça, especialmente os de coordenação, eficiência e estabilidade (Rawls, 1971).

Dentre esses problemas fundamentais, a estabilidade é considerada um dos maiores desafios para sua teoria, segundo autores como Krasnoff (2015). Não obstante os dois princípios de justiça que Rawls propõe lidem diretamente com os problemas de coordenação e eficiência — o primeiro princípio permitindo a coordenação por meio de liberdades básicas iguais, e o segundo permitindo eficiência por meio de incentivos econômicos compatíveis com o benefício dos menos favorecidos —, tais princípios não garantem por si mesmos que a sociedade cooperativa resultante será estável (Krasnoff, 2015). Para Rawls (2001, §55), estabilidade significa conformidade com os princípios e aceitação voluntária das regras básicas; em caso de violações, devem existir forças estabilizadoras que evitem novas infrações e restaurem o arranjo original. A possibilidade de coerção institucional como meio de assegurar a estabilidade é reconhecida, mas Rawls recusa-se a considerá-la suficiente. Como é sabido, Rawls (1971) distingue entre “teoria ideal” e “teoria não ideal”. A teoria ideal busca identificar os princípios e instituições que caracterizam uma sociedade justa, assumindo condições favoráveis à cooperação social e conformidade plena (*full compliance*) com os princípios de justiça. A teoria não ideal, por sua vez, examina circunstâncias em que esses pressupostos não se verificam integralmente, abrangendo tanto situações de conformidade parcial (*partial compliance*) e de não conformidade com as exigências da justiça quanto a contextos marcados por condições históricas, sociais ou econômicas desfavoráveis à realização de uma sociedade bem-ordenada. Ainda assim, Rawls sustenta que o problema da estabilidade deve ser enfrentado também pela teoria ideal. A dificuldade conceitual aqui é clara: se a teoria ideal pressupõe a conformidade com princípios justificados, por que haveria necessidade de provar que esses princípios são estáveis? Rawls responde que pode haver um hiato entre a melhor teoria de justiça concebível e aquilo que é justificável às

peças reais, dotadas de valores e concepções particulares de bem (Krasnoff, 2015). Mesmo que os princípios de justiça sejam justificados em abstrato, os indivíduos ainda precisam encontrar razões internas para aceitá-los e agir de acordo com eles, considerando seus próprios valores. Idealmente, isso deveria ocorrer não apenas por receio de sanções, mas porque agir justamente faz sentido do ponto de vista da racionalidade prática. Esse problema é abordado de maneira distinta ao longo da obra de Rawls: em “Uma teoria da justiça”, Rawls tenta demonstrar uma congruência entre os princípios de justiça e os interesses racionais individuais. Para isso, desenvolve uma “teoria fraca do bem”, que seria comum a todas as concepções de bem, e que define a racionalidade em termos instrumentais: indivíduos racionais desejam alcançar os fins de sua concepção de bem e buscar os meios adequados para tanto. Segundo Rawls (2016, p. 490), “o bem da pessoa é definido pelo que para ela é o projeto de vida mais racional, em circunstâncias razoavelmente favoráveis”. O desafio, então, é mostrar que o desejo de adotar o ponto de vista da justiça pertence ao bem de cada indivíduo à luz dessa teoria fraca. Rawls tenta resolver essa questão com um argumento em psicologia moral, defendendo que os indivíduos em uma sociedade bem-ordenada desenvolveriam normalmente um senso de justiça que coincide com a justiça como equidade (Krasnoff, 2015).

Contudo, Rawls (1993) reconhece posteriormente limitações importantes da formulação original desse argumento. Em “O liberalismo político”, o argumento da congruência é significativamente reformulado. A estabilidade da justiça como equidade deixa de depender principalmente das considerações apresentadas na Parte III de “Uma teoria da justiça” e passa a ser explicada em termos de reciprocidade, razão pública e consenso sobreposto entre doutrinas abrangentes razoáveis. Em uma sociedade caracterizada pelo pluralismo razoável, é irrealista presumir que todas as pessoas verão a justiça como equidade como parte integrante de sua concepção abrangente de bem racional. Por isso, Rawls (1993) reformula a exigência de estabilidade com base no conceito de consenso sobreposto: pessoas portadoras de diferentes doutrinas abrangentes razoáveis podem, a partir de seus próprios fundamentos, apoiar uma concepção política comum de justiça. A estabilidade é alcançada quando essa concepção é sustentada por um consenso sobreposto entre cidadãos razoáveis, cada qual justificando-a com base em suas doutrinas distintas (Krasnoff, 2015). O papel do “razoável” adquire centralidade na justificação política: Rawls argumenta que, em um contexto de pluralismo razoável, a razoabilidade não é apenas uma condição necessária, mas também suficiente para justificar politicamente uma concepção de justiça (Krasnoff, 2015). A tarefa, então, é demonstrar que uma

concepção de justiça politicamente justificada em termos do razoável também pode ser endossada pelos cidadãos à luz de seus interesses e fins racionais.

Tendo em vista essas considerações iniciais, este artigo tem por objetivo discutir questões de viabilidade e estabilidade como condições essenciais para se completar a argumentação em favor de uma concepção de justiça. Na última seção, pretende-se esclarecer mal-entendidos relativos à aplicabilidade de princípios ideais, tanto em uma sociedade bem-ordenada quanto em contextos de injustiça. Para tanto, o princípio de diferença será utilizado como um princípio exemplar para ser o objeto dessa investigação.

2 Problemas de viabilidade: aspectos teóricos e conceituais

Os problemas de viabilidade se conectam com os aspectos “práticos” das construções teórico-normativas do âmbito das teorias da justiça. Segundo Ben Laurence (2021), a filosofia política é marcada por uma tensão conceitual: como pode a filosofia política, enquanto teoria, ser prática? Frequentemente se supõe que a teoria trata da formulação de princípios gerais e da reflexão abstrata, ao passo que a prática diz respeito à ação em contextos particulares. Seria, então, um erro categorial supor que a filosofia política possa ser prática? O senso comum, observa Laurence (2021), sugere que a filosofia ocorre na torre de marfim, distante das ruas e dos corredores de poder — uma atividade contemplativa, não transformadora. Essa perplexidade se acentua quando consideramos o aspecto normativo da filosofia política. Se a tarefa é dizer o que é justo, o que devemos fazer, seria ilegítimo comprometer tais reflexões com considerações sobre realismo político, limitações humanas ou estratégias de ação? Para muitos, introduzir esses elementos no núcleo da teoria normativa seria uma forma de rebaixá-la. O ideal seria preservar a pureza da filosofia, distinguindo rigorosamente entre reflexão normativa e análise estratégica. Essa concepção austera, porém, gera insatisfação: ao preservar a integridade da teoria, a promessa prática que justifica sua existência é sacrificada. Surge, então, uma reação oposta: visões reconfiguram a filosofia política como um conjunto de respostas reativas e pragmáticas a males políticos específicos — como se ela devesse ser uma oficina de soluções experimentais, orientada apenas por viabilidades imediatas. Tais abordagens, contudo, parecem abandonar com demasiada facilidade a ambição sistemática e utópica da filosofia política. Elas pretendem convencer-nos a renunciar ao ideal de uma sociedade justa, em vez de nos orientar em sua direção (Laurence, 2021).

É precisamente contra essa dicotomia — entre a pureza teórica e o pragmatismo cético — descrita por Laurence (2021) que as questões

de viabilidade serão tratadas a seguir. Desse modo, mesmo sem nos comprometermos com as implicações “aristotélicas” de sua argumentação (que não desenvolveremos aqui), seguiremos a posição de Laurence, segundo a qual a filosofia política *pode* ser prática sem abandonar sua integridade teórica; e mais: que ela *deve* ser compreendida como um exercício de razão prática (Cf. Laurence, 2021, pp. 1–19).

Rawls (2001, p. 6) reconhece a dificuldade em “discernir os limites do praticável e quais são, de fato, as condições de nosso mundo social; o problema, a esse respeito, é que os limites do possível não são dados pelo existente, pois podemos, em maior ou menor grau, mudar as instituições políticas e sociais e muito mais”. Em sintonia com essa preocupação, Pablo Gilabert (2023) examina a relação entre justiça e viabilidade, abordando a necessidade de equilibrar a ambição normativa com as limitações práticas. Ele argumenta contra a visão de que a viabilidade deva ser tratada como um limite absoluto para a justiça, e propõe, em vez disso, um modelo dinâmico de análise de viabilidade que reconhece a possibilidade de transformação social. Embora certas restrições sejam realmente intransponíveis — como as “leis da lógica e da natureza” (Gilabert, 2023, p. 141) —, muitas outras são maleáveis e podem ser modificadas ao longo do tempo por meio da ação coletiva e da inovação institucional. Assim, uma teoria da justiça deve focar nessas restrições mutáveis, explorando estratégias para superá-las.

Para esclarecer essa relação entre princípios de justiça e condições práticas, Gilabert (2023) propõe três perguntas fundamentais. Primeiro, ele investiga o significado de viabilidade, descrevendo-a como um conceito relacional que depende do contexto e dos agentes envolvidos. A viabilidade não deve ser vista de maneira binária, como algo simplesmente possível ou impossível, mas em termos de graus e escalas de realização. Em segundo lugar, ele pergunta por que devemos nos preocupar com a viabilidade, rejeitando tanto o “realismo conservador”, que limita o escopo da justiça ao que já é viável no presente, quanto o “idealismo ingênuo”, que ignora completamente as restrições práticas. Para Gilabert (2023), uma teoria da justiça deve considerar a viabilidade não apenas para orientar a ação política de forma eficaz, mas também para preservar o compromisso com a transformação social de longo prazo. Por fim, ele analisa como lidar com alegações de inviabilidade, distinguindo entre juízos ideais, que examinam se um princípio é desejável independentemente das restrições atuais, e juízos prescritivos, que determinam o que pode ser realisticamente exigido em um dado momento. Essa distinção permite manter uma visão normativa de longo prazo sem perder de vista a necessidade de ações pragmáticas no presente.

A ideia central da contribuição de Gilabert (2023) é a proposta de um modelo dinâmico de justiça que articula o pensamento normativo em três dimensões interconectadas. A primeira dimensão (DI) envolve os princípios fundamentais de justiça, por exemplo, igualdade, liberdade e reciprocidade. A segunda dimensão (DII) considera as instituições e práticas sociais que operacionalizam esses princípios. A terceira dimensão (DIII) se concentra nos processos de transformação que podem tornar as instituições mais justas ao longo do tempo. Essa estrutura evita a armadilha de assumir que as condições políticas e sociais são fixas, enfatizando a necessidade de expandir gradualmente as possibilidades de mudança. Para Gilabert (2023), a viabilidade deve ser encarada como um campo em expansão, moldado pela ação coletiva e pela imaginação política. Outro ponto importante em sua análise é a discussão do princípio de que “dever implica poder”. Em sua formulação tradicional, esse princípio sugere que não se pode exigir moralmente aquilo que é impossível de realizar. Conquanto tal ideia pareça intuitiva, Gilabert adverte que ela pode ser interpretada de forma equivocada para justificar injustiças existentes sob a alegação de inviabilidade. Ele defende que o “poder” envolvido nesse princípio seja entendido de maneira dinâmica, levando em conta a capacidade dos agentes de transformar suas próprias condições ao longo do tempo.

Gilabert (2023) ilustra essa ideia com exemplos históricos, como os movimentos pelos direitos civis e pelos direitos trabalhistas, que inicialmente foram considerados inviáveis, mas tornaram-se realidade por meio da mobilização política e da reconfiguração das instituições. Para ele, a viabilidade não é apenas uma medida estática do presente, mas também um critério para ampliar possibilidades futuras. A imaginação política desempenha um papel essencial nesse processo. Gilabert (2023) sugere que agentes políticos não devem apenas reagir às condições existentes, mas também visualizar novas formas de organização social. Essa imaginação permite projetar alternativas que antes pareciam impossíveis, orientando a criação de políticas progressistas. Ele propõe inclusive a existência de deveres dinâmicos, que exigem não apenas ação dentro das condições atuais, mas um esforço contínuo para modificar essas condições. Isso implica que a luta por justiça não deve ser passiva, mas sim um compromisso ativo com a transformação das estruturas sociais.

Uma possível objeção à abordagem de Gilabert é que ela poderia levar a um perfeccionismo impraticável, dificultando a implementação de reformas concretas no curto prazo. Ele responde que sua proposta não exige que todos os ideais sejam imediatamente realizáveis, mas sim que haja um compromisso contínuo com sua concretização gradual (2023). Outra crítica seria que a abordagem dinâmica permitiria adiar indefinidamente mudanças essenciais

sob a justificativa de que ainda não são viáveis. Para evitar esse risco, Gilabert (2023) defende que os processos de transformação devem ter prioridades claras e prazos definidos, garantindo que melhorias reais sejam alcançadas ao longo do tempo e que a inviabilidade não sirva de desculpa para a inércia. Em resumo, Gilabert (2023) conclui reafirmando que a justiça não deve ser vista como um estado fixo, mas como um processo contínuo de transformação.

3 A estabilidade de uma concepção de justiça

O conceito de estabilidade em Rawls ocupa um papel central em sua teoria da justiça, sendo um critério fundamental para avaliar a viabilidade de uma sociedade regulada por princípios justos. Rawls sustenta que uma teoria da justiça não é completa se ela apenas apresentar um conjunto de princípios normativos bem-justificados; ela também deve demonstrar que é estável, ou seja, que é capaz de gerar sua própria sustentação ao longo do tempo. A estabilidade, para Rawls, não é meramente um requisito pragmático ou instrumental, mas uma característica inerente de uma ordem social verdadeiramente justa. Em “Uma teoria da justiça”, Rawls apresenta a estabilidade como um dos três desafios principais que uma teoria política precisa enfrentar, ao lado dos problemas da coordenação e da eficiência. Enquanto os dois princípios de justiça por ele propostos tratam diretamente dos problemas de coordenação e eficiência, a questão da estabilidade requer um aprofundamento distinto. Inicialmente, Rawls concebe a estabilidade em termos de psicologia moral: uma sociedade justa deve ser organizada de modo que os indivíduos desenvolvam, ao longo do tempo, uma disposição para agir conforme os princípios de justiça. Isso significa que uma concepção de justiça precisa estar alinhada com uma psicologia moral plausível, isto é, deve explicar como as pessoas poderiam internalizar um senso de justiça e incorporá-lo às suas próprias concepções do bem. Em sua formulação original, Rawls pressupõe que, em uma sociedade bem-ordenada regida pela justiça como equidade, os cidadãos acabariam adquirindo um forte senso de justiça que os levaria espontaneamente a respeitar as instituições justas.

No entanto, Rawls percebeu posteriormente que essa argumentação inicial era insuficiente. Em “O liberalismo político”, ele reformula sua abordagem, reconhecendo que as sociedades democráticas modernas são marcadas pelo pluralismo razoável, ou seja, pela coexistência de diversas doutrinas filosóficas, religiosas e morais razoáveis. Diante desse fato, a estabilidade não pode depender da suposição pouco realista de que todos os cidadãos internalizarão a mesma concepção abrangente de bem ou desenvolverão o mesmo senso

de justiça. Em vez disso, Rawls propõe o conceito de consenso sobreposto: diferentes concepções do bem podem chegar a um acordo quanto aos princípios políticos fundamentais, independentemente de suas justificações filosóficas particulares. Assim, uma sociedade justa será estável quando seus cidadãos, apesar de professarem doutrinas abrangentes distintas, aceitarem o arcabouço institucional da justiça como equidade, ou alguma concepção liberal semelhante, como base legítima para a cooperação social.

Rawls (1993, IV, §2) distingue, nesse contexto, dois tipos de estabilidade: a estabilidade pelas razões erradas e a estabilidade pelas razões certas. Uma sociedade pode ser estável porque as pessoas são constrangidas a cumprir as regras, seja por coerção direta, seja por incentivos externos, mas esse tipo de estabilidade é frágil, pois depende de mecanismos repressivos ou artificiais contínuos. Em contraste, uma sociedade é estável pelas razões corretas quando as pessoas aderem espontaneamente às normas de justiça porque as reconhecem como legítimas e compatíveis com seus próprios valores e concepções de vida boa. Esse ideal está ligado ao desenvolvimento do que Rawls chama de um “senso de justiça” compartilhado: uma disposição psicológica e moral cultivada por instituições justas e reforçada pela socialização política ao longo do tempo. Em outras palavras, estabilidade pelas razões certas significa que os cidadãos apoiam as instituições justas por convicção, não por mero cálculo ou coerção.

Outro aspecto crucial da estabilidade para Rawls é sua relação com o princípio da reciprocidade. Ele argumenta que as pessoas tendem a responder positivamente quando percebem que estão sendo tratadas de maneira justa; esse comportamento recíproco reforça a estabilidade, ao gerar um ciclo de apoio mútuo às instituições justas e reduzir as tentações de exploração ou de busca de vantagens indevidas. Desse modo, quando os cidadãos veem que os demais estão cumprindo os princípios de justiça, eles próprios se sentem motivados a cumprir as normas, o que consolida ainda mais a ordem justa. Esse elemento psicológico — a disposição de agir justamente quando confrontado com a justiça por parte dos outros — contribui para a autossustentação de uma sociedade bem-ordenada.

Rawls (1993) também explora a estabilidade no contexto da transmissão intergeracional da justiça. Para que uma sociedade seja verdadeiramente estável, ela deve ser capaz de reproduzir em cada nova geração o senso de justiça e o apoio às instituições justas. Isso depende não apenas da estrutura institucional em vigor, mas também de práticas educacionais e culturais que promovam os valores democráticos e igualitários, os chamados *valores políticos*. Em uma sociedade estável e justa, as instituições — família, escolas, comunidade —

inculcam nas novas gerações a compreensão e o apoio aos princípios de justiça, garantindo que a ordem justa perdure ao longo do tempo.

Rawls também enfatiza que a estabilidade não é um objetivo separado da justiça, mas um critério de validação da própria viabilidade de uma concepção de justiça. Uma sociedade que requer repressão constante para manter sua ordem não pode ser considerada genuinamente justa, pois falha em gerar apoio interno aos seus princípios normativos. Assim, a exigência de estabilidade reforça a ideia de que a justiça deve ser autossustentável, cultivando instituições e atitudes que assegurem sua continuidade sem necessidade de coerção extrema. Algumas interpretações das obras de Rawls enfatizam uma suposta ruptura entre “Uma teoria da justiça” e “O liberalismo político”, sugerindo que as mudanças em sua teoria foram respostas pragmáticas a novas críticas e circunstâncias. Contudo, análises como as de Samuel Freeman (2007b) e Paul Weithman (2011) apontam para uma forte coerência interna no desenvolvimento do pensamento rawlsiano. A tese central desses autores é que as modificações na concepção de justiça como equidade podem ser entendidas como um esforço para garantir a estabilidade da teoria, um problema que o próprio Rawls já reconhecia em sua formulação inicial e que se tornou mais evidente com o passar do tempo.

Assim, a relação entre estabilidade e coerência torna-se um eixo central para compreender a evolução da filosofia política de Rawls. No cerne dessa interpretação está a ideia de que as alterações na teoria de Rawls não configuram uma mudança radical, mas sim ajustes internos motivados pelo desafio de sustentar a estabilidade de uma sociedade bem-ordenada diante do fato do pluralismo razoável. A estabilidade, em “Uma teoria da justiça”, estava fundamentada na congruência entre o justo e o bem, isto é, na suposição de que os indivíduos desenvolveriam um senso de justiça que os levaria a apoiar e manter as instituições justas por reconhecê-las também como benéficas para si. No entanto, essa expectativa mostrou-se frágil diante da complexidade do pluralismo contemporâneo e de certos dilemas sociais (Weithman, 2011). Um exemplo é o “dilema do prisioneiro generalizado”, que Rawls inicialmente acreditou ter resolvido, mas posteriormente reconheceu como uma ameaça profunda à estabilidade de sua teoria. Esse dilema refere-se à situação em que, para que cada indivíduo aja justamente, é necessário que ele tenha confiança de que os demais também agirão — caso contrário, a tentação de “pegar carona” (*free-rider*) e violar as regras justas pode minar a cooperação. Rawls percebeu que confiar apenas no senso de justiça de cada um poderia não ser suficiente para superar esse problema de garantias mútuas (Weithman, 2011). A partir desse reconhecimento, Rawls reformula a questão da estabilidade, abandonando a ideia de que a justiça como equidade poderia ser inerentemente estável com

base na congruência psicológico-moral como descrita na Parte III de “Uma teoria da justiça”, ou seja, substantivamente fundamentada no argumento do valor do autorrespeito/autoestima e da sociedade como a união entre as uniões. Ele adota, em vez disso, a noção de consenso sobreposto, na qual diferentes concepções abrangentes convergem em um acordo político sobre princípios de justiça, mesmo sem compartilhar a mesma visão de bem. Weithman (2011) sustenta que essa mudança não foi apenas um deslocamento metodológico, mas uma resposta a uma fragilidade estrutural identificada pelo próprio Rawls na teoria original. Rawls havia presumido que uma sociedade bem-ordenada evitaria o problema do *free-rider* porque os cidadãos desenvolveriam uma motivação moral sólida para agir justamente; porém, ao reconhecer que essa confiança era insuficiente, ele reformulou sua abordagem enfatizando que a estabilidade deve ser conquistada por meio de um consenso baseado em valores políticos compartilhados. A leitura de Freeman (2007b) complementa essa perspectiva ao sugerir que a preocupação com a estabilidade sempre esteve no cerne do projeto rawlsiano, embora muitos intérpretes tradicionais a tenham subestimado. Para Freeman (2007b), Rawls estava interessado não apenas na estabilidade como requisito funcional de uma sociedade, mas também como uma exigência moral: uma teoria da justiça deve ser capaz de gerar sua própria sustentação sem recorrer à coerção ou à manipulação ideológica.

Vale ressaltar que, para Freeman (2023), Rawls não abandonou a ideia de congruência. Em “Uma teoria da justiça”, congruência significa, em primeiro lugar, que, assumindo que os princípios mais razoáveis estejam estabelecidos, a justiça e o correspondente senso de justiça são compatíveis com a nossa busca racional pelos interesses individuais. Além disso, a justiça e o exercício das capacidades morais associadas à justiça podem ser bens fundamentais para as pessoas morais, valendo a pena serem perseguidos e incorporados aos seus planos de vida por si mesmos (Freeman, 2023). Demonstrar que o justo e o bem são “congruentes” é um elemento central do argumento de Rawls a favor da estabilidade da justiça como equidade em uma sociedade bem-ordenada. Esse argumento é igualmente essencial nas reformulações propostas em “O liberalismo político”, nas quais a congruência assume a forma de demonstrar que pode haver um consenso sobreposto em torno da justiça como equidade entre concepções razoáveis do bem, informadas pelas doutrinas abrangentes — religiosas, filosóficas e morais — que pessoas razoáveis professam enquanto membros de uma sociedade bem-ordenada regida pela justiça como equidade (Freeman, 2023).

Lucas Petroni (2017) também propõe uma defesa da continuidade entre a teoria da justiça rawlsiana e a teoria da legitimidade desenvolvida na obra

posterior de Rawls. Ele defende que, em “O liberalismo político”, Rawls não altera o núcleo normativo da justiça como equidade, mas introduz uma nova perspectiva sobre a legitimidade do poder político em sociedades democráticas pluralistas. Para compreender a continuidade entre as duas obras, Petroni investiga a distinção entre os conceitos de justiça e de legitimidade em Rawls. Em “Uma teoria da justiça”, Rawls desenvolve um modelo contratualista no qual os princípios de justiça são escolhidos sob um véu de ignorância, garantindo um sistema equitativo de cooperação social. Todavia, essa teoria enfrentava um problema: como garantir que esses princípios seriam estáveis ao longo do tempo em uma sociedade democrática real? A solução da congruência, apresentada em “Uma teoria da justiça”, não levava plenamente em conta o fato do pluralismo razoável, isto é, a diversidade inevitável de doutrinas filosóficas e religiosas que emerge em sociedades livres. É nesse ponto que “O liberalismo político” surge como uma tentativa de reformular o problema da estabilidade de maneira compatível com o pluralismo. Rawls passa a enfatizar a legitimidade como critério para avaliar concepções de justiça: uma ordem política democrática deve ser estruturada de modo a ser aceitável para cidadãos livres e iguais, independentemente de suas concepções particulares do bem. O princípio liberal de legitimidade estabelece que o poder político deve ser exercido apenas segundo leis e princípios que todos os cidadãos possam, em princípio, endossar por meio da razão pública. Isso significa que, não obstante diferentes grupos tenham razões morais distintas, o Estado deve se apoiar apenas em princípios políticos que possam ser compartilhados por todos numa base comum.

Petroni argumenta que essa mudança de enfoque — da justiça em si para a legitimidade do poder — não implica um recuo do igualitarismo rawlsiano. Em vez disso, ele sustenta que a teoria da legitimidade de Rawls traz consequências distributivas significativas, incompatíveis com o liberalismo clássico e o libertarianismo. Por exemplo, a ideia de que um governo legítimo deve garantir um mínimo social robusto indica que a justiça distributiva não é apenas uma questão de ideal moral, mas também uma exigência de legitimidade política. De acordo com Rawls (2016, p. 343), um mínimo social para atender às necessidades básicas dos cidadãos deve ser garantido constitucionalmente, seja por meio de “benefícios familiares e de transferências especiais em caso de doença e desemprego, seja mais sistematicamente por meio de dispositivos tais como a complementação progressiva da renda (denominado imposto de renda negativo)”. Contudo, essa exigência não deve ser interpretada em termos suficientaristas. O mínimo social constitui um componente de uma concepção mais ampla de justiça social, cujas exigências distributivas vinculam-se,

primordialmente, à “igualdade equitativa de oportunidades” e ao “princípio de diferença”. De acordo com Rawls, um regime que falha em assegurar um nível básico de segurança econômica a seus cidadãos não comete apenas uma injustiça; ele perde também sua legitimidade. E isso traz implicações para a estabilidade. Outro ponto central da teoria da legitimidade de Rawls é a defesa da igualdade política efetiva. Rawls argumenta que os direitos políticos não podem ser meramente formais; seu valor equitativo deve ser assegurado de modo que desigualdades econômicas não comprometam a participação democrática dos cidadãos. Isso significa que medidas como financiamento público de campanhas e a rigorosa regulamentação do papel do dinheiro nas eleições são exigências normativas fundamentais para assegurar que todos os cidadãos tenham influência política comparável. Essa visão conflita com concepções liberais clássicas que tratam a participação política como um direito formal, ignorando como a desigualdade material de recursos pode distorcer o exercício dessa participação.

Por fim, Petroni (2017) discute a concepção rawlsiana do direito de propriedade. Rawls distingue a propriedade pessoal, que é essencial para o desenvolvimento individual, da propriedade dos meios de produção, que não é garantida de modo absoluto por seus princípios de justiça. Diferentemente de tradições liberais que tratam a propriedade privada como direito natural intocável, Rawls a considera uma instituição a ser moldada por parâmetros normativos. Em particular, ele defende que sistemas econômicos compatíveis com a justiça como equidade devem assegurar uma ampla difusão da propriedade produtiva, evitando a concentração excessiva de riqueza. Isso se reflete em sua preferência por modelos como a democracia de cidadãos-proprietários, nos quais o Estado atua ativamente para dispersar a riqueza e proteger a igualdade de condições econômicas. Em conclusão, Petroni sustenta que “O liberalismo político” não representa um recuo da justiça como equidade, mas sim uma expansão da teoria rawlsiana para lidar com questões de legitimidade política em uma sociedade pluralista. A mudança de foco, da busca pelos princípios ideais de justiça para a preocupação em como torná-los aceitáveis a todos os cidadãos, não enfraquece as exigências igualitárias, mas as reforça, ao demonstrar que certas condições materiais são necessárias não apenas para a justiça, mas para a própria legitimidade de uma ordem democrática e, consequentemente, para a sua estabilidade.

Uma crítica relevante à concepção rawlsiana de estabilidade foi formulada por Gerald Gaus (2011) a partir de uma perspectiva liberal centrada na razão pública como processo deliberativo entre concepções morais razoáveis. Gaus argumenta que Rawls permanece excessivamente confiante na ideia de que a

razão pública produziria convergência entre cidadãos razoáveis em torno de princípios compartilhados, subestimando a persistência do dissenso racional mesmo num contexto de respeito mútuo e reciprocidade. Para Gaus, uma concepção de estabilidade que dependa de consenso substantivo — ainda que limitado aos princípios políticos — tende a colidir com a diversidade profunda das sociedades contemporâneas. A exigência de justificação compartilhada, na forma de um consenso sobreposto, seria demasiadamente forte e pouco realista. Em vez disso, Gaus defende um modelo de ordem política baseado na sobreposição dinâmica de justificações públicas distintas, em que a legitimidade de uma norma dependeria de ser aceitável a partir de múltiplos pontos de vista razoáveis, mesmo que não haja consenso substantivo entre eles. Essa crítica não apenas questiona a suficiência do consenso rawlsiano para sustentar a estabilidade democrática, mas também sugere que a própria estrutura da justiça como equidade poderia ser excessivamente dependente de premissas epistemicamente homogêneas. Assim, Gaus reposiciona o problema da estabilidade como um desafio mais profundo de governança em contextos de dissenso persistente e racionalmente inevitável — uma visão que reforça a dificuldade de sustentar concepções de justiça robustas sem recorrer a mecanismos mais flexíveis de acomodação institucional da divergência moral. Uma possível resposta rawlsiana a essa objeção seria enfatizar que o liberalismo político já incorpora o pluralismo razoável como dado constitutivo das democracias contemporâneas. E o conceito de consenso sobreposto não exige acordo integral sobre fundamentos morais, mas apenas a aceitabilidade política dos princípios de justiça por parte de diferentes doutrinas razoáveis. A razão pública, por sua vez, é limitada a questões constitucionais essenciais e de justiça básica, portanto não abrange todos os aspectos da deliberação democrática. Nesse sentido, Paul Weithman (2011) defende que a estabilidade rawlsiana não depende da eliminação do dissenso, mas da possibilidade de que diferentes cidadãos, motivados por razões próprias, vejam a estrutura básica como justa e digna de apoio. Embora a crítica de Gaus reforce a tensão entre estabilidade normativa e diversidade epistêmica, isso não invalida necessariamente a proposta rawlsiana, mas apenas desloca o debate para a viabilidade da razão pública como dispositivo de justificação política em sociedades profundamente pluralistas.

Mesmo supondo um cenário viável e estável, surge uma outra preocupação: a teoria ideal de Rawls seria aplicável às situações reais de injustiça flagrante ou ela se restringe a sociedades já justas ou quase justas? A sequência do artigo aborda essa questão por meio da crítica de Charles Mills à distinção entre teoria ideal e não ideal. Veremos a seguir os argumentos de Mills, que acusa a

justiça como equidade de ignorar contextos marcados por racismo persistente, e examinaremos as respostas possíveis a essa objeção — notadamente a defesa metodológica de que a teoria ideal serve como um referencial crítico, e não de indiferença, em relação às injustiças existentes. Essa transição contextualiza o debate sobre aplicabilidade, preparando o terreno para reafirmar a relevância prática da perspectiva rawlsiana mesmo diante de sociedades imperfeitas.

4 Questões de aplicabilidade de uma teoria ideal de justiça

A aplicabilidade de uma concepção de justiça deve ser clarificada, pois ela costuma ser uma fonte quase inesgotável de interpretações equivocadas. É fato que uma das objeções mais frequentes às teorias normativas como as de Rawls dizem respeito a questões de aplicabilidade. Para discutir questões de aplicabilidade, iremos recorrer a dois casos exemplares: primeiramente, será discutida a aplicação do princípio de diferença em uma sociedade bem-ordenada, e em situações de injustiça. Em seguida, serão tratadas as objeções de Charles Mills (2017), segundo as quais a teoria de Rawls só se aplicaria às sociedades ideais e, portanto, careceria de qualquer valor prático. Isso nos ajudará a compreender melhor as suas implicações normativas e institucionais, além de dirimir mal-entendidos quanto à aplicação desse princípio e, por extensão, da teoria como um todo.

4.1 A aplicabilidade do princípio de diferença

O princípio de diferença de Rawls estabelece que as desigualdades são justas *somente se* beneficiarem os menos favorecidos da sociedade. É sabido que Rawls considera o princípio de diferença como um “princípio para instituições”, que se aplica diretamente apenas à estrutura básica da sociedade. Ele não se destina a ser aplicado a situações de pequena escala, mas é um “macroprincípio” (Freeman, 2018). O princípio de diferença, portanto, não é uma regra individual de conduta que seguimos ao tomar decisões econômicas; não há um dever natural de escolher maximizar as condições dos menos favorecidos. Em vez disso, o princípio de diferença se aplica ao arcabouço econômico e às políticas, leis e demais normas que regulam as transações econômicas. Rawls foi amplamente criticado por restringir a aplicação do princípio de diferença à estrutura básica. Seus motivos são complexos; eles se relacionam com o liberalismo de Rawls, que garante os compromissos pessoais dos indivíduos e de sua liberdade para perseguir suas concepções de bem, além

do fato de o princípio de diferença ser um princípio de reciprocidade, e não “*end-state*”¹ (Freeman, 2018).

Por “maior benefício aos menos favorecidos”, Rawls entende que as “desigualdades inevitáveis” — de renda, riqueza, poderes e posições — devem ser organizadas de modo a tornar a classe menos favorecida melhor do que estaria sob qualquer outro esquema alternativo de cooperação social e econômica (Rawls, 2001, §17). Um papel fundamental do princípio de diferença, então, é a comparação e a avaliação da justiça de diferentes sistemas econômicos. Rawls afirma que “a justiça como equidade é uma concepção para classificar formas sociais vistas como sistemas fechados” (Rawls, 1971, p. 229, tradução própria). Os sistemas econômicos existentes devem ser avaliados criticamente de acordo com o grau em que se aproximam do ideal de uma sociedade bem-ordenada. O segundo princípio de justiça deve, portanto, ser utilizado como padrão último para elaborar e reformar leis, regulamentos e outras normas econômicas, com o objetivo de realizar, em algum momento, o sistema econômico que o próprio segundo princípio prescreve (Freeman, 2018). Isso levanta a complexa questão de como o princípio de diferença deve ser aplicado em circunstâncias ideais e não ideais.

Para se ter uma ideia precisa de sua aplicação, a pergunta que se coloca é o que, exatamente, o princípio de diferença nos orienta a fazer? Conforme Laden (2015), há seis pontos a serem considerados. Primeiro, o objetivo geral é avaliar a justiça das estruturas básicas com base em como se saem os que estão em pior situação. Dadas duas sociedades que satisfazem igualmente as demais partes dos dois princípios de justiça, a mais justa é aquela que oferece um padrão de vida mais elevado aos que estão na base. Segundo, não se deve observar o indivíduo específico que está na pior situação, mas sim o membro representativo do grupo que se encontra na base. Ao definir esse grupo, pode ser necessário utilizar critérios distintos, a depender do aspecto do sistema social que está sendo avaliado. Assim, se a questão envolve a estrutura da economia e a distribuição de renda e riqueza, pode-se perguntar como se saem os trabalhadores não qualificados. Se o foco recai sobre instituições sociais que criam e perpetuam desigualdades de gênero ou raça, então deve-se observar os membros do gênero ou da raça que estão em desvantagem (mulheres, pessoas não brancas). Em “Uma teoria da justiça”, Rawls concentra-se em questões de desigualdade econômica e, por isso, utiliza critérios econômicos para identificar o grupo menos favorecido. Ainda assim, esse grupo não é definido por outra coisa

1 Um princípio *end-state* avalia a justiça de uma distribuição com base apenas no padrão final de resultados, independentemente do processo que o gerou.

senão sua posição na distribuição. Terceiro, devemos considerar as expectativas das pessoas ao longo de uma vida inteira, e não em um momento específico em que estejam em situação desfavorável. Quarto, o princípio de diferença não é o único princípio que regula a distribuição de benefícios sociais e econômicos. O primeiro princípio — que exige a proteção igual de um conjunto de liberdades básicas e a garantia do valor equitativo das liberdades políticas — e o princípio da igualdade equitativa de oportunidades — que requer que a origem familiar não influencie as chances de um indivíduo alcançar posições de prestígio e poder na sociedade — apresentam enormes implicações distributivas. O princípio de diferença só entra em cena depois que esses dois primeiros foram satisfeitos. Quinto, as vantagens sociais e econômicas devem ser medidas em termos do que Rawls chama de bens primários. Os bens primários devem ser coisas úteis para qualquer plano de vida razoável e racional. Além disso, eles oferecem uma medida objetiva em termos de recursos, que permite comparações públicas de vantagem. Como os princípios de justiça devem fundamentar uma concepção pública de justiça, eles precisam referir-se a características das pessoas que sejam objetivas e publicamente acessíveis (Laden, 2015). Por fim, o princípio de diferença, juntamente com os demais princípios da justiça, é um princípio do que Rawls chama de justiça procedimental pura. Isso significa que ele não deve ser aplicado diretamente a estados de coisas particulares produzidos pelas instituições, mas sim à estrutura das próprias instituições (Laden, 2015). Para criticar as instituições da estrutura básica por não atenderem ao princípio de diferença, é necessário examinar não apenas a distribuição de renda e riqueza em dado momento da sociedade, mas também como as instituições sociais operam conjuntamente para moldar essa distribuição (Laden, 2015).

Sob condições ideais, em uma sociedade bem-ordenada, presume-se que já se tenha estabelecido o sistema institucional mais eficaz, segundo o qual os menos favorecidos estão tão bem quanto poderiam estar em qualquer alternativa (Freeman, 2018). Nessa configuração, não há uma classe privilegiada que concentre a maior parte da riqueza produtiva, e os mercados são regulados de maneira a evitar abusos e ineficiências. Ainda que existam desigualdades moderadas, estas são justificadas pela existência de uma igualdade equitativa de oportunidades e de um mínimo social adequado, o que permite afirmar que tais desigualdades realizam uma forma profunda de reciprocidade. Nessas condições, as políticas econômicas devem continuar a beneficiar os menos favorecidos até o ponto em que novos aumentos no mínimo social deixem de ampliar o exercício efetivo das liberdades básicas ou o acesso equitativo a oportunidades educacionais, profissionais e culturais. A partir desse ponto, o excedente social pode ser utilizado para finalidades públicas legitimadas democraticamente, como o apoio a sociedades mais pobres ou investimentos

em infraestrutura. Alternativamente, se todas as obrigações de justiça forem satisfeitas, a sociedade pode optar por manter o *status quo* econômico e priorizar a manutenção de bens públicos e capital existente, em uma espécie de estado estacionário nos termos de J. S. Mill (Freeman, 2018).

É importante ressaltar que o argumento de Rawls (2001, p. 94–95) em favor do princípio de diferença não se apoia, em última instância, na regra *maximin* como critério de escolha adotado por partes racionais na posição original. Em vez disso, seu argumento é contratualista, no sentido de que apela à perspectiva de pessoas razoáveis em uma sociedade bem-ordenada: “o princípio de diferença é justificável e geralmente aceitável para pessoas morais livres e iguais, motivadas moralmente por um senso de justiça e de autorrespeito” (Freeman, 2018, p. 126, tradução própria). Os argumentos de Rawls em defesa do princípio de diferença baseiam-se em considerações de reciprocidade democrática, nas exigências do comprometimento (*strains of commitment*) e nas condições necessárias para manter o autorrespeito dos cidadãos livres e iguais (Freeman, 2018). Em vista disso, críticos afirmam que a teoria ideal de Rawls implicaria que os princípios de justiça se aplicam apenas a uma sociedade “perfeitamente justa”, e não a circunstâncias não ideais marcadas pela injustiça. No entanto, como salienta Freeman (2018, p. 127, tradução própria), “a justiça como equidade não teria sentido se os princípios de justiça não fossem aplicáveis à avaliação das injustiças em contextos não ideais”. Além disso, Rawls (1971) é explícito ao afirmar que a teoria ideal orienta a aplicação dos princípios de justiça em situações não ideais. A questão relevante, portanto, é: como o princípio de diferença deve ser aplicado na avaliação de circunstâncias não ideais e na orientação de reformas sob condições injustas, nas quais uma estrutura básica justa ainda não está estabelecida? As questões relativas às reformas necessárias para realizar uma estrutura básica justa são, em grande parte, empíricas e estratégicas — especialmente diante de interesses arraigados que se beneficiam da injustiça —, mas ainda assim certas regras de aplicação permanecem válidas (Freeman, 2018). A seguir, com base na interpretação de Freeman², reconstruo, em linhas gerais, uma proposta sucinta para aplicar o princípio de diferença de Rawls em circunstâncias não ideais.

Sob condições não ideais, como nas sociedades capitalistas com redes mínimas de proteção social ou em arranjos libertários sem qualquer tipo de amparo institucional, a aplicação do princípio de diferença enfrenta desafios

2 Cf. FREEMAN, 2018, pp. 127–130.

substanciais. Freeman (2018) destaca que, nesses contextos, além de problemas estratégicos — decorrentes da rejeição generalizada aos princípios de justiça —, há uma tensão entre medidas locais que maximizam os benefícios imediatos aos menos favorecidos e reformas estruturais mais amplas que visam instaurar um sistema econômico efetivamente justo, como a democracia de cidadãos-proprietários. A pergunta central passa a ser: até que ponto uma sociedade pode sacrificar ganhos imediatos para os menos favorecidos em nome de reformas estruturais mais amplas? Freeman (2018) argumenta que, embora o princípio de diferença não autorize sacrifícios das liberdades básicas ou do bem-estar dos menos favorecidos atuais para maximizar benefícios a futuras gerações, pode ser aceitável abdicar de medidas imediatas que melhorem a situação dos menos favorecidos se estas forem a única via realista para reformas necessárias. Isso inclui, por exemplo, casos em que medidas que aumentam a desigualdade são politicamente indispensáveis para iniciar uma transição institucional. Por outro lado, medidas que perpetuam desigualdades extremas e dificultam reformas futuras podem ser consideradas injustas mesmo que beneficiem os menos favorecidos no curto prazo.

Freeman (2018) propõe duas diretrizes para guiar a aplicação do princípio de diferença em contextos não ideais. Primeiramente, a sociedade deve adotar políticas econômicas locais que reduzam desigualdades ou maximizem os benefícios aos menos favorecidos ao longo de suas vidas, a menos que tais medidas solidifiquem desigualdades substanciais e impeçam reformas. Em segundo lugar, há uma obrigação de implementar reformas econômicas mais amplas que sejam eficazes — ou ao menos moderadamente eficazes — na construção de uma estrutura básica justa, contanto que essas reformas melhorem, em algum momento, a situação dos menos favorecidos nas gerações atuais.

4.2 As objeções de Mills à aplicabilidade da teoria de Rawls

Segundo Charles Mills, a estrutura metodológica da justiça como equidade indica uma restrição fundamental de escopo: “a teoria da justiça de Rawls, apesar de toda a sua abstração, foi concebida para se aplicar apenas a sociedades ideais” (Mills, 2017, p. 167, tradução própria). Como Rawls distingue explicitamente entre as partes ideal e não ideal de sua teoria, e afirma que só faz sentido passar à teoria não ideal depois que a parte ideal estiver resolvida, então “a sua teoria não se aplica de modo algum a sociedades não ideais — como as sociedades supremacistas brancas” (Mills, 2017, p. 167, tradução própria). Vale observar que a objeção aqui não é dirigida à plausibilidade dos princípios ou à coerência interna da construção normativa, mas ao fato de que, de acordo com o próprio Rawls, a teoria ideal precede e condiciona o desenvolvimento

da teoria não ideal. Isso significa que, até que as instituições básicas estejam organizadas de modo a satisfazer plenamente os dois princípios de justiça, os critérios normativos da teoria permaneceriam, na prática, inaplicáveis. Para Mills, essa orientação metodológica tem consequências significativas: ao limitar-se às condições ideais de uma sociedade bem-ordenada, a teoria não forneceria os instrumentos conceituais necessários para avaliar ou transformar sociedades estruturalmente injustas — como aquelas marcadas pela dominação racial persistente.

Esse ponto é reiterado em várias passagens. Em um momento particularmente enfático, Mills (2017, p. 168, tradução própria) afirma: “[a] teoria ideal rawlsiana não é uma teoria que possa ser aplicada ao mundo não ideal para nos dizer o que devemos fazer agora em sociedades como os Estados Unidos e a África do Sul, que por muito tempo foram caracterizadas por injustiça racial e que, muito provavelmente, ainda o são”. A crítica, portanto, é de ordem metodológica: a teoria seria normativamente impotente diante das injustiças persistentes, justamente porque sua arquitetura metodológica a restringe ao domínio do ideal. Em outras palavras, por se orientar por pressupostos de justiça para uma sociedade bem-ordenada, a teoria não poderia orientar ações corretivas nem avaliar a legitimidade de instituições existentes em contextos nos quais os princípios de justiça estão sistematicamente violados. A objeção pode ser reconstruída da seguinte forma: a teoria rawlsiana, por definição, se aplica apenas a uma sociedade ideal, isto é, uma sociedade cujas instituições básicas satisfazem plenamente os dois princípios de justiça. Isso implica que seus critérios normativos só ganham aplicabilidade real depois da realização de condições institucionais ideais — ou, no mínimo, suficientemente justas. Consequentemente, a teoria não estaria equipada para lidar com os problemas morais e políticos urgentes enfrentados por sociedades marcadas por desigualdade estrutural, racismo sistêmico ou injustiça institucionalizada. Nessas sociedades, os princípios da justiça como equidade seriam ou inoperantes ou irrealizáveis, e a própria teoria não forneceria critérios normativos úteis para orientar políticas públicas, protestos, reformas ou julgamentos morais. Seu domínio de validade seria, portanto, normativamente restrito: útil como modelo de referência em contextos ideais, mas silencioso ou irrelevante diante de violações concretas da justiça. Em palavras mais diretas, como o mundo da teoria ideal não existe, seu objeto corresponderia a um conjunto vazio e, dessa forma, ela seria inútil para lidar com os problemas reais. Essa linha de crítica é reiterada por Mills (2017, p. 168, tradução própria) quando afirma que “a teoria foi concebida para uma sociedade perfeitamente justa e não deveria ser aplicada às sociedades supremacistas brancas reais”. Ainda que se

reconheça o papel de modelos ideais como referência regulativa, a crítica de Mills recai sobre a insuficiência desses modelos em orientar transformações em sociedades reais. A teoria não seria simplesmente incompleta, como se faltassem partes ainda não desenvolvidas, mas inapropriada para enfrentar diretamente as exigências normativas que surgem em contextos marcados por opressão racial e desigualdade sistêmica. Em síntese, Mills propõe um dilema para a teoria rawlsiana: ou ela se mantém no domínio da idealidade, sem aplicabilidade concreta a contextos de injustiças persistentes, tornando-se normativamente inoperante; ou ela pretende se estender a esses contextos, mas, nesse caso, carece dos instrumentos conceituais adequados para analisar relações de poder reais e dinâmicas de exclusão social. A crítica se concentra, assim, na aplicabilidade da teoria: seu suposto compromisso com a idealização restringiria sua utilidade normativa diante de problemas urgentes, tornando-a uma teoria apropriada apenas para um mundo inexistente.

Para responder a essa crítica, apoio-me nas argumentações de Freeman (2022) que tratam das objeções de Mills à aplicabilidade da teoria ideal de Rawls em contextos de injustiça. Em vista das críticas de Mills, Freeman (2022) contesta de modo direto e sistemático a objeção de que a teoria da justiça de Rawls seria inaplicável a contextos de injustiça como o das sociedades racialmente estruturadas. Conforme Freeman, essa crítica incorre numa má compreensão dos propósitos, da estrutura e do escopo da teoria ideal tal como formulada por Rawls. Em vez de ser um artifício escapista ou um método para evitar a consideração das injustiças reais, a teoria ideal é, para Rawls, uma etapa preliminar necessária da investigação normativa: ela estabelece as condições de uma sociedade justa para que, a partir desse modelo, seja possível avaliar com clareza as deficiências do mundo real e propor reformas orientadas por princípios de justiça. Nesse sentido, como apontou Rawls (1971), a teoria ideal fornece um ponto de referência essencial para julgar o que é injusto e por que é injusto. Freeman reforça essa perspectiva ao defender que a teoria ideal é parte de uma estratégia metodológica mais ampla, cuja função é ordenar e dar coerência às exigências normativas de justiça.

Freeman (2022) argumenta que Rawls jamais sustentou que a justiça como equidade só seria válida em condições ideais, nem que seria incapaz de orientar a avaliação de sociedades reais marcadas por injustiças profundas. Ao contrário, ele destaca que Rawls concebe sua teoria como composta por duas partes *complementares*: a ideal e a não ideal. A primeira define os princípios de justiça sob condições favoráveis — incluindo uma estrutura básica justa, cidadãos razoáveis e instituições funcionando adequadamente —, enquanto a segunda trata das estratégias para lidar com injustiças específicas, identificar violações

sistemáticas da justiça e orientar transições para uma ordem institucional mais justa. O que distingue essas partes não é uma separação rígida de domínios normativos, mas uma diferença de perspectiva analítica: a teoria ideal fornece os parâmetros de avaliação, e a teoria não ideal lida com as contingências da aplicação. Negar a utilidade da teoria ideal com base em sua abstração seria, segundo Freeman, equivalente a descartar os próprios critérios de justiça que tornam possível avaliar as injustiças concretas.

Essa leitura é reforçada pela argumentação de Freeman, segundo a qual a estrutura da justiça como equidade seria compatível com a análise crítica de sistemas injustos, inclusive aqueles baseados em hierarquias raciais. Ele observa que Rawls condena explicitamente, ao longo de sua obra, práticas como o *apartheid*, o colonialismo, o racismo institucional e a escravidão — e o faz com base nos próprios princípios da teoria. Que tais condenações não apareçam no núcleo sistemático de “Uma teoria da justiça” não significa que estejam excluídas do escopo da teoria, mas sim que pertencem à aplicação posterior dos princípios em contextos não ideais. A ausência de exemplos históricos específicos ou de análises detalhadas de injustiças concretas não decorre de negligência, mas de uma escolha metodológica: Rawls busca formular princípios normativos gerais que possam orientar, de modo sistemático, juízos sobre o mundo real. Freeman sugere que exigir que a teoria ideal contenha desde o início diagnósticos empíricos detalhados equivaleria a confundir os níveis da análise normativa com os da aplicação contextual.

Freeman também contesta a acusação de que a teoria de Rawls seria indiferente à história, à dominação social ou às formas sistemáticas de exclusão. Ele observa que a ideia de “estruturas básicas” abrange exatamente as instituições sociais que organizam as principais desigualdades de poder, renda, *status* e oportunidades. Não obstante a justiça como equidade não se dirija especificamente à raça como categoria normativa, ela condena qualquer forma de desigualdade institucional que viole os princípios de liberdades iguais, igualdade equitativa de oportunidades e benefício mútuo. Nesse sentido, a teoria tem recursos normativos robustos para criticar e corrigir injustiças raciais, mesmo que elas não tenham sido tematizadas diretamente no momento da construção da teoria ideal.

Um dos pontos centrais da defesa de Freeman diz respeito ao papel metodológico da abstração na filosofia política normativa. Ele argumenta que o ideal de imparcialidade, incorporado na posição original, exige precisamente que os princípios de justiça sejam escolhidos sob o véu de ignorância, sem conhecimento das identidades sociais, étnicas ou econômicas das partes. Essa estratégia visa garantir que os princípios escolhidos sejam aceitáveis para todos

os cidadãos livres e iguais, independentemente de sua posição contingente na sociedade. Para Freeman, essa abstração não é um apagamento da realidade, mas uma exigência de justiça: somente quando não se leva em conta os interesses particulares é possível chegar a princípios de justiça válidos para todos. É precisamente esse procedimento que permite que a teoria sirva como critério para julgar as injustiças do mundo real, incluindo aquelas baseadas em raça e gênero.

Freeman reconhece que a crítica de Mills acerta ao apontar lacunas na aplicação empírica da teoria de Rawls, mas sustenta que isso não invalida sua estrutura normativa. Ele admite que Rawls dedicou pouca atenção às injustiças raciais em sua obra, mas insiste que essa limitação temática não compromete a aplicabilidade da teoria. Em vez de um erro de teorização, trata-se de uma limitação contingente, que pode e deve ser superada por meio da extensão da teoria a novos domínios. Freeman sugere que a justiça como equidade é inteiramente compatível com políticas de reparação, ação afirmativa e reconhecimento de injustiças históricas, desde que essas medidas sejam formuladas de modo compatível com os princípios centrais da teoria. Em outras palavras, a teoria não exige passividade diante da injustiça, mas sim coerência normativa no seu enfrentamento.

A crítica de Mills, de acordo com Freeman, parte de uma interpretação equivocada do alcance e da função da teoria ideal. Ao sugerir que a teoria de Rawls é inútil para avaliar sociedades marcadas pelo racismo, Mills ignora o papel da teoria ideal como instrumento normativo e pressupõe que apenas diagnósticos empíricos detalhados podem oferecer orientação política. Freeman responde que a teoria ideal fornece os critérios para determinar o que deve ser reformado, enquanto a teoria não ideal trata das estratégias para realizar essas reformas. Ambas são necessárias e complementares. A falha não está na abstração da teoria, mas em confundir sua função com a de uma teoria empírica ou sociológica. Rawls nunca pretendeu descrever a realidade social tal como ela é, mas estabelecer os princípios que devem orientar sua transformação.

Por fim, Freeman ressalta que a justiça como equidade é uma teoria de princípios. Sua validade normativa não depende da inclusão de exemplos históricos ou de análises contextuais específicas. É precisamente por isso que ela pode ser aplicada a uma ampla gama de contextos de injustiça, inclusive os que envolvem dominação racial, desigualdade de gênero, pobreza ou exclusão política. A acusação de inaplicabilidade, nesse sentido, resulta de uma leitura literal e excessivamente restritiva da teoria, que não distingue entre a formulação dos princípios e sua posterior aplicação. Ao esclarecer essa distinção, Freeman não apenas responde à objeção metodológica de Mills, como

também reforça a relevância da filosofia política normativa como disciplina voltada à crítica das instituições reais, a partir de parâmetros conceitualmente bem-fundamentados³.

5 Considerações finais

O artigo procurou analisar os critérios normativos e as exigências pragmáticas que condicionam a viabilidade e a estabilidade de uma concepção de justiça. A partir das contribuições de Gilabert, distingui a viabilidade como um conceito relacional e dinâmico, que envolve tanto o que é possível realizar no presente quanto o que pode se tornar realizável mediante transformação institucional e mobilização coletiva. Essa leitura permite incorporar à teoria da justiça um compromisso com a superação progressiva de obstáculos estruturais, sem abdicar da ambição normativa. Em seguida, discuti o conceito de estabilidade como requisito normativo fundamental da justiça como equidade. Demonstrei que, para Rawls, uma concepção de justiça deve ser capaz de gerar sua própria sustentação de modo estável e pelas razões certas, isto é, com base no reconhecimento de sua legitimidade por parte dos cidadãos. Ao constatar a insuficiência de modelos baseados apenas na psicologia moral ou na racionalidade individual, Rawls incorpora o pluralismo razoável como dado constitutivo das democracias modernas e reformula sua teoria para lidar com o problema da estabilidade nos termos de um consenso sobreposto e da ideia de razão pública. Procurei demonstrar que a estabilidade, para Rawls, não é uma exigência meramente funcional, mas uma condição de validade normativa. Nesse contexto, discuti a crítica de Gaus à posição de Rawls, bem como as defesas de Freeman, Weithman e Petroni.

Para ilustrar a aplicabilidade da concepção rawlsiana, com base em Freeman, mostrei como o princípio de diferença pode operar como padrão regulativo em contextos não ideais e, portanto, orientar reformas institucionais e políticas públicas sob condições de injustiça. Como procurei demonstrar, a teoria ideal pode ser utilizada satisfatoriamente para fornecer um padrão normativo a partir do qual as injustiças possam ser identificadas e enfrentadas. Tratei ainda da crítica de aplicabilidade dirigida por Charles Mills à justiça

3 Curiosamente, Mills propõe a formulação de princípios para uma teoria não ideal da justiça racial — o que pressupõe, metodologicamente, a mesma distinção entre teoria ideal e não ideal que ele criticara anteriormente. A proposta de um liberalismo radical negro pode ser interpretada não como uma rejeição estrutural do modelo rawlsiano, mas como um esforço de preenchê-lo com elementos sensíveis à história da opressão racial. A crítica à “inaplicabilidade” da teoria parece, assim, menos uma refutação da metodologia e mais um apelo por sua extensão crítica a contextos de injustiça sistemática. Cf. Mills, 2017, pp. 201–216.

como equidade. Segundo essa leitura, a distinção rawlsiana entre teoria ideal e não ideal excluiria a análise normativa de contextos efetivamente injustos, como sociedades marcadas pelo racismo persistente. Essa objeção foi respondida com base nas argumentações de Samuel Freeman. Em suma, os elementos analisados neste artigo reforçam a tese de que uma concepção rawlsiana de justiça pode ser considerada viável e politicamente estável. Do mesmo modo, a estabilidade política é concebida “pelas razões certas”, a partir do reconhecimento da legitimidade das instituições por cidadãos livres e iguais. O debate entre Mills e Freeman, por sua vez, permitiu ilustrar os limites e as possibilidades da teoria ideal em contextos de injustiça estrutural, sinalizando que o modelo rawlsiano pode oferecer diretrizes normativas mesmo em situações de injustiça.

Disponibilidade de dados:

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

Ausência de conflito de interesses:

O autor declara que não há conflito de interesses.

Editor responsável:

Mauro Luiz Engelmann

Referências

- FREEMAN, S. “Rawls”. New York: Routledge, 2007a.
- _____. “Justice and the social contract: essays on rawlsian political philosophy”. New York: Oxford University Press, 2007b.
- _____. “Ideal theory and racial justice”. *Seminário do Programa de Pós-Graduação em Lógica e Metafísica – PPGLM/UFRJ*, 2022. Disponível em <https://www.youtube.com/watch?v=IcDXvyLwqTw> Acessado em 26 de Agosto de 2024.
- _____. “Original position”. In: ZALTA, E. N., NODELMAN, U. (eds.). *The Stanford Encyclopedia of Philosophy*. Stanford: Stanford University, 2023. Disponível em <https://plato.stanford.edu/archives/win2023/entries/original-position/> Acessado em 4 de Julho de 2024.
- GAUS, G. “The order of public reason”. New York: Cambridge University Press, 2011.
- GILABERT, P. “Human dignity and social justice”. New York: Oxford University Press, 2023.

- KRASNOFF, L. "Stability". In: MANDLE, J., REIDY, D. A. (eds.). Cambridge: Cambridge University Press, 2015, pp. 804-810.
- LADEN, A. S. "Difference principle". In: MANDLE, J., REIDY, D. A. (eds.). 2015, pp. 211-216.
- LAURENCE, B. "Agents of change: political philosophy in practice". Cambridge: Harvard University Press, 2021.
- MANDLE, J., REIDY, D. A. (eds.). "The Cambridge Rawls lexicon". Cambridge: Cambridge University Press, 2015.
- MILLS, C. W. "Black rights/ white wrongs: the critique of racial liberalism". Oxford: Oxford University Press, 2017.
- PETRONI, L. "O argumento da estabilidade no contratualismo de John Rawls". *Kriterion: Revista de Filosofia*, Vol. 58, Nr. 136, pp. 139-161, 2017.
- RAWLS, J. "A theory of justice". Oxford: Oxford University Press, 1971.
- _____. "Justice as fairness: a restatement". Cambridge: Harvard University Press, 2001.
- _____. "Political liberalism". New York: Columbia University Press, 1993.
- _____. "Uma teoria da justiça". São Paulo: Martins Fontes, 2016.
- WEITHMAN, Paul J. "Why political liberalism? On John Rawls's political turn". New York: Oxford University Press, 2011.