

PUCRS

ESCOLA DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA
MESTRADO EM FILOSOFIA

GUSTAVO OLIVA DE OLIVEIRA

CONTRATUALISMO E A NOÇÃO MORAL DE AGREGAÇÃO

Porto Alegre

2022

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

GUSTAVO OLIVA DE OLIVEIRA

CONTRATUALISMO E A NOÇÃO MORAL DE AGREGAÇÃO

Dissertação apresentada como requisito para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Filosofia da Escola de Humanidades da Pontifícia Universidade Católica do Rio Grande do Sul.

Área de concentração: Estado e teorias da justiça

Orientador: Prof. Dr. Thadeu Weber

Porto Alegre

2022

AGRADECIMENTOS

Agradeço, em primeiro lugar, à minha família próxima, minha mãe, meu pai e meu irmão, por serem os motivos óbvios pelos quais eu consigo fazer qualquer coisa.

Ao meu orientador, o Prof. Dr. Thadeu Weber, agradeço a imensa ajuda, confiança, colaboração e companhia durante a jornada que começou na graduação e passou pelo mestrado inteiro.

Agradeço a meus amigos Pedro Moser e Danielli Marafigo Arpino pela companhia intelectual frutífera e sinceramente divertida.

Aos membros do grupo de estudos Filosofia da Economia, especialmente André Pontes, André Neiva, Ramiro e Mariana, pelas excelentes discussões e sugestões para o meu primeiro capítulo.

A meu amigo Eduardo Alves pela companhia acadêmica e intelectual mais importante que descobri na PUCRS.

Ao Library Genesis e ao Sci-Hub, me sinto na obrigação de agradecer pela disponibilização de inúmeros artigos e livros acadêmicos sem os quais a presente dissertação e muitas outras pelo mundo seriam muito menos sofisticadas.

A todos os colegas de departamento e profissão que ajudaram, direta ou indiretamente, na discussão sobre Filosofia Moral que me levou até o fim desta dissertação.

A todos os membros do Bosteam, Benhur, Diego, P.F., Silvinha e Vini, agradeço a motivação e criação de uma sensação geral de pertencimento (vamos pra cima).

Agradeço à Luiza pela fé inabalável na minha capacidade de fazer quase qualquer coisa, que por menos justificada que seja, volta e meia me convence momentaneamente.

À Caroline agradeço por fazer com que os últimos dois anos tenham sido, contrariando todas as estatísticas, pouco solitários.

“I got so many theories and suspicions [...] today is the day I follow my intuition”

Kendrick Lamar.

RESUMO

O tema desta dissertação é a relação entre contratualismo moral e a noção de agregação moral. A dissertação possui três objetivos principais. Em primeiro lugar, fornecer uma explicação detalhada da agregação moral e dos seus principais problemas. A principal objeção é a de que teorias agregativas são incapazes de fazer distinções morais relacionadas à separação normativa das pessoas. Em segundo lugar, oferecer uma teoria não-agregativa da moralidade, neste caso, o contratualismo moral, como solução para os problemas da agregação. O contratualismo é centrado na noção de justificação moral para cada pessoa, e por isso encontra sucesso nos casos em que a agregação falha. Por último, defender o contratualismo moral de críticas relacionadas à sua exclusão das cláusulas agregativas. Os dois tipos de objeção analisadas são: (i) o Problema dos Números, relacionado à dificuldade que teorias não-agregativas encontram para explicar casos de salvamento de vidas, nos quais os números importam do ponto de vista moral; e (ii) os casos de risco social, nos quais ações arriscadas são justificadas graças à relação entre benefícios e graus de risco. O contratualismo é modificado para ser defendido dos dois tipos de objeção.

Palavras-chave: Filosofia moral; Contratualismo; John Rawls; Teoria da justiça; Agregação.

ABSTRACT

The subject of this dissertation is the relationship between moral contractualism and the notion of moral aggregation. Three main objectives follow. First, providing a detailed explanation of moral aggregation and its main problems. The main objection is that aggregative theories are incapable of making moral distinctions related to the separateness of persons. The second objective is offering a non-aggregative theory of morality, which will be moral contractualism, as a solution for aggregation's problems. Contractualism is centered on moral justification and that is why it can account for cases in which aggregative theories fail. The third and last objective is defending moral contractualism from objections concerning its exclusion of aggregative clauses. Two forms of critique are examined: (i) the Numbers Problem, related to the troubles non-aggregative theories have when trying to explain cases that involve saving lives, in which numbers matter from a moral standpoint; and (ii) social risk cases, in which risky actions are justified thanks to the relationship between benefits and degrees of risk. Contractualism is properly modified so that it can be defended from both kinds of objection.

Key-words: Moral philosophy; Contractualism; John Rawls; Theory of justice; Aggregation.

SUMÁRIO

1 INTRODUÇÃO	8
2 AGREGAÇÃO MORAL	12
2.1 UMA DEFINIÇÃO GERAL DE AGREGAÇÃO.....	12
2.2 AS PROPRIEDADES ESTRUTURAIS DA AGREGAÇÃO	14
2.3 SEPARAÇÃO NORMATIVA DAS PESSOAS: UMA OBJEÇÃO À AGREGAÇÃO	17
2.4 RESPOSTAS AGREGATIVAS	21
3 CONTRATUALISMO, MORALIDADE E JUSTIÇA	31
3.1 O CONTRATUALISMO DE RAWLS NA TEORIA DA JUSTIÇA: ORIGEM DO CONTRATUALISMO CONTEMPORÂNEO	31
3.2 O CONTRATUALISMO MORAL DESDE SCANLON	39
3.3 ALGUMAS DIFERENÇAS ENTRE AS TEORIAS DE RAWLS E SCANLON ..	45
3.4 AS VANTAGENS DE ABANDONAR A AGREGAÇÃO.....	50
3.5 CONTRATUALISMO, MAXIMIN, LEXIMIN.....	52
3.6 TEORIAS AGREGATIVAS, PRINCÍPIOS AGREGATIVOS E A SEPARAÇÃO ENTRE AS PESSOAS.....	58
4 O PROBLEMA DOS NÚMEROS	62
4.1 NÚMEROS E AGREGAÇÃO	62
4.2 OS NÚMEROS CONTAM?	63
4.3 O ARGUMENTO DO DESEMPATE	67
4.4 VÉU DA IGNORÂNCIA.....	73
4.5 O ARGUMENTO DO PRINCÍPIO MAIS RAZOÁVEL.....	75
4.6 REIVINDICAÇÕES RELEVANTES AGREGADAS	78
5 CONTRATUALISMO E RISCO.....	82
5.1 CONTRATUALISMO E O PROBLEMA MORAL DO RISCO	82
5.2 CONTRATUALISMO EX POST E EX ANTE	85
5.3 O ARGUMENTO DA INFORMAÇÃO IRRELEVANTE E O ARGUMENTO DO CASO INDIVIDUAL	86
5.4 O PROBLEMA DAS REGRAS EX ANTE.....	93
5.5 VIDAS ESTATÍSTICAS E VIDAS IDENTIFICADAS	94
6 CONCLUSÕES.....	99
REFERÊNCIAS	102

1 Introdução

O tema desta dissertação é a relação entre contratualismo moral e a noção de agregação moral. A dissertação possui três objetivos principais. Em primeiro lugar, fornecer uma explicação detalhada da agregação moral e dos seus principais problemas. Em segundo lugar, oferecer uma teoria não-agregativa da moralidade, neste caso, o contratualismo moral, como solução para os problemas da agregação. Por último, defender o contratualismo moral de críticas relacionadas à sua exclusão das cláusulas agregativas.

O segundo capítulo é dedicado à noção moral de agregação. Partindo de uma definição geral desta ideia como uma junção de tudo que é digno de consideração moral em um valor único, sigo a discussão de Hirose (2014) em seu *Moral aggregation* para destacar as quatro propriedades estruturais da agregação moral interpessoal: comparabilidade interpessoal, imparcialidade, Pareto e continuidade. A conjunção destas quatro será tomada, neste texto, como condição necessária e suficiente para que uma teoria ou princípio seja considerado agregativo — o que quer dizer que qualquer princípio ou teoria que fira alguma das quatro condições *automaticamente não é* agregativo, e qualquer princípio ou teoria que possua as quatro *automaticamente é* agregativo.

Depois de definir a noção de agregação moral, uma objeção central à agregação será discutida: a de que teorias agregativas não respeitam a separação entre as pessoas. Esta objeção foi famosamente trazida por John Rawls em *Uma teoria da justiça* como uma maneira de rejeitar o utilitarismo na filosofia política. Argumento, no entanto, que esta reclamação se aplica a todas as teorias agregativas em sentido estrito — de fato, o motivo pelo qual ela se aplica ao utilitarismo é porque o utilitarismo é um tipo de teoria agregativa, a teoria agregativa *por excelência*. Algo na junção de diferentes fatores morais em um único valor apaga a distinção entre diferentes indivíduos, o que leva a consequências indesejáveis do ponto de vista da distribuição de recursos e bem-estar. A última seção aborda algumas tentativas de resposta do campo dos agregativistas, passando pela rejeição do argumento de Rawls e pela proposição de diferentes teorias agregativas como o Igualitarismo Teleológico e o Prioritarismo. Com inspiração no argumento de Voorhoeve e Fleurbaey, será mostrado como mesmo as teorias agregativas preocupadas com a ideia de distribuição, como é o caso de Igualitarismo Teleológico e Prioritarismo, falham em traçar as

distinções morais relevantes em diferentes casos que salientam a separação das pessoas. Estas considerações são suficientes para motivar a busca de uma teoria não-agregativa da moralidade.

O terceiro capítulo é dedicado ao estabelecimento de uma teoria deste tipo: o contratualismo moral. Depois de passar pelas origens deste tipo de teoria encontradas na teoria da justiça de John Rawls, exponho a versão de contratualismo que será base do resto da dissertação: o contratualismo moral de T.M Scanlon. Baseado na ideia de que a correção moral depende da justificabilidade de princípios a cada pessoa, o contratualismo moral incorpora restrições que garantem que a teoria não se torne agregativa. A conjunção da *Restrição Individualista*, a ideia de que razões para rejeitar princípios morais são sempre individuais, e *Princípio do Fardo Maior*, segundo o qual é irrazoável rejeitar um princípio com base em razão individual se suas alternativas impõem fardos maiores a outros, fazem com que o contratualismo bloqueie reivindicações agregativas no raciocínio moral — o contratualismo fere a condição de Continuidade da agregação moral. Assim, a teoria é capaz de obter os resultados corretos nos casos discutidos no capítulo anterior, porque existem dispositivos teóricos que garantem que a separação das pessoas importa. O capítulo ainda discute a relação entre este modelo de contratualismo e outros, a relação entre contratualismo e diferentes princípios de escolha racional e as diferenças entre teorias agregativas e princípios agregativos.

Os próximos dois capítulos discutem dois tipos de problema que o contratualismo enfrenta a partir de sua exclusão da noção de agregação moral — isto é, trato de lidar com os casos em que a agregação parece fazer trabalho positivo no raciocínio moral. O quarto capítulo é sobre o Problema dos números. Quando excluimos considerações agregativas de uma teoria, como explicar os casos em que os números importam? Situações simples como salvar uma vida vs salvar cinco vidas parecem se tornar problemáticas para o contratualismo, que não pode basear sua resposta na ideia de que devemos salvar mais vidas porque somamos as considerações de todas as cinco pessoas. Para responder a este problema, me apoio no insight de que devemos escolher o número maior quando os fardos em questão estão *no mesmo nível de relevância moral*. Esta consideração é base do argumento do desempate de Kamm e Scanlon, da ideia do véu da ignorância de Rawls e do argumento probabilístico de Schelling, do argumento do princípio mais razoável inspirado em Kamm e Hirose, e das propostas de agregação limitada como

Reivindicações relevantes agregadas, de Voorhoeve. O problema dos números pode ser resolvido de diversas maneiras se este insight principal for reconhecido.

O último capítulo trata da questão do risco social. Uma objeção comum ao contratualismo é que a teoria é incapaz de lidar com casos de risco social porque sua exclusão da agregação faz com que ações justificáveis sejam proibidas em função de possuírem um grau baixo de risco. Este é um problema sério para teorias de vertente deontológica, como é o caso do contratualismo. Começo traçando a distinção entre contratualismo *ex post* e contratualismo *ex ante*, baseada na perspectiva temporal/informacional através da qual agentes podem basear suas rejeições a princípios morais. Segundo a versão *ex post*, as rejeições devem ser feitas com informação total, o que significa que o que importa na avaliação de um princípio é quais serão seus resultados *reais* e como eles impactam a justificação a cada pessoa; na versão *ex ante*, o que importa são *as chances* que os princípios oferecem a cada pessoa, o que significa que, em casos de risco, nos quais não possuímos informação completa sobre os resultados, as decisões são feitas sem que cada pessoa saiba exatamente como será afetada. A versão *ex ante* lida de maneira efetiva com os casos de risco social e será defendida nesta dissertação.

O principal argumento a favor do contratualismo *ex post*, o Argumento da Informação Irrelevante, é rejeitado depois de um exame cuidadoso. O contratualismo *ex ante*, apoiado no Argumento do Caso Individual, fornece uma explicação mais simples, direta e com menos consequências problemáticas da moralidade do risco. As últimas duas seções lidam com objeções ao contratualismo *ex ante*. Primeiro, o Problema das Regras Ex Ante é considerado. Segundo este problema, a postura *ex ante* permite ações injustificáveis que teriam impacto geral benéfico e risco baixo, como experimentação médica involuntária em pessoas escolhidas aleatoriamente, apenas porque este tipo de ação está no interesse *ex ante* de todos. A resposta para este problema, proposta por Johann Frick, é que um teste de decomposição de estágios de ações como a experimentação médica revela que as ações devem ser justificadas *ex ante* em cada fase de implementação. Esta consideração defende o contratualismo *ex ante*. A última objeção ao contratualismo *ex ante* é a de que a teoria confere peso inadequado ao salvamento de vidas identificadas (isto é, prevenir mortes quando sabemos quem vai morrer) vs. vidas estatísticas (quando sabemos que as mortes ocorrerão, mas não sabemos quem vai morrer). Este é o desafio mais difícil para o contratualismo *ex ante*. Sugiro, seguindo Frick, que este viés é justificado moralmente em muitos casos em que os números

não diferem muito. Além disso, proponho um princípio de urgência que pode explicar os motivos pelos quais salvar vidas identificadas possui valor superior em muitos casos. O contratualismo possui explicações interessantes de diferentes fenômenos da moralidade do risco. A versão *ex ante* de contratualismo, no entanto, ainda precisa de uma explicação melhor de casos que envolvem esta característica do risco social se pretender ser uma teoria completa da moralidade.

2 Agregação moral

O assunto da presente dissertação é a relação entre contratualismo e a noção moral de agregação. Este capítulo será uma introdução dessa noção. Começo apresentando uma definição geral da agregação moral e alguns exemplos de juízo moral agregativo para depois introduzir suas características estruturais de maneira mais detalhada. Será de importante consideração, aqui, que existem diferentes modos de agregação. Em seguida, examinarei o que acredito ser a principal crítica à noção de agregação no âmbito moral: a de que teorias agregativas não respeitam a “separação normativa entre as pessoas”, o que tem implicações normativas indesejáveis, especialmente no que se refere à justiça distributiva. Principalmente a agregação total, que é a forma de agregação envolvida em teorias como o utilitarismo, está sujeita a esta crítica. No entanto, teorias que usam a agregação interpessoal de outras formas também sofrem, embora não tanto quanto o utilitarismo, do problema de ignorar a diferença entre indivíduos.

2.1 Uma definição geral de agregação

Quando estamos decidindo entre diferentes ações, uma das considerações mais importantes é qual será o resultado de cada ação, isto é, quais estados de coisas são trazidos por sua performance. Boa parte das decisões morais que precisamos tomar afetam muitas pessoas. Algumas dessas pessoas possuem objetivos conflitantes, e teorias morais devem nos fornecer princípios para decidir em favor de uma ou outra ação, baseando-se em seus resultados¹. Uma das maneiras de avaliar os resultados de uma ação é através de uma junção de todo seu impacto em diferentes fatores moralmente relevantes, obtendo um valor único que representa o resultado, e que pode ser comparado com os valores de outros resultados, a fim de decidir qual é o melhor. Esta junção é o que chamamos de *agregação moral*.

¹ Embora exatamente qual é o papel dos resultados na teoria moral seja bastante controverso. Como Rawls (2016, p. 30- 37) explica, teorias teleológicas da moralidade sugerem, grosso modo, que o objetivo da ação moral se resume em trazer algum resultado desejável; teorias deontológicas definem a correção moral de maneira diferente, sem reduzi-la a alguma função da quantidade de bem em um estado de coisas. Ainda assim, é importante lembrar que, mesmo para teorias deontológicas, as consequências são de importante consideração: nas palavras de Rawls, “Todas as doutrinas éticas dignas de atenção levam em conta as consequências ao julgar o que é certo. Aquela que não o fizesse seria simplesmente irracional, insana” (2016, p. 36). A teoria deontológica que será discutida nesta dissertação, por exemplo, o contratualismo, analisa a correção moral em termos da justificabilidade dos princípios por trás de cada ação. Um dos fatores que influencia a justificabilidade dos princípios é quais são suas consequências. No entanto, a maneira em que as consequências figuram no raciocínio moral contratualista é não-teleológica, já que não se preocupa com o valor das consequências em si, e sim com a maneira como estas afetam a justificabilidade dos princípios.

Um tipo de agregação está envolvido em decisões que afetam nossa vida em diferentes estágios. Por exemplo, a decisão de estudar para uma prova tem o resultado imediato de causar algum incômodo ou tédio. Esta decisão também, no entanto, tem o resultado de aumentar as chances de sucesso na prova, o que pode ter impacto relevante na vida de alguém. Combinar esses dois fatores do resultado em um só é um exemplo de *agregação intrapessoal* (HIROSE, 2014, p. 4), uma forma de agregação que envolve os impactos de uma ação entre diferentes estágios da vida de uma mesma pessoa. Este tipo de agregação não é o foco da discussão desta dissertação.

Quando a agregação combina os impactos de uma ação entre diferentes pessoas, é um caso de *agregação interpessoal* (HIROSE, 2014, p. 4). Imagine, por exemplo, o seguinte problema moral proposto por Taurek:

A situação é a de que eu tenho um estoque de algum remédio que salva vidas. Seis pessoas vão certamente morrer se não forem tratadas com o remédio. Mas uma das seis pessoas precisa de todo o remédio para sobreviver. Cada um dos outros cinco precisa apenas de um quinto do remédio. O que devo fazer? (TAUREK, 1977, p. 294, Tradução minha.).

Uma resposta para esta pergunta é: devemos salvar as cinco pessoas, porque é *melhor* que aconteça apenas uma morte. O raciocínio é feito através da combinação dos fatores relevantes do ponto de vista moral (nesse caso, o valor negativo da morte de cada uma das seis pessoas). Esta forma de agregação será o foco de nossa conversa. A partir daqui, quando me referir a “agregação”, a não ser que explicito que me refiro à agregação intrapessoal ou ao sentido mais geral de agregação, estarei me referindo à agregação interpessoal.

A agregação exige, para a combinação dos fatores moralmente relevantes, a conversão deles em um valor numérico. Isso é importante porque permite unir os impactos de uma ação através de diversas pessoas em um único valor. Considere o utilitarismo, por exemplo: segundo esta teoria, ações corretas são aquelas que maximizam o balanço geral de bem-estar. Para obter esse balanço, os utilitaristas devem *somar* os valores que representam o impacto da ação no bem-estar de cada pessoa. O maior valor é o que maximiza o balanço de bem-estar. Sem a conversão dos fatores moralmente relevantes (no caso do utilitarismo, o bem-estar) em um valor numérico, é difícil agregá-los. No entanto, não devemos pensar que o único método de agregação é a soma. Uma teoria moral poderia usar qualquer função para agregar os fatores moralmente relevantes. Existem formas de utilitarismo como o utilitarismo médio, que soma o bem-estar de todos os indivíduos e

depois divide este valor pelo número de indivíduos (HIROSE, 2014, p. 21). É possível criar teorias baseadas na multiplicação dos valores dos fatores moralmente relevantes — como é o caso, por exemplo, de alguns tipos de prioritarismo (HIROSE, 2014, p. 22).

Com algumas das características gerais da noção de agregação, posso avançar uma definição geral proposta por Iwao Hirose em *Moral Aggregation*. Segundo ele, a agregação é a “combinação de fatores moralmente relevantes como bem-estar, felicidade, prazer, satisfação de desejos, reivindicações, razões etc., em um valor objetivo. Por valor objetivo, quero dizer um valor em uma escala numérica [...]” (HIROSE, 2014, p. 23. Tradução minha.)². A agregação interpessoal, portanto, é uma combinação de fatores moralmente relevantes de diferentes indivíduos em um único valor numérico. É importante ressaltar que teorias agregativas avaliam estados baseados neste valor agregado, e não conferem peso a como cada indivíduo é afetado (FRICK, 2015, p. 175). Isto não quer dizer que, em algum sentido, todos indivíduos sejam considerados, afinal, como veremos na próxima seção, teorias agregativas são imparciais. Mas estes indivíduos nunca são considerados *como indivíduos*, isto é, individualmente. São considerados apenas indiretamente, ou seja, através da contribuição que seus fatores moralmente relevantes, já convertidos em valor numérico e agregados aos dos outros indivíduos, oferecem ao saldo total. Esta observação será importante na discussão da terceira seção.

2.2 As propriedades estruturais da agregação

Na seção passada, apresentei uma definição geral de agregação segundo a qual esta é a combinação dos fatores moralmente relevantes em um único valor numérico. A agregação interpessoal acontece quando este tipo de combinação utiliza os fatores de diferentes pessoas. Nesta seção, apresentarei as quatro propriedades estruturais da agregação interpessoal identificadas por Hirose (2014). São elas: comparabilidade interpessoal, imparcialidade, Pareto e continuidade. Estas propriedades serão importantes porque será a elas que recorreremos para

² As definições gerais de agregação apresentadas neste capítulo são válidas para todas ou quase todas as formas de agregação. Acredito que a consideração de Rawls sobre a separação normativa das pessoas, que será apresentada na terceira seção, se aplica não apenas à agregação total, que é a agregação baseada somente na soma, mas também à agregação interpessoal de modo geral.

identificar quais teorias ou princípios são agregativos. Quando quiser mostrar que uma teoria é anti-agregativa, demonstrarei qual ou quais dessas propriedades não estão presentes na teoria. Isto será especialmente útil nos casos em que é difícil dizer se alguma teoria ou princípio envolve agregação. Neste capítulo, no entanto, os exemplos de teoria agregativa que oferecerei serão razoavelmente simples.

Começo pela comparabilidade interpessoal. Segundo esta propriedade, “é possível comparar os fatores moralmente relevantes de uma pessoa com os de outra” (HIROSE, 2014, p. 32. Tradução minha.). Para que possamos agregar os ganhos e perdas de fatores moralmente relevantes, é necessário que eles possam ser comparados. Caso contrário, seria impossível conferir a eles valores em um saldo total. A comparabilidade interpessoal exige que sejamos capazes de dizer que um ou outro fator possui valor maior que outros no contexto relevante. Imagine dois estados distributivos $x = (8, 9, 10)$ e $y = (3, 5, 20)$. Segundo o utilitarismo, y é melhor que x porque seu valor total é maior (o valor de x é 27, enquanto o valor de y é 28). A aceitação da comparabilidade independe da concepção de valor adotada pela teoria. Utilitaristas clássicos, por exemplo, devem comparar estados através da quantidade total de bem-estar. Contudo, outras teorias morais podem incluir diferentes fatores na comparação. Considere, por exemplo, teorias que incluem uma cláusula de que a desigualdade de uma distribuição conta negativamente na comparação. Uma teoria igualitária segundo a qual a desigualdade é indesejável preferiria x a y . Esta preferência se baseia em uma comparação de fatores de diferentes indivíduos. A comparabilidade interpessoal é necessária, mas não é suficiente para a agregação. Isso quer dizer que mesmo teorias não-agregativas podem possuir a propriedade. O contratualismo moral, que será explorado no próximo capítulo, é um exemplo. Sua análise da correção moral depende da identificação da força das razões que os indivíduos possuem para rejeitar princípios morais. Sem comparabilidade interpessoal, é impossível dizer quais razões possuem maior força.

A segunda propriedade da agregação interpessoal identificada por Hirose é a Imparcialidade. Segundo esta propriedade, “duas alternativas não diferem moralmente se elas diferem apenas no que diz respeito à identidade das pessoas” (HIROSE, 2014, p. 36. Tradução minha.). A ideia geral da imparcialidade é que as identidades das pessoas não importam quando avaliamos estados moralmente. O utilitarismo clássico, por exemplo, é indiferente entre $x = (2, 3, 5)$ e $y = (5, 2, 3)$. Como Hirose ressalta, essa é uma definição técnica de imparcialidade,

potencialmente divorciada da discussão mais geral sobre imparcialidade na ética ou na filosofia política (HIROSE, 2014, p. 37). Dizer que uma teoria possui essa propriedade não quer dizer, necessariamente, que a teoria possui resultados ou pressuposições morais neutras em qualquer sentido, ou que possua imparcialidade sobre outra definição de imparcialidade. O que esta propriedade expressa é simplesmente que as identidades de indivíduos não afetam os juízos que fazemos sobre uma ação, princípio ou estado.

A terceira propriedade da agregação interpessoal é Pareto. Em teorias que aceitam Pareto, se uma opção x é melhor (mais justa/mais razoável etc.) para alguém que outra opção y e é pelo menos tão boa quanto y para outras pessoas, x é melhor que y (HIROSE, 2014, p.38). A ideia é de que ganhos para alguém que não pioram a situação de ninguém melhoram estritamente alguma distribuição. Embora tenha apelo intuitivo forte, existem princípios de escolha que rejeitam Pareto. No terceiro capítulo, apresento uma discussão da regra Maximin para decisões sob ignorância. Este é um princípio focado exclusivamente no nível do indivíduo menos favorecido em um estado, o que faz com que ganhos para outros indivíduos sejam ignorados. Maximin não é, portanto, agregativo. Isto será demonstrado com maior calma no capítulo seguinte.

A última propriedade necessária para a agregação interpessoal é a Continuidade, segundo a qual fatores moralmente relevantes de diferentes indivíduos podem ser representados por um valor numérico real (HIROSE, 2014, p. 39-40). Quando dois valores são contínuos entre si, podem ser convertidos para a mesma escala. Por exemplo, suponha que exista continuidade entre felicidade e tristeza. Isto quer dizer que qualquer impacto na felicidade ou tristeza de indivíduos pode ser representado com um valor numérico na mesma escala. O que significa que diferentes quantidades de felicidade podem compensar diferentes quantidades de tristeza, mesmo que você ache diferentes casos de felicidade e tristeza possuem pesos diferentes. Suponha que uma risada vale 5, enquanto evitar um choro vale 30. Isso significa que sete risadas possuem valor maior que evitar um choro, mesmo que o valor de evitar um choro seja maior que o de uma risada. Quando os valores são contínuos, tudo é uma questão de *quantidade*. Isso pode ser visto já que é uma propriedade dos números. Como Hirose aponta, “Continuidade é como a propriedade arquimediana dos números reais. Para quaisquer números positivos x e y , existe um número natural n tal que nx é maior que y ” (HIROSE, 2014, p. 40. Tradução minha.). Como continuidade permite

que os fatores sejam representados por valores numéricos reais, os fatores contínuos entre si possuem também essa propriedade.

Segundo Continuidade, *todos* os fatores moralmente relevantes são contínuos entre si, inclusive fatores de diferentes indivíduos. A principal implicação moral da continuidade é que para quaisquer dois fatores moralmente relevantes a e b , existe uma quantidade n de b que possui valor objetivo maior que a . É possível, então, que uma ação que tenha como consequência a morte de um indivíduo tenha valor positivo por causa das pequenas felicidades que traz a uma grande quantidade de pessoas. Isso acontece porque, mesmo que uma morte possua valor -5000 , se, por exemplo, a ação causar risadas de valor $+1$ a 5001 pessoas, se usarmos um método agregativo de soma, o valor da ação será $+1$. Esta questão é discutida por aqueles que rejeitam a propriedade, como veremos ainda neste capítulo. Devemos, entretanto, lembrar que este mecanismo de soma não é implicado pela continuidade. Continuidade implica apenas que os fatores podem ser convertidos em valores numéricos, deixando em aberto a questão sobre como devemos agregá-los.

Qualquer teoria, princípio ou método de escolha que aceitar as quatro propriedades é agregativa no sentido interpessoal. Como ressaltai antes, o método que será usado para agregar os fatores moralmente relevantes não é implicado pelo fato de que uma teoria é agregativa no sentido geral, o que faz com que existam diferentes tipos de teoria agregativa no sentido interpessoal. No que se segue, discutirei a plausibilidade de teorias morais baseadas nesta forma de agregação.

2.3 Separação normativa das pessoas: uma objeção à agregação

Em *Uma teoria da justiça*, de 1971, John Rawls forneceu um argumento contra o utilitarismo baseado na “separação normativa das pessoas”. Defenderei, nesta seção, que este argumento é forte o suficiente para nos fazer rejeitar não apenas o utilitarismo, mas também a noção moral de agregação interpessoal.

A teoria moral agregativa mais popular na história da ética é o utilitarismo clássico. Com “utilitarismo clássico”, me refiro à ideia de que uma ação é moralmente correta se e somente se é aquela que possui as consequências que maximizam o “saldo geral” de bem-estar no mundo. Observe, por exemplo, esta passagem de Mill, descrevendo o utilitarismo:

O credo que aceita a utilidade, ou o Princípio da Maior Felicidade, como fundamento da moralidade, defende que as ações estão certas na medida em que

tendem a promover a felicidade, erradas na medida em que tendem a produzir o reverso da felicidade. Por felicidade, entende-se o prazer e a ausência de dor; por infelicidade, a dor e a privação de prazer (MILL, 2005, p. 48.)

É fácil ver como o utilitarismo é agregativo: os fatores moralmente relevantes — neste caso, o bem-estar — de diferentes pessoas é combinado em uma mesma escala, e as ações, princípios, instituições com resultados que maximizam este valor são moralmente corretas. O utilitarismo clássico emprega uma forma de agregação total porque usa o método da soma, o que fica claro pela ideia de “saldo geral” de bem-estar. Isto significa que, no utilitarismo, o valor de um estado $x = (w_1, w_2, w_3)$, onde w_n representa o bem-estar de n , é igual à soma de cada elemento, isto é, $(\text{valor de } x) = (w_1 + w_2 + w_3)$.

Existe um problema nesta forma de agregação, apontado por Rawls (2016, p. 33). Este processo ignora o fato de que cada pessoa é um indivíduo separado, no sentido normativo, dos outros. O processo de decisão utilitarista envolve tratar o mundo inteiro como se fosse um único indivíduo decidindo entre alternativas, escolhendo aquela que maximiza seu bem-estar. Mas talvez a ideia de que as pessoas são separadas e possuem direitos a serem considerados seja ameaçada por esse raciocínio. Ao avaliar consequências de ações apenas através do resultado de uma soma, teorias agregativas permitem ações que impõem um grande fardo a uma minoria para o pequeno benefício de um grupo maior. Para enxergar isto, examine este caso, oferecido por T.M Scanlon:

Suponha que Jones sofreu um acidente na sala de transmissão de uma estação de televisão. O equipamento eletrônico caiu no seu braço e não podemos salvá-lo sem desligar o transmissor por quinze minutos. Uma partida da Copa do Mundo de futebol está acontecendo, sendo assistida por muitas pessoas, e não vai acabar por mais uma hora. O machucado de Jones não vai piorar muito se esperarmos, mas sua mão foi esmagada e ele está recebendo choques elétricos extremamente dolorosos. Devemos salvá-lo agora ou esperar a partida acabar? A coisa certa a se fazer depende de quantas pessoas estão assistindo - se é um, dois ou cem milhões? (SCANLON, 1998, p. 235. Tradução minha.)

O caso Jones desperta a seguinte intuição: é errado deixar Jones agonizando apenas para que os espectadores possam continuar assistindo ao jogo, não importa quantas pessoas estão no grupo dos espectadores. No entanto, o utilitarismo parece obter o resultado de que, dependendo do número de espectadores, temos a obrigação moral de manter a transmissão. Para ver como isso acontece, suponha que a dor de Jones tem um impacto de -10000 em seu bem-estar, enquanto o prazer de continuar assistindo à partida confere a cada espectador 1 unidade de bem-estar. Neste caso, se existirem 10001 espectadores, o utilitarismo recomenda que continuemos a transmissão, já que

esta ação produziria um saldo total superior ao de salvar Jones. Repare que os valores exatos são irrelevantes: o ponto é que, para qualquer valor a do impacto no bem-estar que conferirmos ao sofrimento de Jones e b ao impacto de assistir à partida no bem-estar dos espectadores, existirá um número n tal que $nb > a$. Mesmo o pequeno aumento de bem-estar de cada espectador pode, se multiplicado por um número suficientemente alto de espectadores, compensar o sofrimento de Jones. Este resultado acontece por causa da agregação: a ideia de que podemos combinar pequenos e grandes impactos em um só valor e avaliar resultados de ações a partir daí.

A objeção de Rawls explica o motivo de nos sentirmos incomodados com essa conclusão: teorias que tratam a agregação desta maneira preocupam-se apenas com a *quantidade total* de bem-estar produzido, sem considerar a maneira em que este bem-estar é *distribuído*. No entanto, este é um fator que deve ser considerado por qualquer teoria moral desejável. Além disso, a noção de “separação das pessoas” aponta que a distribuição dos fatores moralmente relevantes não deve ser levada em conta do ponto de vista de todos considerados em combinação, e sim daquilo que é justificável *para cada um* de sua própria perspectiva. O utilitarismo permite e, em alguns casos, exige distribuições de bem altamente desiguais (como as do caso Jones) onde um pequeno grupo é altamente sacrificado para que um grupo maior de pessoas receba um pequeno benefício. Acredito que este resultado é indesejável para uma teoria moral.

Esta é uma consideração forte contra o utilitarismo. Creio, no entanto, que a objeção de Rawls pode se aplicar à agregação de modo mais amplo. O que Rawls tem em mente quando fala na separação entre as pessoas parece ser que uma teoria moral adequada não pode combinar a perspectiva de indivíduos em uma só, já que isso borra os limites entre cada pessoa. Mas este tipo de combinação não é exclusivo do utilitarismo: está associado, na verdade, à noção de agregação moral. É possível, portanto, entender a objeção de Rawls como se aplicando não apenas ao utilitarismo, mas também a outras teorias agregativas. Como Hirose aponta,

A objeção da separação, baseada em um entendimento amplo da separação entre as pessoas, parece fazer sentido apenas se for entendida como uma objeção direta à agregação. Ao juntar os ganhos e perdas de diferentes pessoas aos ganhos e perdas de um grupo, a afirmação é a de que estamos reduzindo as perspectivas de diferentes pessoas a uma única perspectiva impessoal, assim ignorando os limites entre estas pessoas (HIROSE, 2014, p. 72. Tradução minha.).

Estou marcando este ponto porque existem teorias agregativas que colocam *algum* peso na maneira em que o bem é distribuído. Considere, por exemplo, o igualitarismo teleológico

(PARFIT, 1997; HIROSE, 2014, p. 26-27). A teoria tem dois princípios: o princípio da igualdade, segundo o qual é *em si* ruim que algumas pessoas estejam em pior situação que outras; e o princípio da utilidade, segundo o qual é *em si* melhor se as pessoas estiverem em melhor situação, comparadas a estados de coisas alternativos (PARFIT, 1997, p. 204-205). O igualitarismo teleológico confere peso negativo à desigualdade, o que quer dizer que se importa com a noção de distribuição. No entanto, continua agregativa:

Claramente, o igualitarismo teleológico é não-utilitarista porque se importa com como o bem-estar das pessoas é distribuído. Entretanto, é também claramente agregativo. Ele junta o valor do bem-estar das pessoas e o desvalor da desigualdade em um valor numérico que representa a bondade relativa de estados de coisas (HIROSE, 2014, p. 27. Tradução minha.).

O igualitarismo teleológico confere algum valor à distribuição. Ainda assim, acredito que esta teoria não respeita a separação entre as pessoas da maneira correta. Ressalto este ponto para esclarecer que a crítica de Rawls não se aplica apenas ao utilitarismo, e sim, se interpretada de maneira ampla, a diferentes formas de agregação interpessoal. A ideia da separação entre pessoas é que não basta que a teoria confira *qualquer tipo* de peso à distribuição: uma teoria moral adequada deve ser justificável do ponto de vista de cada pessoa, sem eliminar este ponto de vista, reduzindo-o ao valor agregado final, o que reflete um tipo específico de preocupação com a distribuição. O igualitarismo teleológico não reflete a distribuição no sentido em que cada pessoa tem uma reivindicação moralmente relevante a algum nível de bem-estar, ele apenas levanta um princípio sobre como a desigualdade afeta o valor de estados de coisas agregados. Este modo de considerar a distribuição me parece insuficiente. A preocupação moral com a distribuição não pode estar fundamentada meramente na ideia de que o mundo é pior quando é desigual; alguma consideração sobre como ela afeta cada indivíduo separadamente, especialmente os que estão em pior situação, deve estar por trás da relevância moral da distribuição. Discutirei esta teoria em mais detalhes na próxima seção, mas trago aqui o assunto porque é por isso que a objeção se aplica a diferentes formas de agregação interpessoal: mesmo que um método diferente da soma seja utilizado, a perspectiva individual é sempre apagada pela ideia da combinação dos fatores morais de diferentes pessoas em um valor único atribuído a um estado de coisas. Acredito que o igualitarismo teleológico é uma teoria agregativa muito mais plausível que o utilitarismo clássico. O utilitarismo clássico sequer confere algum peso à distribuição. Mesmo assim, se a separação entre pessoas é uma consideração moralmente relevante, e agregação envolve uma combinação

interpessoal que apaga essa separação, então a agregação deve ser rejeitada no caso do igualitarismo teleológico também.

2.4 Respostas agregativas

A consideração de que a separação entre as pessoas importa do ponto de vista moral é uma objeção à noção de agregação, já que a agregação elimina essa divisão, considerando cada pessoa apenas indiretamente, através do valor final de um estado de coisas. Existem diferentes maneiras de reagir a esta objeção. Examinamos algumas respostas nesta seção.

A primeira estratégia que considero é a que acho menos promissora. Alguns utilitaristas apontam para o caso Jones como um contraexemplo não ao utilitarismo de modo geral, mas apenas ao hedonismo — uma teoria sobre valor segundo a qual a única coisa valorosa em si é o prazer. A ideia é a seguinte: o caso Jones revela a implausibilidade do hedonismo porque nos mostra que o sofrimento de Jones é mais relevante que os pequenos prazeres dos espectadores. O utilitarismo precisa, segundo este raciocínio, de uma análise do bem mais promissora. Considere, por exemplo, o utilitarismo negativo, avançado por David Pearce (1995). Segundo esta teoria, em vez da maximização do bem-estar, o utilitarismo deve buscar ações que *minimizem o sofrimento*. Existe uma assimetria entre prazer e sofrimento: o sofrimento importa *mais*³. Se aceitarmos o utilitarismo negativo, obtemos a resposta correta, já que é mais importante diminuir o sofrimento de Jones que aumentar o prazer dos espectadores.

Esta e outras sugestões de modificação na análise do bem não surtirão efeito. O caso Jones não revela uma falha na análise do bem utilitarista; revela uma falha estrutural na maneira em que o utilitarismo realiza a agregação moral. Segundo o utilitarismo, a distribuição daquilo que é valoroso é irrelevante do ponto de vista moral, e devemos nos preocupar apenas com a quantidade total de bem que é produzida pelas ações. O problema no caso Jones não é o sofrimento de Jones: é, sim, o fato de que existe uma desigualdade imensa produzida por um processo de avaliação das consequências de ações que ignora a perspectiva individual e os limites entre cada pessoa. Para

³ Exatamente qual é a relação entre prazer e sofrimento é matéria de discussão entre utilitaristas negativos. Existem muitas versões de utilitarismo negativo, mas especificá-las aqui seria improdutivo. Ver ORD, Toby, **Why I'm Not a Negative Utilitarian**, disponível em: <<http://www.amirrorclear.net/academic/ideas/negative-utilitarianism/>>, acesso em: 5 abr. 2021;

ver que a substituição de análise do bem não funciona, considere uma leve alteração do caso Jones que ofereci:

Caso Jones-negativo: Imagine que Jones sofre um acidente na sala de transmissão da Copa do Mundo. O equipamento cai em seu braço e ele começa a sofrer choques dolorosos. A única maneira de salvá-lo é interromper o equipamento que causa os choques. No entanto, se isso for feito, a transmissão do jogo será danificada, causando um pequeno desconforto nos olhos e ouvidos dos milhões de espectadores (OLIVEIRA, 2019a, p. 24-25).

Esta simples diferença na descrição do caso mostra que o problema não é resolvido pelo utilitarismo negativo: mesmo que seja preferível eliminar o sofrimento, se o utilitarismo negativo for entendido como agregativo, deve ter o resultado de que teremos menos sofrimento *no geral*, se ignorarmos Jones — ou, pelo menos, que existe, em princípio, um número de espectadores que faria ignorar Jones ser a ação correta. Para seguirmos em frente, quero mostrar que estas mudanças superficiais na análise do bem não serão efetivas usando distribuições neutras em relação à análise do bem⁴. Considere distribuições de bem $x = (20, 25, 30)$ e $y = (5, 5, 80)$. Independentemente do que uma determinada versão de utilitarismo considera como bem (se os números estão representando bem-estar, sofrimento evitado, ou outro tipo de estado), deve escolher y , que possui valor total maior que x (o valor de y , para o utilitarismo, é 90, enquanto o de x é 75). Essa preferência por distribuições desiguais, desde que aumente o valor total, é um problema que não passa pela definição de bem.

A segunda estratégia é mais popular, mas não acredito que tenha muito apelo. Se trata da rejeição de intuições. O raciocínio é que devemos rejeitar o papel de intuições morais sobre casos complexos como o caso Jones. Esta é uma maneira do utilitarista se defender das implicações implausíveis de sua teoria e pode ser notada. Um exemplo é J. J. C. Smart:

[...] o utilitarismo tem consequências que são incompatíveis com a consciência moral comum, mas tendo a aceitar a visão “pior para a consciência moral comum”. Isto é, estou inclinado a recusar a metodologia de testar princípios éticos gerais vendo como eles se adaptam a nossos sentimentos em casos particulares (SMART, J. J. C.; WILLIAMS, B., 1973, p. 68. Tradução minha.).

⁴ Como foi mencionado acima e será discutido em maior detalhe abaixo, a única modificação possível para evitar este problema seria incluir algum peso moral à própria distribuição de bem (que é o que faz, por exemplo, o igualitarismo teleológico). Argumento, no entanto, que mesmo uma teoria assim falha em considerar a separação das pessoas.

Outro é Peter Singer, em “Ethics and Intuitions”, sobre a relação entre princípios e intuições morais sobre casos particulares:

Rachels rejeitou a ideia de que o papel de filósofos morais é enxergar nossas intuições morais comuns como evidência, e tentar desenvolver uma teoria moral que se adeque a essas evidências. Ao contrário, ele mantém, deveríamos estar prontos para desafiar as intuições que primeiro nos surgem quando somos perguntados sobre uma questão moral. Esta é uma visão que compartilho com ele, e uma sobre a qual escrevi em diversas ocasiões [...] (SINGER, 2005, p. 332. Tradução minha.).

A sugestão parece ser que um certo tipo de raciocínio sobre teorias morais, conhecido como “equilíbrio reflexivo”, seja abandonado. O equilíbrio reflexivo, basicamente, é a ideia de que devemos construir teorias morais pesando princípios mais gerais e intuições sobre casos particulares, escolhendo aqueles que mais se adequam a nossos juízos ponderados sobre a justiça, fazendo ajustes nos princípios ou nos juízos, com o objetivo de entrar em “equilíbrio”, isto é, encontrar certa coerência entre os juízos particulares e os princípios. Nas palavras de Rawls,

É possível que se considere a teoria moral, a princípio [...], como a tentativa de descrever a nossa capacidade moral [...] o que é necessário é a formulação de um conjunto de princípios que, quando conjugados com nossas convicções e nossos conhecimentos das circunstâncias, nos levam a emitir esses juízos com os motivos que os respaldam, se tivermos de aplicar esses princípios de maneira consciente e inteligente. Uma concepção de justiça caracteriza nossa sensibilidade moral quando nossos juízos cotidianos estão de acordo com os princípios dessa concepção (RAWLS, 2016, p. 56).

Esta maneira de pensar sobre teorias morais é extremamente popular. Intuições têm, nas teorias filosóficas, um papel similar ao da evidência nas teorias científicas: devemos adequar as leis (princípios) à evidência (intuições), ou fornecer explicações que adequem a evidência às leis. Não acredito que a objeção de que não devemos basear nossas teorias morais em respostas psicológicas imediatas seja suficiente para derrotar a importância das intuições, já que não se trata de qualquer tipo de juízo, mas de juízos *ponderados*. Além disso, o fato de que intuições possuem *algum* papel na teoria moral não significa que devemos desistir imediatamente de qualquer princípio por causa de uma intuição: elas apenas podem contar contra o princípio, assim como uma intuição sobre um princípio pode também contar contra uma intuição sobre um caso particular. Podemos ver isso pelo fato de que o princípio de utilidade é intuído, embora não seja uma intuição *prática*, isto é, sobre casos particulares, é uma intuição *teórica*, o tipo de intuição por trás de princípios gerais (SANDBERG; JUTH, 2011, p. 213). Mas se os utilitaristas se fecham a qualquer evidência de um

problema com o princípio, não estão rejeitando intuições; estão simplesmente defendendo dogmaticamente uma intuição teórica:

Singer e Smart não estão rejeitando intuições, apenas preferindo uma (o princípio de utilidade) a outras (as que conflitam com o princípio) [...] o equilíbrio reflexivo é a melhor maneira de “desafiar” intuições morais, já que todos os juízos podem ser desafiados, ainda que os mais fortes tenham a tendência de não ser derrubados facilmente. O ponto do equilíbrio reflexivo é que nenhum juízo ou crença moral é intocável; o que Singer e Smart sugerem é que o princípio utilitarista seja pétreo — ironicamente, sugerem uma intuição que não pode ser questionada, justamente o que gostariam de rejeitar (OLIVEIRA, 2019a, p. 23).

Se devemos questionar nossas intuições morais, mesmo as mais profundas, então faz mais sentido escolher o equilíbrio reflexivo, que é capaz de questionar *qualquer* intuição, seja prática ou teórica, do que escolher a metodologia de adotar uma intuição teórica e ignorar qualquer consideração que vá contra esta intuição. É por isso que o equilíbrio reflexivo é um método tão importante na ética: o fato é que os princípios gerais (como o de utilidade) que informam as teorias morais são também intuições, então é melhor não se apegar demais a uma intuição específica.

Além disso, devemos levar em consideração que o argumento da separação normativa das pessoas não está fundamentalmente baseado numa intuição sobre o caso Jones. O caso Jones desperta uma intuição particular que ilustra um problema teórico estrutural no utilitarismo: o fato de que ele não confere peso moral à maneira em que o bem é distribuído. A objeção da separação das pessoas revela que a ideia de agregação envolve um apagamento dos limites de cada um e do fato de que devemos ser capazes de justificar princípios morais a cada pessoa, em vez de apenas escolher os princípios olhando para seu resultado agregado. Estas são considerações teóricas, apoiadas por intuições sobre casos particulares e princípios. Não se trata de rejeitar decisivamente uma teoria com base apenas em intuições sobre um contraexemplo; o uso das intuições, nesse caso, é mais sofisticado. Embora seja o caminho de resistência de muitos utilitaristas, rejeitar totalmente o papel das intuições na filosofia moral é improdutivo.

Por último, quero considerar duas teorias agregativas que conferem algum tipo de peso moral à distribuição: o igualitarismo teleológico e o prioritarismo. Alguém poderia reconhecer o valor da separação entre as pessoas mas argumentar que teorias agregativas deste tipo respeitam adequadamente a distinção. Acredito que essas são as teorias agregativas mais plausíveis que o

utilitarismo. Entretanto, existe ainda algum conflito entre estes princípios e a noção de separação entre as pessoas. Primeiro, vejamos como estas visões lidam com o caso Jones.

O igualitarismo teleológico apresentado por Parfit (1997) possui dois princípios, um de igualdade e outro de utilidade. O princípio da igualdade sugere que desigualdades são ruins em si. Diferentes versões de igualitarismo podem colocar um peso maior ou menor no desvalor da desigualdade — os valores exatos são irrelevantes aqui. O fato é que, segundo o igualitarismo teleológico, dependendo do valor da desigualdade entre o bem-estar de Jones e dos espectadores, devemos salvar Jones: basta que o desvalor da desigualdade, somado ao desvalor do sofrimento de Jones, supere o valor da utilidade produzida pelos minutos de partida que os espectadores assistiriam caso ignorássemos o acidente de Jones. É claro que é possível que o igualitarismo prescreva a ação contrária se o peso dado à desigualdade for baixo demais, mas o igualitarista pode apenas recomendar um peso maior ao desvalor da desigualdade.

O prioritarismo é a visão de que “beneficiar alguém importa mais quanto pior é a sua situação” (PARFIT, 1997, p. 213, Tradução minha.). Isso significa, basicamente, que aumentos do bem-estar das pessoas que estão em níveis mais baixos de bem-estar recebem algum tipo de prioridade. É possível construir uma função prioritarista segundo a qual a ação recomendada no caso Jones é interromper a transmissão: como Jones está em posição pior que os espectadores, ele deve ser salvo, porque benefícios a Jones importam mais. Novamente, como a prioridade dada àqueles que estão em pior situação não é absoluta, talvez o prioritarismo não defenda o salvamento de Jones se a prioridade for muito baixa. No entanto, os defensores do prioritarismo podem apenas sugerir uma prioridade maior.

Embora os princípios considerem de alguma maneira a relevância moral da distribuição e possam, em algum sentido, obter o resultado de que devemos salvar Jones, não acredito que a aplicação destes princípios a casos como esse seja satisfatório. Considere um caso Jones levemente modificado, no qual o bem-estar dos espectadores é muito pequeno, bem menor, inclusive, do que o bem-estar inicial de Jones. Nesse caso, se permitirmos a continuação da transmissão, o grande sofrimento causado a Jones, junto do pequeno benefício aos espectadores *diminuiria* a desigualdade da distribuição e, além disso, aumentaria a utilidade total. Segundo o igualitarismo teleológico, deveríamos ignorar Jones. Esta maneira de diminuir a desigualdade de uma distribuição, no entanto, parece pouco adequada, já que ela faz muito pouco para melhorar a

situação dos que estão mal e consiste, primeiramente, em piorar a situação de Jones. O mesmo acontece com o prioritarismo. Se o nível de bem-estar dos espectadores for mais baixo que o de Jones, benefícios a eles importam mais. Isso significa que existem configurações que permitem continuar a transmissão às custas de Jones, mesmo que o benefício a cada um dos milhões de espectadores seja muito pequeno, porque sua relevância aumenta consideravelmente por estar em um nível baixo, enquanto a relevância de Jones diminui consideravelmente por estar em um nível maior. Parfit (1997) considerou que uma implicação negativa do princípio de igualdade é o fato de que a diminuição das desigualdades é uma coisa boa em si, não importando a maneira em que é realizada, uma objeção conhecida como *Levelling Down Objection*. Minha modificação no caso Jones aponta para algo parecido com esta objeção. No entanto, a modificação também revela um problema no prioritarismo, que não sofre deste mesmo problema. O prioritarismo não confere valor intrínseco à igualdade; apenas confere prioridade aos níveis mais baixos. Mas em casos como na minha modificação do caso Jones, o fato de que os espectadores estão em níveis mais baixos não deveria fazer com que o sofrimento de Jones tenha valor semelhante ao dos espectadores apenas em razão do seu nível de bem-estar ser maior. Quando os dois fardos ou benefícios são iguais ou similares, é claro que o nível de bem-estar de cada pessoa influencia o valor dos fardos ou benefícios. Entretanto, nos casos Jones, o valor do fardo de Jones e o do benefício a cada espectador é tão diferente que conferir valor maior ao conjunto de pequenos prazeres dos espectadores que ao sofrimento de Jones neste caso modificado e não no caso original parece injustificado.

Depois de considerar porque não acho que igualitarismo teleológico e prioritarismo se saem muito bem no caso Jones, que é uma espécie de contraexemplo à agregação, quero voltar à questão da separação entre as pessoas, já que não pretendo basear toda a rejeição da agregação em um contraexemplo, especialmente porque as reações à última forma de contraexemplo podem ser variadas. Quero demonstrar porque o igualitarismo e o prioritarismo não conseguem dar conta da separação normativa das pessoas de maneira independente do caso Jones. O argumento é o seguinte: por causa de sua estrutura agregativa, o igualitarismo e o prioritarismo ignoram a separação das pessoas em casos em que o bem-estar inicial de dois indivíduos é igual e é oferecida uma aposta que pode beneficiar um às custas do outro. A falha está no fato de que estas apostas deveriam ser justificadas apenas se o benefício e o dano se aplicassem ao mesmo indivíduo, mas ambos igualitarismo e prioritarismo ignoram a distinção entre os casos de aposta individual e

interpessoal. Esta observação foi feita por Alex Voorhoeve, Marc Fleurbaey e Michael Otsuka (OTSUKA, 2012; OTSUKA; VOORHOEVE, 2009; VOORHOEVE; FLEURBAEY, 2012). Considere a seguinte comparação entre dois casos, adaptados de Otsuka, Voorhoeve e Fleurbaey.

Caso Intrapessoal: Diego e Rafael possuem um problema de dor na perna. Benhur pode impactar a situação da seguinte maneira: se intervir, Diego recebe 50% de chance de ser curado completamente, porém existe uma probabilidade de 50% de que seu problema piore levemente. A situação de Rafael não muda. Se Benhur não intervir, os dois continuam com o problema inicial.

Caso interpessoal: Diego e Rafael possuem o mesmo problema. Desta vez, se Benhur intervir, Rafael recebe 50% de chance de ser curado completamente, porém existe uma probabilidade de 50% de que o problema de Diego piore levemente, de maneira análoga à piora no primeiro caso. Se Benhur não intervir, os dois continuam com o problema inicial.

Suponha que a utilidade de viver com o problema inicial é 0.7, a de estar completamente curado é 1, e a de viver com o problema piorado é 0.6. Os estados de benefício e piora são representados, respectivamente, por “s1” e “s2”. Podemos construir as seguintes tabelas para ilustrar os dois casos:

Caso Intrapessoal:

Ação	Pessoa	s1	s2
Intervir	Diego	1	0.6
	Rafael	0.7	0.7
Não intervir	Diego	0.7	0.7
	Rafael	0.7	0.7

No Caso Interpessoal, temos:

Ação		s1	s2
Intervir	Diego	0.7	0.6

	Rafael	1	0.7
Não intervir	Diego	0.7	0.7
	Rafael	0.7	0.7

A intervenção de Benhur é justificada no caso Intrapessoal. Se ele intervir, pode se justificar para Diego dizendo que fez uma aposta racional em seu próprio benefício: como o benefício de ser completamente curado (0.3) é maior que o malefício de ter a situação levemente piorada (-0.1), e a probabilidade de cada evento é 50%, Diego não pode reclamar da razoabilidade da intervenção de Benhur, que considerou apenas o melhor interesse de Diego⁵. A mesma justificação não pode ser oferecida a Diego no que diz respeito à intervenção no caso Interpessoal. Isso acontece porque a chance de benefício no caso Interpessoal não é para Diego, e sim para Rafael. Ao decidir pela intervenção, Benhur estaria impondo um fardo em Diego que é diferente do fardo imposto no caso Intrapessoal, já que a chance de melhora é para Rafael.

O que a consideração destes dois casos revela é o seguinte: existe uma diferença moralmente relevante entre os dois, fundada na separação normativa de Diego e Rafael. Como os dois são pessoas separadas, o risco é justificado para Diego no caso Intrapessoal porque é racional de acordo com seus interesses; o mesmo não pode ser dito do caso Interpessoal. Mesmo se a intervenção de Benhur fosse justificada no caso Interpessoal, ela não poderia ter *a mesma* justificação que no caso Intrapessoal. Como Voorhoeve e Fleurbaey (2012, p. 388) apontam, em casos como esse, a chance de um benefício para Rafael não pode justificar a intervenção no caso Interpessoal, mesmo que a aposta envolva as mesmas probabilidades e os mesmos benefícios e perdas.

Teorias morais que ignoram a diferença de justificação entre os casos Interpessoal e Intrapessoal não são capazes de respeitar a separação entre as pessoas. Acredito que esse é o caso do igualitarismo teleológico. Esta forma de igualitarismo se importa com a quantidade de utilidade

⁵ Esta consideração se baseia na ideia amplamente aceita de que a atitude racional em casos de decisão sob risco, isto é, quando somos capazes de atribuir probabilidades aos estados relevantes, é a maximização da utilidade esperada. No entanto, o argumento não precisa que esta seja a teoria verdadeira sobre decisões deste tipo; o importante é imaginar que a diferença entre o benefício e o malefício é tal que seria racional, para Diego, aceitar a aposta.

e com o nível de igualdade produzido por uma ação. Compare as tabelas do caso Interpessoal e do caso Intrapessoal, no que diz respeito aos resultados da intervenção. Tanto a utilidade quanto a igualdade são iguais nos dois casos: se s_1 obter, uma pessoa está com 1 e outra com 0.7; se s_2 obter, uma pessoa estará com 0.6 e a outra com 0.7 (VOORHOEVE; FLEURBAEY, 2012, p. 390). Isto vale para os dois casos. Então, se a igualdade e utilidade finais são tudo que importa, como sugere o igualitarismo teleológico, não deve existir diferença entre os dois casos: se a intervenção é justificada no caso Intrapessoal, deve ser justificada no caso Interpessoal. Isto quer dizer que o igualitarismo teleológico não é sensível à separação entre as pessoas.

E o prioritarismo? Segundo esta teoria, benefícios àqueles que estão em níveis mais baixos importam mais. A situação inicial de Diego e Rafael é a mesma. Então, segundo o prioritarismo, benefícios a cada um dos dois importam da mesma maneira. Se isso é verdade, a intervenção é justificada nos dois casos da mesma maneira. Repare que esta versão de prioritarismo se importa apenas com “quão grande é a soma de utilidade pesada-atraves-de-prioridade que esta ou aquela linha de ação produz” (OTSUKA, 2012, p. 368. Tradução minha.). Isto quer dizer que, se duas pessoas estão inicialmente no mesmo nível de bem-estar, não importa se as chances de benefício e malefício que a intervenção proporciona são dadas à mesma pessoa ou a indivíduos distintos. Portanto, o prioritarismo também é insensível à separação entre as pessoas

É claro que estas considerações não se aplicam a todas as formas de igualitarismo e prioritarismo. É possível construir versões destas teorias que levam em conta a separação entre as pessoas⁶. Acredito que estas versões, no entanto, serão anti-agregativas, pelo menos no sentido de que devem incluir alguma limitação para a agregação interpessoal. A ideia geral de que podemos combinar fatores morais em um único valor e escolher ações baseadas neste valor é o que “borra” a separação entre as pessoas e leva à cegueira moral que produz os problemas e casos discutidos acima. Repare que tanto o igualitarismo quanto o prioritarismo consideram os indivíduos apenas *indiretamente*, através de seu impacto no valor agregado total. Adicionar cláusulas de igualdade ou prioridade aos menos favorecidos sem se livrar deste processo de combinação não vai resolver de maneira definitiva os problemas da agregação interpessoal. Embora as diversas versões de

⁶ O contratualismo do próprio John Rawls, por exemplo, é considerado por muitos uma forma deontológica de igualitarismo.

utilitarismo sejam exemplo dos usos menos adequados da agregação interpessoal na filosofia moral, não acredito que teorias como o igualitarismo teleológico ou o prioritarismo se deem muito melhor no que diz respeito à objeção de Rawls. A única maneira de produzir uma teoria que adequadamente reflita esta característica da moralidade, isto é, que respeite a separação entre as pessoas, é ou limitar a agregação interpessoal ou abandoná-la. No próximo capítulo, discuto o Contratualismo, uma teoria que abandona a agregação interpessoal por completo.

3 Contratualismo, moralidade e justiça

O objetivo deste capítulo é introduzir o contratualismo contemporâneo e explicar como (1) ele se estabelece como uma teoria moral não-agregativa e (2) como resolve os problemas normativos associados à ideia de agregação. Começaremos com uma breve exposição do contratualismo político de John Rawls, que pode ser encarado como a origem e a maior inspiração teórica do contratualismo moral contemporâneo. A versão de contratualismo desenvolvida por Scanlon será então explorada de maneira mais detalhada, com foco especial no fato de que é uma teoria *não-agregativa*, mostrando em que sentidos o contratualismo rejeita a agregação e quais são as vantagens normativas da teoria em relação às formas de consequencialismo discutidas anteriormente. Esclareço, logo após, as principais diferenças entre a teoria de Rawls e a de Scanlon, e a relação da teoria de Scanlon com as diferentes características estruturais da agregação interpessoal e as semelhanças entre contratualismo e princípios não-agregativos de escolha como Maximin e Leximin. Por último considero uma objeção à ideia de que o contratualismo pode rejeitar totalmente a agregação já que é possível, em princípio, que defenda alguns princípios agregativos; defenderei o contratualismo dessa objeção fazendo a distinção entre processos e princípios agregativos.

3.1 O contratualismo de Rawls na teoria da justiça: origem do contratualismo contemporâneo

O objetivo do segundo capítulo é apresentar o tipo de teoria não-agregativa que discutiremos no resto da dissertação: o contratualismo. Começamos com a teoria contratualista da justiça de John Rawls. A teoria de Rawls é política e, portanto, tem como objeto “a estrutura básica da sociedade”, isto é, “o modo como as principais instituições sociais distribuem os direitos e os deveres fundamentais e determinam a divisão das vantagens decorrentes da cooperação social” (RAWLS, 2016, p. 8). Esta restrição a difere de uma teoria moral mais abrangente, como as formas de consequencialismo discutidas no capítulo anterior, que tinham o objetivo de levantar princípios gerais sobre conduta moral — embora, naturalmente, o consequencialismo exista também na teoria da justiça; basta aplicar o princípio somente a instituições sociais. Apesar das diferenças de escopo, alguns dos insights de Rawls foram usados por T.M Scanlon no desenvolvimento do

contratualismo moral. Neste sentido, a teoria da justiça de John Rawls pode ser encarada como a principal origem do contratualismo moral contemporâneo.⁷

Rawls acredita que os princípios políticos que podem ser justificados são aqueles que agentes racionais escolheriam caso fossem ignorantes em relação a que posição ocupam na sociedade. Em *Uma teoria da justiça*, por exemplo, ele escreve que os princípios da justiça são “os princípios que pessoas livres e racionais, interessadas em promover seus próprios interesses, aceitariam em uma situação inicial de igualdade” (RAWLS, 2016, p. 14). Imagine partes decidindo quais são os princípios que as instituições sociais devem adotar. Agentes racionais e auto interessados escolherão princípios que maximizem a satisfação de suas preferências. Naturalmente, como os recursos são limitados e muitos objetivos conflitam, nenhum acordo poderia ser feito. No entanto, ao introduzir um elemento de imparcialidade, o *véu da ignorância*, isto é, a condição de que as partes escolhendo não conhecem sua posição social, o acordo se torna, para Rawls, possível. Isso acontece porque o véu da ignorância coloca, de certa maneira, todas as partes na mesma posição, o que faz com que a decisão racional seja a mesma para todas as partes. O filósofo chama esta situação de posição original, na qual “ninguém conhece seu lugar na sociedade, sua classe ou seu *status* social; e ninguém conhece sua sorte na distribuição dos recursos e das habilidades naturais, sua inteligência, força e coisas do gênero” (RAWLS, 2016, p. 14-15).

Existe muito debate sobre quais são os princípios que agentes racionais escolheriam se colocados sob o véu da ignorância. Este assunto não é muito importante para os propósitos deste capítulo, mas é interessante apresentar o raciocínio e os resultados de Rawls brevemente, para servir como um exemplo de raciocínio contratualista sobre a justiça. Uma vez que as partes na posição original são agentes racionais sem informação sobre sua situação na sociedade, estes agentes devem tomar uma decisão sob ignorância. Na teoria da decisão racional, decisões sob ignorância são escolhas que um agente deve fazer quando tem informação sobre suas alternativas e os respectivos resultados, mas desconhece e é incapaz de atribuir “[...] a probabilidade dos estados de coisas que correspondem aos resultados” (PETERSON, 2017, p. 41. Tradução minha.).

⁷ A versão da teoria de Rawls que será exposta aqui diz respeito principalmente à formulação de *Uma teoria da justiça*, principalmente porque é a versão da teoria de Rawls mais próxima da Filosofia Moral e a que mais diretamente está conectada ao contratualismo moral. Existem reformulações e atualizações na teoria política de Rawls a partir de *Liberalismo Político*, mas me abstenho dessa discussão porque o propósito de expor Rawls aqui é puramente explicitar a ligação entre o contratualismo rawlsiano e o scanloniano.

Filósofos trabalhando na teoria da decisão racional levantaram diferentes princípios para uma decisão racional em situações de ignorância. O princípio defendido por Rawls para casos como a posição original é conhecido como *maximin*, e postula o seguinte: agentes racionais devem “*maximizar* o valor *mínimo* possível de obtenção através de cada ato” (PETERSON, 2017, p. 44. Tradução minha.); isto é, os atos racionais são aqueles cujo pior resultado possível é melhor que os piores resultados de atos alternativos. O princípio também serve como uma regra para preferir distribuições de valor, exigindo a escolha daquelas distribuições que possuem o maior valor mínimo. Para ver maximin aplicado, examine este exemplo de decisão sob ignorância:

Imagine que você está escolhendo um sabor de pizza para pedir por tele entrega. Você conhece a pizzaria X porque já pediu uma vez a pizza marguerita deles, e achou uma pizza decente: nada especial, mas uma boa pizza. No entanto, antes de pedir a pizza, lembra que sua amiga Danielli comeu uma pizza de pesto da pizzaria X e recomendou fortemente, dizendo que é uma pizza espetacular. Você acha que uma boa pizza de pesto é muito superior à pizza marguerita que já experimentou, mas acredita também que uma pizza de pesto malfeita é muito pior e estragaria sua noite por estar gastando dinheiro com uma janta ruim. Suponha que você não seja capaz de dizer exatamente o quanto o gosto de sua amiga é próximo do seu, então não possui informação o suficiente para determinar uma probabilidade de obter uma pizza de pesto boa ou ruim.

Suponha que o valor da pizza marguerita é 4, o da pizza de pesto boa é 10 e o da pizza de pesto ruim é - 3. Suas alternativas são pedir a pizza marguerita (a_1) ou a de pesto (a_2). Assim, se, por exemplo, o estado de coisas s_1 (Pizza de pesto da pizzaria X é boa) obter, isto é, for verdadeiro, o valor do resultado da ação a_1 é 4, e da ação a_2 é 10. Uma matriz de decisão representando o problema seria algo como:

	Pizza de pesto da pizzaria X é boa (s_1)	Pizza de pesto da pizzaria X é ruim (s_2)
Pedir a pizza marguerita (a_1)	4	4
Pedir a pizza de pesto (a_2)	10	-3

Qual é a decisão racional nesta situação? Segundo o princípio *maximin*, você deve pedir a pizza marguerita, porque seu pior resultado possível (4) é melhor que o pior resultado possível de pedir a pizza de pesto (-3). Maximin ignora os valores máximos de cada alternativa, concentrando-

se sempre nos piores resultados possíveis. A intuição de que este princípio é racional vem do fato de que ele parece conferir alguma segurança ou certeza em decisões sob ignorância: diferente de decisão sob risco, onde sabemos as probabilidades envolvidas e podemos assumir riscos baseados no valor esperado, em ignorância o agente não tem informação suficiente para fazer apostas e, portanto, escolhe a opção que maximiza o menor valor, conferindo uma espécie de “certeza” sobre o resultado, como sugere Martin Peterson (2017, p. 46).

Rawls acredita que agentes racionais seguem *maximin* na posição original e, portanto, as partes na posição original devem escolher princípios que façam com que os menos favorecidos na sociedade estejam na melhor posição possível, se comparada à posição em que estariam sob outros princípios. Eis os dois princípios que ele defende que seriam escolhidos:

Primeiro: cada pessoa deve ter um direito igual ao sistema mais extenso de iguais liberdades fundamentais que seja compatível com um sistema similar de liberdades para as outras pessoas.

Segundo: as desigualdades sociais e econômicas devem estar dispostas de tal modo que tanto (a) se possa razoavelmente esperar que se estabeleça em benefício de todos como (b) estejam vinculadas a cargos e posições acessíveis a todos (RAWLS, 2016, p. 73).

O segundo princípio, conhecido como princípio da diferença, incorpora bem o espírito *maximin*, já que estabelece que desigualdades são aceitáveis *apenas se* melhorarem a situação também daqueles que estão nas piores posições da sociedade — lembre-se que *maximin* é um princípio que se concentra em maximizar o valor dos *piores* resultados; em termos de princípios de justiça, isto significa escolher princípios que melhorem o máximo possível a situação daqueles que estão na *pior* condição social e econômica⁸.

Há quem discorde do uso de *maximin* sugerido por Rawls. O economista John Harsanyi, por exemplo, sugere uma teoria moral contratualista com resultados alternativos. A teoria de Harsanyi começa muito semelhante à de Rawls. Segundo o economista, juízos morais sobre sistemas sociais possuem a característica marcante de que devem ser realizados de maneira em

⁸ Devemos ressaltar, no entanto, que o princípio da diferença não é implicado por *maximin*, já que *maximin* não confere valor nenhum a melhorias no padrão de vida daqueles que não estão na pior situação; esse detalhe teórico será discutido a frente quando demonstro a rejeição da ideia de Pareto por *Maximin*. O princípio de escolha mais parecido e adequado para a teoria de Rawls é uma versão estendida de *Maximin*, conhecida como *Leximin*, que será apresentado na seção 2.4.

que o agente pelo menos tente desconsiderar sua posição social. Quando fazemos um juízo moral, devemos assumir que temos a mesma probabilidade de ocupar qualquer uma das posições sociais disponíveis. Os princípios que seriam escolhidos por agentes racionais sob essa pressuposição são justos. Este modelo é o que Harsanyi chama de “modelo de equiprobabilidade dos juízos de valor moral” (HARSANYI, 1982, p. 45). A semelhança com o argumento da posição original é marcante, e reconhecida pelo próprio Harsanyi:

“[...] John Rawls [...] de maneira independente propôs um modelo muito similar, que ele chamou de “posição original”, baseado no “véu da ignorância” [...] a diferença não está na natureza dos dois modelos, que são baseados em pressuposições qualitativas quase idênticas (HARSANYI, 1982, p. 47. Tradução minha.)

Devo ressaltar que os modelos não são exatamente *idênticos*. A posição original é uma situação na qual os agentes *não conhecem* sua posição social. No modelo de equiprobabilidade, os sujeitos assumem que possuem a mesma chance de estarem em qualquer posição social. Esta diferença sutil será importante a seguir.

Harsanyi tira conclusões utilitaristas de seu modelo. Para ele, o princípio que deve ser escolhido pelos agentes racionais é o utilitarismo médio, segundo o qual devemos maximizar o bem-estar *médio* da sociedade. Nesta forma de utilitarismo, o bem-estar de cada indivíduo é agregado através de uma soma seguida de uma divisão pelo número total de indivíduos. Isto é, o valor de um estado $x = (w_1, w_2, w_3)$, onde w_i representa o bem-estar de i , é igual à soma de cada elemento dividida pelo número de elementos, isto é, neste caso, $v(x) = (w_1 + w_2 + w_3)/3$.

A diferença entre a teoria de Rawls e Harsanyi não está tanto na ideia de imparcialidade que está por trás dos modelos; está mais na caracterização da racionalidade dos agentes. Como vimos, Rawls acredita que os sujeitos racionais seguem, na posição original, a regra maximin para decisões sob ignorância. Harsanyi rejeita esta regra, afirmando que ela tem resultados altamente irracionais. Harsanyi oferece um caso para ilustrar o problema:

Suponha que você mora em Nova Iorque e lhe são oferecidos dois empregos ao mesmo tempo. Um é um emprego tedioso e mal pago na própria Nova Iorque, enquanto o outro é um emprego muito interessante e bem pago em Chicago. O problema é que, se você quiser o emprego de Chicago, você teria que pegar um avião de Nova Iorque para Chicago [...] Então haveria uma chance muito pequena mas positiva de que você morreria em um acidente de avião [...] O princípio maximin diz que você deve avaliar cada opção disponível em termos da *pior*

possibilidade que pode ocorrer a você se escolher essa opção [...] Consequentemente, se você quiser seguir o princípio maximin então deve escolher o emprego em Nova Iorque [...] Claramente, essa é uma conclusão altamente irracional. (HARSANYI, 1975, p. 595. Tradução minha. Grifo do autor).

O princípio de decisão sob ignorância sugerido pelo economista é conhecido como Princípio da Razão Insuficiente. A ideia do princípio é transformar decisões sob ignorância em decisões sob risco, atribuindo equiprobabilidade a cada estado da decisão, isto é, considerando cada estado igualmente provável. Como Peterson coloca, “O princípio da razão insuficiente prescreve que se alguém não possui razão para pensar que um estado do mundo é mais provável que outro, então probabilidades *iguais* devem ser atribuídas a todos os estados” (PETERSON, 2017, p. 54. Tradução minha.). Uma vez que as probabilidades são atribuídas, a decisão passa a ser uma decisão sob *risco*. Em casos deste tipo, o princípio quase unânime para decisão racional é o da maximização da utilidade esperada, segundo o qual devemos considerar, além das utilidades associadas a cada resultado, as probabilidades de cada estado. A maximização da utilidade esperada, dada a igual probabilidade de ocuparmos qualquer posição social, se torna, no modelo de equiprobabilidade de Harsanyi, idêntica à maximização da utilidade média (para ver isso, basta lembrar que a utilidade média é a utilidade total dividida pelo número de indivíduos). Assim, a teoria contratualista de Harsanyi chega ao utilitarismo médio.

Acredito que a teoria de Harsanyi não é uma construção adequada do contratualismo. Em primeiro lugar, sua rejeição do princípio maximin para a posição original é problemática. Recorde o caso dos dois empregos. Uma decisão deste tipo é vastamente diferente da decisão proposta por Rawls na posição original. No caso dos dois empregos, somos capazes de atribuir uma probabilidade ao estado que envolve o acidente de avião. Sabemos que a chance deste estado obter é baixíssima. Uma vez que possuímos esse conhecimento, a decisão se transforma de uma decisão sob ignorância em uma decisão sob risco. Neste caso, é bem plausível que a maximização da utilidade esperada seja racional. A decisão das partes, no entanto, não é assim: é uma decisão sob ignorância, na qual somos incapazes de atribuir probabilidades aos eventos. Maximin é um princípio justamente para este tipo de decisão. Imagine o caso trazido por Harsanyi *sem saber* que a probabilidade de acidentes de avião é baixa. Nesta situação, Maximin deixa de parecer tão irracional, claro. O caso trazido por Harsanyi não pode ser um contraexemplo a maximin porque

é um caso de decisão sob risco, e maximin é um princípio de decisão sob ignorância. Existe em Harsanyi uma confusão entre decisões sob risco e decisões sob ignorância.

Em segundo lugar, a estratégia sugerida pelo economista encontra problemas graves. Harsanyi aplica à posição original um princípio para decisões sob ignorância que *transforma* a decisão em um caso de risco (o princípio da Razão Insuficiente). Esse passo da argumentação deve ser explicitado, porque ignorá-lo faz parecer que a posição original envolve probabilidades desde o começo. A maximização da utilidade esperada só se aplica à posição original se o uso do princípio da Razão Insuficiente for adequado. No entanto, existe razão para duvidar desta adequação. O Princípio da Razão Insuficiente tem um problema central quando aplicado a decisão sob ignorância. Como Peterson aponta:

Se alguém não possui razão para pensar que um estado é mais provável que outro, parece estranho concluir que eles são igualmente prováveis [...] o problema é que parece completamente arbitrário inferir que todos os estados são igualmente prováveis. Qualquer outra distribuição de probabilidades parece ser igualmente justificada (isto é, nem um pouco). (PETERSON, 2017, p. 56, grifo do autor. Tradução minha.)

Rawls levanta o mesmo problema. Segundo ele,

[C]omo não há fundamentos objetivos na posição original para aceitar probabilidades iguais, ou, na verdade, de nenhuma outra distribuição de probabilidades, tais probabilidades são meras suposições [...] é como se continuássemos a usar argumentações probabilísticas [...] muito depois que as condições para seu uso legítimo tivessem sido eliminadas pelas circunstâncias da posição original (RAWLS, 2016, p. 214-215).

O princípio da Razão Insuficiente parece inadequado para a posição original. É claro que Maximin tem também seus problemas. Examine a tabela abaixo, que ilustra um problema de decisão adaptado de Peterson (2017, p. 46-47).

	s ₁	s ₂
a ₁	1	100

a ₂	1.1	1.1
----------------	-----	-----

O princípio Maximin sugere que devemos escolher a₂, já que seu pior resultado é melhor que o pior resultado de a₁. Naturalmente, essa consequência é indesejável para maximin se o princípio pretende ser o princípio verdadeiro para todas as decisões sob ignorância. Contudo, não acho que contraexemplos como esse derrubam a aplicação de maximin à posição original. Não existe nenhum tipo de consenso sobre qual princípio deve ser usado para decisões sob ignorância. Como Resnik sugere em *Choices*, seria mais proveitoso entender melhor como cada regra se aplica a diferentes tipos de decisão sob ignorância em vez de procurar um princípio verdadeiro que se aplica a todas as decisões (RESNIK, 1987, p. 40). A discussão não precisa ser “qual princípio melhor se aplica a todos os tipos de decisão sob ignorância?”; pode ser limitada a “qual princípio melhor se aplica à posição original?”.

Em *Uma teoria da justiça*, Rawls destaca três características que defendem a aplicação da regra maximin a decisões como a posição original. A primeira diz respeito ao uso de probabilidades, que não se adequa a esta decisão; A segunda é a consideração de que é muito mais importante, ao decidir arranjos sociais, garantir direitos básicos do que ter alguma chance de acumular grandes poderes ou riquezas. A terceira diz respeito à possibilidade (não probabilidade) de perder o que é mais importante, algo que suporta a preocupação com os piores resultados e sugere um processo conservador de decisão. Diz Rawls:

Em primeiro lugar, já que a regra não leva em conta as probabilidades das circunstâncias possíveis, deve haver algum motivo para que se descartem sumariamente as estimativas dessas probabilidades [...] seria irracional não ser cético em relação a cálculos probabilísticos [...] A segunda característica que indica o emprego da regra maximin é a seguinte: a pessoa que escolhe tem uma concepção do bem que a leva a preocupar-se muito pouco, quando muito, com o que possa ganhar acima [...] Para ela não vale a pena arriscar-se em nome de uma vantagem a mais [...] Esta última possibilidade apresenta a terceira característica: que as alternativas rejeitadas têm resultados que dificilmente seriam aceitos. A situação envolve sérios riscos (RAWLS, 2016, p. 188-189).

Acredito que a caracterização da posição original é tal que maximin é um princípio de escolha adequado — com certeza, mais que o princípio sugerido por Harsanyi. O contratualismo de Harsanyi encontra resultados utilitaristas em parte porque é mal construído. Seu modelo de equiprobabilidade, embora muito semelhante ao de Rawls, é inadequado para a construção de

princípios morais, porque ignora a diferença entre ignorância e equiprobabilidade, importante nesta discussão.

A teoria de Rawls é contratualista no sentido de que postula que os princípios que fazem uma sociedade ser justa o fazem *em virtude de* serem os princípios que podem ser encontrados por um procedimento baseado em um acordo na posição original — este acordo é o “contrato”. Isso quer dizer que, para Rawls, o que faz os princípios serem justificados é o fato de que eles seriam escolhidos na posição original, e não, por exemplo, o conteúdo substantivo dos princípios.⁹ Dito de outra forma: se Rawls adotasse, por exemplo, outro princípio para decisões racionais sob ignorância, é muito provável que sua teoria implicaria que outros princípios de justiça, como ilustrei ao apresentar a teoria de Harsanyi. Esta é uma distinção importante: a parte contratualista da teoria de Rawls não é seu uso de *maximin* ou os princípios de justiça que ele defende; é, sim, a afirmação de que os princípios justos o são porque podem ser aceitos por partes na posição original.¹⁰ A introdução de um procedimento para a justificação de princípios em uma teoria da justiça reflete o insight de que princípios morais devem poder ser aceitos por todos. Essa ideia está presente na valorização da separação normativa das pessoas, que figura na objeção de Rawls à agregação: no caso Jones, discutido no capítulo anterior, por exemplo, um princípio que permita ignorar o sofrimento de Jones não pode ser justificado *para Jones*. De certa maneira, a parte contratualista da teoria de Rawls pode ser resumida na ideia de que princípios justos são aqueles que podem ser aceitos por todos de maneira imparcial — a imparcialidade é garantida, segundo o filósofo, pelo véu da ignorância.

3.2 O contratualismo moral desde Scanlon

Inspirado pelos avanços de Rawls na teoria da justiça, o filósofo norte-americano T.M. Scanlon criou uma teoria moral centrada na ideia de que princípios morais devem ser justificáveis para cada um. Esta teoria é conhecida como “contratualismo moral”, e será o principal objeto de discussão desta dissertação. Segundo o contratualismo moral, ações moralmente corretas são

⁹ Embora, é claro, o conteúdo dos princípios seja o que faz com que sejam escolhidos pelas partes na posição original.

¹⁰ É por isso que considero a teoria de Harsanyi também uma forma de contratualismo, embora, como foi discutido acima, sofra de problemas de construção.

incorporadas por princípios aos quais ninguém pode apresentar objeções razoáveis. Como o autor coloca, em “Contractualism and Utilitarianism”:

Uma ação é moralmente errada [se e somente] se sua performance sob as circunstâncias não seria permitida por qualquer sistema de regras reguladoras do comportamento geral que ninguém poderia razoavelmente rejeitar como base de um acordo informado e não-coercitivo (SCANLON, 1982, p. 110. Tradução minha.).

Princípios que não podem ser razoavelmente rejeitados nada mais são do que princípios que podem ser justificados a cada um, separadamente. Assim como na teoria de Rawls, o que confere correção a um princípio moral não é algum fato sobre o conteúdo do princípio, mas o próprio fato de que não pode ser rejeitado de maneira razoável.

Talvez todas as teorias morais acreditem que atos moralmente errados são injustificáveis — mas a injustificabilidade costuma ser de segunda ordem. Considere, por exemplo, uma explicação utilitarista da imoralidade da agressão física por diversão. Um utilitarista diria que agressão física por mera diversão é imoral na maioria dos casos porque causa mais sofrimento às vítimas que prazer ao agressor, tendo, então, impacto negativo no balanço de bem-estar. O mesmo utilitarista afirmaria, então, que por ser imoral, esse ato é injustificável. O que Scanlon quer dizer, no entanto, é mais forte: ele identifica as propriedades de “ser moralmente errado” e “ser injustificável”. De certa maneira, o contratualismo inverte a relação entre justificabilidade e moralidade dos atos: em vez da imoralidade ser fundamento da injustificabilidade, a injustificabilidade é fundamento da imoralidade. A ideia de justificabilidade é básica para Scanlon, porque ela nos dá uma explicação normativa do “certo” e do “errado”, assim como é uma caracterização geral da moralidade: para contratualistas, a moralidade é um espaço de deliberação sobre razões que justificam (ou não) certas ações (SCANLON, 1998, p. 189). É isto que faz do contratualismo uma teoria que dá prioridade à noção de justificação. Como o próprio Scanlon coloca,

De acordo com o contratualismo, quando dirigimos nossas mentes a questões de certo e errado, o que estamos tentando decidir é, em primeiro lugar, se certos princípios são princípios que ninguém, se motivado da maneira correta, poderia razoavelmente rejeitar (SCANLON, 1998, p. 189. Tradução minha.).

Desta forma, o contratualismo não é apenas uma teoria sobre o significado dos termos morais básicos, mas também fornece uma explicação do conteúdo da moralidade. Scanlon não pretende afirmar que agentes morais sempre raciocinam desta maneira; afinal, o contratualismo moral não é uma teoria empírica sobre a psicologia dos seres humanos. O que a afirmação acima significa é que o conteúdo de nossos estados mentais, quando direcionados à moralidade, é este tipo de “decisão” sobre a justificabilidade de princípios que estão por trás de atos. Isto quer dizer que o conceito de justificação exerce papel fundamental mesmo quando os agentes não estão conscientes do fato de que estão pensando sobre a justificação.

A fórmula de Scanlon também sugere que o predicado moral fundamental para o contratualismo não é a correção moral, mas o erro moral. A rigor, a fórmula contratualista nos dá uma análise do que significa para uma ação ser “moralmente errada” (a partir da qual, é claro, inferimos uma análise da correção moral). Esta observação é relevante porque revela o fato de que, no contratualismo, ações moralmente corretas são permissíveis, mas não necessariamente obrigatórias (PETTIT, 2009, p. 78). As ações corretas são, no final das contas, apenas aquelas que não são erradas. Evidentemente, se todas as alternativas a uma ação são erradas, esta ação é correta e obrigatória. O contratualismo enxerga a moralidade, primeiramente, como um sistema de proibições a certos tipos de comportamento, em oposição a um sistema de encorajamento de certos tipos de comportamento.

É importante ressaltar que a teoria de Scanlon é restrita à moralidade interpessoal. Isto quer dizer que existem questões morais que podem estar fora da abrangência do contratualismo; discussões como o valor intrínseco da preservação da natureza e dos animais (isto é, seus valores separados de seu impacto na vida dos seres humanos) ou discussões sobre costumes e virtudes não dão origem ao tipo de consideração relevante para rejeição de princípios morais. O contratualismo moral pretende ser uma teoria sobre o sentido mais básico de “certo” e “errado” aplicados a seres com capacidade de justificação — não é, por exemplo, uma teoria sobre o significado de termos morais impessoais como “bem” ou “mal”.

O contratualismo moral scanloniano possui três características estruturais que, quando levadas em conta, oferecem critérios para julgar quais rejeições de princípios são razoáveis e quais não o são. As duas primeiras são restrições em relação a que tipos de razão podem ser usados para uma rejeição razoável. Em primeiro lugar, a *Restrição Personalista* diz que não podemos rejeitar

princípios morais com base em razões impessoais, isto é, razões referentes ao fato de que o princípio poderia causar resultados ruins sem considerar seu efeito em alguma pessoa (por exemplo, não podemos rejeitar princípios diretamente porque afetam negativamente uma rocha, árvore ou riacho; podemos, no entanto, rejeitá-los baseando-nos em como este efeito negativo impacta a vida das pessoas relevantes). A *Restrição Individualista*, por sua vez, dita que rejeições razoáveis só podem ser feitas com apelos a razões individuais, isto é, devemos rejeitar princípios baseados nas implicações que possuem para cada pessoa individualmente, sem fazer referência, por exemplo, a grupos de pessoas. O propósito das restrições é evitar que a agregação aconteça, garantindo que todas as razões utilizadas para rejeitar princípios digam respeito à justificação dos mesmo *a cada pessoa separadamente*. O efeito conjunto das duas postula, basicamente, que razões para rejeição razoável de princípios morais são sempre relativas a uma pessoa (Restrição Personalista), e não mais de uma pessoa (Restrição Individualista).

Além das duas restrições, o contratualismo moral conta com o *Princípio do Fardo Maior*, uma espécie de princípio para seleção de princípios. Uma vez aplicadas as restrições, o Princípio do Fardo Maior diz que devemos comparar as razões para rejeição uma a uma, escolhendo as mais fortes baseadas no tamanho dos fardos que cada princípio impõe àqueles que rejeitam as razões. Segundo Scanlon, “seria irrazoável, por exemplo, rejeitar um princípio porque ele impõe um fardo em você quando todos os princípios alternativos impõem um fardo muito maior em outros” (SCANLON, 2003, p. 132. Tradução minha.). Estas características fazem com que o contratualismo seja uma teoria moral concentrada nas razões daqueles que estarão em pior situação caso um princípio seja adotado (o que lembra a intuição da teoria da justiça de Rawls). Isto acontece porque graças à Restrição Individualista, olhamos apenas para razões individuais de cada pessoa separadamente, e graças ao Princípio do Fardo Maior, buscamos sempre aqueles que possuem as razões mais fortes para rejeitar um princípio.

Considere, por exemplo, a imoralidade de quebrar promessas em casos simples. Quando prometemos algo a alguém, damos a essa pessoa, na maioria dos casos, boas razões para rejeitar princípios que nos permitam quebrar essa promessa. O contratualismo deixa aberta a possibilidade de que *algumas* promessas possam ser quebradas, em casos extremos. Entretanto, de modo geral, aqueles a quem algo foi prometido têm razão mais forte do que os que pretendem quebrar uma promessa por mera conveniência, o que faz dos princípios que permitem a quebra de promessas

nessa situação razoavelmente rejeitáveis e, portanto, imorais. Suponha que você prometeu a um amigo que ia ajudá-lo a estudar para uma prova de matemática. Ele estava muito preocupado com o resultado da prova e talvez fosse reprovado se não alcançasse uma boa nota. No entanto, no dia em que combinaram estudar, você não cumpriu a promessa, baseando-se na consideração de que prefere assistir a um episódio de sua série favorita. O contratualista explica a imoralidade desta ação apontando para o fato de que sua razão para rejeitar um princípio que exige cumprir promessas é mais fraca que a de seu amigo para rejeitar um princípio que permita a quebra de promessas. Aplicando as Restrições Personalista e Individualista, ficamos com a reivindicação do cumprimento de uma promessa importante, do lado de seu amigo, e a reivindicação da conveniência, de seu lado. Pelo Princípio do Fardo Maior, a consideração de seu amigo é mais forte, o princípio que permite quebrar promessas pode ser razoavelmente rejeitado e, portanto, quebrar a promessa é uma ação imoral.

Por causa da maneira em que o raciocínio moral contratualista acontece, ele pode ser descrito como um “modelo de competição de reivindicações”. A ideia é simples: para decidir quais são as objeções razoáveis, primeiro separamos as reclamações através das duas restrições e depois aplicamos o Princípio do Fardo Maior, comparando as razões de cada pessoa para rejeitar os princípios relevantes. A objeção “vencedora” é a que tiver maior peso. Este procedimento de comparação de razões é o “contrato”.

Razões, segundo Scanlon, são considerações que contam a favor de algo. Esta visão é conhecida como fundamentalismo sobre razões, porque defende que razões são conceitualmente fundamentais, isto é, não podem ser reduzidas a outro conceito — como, por exemplo, desejos (SCANLON, 2014). Além disso, ele é um pluralista sobre valores, o que significa que acredita que existem múltiplos fundamentos para o valor, em vez de um valor-mestre — como é o caso do prazer no hedonismo ou do bem-estar em versões mais amplas de consequencialismo. Para Scanlon, então, as razões utilizadas no procedimento hipotético podem ser baseadas em muitos fatores diferentes, como, por exemplo, bem-estar, planos para o futuro ou respeito por seus direitos. No entanto, alguém poderia rejeitar o fundamentalismo de razões ou o pluralismo sobre valores e continuar contratualista. O contratualismo não é uma teoria sobre o conteúdo das razões ou sobre valores; é uma teoria sobre o que é certo e errado fazer no contexto interpessoal. Naturalmente, Scanlon acredita que estas visões sobre razões e valores se encaixam perfeitamente com o

contratualismo. Contudo, elas não são necessárias a um contratualista: ele pode preencher os critérios do que conta como uma razão para rejeição de diferentes maneiras. O essencial do contratualismo é a ideia de que a correção moral das ações é definida por quais princípios poderiam ser rejeitados em um acordo hipotético.

O raciocínio moral contratualista envolve um procedimento de competição entre razões para rejeitar princípios. Segundo a análise contratualista da moralidade, o fato de que alguns princípios podem ser razoavelmente rejeitados é o fundamento de sua imoralidade. Alguns filósofos se mostraram insatisfeitos com essa explicação. Afinal, não seriam as razões levantadas para rejeitar um princípio o verdadeiro fundamento da imoralidade de uma ação? Nas palavras de Blackburn,

[...] o papel do contrato diminui ainda mais se voltarmos nossa atenção para as tais razões para rejeição. Por que não supor que elas em si são o que fazem uma ação errada, eliminando qualquer apelo residual a contratos com outros? Suponha que é razoável rejeitar meus princípios porque, por exemplo, eles levam a grandes desigualdades em riquezas. Por que não é essa própria característica que faz meus princípios errados? (BLACKBURN, 1999. Tradução minha.).

O contratualismo scanloniano introduz, seguindo esta linha de críticas, um elemento desnecessário no raciocínio moral. Se as razões que usamos para rejeitar um princípio são boas, então é plausível que sejam elas diretamente que fazem a ação discutida ser imoral, em vez do fato de que o princípio pode ser rejeitado, o que faria o procedimento contratualista ser supérfluo. Esta objeção, levantada também por Philip Pettit (1993) e Colin McGinn (1999), ficou conhecida na literatura como “Problema da redundância”.

A discussão dos críticos sobre a redundância do contrato ignora um aspecto importante do contratualismo que já foi mencionado: as restrições — em especial, a Restrição Individualista. Para que as razões de rejeição fossem *diretamente* o fundamento da imoralidade de ações, elas precisariam ser razões morais. No entanto, as rejeições razoáveis de princípios não são feitas através de razões morais; segundo a Restrição Individualista, apenas razões *individuais* podem ser levantadas para rejeitar princípios. Razões individuais são considerações de um indivíduo relativas a seus próprios interesses. Razões morais não poderiam ser usadas em modelos de reivindicação como o contratualismo, já que são, em muitos casos, impessoais e imparciais.

Os contratualistas acreditam, é claro, que sua teoria nos informa sobre o que temos razão moral para fazer; por isso, em algum momento do raciocínio contratualista devem surgir razões morais. O papel essencial do contrato na teoria é justamente ser um mecanismo que transforma razões meramente individuais em razões morais, através da aplicação do Princípio do Fardo Maior e da Restrição Individualista. Como Michael Ridge comenta, para os contratualistas,

[Sua] reclamação relativa-a-um-agente de que minha ação lhe causaria séria dor física pode qualificar minha ação como errada mesmo se essa ação for necessária para prevenir um número enormemente maior de outras pessoas de sofrer uma dor menos severa [...] a ideia de Scanlon é que razão morais “pegam carona” nas razões relativas-a-um-agente das pessoas. De fato, o princípio contratualista simplesmente expressa a maneira em que razões morais “pegam carona” dessa forma. Desde que suas razões relativas-a-um-agente possam fundamentar uma rejeição razoável a qualquer princípio que permita minha ação, *este fato* me dá razão moral para não performar a ação [...] somente com a fórmula contratualista (ou um princípio de ponte deste tipo) nossas razões morais podem pegar carona nas razões relativas-a-um-agente dos outros. (RIDGE, 2001, p. 478. Grifo meu. Tradução minha.).

Segundo a análise contratualista da moralidade, então, não são as razões individuais (que Ridge chama de “razões relativas-a-um-agente”) que fazem uma ação ser errada, mas sim o fato de que essas razões podem “fundamentar uma rejeição razoável” aos princípios que permitam a ação. Este fato de alta ordem sobre quais princípios poderiam ser bem rejeitados por razões individuais nos dá razão moral para realizar ou não certos atos. Nas palavras do próprio Pettit, que mudou de ideia sobre a redundância do contratualismo, “o que importa para a ação ser errada [...] não são as diversas bases por trás das reclamações, mas o simples fato de que elas podem ser razoavelmente feitas”(2006, p. 81. Tradução minha.). Por ser uma teoria que identifica “certo” e “errado” com uma propriedade abstrata que depende de rejeições de princípios, o contratualismo precisa do “contrato”.

3.3 Algumas diferenças entre as teorias de Rawls e Scanlon

Embora as duas versões de contratualismo sejam baseadas na ideia de construir uma teoria normativa fundamentada na noção de justificação, algumas diferenças devem ser notadas entre o contratualismo de Rawls e o de Scanlon. A partir desta seção, quando usar o termo “contratualismo”, apenas, me refiro à versão scanloniana de contratualismo moral, que será a versão da teoria mais discutida nesta dissertação. Quando estiver comentando a teoria de Rawls, farei esta especificação.

A primeira diferença entre as duas teorias é, de certa maneira, superficial, mas deve ser feita. Está relacionada ao escopo dos dois tipos de contratualismo: a teoria de Rawls é uma teoria da justiça, isto é, uma teoria política, enquanto a teoria de Scanlon é uma teoria moral no sentido mais abrangente. Isto significa que enquanto o objetivo da teoria de Rawls é fornecer uma justificação de princípios políticos que se apliquem à estrutura básica da sociedade (suas principais instituições sociais), o contratualismo de Scanlon pretende fornecer uma análise do “assunto” da moralidade, isto é, do que estamos falando quando falamos de moralidade, bem como do significado dos termos morais *em geral* (SCANLON, 1982, p. 104-107). Devemos lembrar que a definição de ação errada fornecida por Scanlon faz referência a um “sistema de regras reguladoras do *comportamento geral*”(SCANLON, 1982, p. 110. Grifo meu). A teoria de Scanlon é, portanto, mais abrangente que a de Rawls, se aplicando não apenas à normatividade de princípios políticos e das instituições sociais, mas também a ações como mentiras, promessas, escolhas distributivas em pequena escala, relacionamentos, ética da saúde etc.

As diferenças número dois e três entre as teorias de Rawls e Scanlon são mais profundas e estão relacionadas. A segunda diferença é que o contratualismo de Rawls faz uso da noção de racionalidade, enquanto o de Scanlon aponta para a noção de razoabilidade. A terceira diz respeito ao fato de que o contratualismo de Rawls emprega um dispositivo de imparcialidade, o véu da ignorância, que não aparece na definição contratualista scanloniana. Examinemos abaixo as motivações e implicações dessa distinção.

A teoria de Rawls se apoia na ideia de escolha racional sob o véu da ignorância. Uma das ideias por trás desta construção é a de explicar o que faz os princípios serem justificáveis reduzindo termos morais/políticos a termos não-morais. Por isso usar a racionalidade (autointeressada) de sujeitos “filtrada” por um mecanismo de imparcialidade (que é o véu da ignorância, ou, no caso de Harsanyi, a pressuposição de equiprobabilidade) como modelo para juízos normativos do tipo relevante é atrativo: a teoria explica a normatividade moral/política em termos não-morais¹¹.

¹¹A noção de razoabilidade aparece na teoria da justiça de Rawls, mas em momentos posteriores ao argumento da posição original. Seu modelo da posição original depende da noção de racionalidade. É verdade que a relação entre as duas ideias é interessante, mas para os propósitos desta exposição, serve diferenciar os dois pela direção em que se concentram: a racionalidade considera, de forma direta, apenas razões do próprio agente para agir (Rawls busca este conceito de racionalidade na teoria da decisão), enquanto razoabilidade parece considerar, como comentário a seguir, as razões dos outros. Os cidadãos em uma sociedade bem ajustada de Rawls devem ser racionais e razoáveis, mas as partes na posição original precisam apenas ser racionais, porque o véu da ignorância “cuida do resto”.

Scanlon rejeita a ideia de que o contratualismo é bem construído através do modelo “escolha racional + mecanismo de imparcialidade”, por três motivos diferentes: (i) não deseja que seu procedimento contratualista pareça uma questão de estratégia; (ii) acredita que existem casos em que os princípios que seriam escolhidos nesse modelo são injustificáveis; e (iii) acredita que a moralidade de princípios e ações deve ser decidida fazendo referência às razões das pessoas a partir de sua perspectiva real, e não de uma perspectiva hipotética.¹²

Scanlon acredita que a construção de uma teoria moral baseada na ideia de racionalidade poderia passar a impressão de que a moralidade depende de qual seria a melhor estratégia para os agentes conseguirem algum tipo de acordo, já que a racionalidade é comumente definida como a melhor maneira de alcançar certos objetivos. Evidentemente, sabemos que a moralidade não pode ser simplesmente reduzida à racionalidade, já que existem muitas situações em que alcançar nossos objetivos de maneira eficiente seria imoral (considere, por exemplo, um sujeito que precisa passar em uma entrevista de emprego em que suas chances são muito baixas, e a melhor maneira de realizar o objetivo é assassinando seus concorrentes sem ser pego). É por isso que teorias como a de Rawls adicionam um mecanismo como o véu da ignorância, para que a escolha racional possa simular o que seria um juízo moral. Scanlon, por sua vez, prefere substituir a noção de racionalidade pela de “razoabilidade”. Segundo ele, esta é uma noção com um significado “muito mais próximo do que considero ser o pensamento moral básico” (SCANLON, 1998, p. 192. Tradução minha.). Scanlon não define “razoabilidade”, apenas apresenta exemplos em que razoabilidade e racionalidade se diferenciam. O papel que a razoabilidade encontra na teoria aponta para a ideia de que “razoável” é uma espécie de racionalidade que envolve considerar as razões das outras pessoas, além da do próprio agente. Podemos ver isso pelo Princípio do Fardo Maior, por exemplo, segundo o qual minhas objeções a um princípio são irrazoáveis se o fardo que o princípio impõe a mim é menor que o fardo que seria imposto a outras pessoas por princípios alternativos. Isso significa que, diferente de “racionalidade”, a razoabilidade não é auto-interessada: mesmo que também esteja preocupada com o que temos as melhores razões para fazer,

¹²Embora, devo ressaltar, exista cisão entre contratualistas sobre a aplicação desta característica em alguns casos. Em especial, existem contratualistas que acreditam que devemos introduzir um véu da ignorância em casos de decisão moral sob risco. No entanto, esta é uma construção do contratualismo aplicado a um tipo específico de juízo moral; no geral, o contratualismo moral não usa mecanismos como o véu da ignorância. Esta discussão será feita em maiores detalhes no capítulo 4, que é justamente sobre a aplicação do contratualismo moral a casos de risco.

a preocupação não aparece, como no caso da racionalidade, somente do ponto de vista da promoção dos objetivos do agente, mas do ponto de vista das razões de outros agentes individuais também. Segundo Scanlon, um juízo sobre a razoabilidade de uma rejeição a um princípio

[...] não é um juízo sobre o que teria a maior probabilidade de avançar [...] interesses ou produzir acordo em suas circunstâncias reais ou em qualquer situação idealizada, mas sim um juízo sobre adequação de certos princípios para servir como a base de reconhecimento e acomodação mútua (SCANLON, 1998, p. 194. Tradução minha.).

Esta descrição sugere que o conceito “razoável” possui algum conteúdo moral, o que Scanlon (1998, p. 194) reconhece. O contratualismo de Scanlon perde o apelo de definir a moralidade em termos não-morais, já que uma de suas noções centrais, a razoabilidade, não é explicada em termos não-morais. A reação de Scanlon a esta objeção é afirmar que seu contratualismo moral “[...] não procura explicar a noção [de razoabilidade]. A teoria tenta apenas descrevê-la claramente e mostrar como outras características da moralidade podem ser entendidas em termos desta noção” (SCANLON, 1982, p. 122. Tradução minha.). Acredito que a insatisfação de alguns filósofos com este tipo de resposta está relacionada ao fato de que costumeiramente existem expectativas reducionistas em nossas teorias morais, isto é, geralmente esperamos que uma teoria moral não inclua nenhuma noção moral em sua análise da moralidade, porque isto seria um tipo indesejável de circularidade. Acredito, no entanto, que esta preocupação não se aplica no caso do contratualismo moral. A teoria explica conceitos como “certo” e “errado” no sentido moral fazendo referência a uma noção com conteúdo moral, a razoabilidade, que é tomada como básica. Se a razoabilidade, no contratualismo, é uma noção moral básica, preocupada com a consideração das razões de diferentes indivíduos, ela não precisa ser reduzida em partes não-morais. Ainda assim, o contratualismo fornece uma análise da correção moral que não é circular, visto que a ideia de razoabilidade não faz referência ao que é certo ou errado diretamente. Usar noções com algum conteúdo moral para explicar outras noções morais não é, necessariamente, circular.

Espero que agora esteja claro porque eu disse que as duas diferenças estão relacionadas. Ao usar razoabilidade em vez de racionalidade como critério para seleção de princípios, Scanlon também exclui a necessidade de um mecanismo de imparcialidade como o véu da ignorância. Isso acontece porque a noção de razoabilidade já é uma espécie de racionalidade expandida às considerações alheias; isto faz com que ela já inclua algum tipo de imparcialidade, que era o papel do véu da ignorância. No caso da teoria de Rawls, o véu da ignorância é o que introduz algum

caráter moral na decisão das partes na posição original, já que ele garante, de certa forma, alguma imparcialidade e igualdade na decisão, porque transforma, do ponto de vista da escolha racional, todas as pessoas na mesma (isto é, cada um possui as mesmas informações e escolhem da mesma maneira, a maneira racional).

Por último, quero ressaltar uma distinção normativa. Scanlon acredita que existem princípios que seriam escolhidos por agentes racionais sob o véu da ignorância que, mesmo assim, seriam irrazoáveis. Como Hirose aponta, uma vez que as rejeições razoáveis são feitas do ponto de vista *real* das pessoas, alguns princípios que seriam escolhidos por agentes racionais podem acabar sendo injustificáveis para alguém (HIROSE, 2014, p. 86). Por exemplo, considere a escolha entre duas distribuições $x = (4, 4)$ e $y = (2, 8)$. O indivíduo 1 possui, segundo o contratualismo, razões para rejeitar um princípio que prefira y , como o utilitarismo médio, mesmo se a escolha do utilitarismo médio fosse racional sob o véu da ignorância. Isso vale também para distribuições como $x' = (5, 10)$ e $y' = (6, 120)$, na qual o indivíduo 2 possuiria rejeições razoáveis a princípios como maximin, que requerem a escolha de x' . É claro que o uso destas distribuições serve apenas como ilustração, já que os detalhes de cada processo de justificação podem envolver outros fatores. Mas o ponto aqui, que será retomado na seção 2.5, é que o contratualismo não pode ser reduzido a uma versão moral de processos de decisão racional e, por isso, rejeita alguns princípios que seriam escolhidos por pessoas racionais imparciais. É comum achar que, assim como o de Rawls, o contratualismo de Scanlon guarda uma relação muito próxima com o princípio Maximin, mas existe uma diferença na relação entre o contratualismo de Scanlon e aqueles que estão em pior situação. Segundo o autor,

No contratualismo, quando consideramos um princípio nossa atenção é naturalmente dirigida primeiro àqueles que estariam em pior situação sob este princípio. Isso porque se alguém possui fundamentos razoáveis para objetar ao princípio é *provável* que sejam estas pessoas. Não se segue, contudo, que o contratualismo sempre nos obriga a selecionar o princípio em que as expectativas dos que estão em pior situação são maiores. (SCANLON, 1982, p. 123. Grifo do autor. Tradução minha.)

Essa dissociação ocorre porque o contratualismo não depende de nenhum princípio de decisão racional, nem mesmo um princípio não-agregativo preocupado exclusivamente com os indivíduos que estão em pior nível. O contratualismo é uma teoria moral preocupada com a justificação de princípios para cada um, o que no caso da versão scanloniana envolve considerar a força das razões

de cada indivíduo, e não apenas sua situação. Na teoria de Rawls, em contraste, o que fundamenta a rejeição de princípios é uma decisão racional sob ignorância realizada através do princípio Maximin, centrado em melhorar a situação daqueles que estão pior (as razões daqueles que estão melhor, como o indivíduo 2 na comparação x' e y' , nem são consideradas). Essa diferença vem do que é a principal mudança entre a teoria de Rawls e a de Scanlon: a passagem de “racionalidade + véu da ignorância” para “razoabilidade”.

3.4 As vantagens de abandonar a agregação

Quando apresentei as duas restrições do contratualismo moral (Impessoal e Individualista), comentei que seu principal objetivo era excluir considerações agregativas da teoria. Como foi discutido no primeiro capítulo, teorias que adotam agregação total têm consequências moralmente indesejáveis no que diz respeito à distribuição de bem. Quando o único fator relevante para a avaliação de princípios e ações é o saldo total de bem gerado por suas consequências, a maneira em que o bem é dividido é ignorada, o que acaba sancionando distribuições altamente desiguais. Esta objeção foi levantada por Rawls em *Uma teoria da justiça* e Scanlon nos forneceu um caso ilustrativo para pensar o problema, conhecido como “caso Jones”. Lembremos do caso:

Suponha que Jones sofreu um acidente na sala de transmissão de uma estação de televisão. O equipamento eletrônico caiu no seu braço e não podemos salvá-lo sem desligar o transmissor por quinze minutos. Uma partida da Copa do Mundo de futebol está acontecendo, sendo assistida por muitas pessoas, e não vai acabar por mais uma hora. O machucado de Jones não vai piorar muito se esperarmos, mas sua mão foi esmagada e ele está recebendo choques elétricos extremamente dolorosos. Devemos salvá-lo agora ou esperar a partida acabar? A coisa certa a se fazer depende de quantas pessoas estão assistindo - se é um, dois ou cem milhões? (SCANLON, 1998, p. 235. Tradução minha.)

Se aceitarmos a agregação total, obteremos a resposta moralmente duvidosa de que existe um número de espectadores que faria ignorar o sofrimento de Jones a ação correta. Como foi discutido no primeiro capítulo, este problema vem da condição de Continuidade da agregação, segundo a qual qualquer consideração moral pode ser representada por uma função real, implicando que “para quaisquer dois fatores moralmente relevantes a e b , existe uma quantidade n de b que possui valor objetivo maior que a ” (HIROSE, 2015, p. 40. Tradução minha.). Se uma teoria satisfaz essa condição de maneira completa, se compromete com a ideia de que deve existir, ao menos em princípio, um número de pequenas frustrações dos espectadores que possua maior relevância moral que o sofrimento de Jones. Isso acontece porque, a partir do momento em que permitimos que o

valor moral do sofrimento de Jones seja convertido em um valor numérico na mesma escala que as frustrações dos espectadores, se o número de frustrações for suficientemente alto, a junção de todas as frustrações possuirá maior valor.

Ao rejeitar a agregação interpessoal, os contratualistas pretendem criar uma teoria que dê o devido peso à noção de distribuição justa. Uma rápida análise contratualista do caso Jones nos mostra que o contratualismo leva à resposta correta. Em primeiro lugar, aplicamos a Restrição Individualista, fazendo com que as razões consideradas sejam restritas a razões de uma pessoa apenas. Neste caso, a comparação será feita entre as razões de Jones para rejeitar um princípio que permite ignorá-lo e as razões de uma das pessoas que assiste à Copa do Mundo para rejeitar um princípio que obrigue interromper a transmissão. Depois disso, aplicamos o Princípio do Fardo Maior para decidir qual razão é mais forte. Naturalmente, nenhum dos espectadores está em posição de rejeitar o princípio que exige o salvamento de Jones, porque o fardo colocado em Jones por um princípio alternativo seria muito maior que o colocado em qualquer um dos espectadores individuais, uma vez que o sofrimento dos choques elétricos é mais oneroso que perder a transmissão por quinze minutos. Lembre-se que nenhum espectador pode apontar para o fato de que o salvamento de Jones afetaria milhões de pessoas, já que a rejeição de princípios exige apenas razões individuais.

A combinação de Restrição Individualista e Princípio do Fardo Maior dão ao contratualismo uma característica normativa importante: como a teoria se preocupa com aqueles que estão na *pior* posição dada a aceitação de um princípio, o contratualismo consegue explicar o papel da noção de distribuição justa de bem. Ao excluir a agregação total, a teoria evita os problemas mencionados anteriormente. O mesmo vale, é claro, para a teoria de Rawls: a aplicação do Princípio Maximin, unida a sua caracterização da posição original, garante que a distribuição de bens escolhida sempre beneficie os cidadãos menos privilegiados. Perceba que nenhuma das duas teorias implica, necessariamente, a escolha dos princípios (morais ou políticos) mais igualitários possíveis. Os princípios escolhidos são aqueles que melhoram a posição dos que se sairão pior. Considere, por exemplo, o princípio da diferença, defendido por Rawls:

[...] as desigualdades sociais e econômicas devem estar dispostas de tal modo que tanto (a) se possa razoavelmente esperar que se estabeleça em benefício de todos como (b) estejam vinculadas a cargos e posições acessíveis a todos (RAWLS, 2016, p. 73).

Repare que o princípio não diz que desigualdades sociais devem *necessariamente ser evitadas*, apenas que elas devem melhorar, de alguma forma, a vida de todos. O que isso quer dizer é que se uma desigualdade social ou econômica promove a melhora das condições dos cidadãos menos favorecidos, não existe motivo para rejeitá-la. Lembre-se que Maximin exige escolhermos distribuições nas quais obtemos o maior valor mínimo. Nesse sentido, uma distribuição desigual $Y = (4, 8, 100)$ é superior a uma distribuição igualitária como $Z = (3, 3, 3)$, já que 4, o valor mínimo de Y , é superior a 3, o valor mínimo de Z . O contratualismo de Scanlon apoia essa decisão, já que nenhum indivíduo teria, supomos, uma rejeição razoável à adoção de Y — lembre-se que os números não estão representando um bem específico, como dinheiro ou privilégios sociais, e sim a noção mais abstrata de valor. Isto faz com que a distribuição seja mais *justa*, embora menos igualitária: as desigualdades de Y são justificadas pelo fato de que todos estão em melhor posição em Y , especialmente, o que é essencial para a escolha de Y , os menos favorecidos. As desigualdades envolvidas em problemas da agregação, no entanto, como a do caso Jones, não apenas não melhoram a situação do menos privilegiado como ativamente a piora, sob o pretexto de aumentar o saldo de bem. O problema da agregação não é meramente um problema de desigualdade, é um problema de *distribuição*.

As duas formas de contratualismo discutidas são teorias não-agregativas da moralidade e da justiça, o que as permite dar um papel adequado à ideia de distribuição justa, o que coloca o contratualismo em vantagem normativa em relação a teorias que aceitam a agregação interpessoal total, como é o caso das formas dominantes de consequencialismo. Antes de discutir, nos próximos capítulos, as potenciais desvantagens de uma teoria moral sem agregação, gostaria de examinar melhor a relação entre o contratualismo e agregação, mostrando exatamente como o contratualismo rejeita a cláusula.

3.5 *Contratualismo, Maximin, Leximin*

Mencionei que o problema da separação normativa entre as pessoas surge da agregação moral, e que o contratualismo pode evitar este problema por ser uma teoria anti-agregativa. O contratualismo de Scanlon, no entanto, é compatível com três características estruturais da agregação, rejeitando apenas uma. Recordo a caracterização dos quatro pilares da agregação interpessoal feita por Hirose em *Moral Aggregation* (HIROSE, 2014, p. 31-41):

Comparabilidade interpessoal: é possível comparar os fatores moralmente relevantes de uma pessoa com os de outra.

Imparcialidade: duas alternativas são igualmente boas se elas diferem apenas no que diz respeito à identidade das pessoas.

Pareto: se uma opção x é melhor (mais justa/mais razoável, etc.) para alguém que outra opção y e é pelo menos tão boa quanto y para outras pessoas, x é melhor que y .

Continuidade: ordenamentos podem ser representados por um valor numérico real. Isto implica que, para quaisquer dois fatores moralmente relevantes a e b , existe uma quantidade n de b que possui valor objetivo maior que a .

Começo pela Imparcialidade. Embora o contratualismo exija que, no momento do contrato, as razões levantadas sejam individuais, e, portanto, em muitos casos, *parciais*, as razões propriamente morais resultantes do contrato são imparciais. O próprio Princípio do Fardo Maior aponta também nessa direção, já que rejeições, segundo este princípio, não podem ser razoáveis se impuserem um fardo muito maior a outra pessoa, mesmo que esteja rejeitando um princípio que nos impõe um fardo alto. Os procedimentos contratualistas possuem ferramentas de imparcialidade — como é o caso do Princípio do Fardo Maior, que ajuda a construir o lugar da razoabilidade na teoria de Scanlon, e do véu da ignorância, na teoria de Rawls — que garantem que os resultados normativos sejam também compatíveis com a Imparcialidade.

Comparabilidade Interpessoal pode parecer incompatível com o contratualismo moral, já que em algum sentido pode parecer que a teoria implica que é “impossível comparar”, no caso Jones, por exemplo, o sofrimento de Jones com o dos espectadores. No entanto, o que o contratualismo implica é que, uma vez comparados, percebemos que o sofrimento de Jones é muito maior. Comparar aqui não significa, como às vezes usamos no cotidiano, “equiparar”; a comparação é justamente um procedimento para explicitar as diferenças de relevância moral entre os fardos de diferentes pessoas. Na verdade, Comparabilidade Interpessoal é uma característica exigida pelo Princípio do Fardo Maior. Para enxergar o motivo, basta considerar que a aplicação deste princípio envolve *comparar* os fardos de diferentes pessoas para que possamos decidir qual razão é mais forte e, portanto, que rejeições são razoáveis. Se o contratualismo rejeitasse Comparabilidade Interpessoal, seria impossível decidir qual fardo é maior, já que o sofrimento dos

choques elétricos para Jones, por exemplo, não poderia ser comparado de maneira significativa à frustração de perder quinze minutos da transmissão para um espectador da Copa do Mundo; ficaríamos sem poder dizer qual fardo é maior.

Pareto não é exigido por nenhuma das características estruturais do contratualismo de forma tão direta quanto Imparcialidade ou Comparabilidade Interpessoal, mas acredito que é uma consequência do contratualismo moral. Pareto diz que uma opção é melhor que outra quando melhora a situação de alguém sem piorar a situação de ninguém. Suponha que um princípio G leve a uma distribuição de valor $g = (5, 5, 8)$, enquanto um princípio H entrega uma distribuição de valor $h = (5, 7, 8)$. Segundo o contratualismo, o sujeito que teria sua parte do valor aumentado de 5 para 7 sob o princípio H pode rejeitar razoavelmente o princípio G , e nenhuma das outras duas partes possui fundamento para rejeitar o princípio H , já que o fardo do sujeito que deixa de ganhar valor é, naturalmente, maior do que o prejuízo inexistente para as outras partes. É claro que a discussão em termos de distribuição de valores é bastante abstrata e pode ficar complicada se substituirmos *valor* por algo como dinheiro, já que é possível que, em alguns casos, a mudança de 5 para 7 milhões de reais, por exemplo, possa criar algum tipo de situação onerosa para as outras partes. Entretanto, nos casos em que a situação de alguém é piorada, de qualquer forma, pela melhora de outro, Pareto não se aplica. A consideração dos princípios G e H revela que o contratualismo sempre segue Pareto, já que se um princípio, comparado a suas alternativas, melhorar a situação de alguém sem piorar a dos outros, este alguém sempre poderá rejeitar razoavelmente as alternativas deste princípio.

O contratualismo, no entanto, rejeita Continuidade. Isso acontece porque, da maneira em que o contratualismo é formulado, ele confere status moral especial às razões daqueles que serão mais fortemente afetados de maneira negativa pela adoção de um princípio. Por causa da combinação de Restrição Individualista e Princípio do Fardo Maior, não importa quantas pessoas possuem razões para rejeitar um princípio se estas considerações são mais fracas que as de alguém que seria afetado de maneira mais profunda. Este resultado conflita com Continuidade, que exige que os dois tipos de consideração possam ser representados por um valor numérico real. As considerações daqueles que estão em pior situação e as dos outros são descontínuas porque não podem ser convertidas para um valor numérico real, o que implicaria que para cada consideração forte exista um número possível de considerações mais fracas que ultrapasse a consideração forte

em valor (e, junto da aceitação de Pareto, implica que este valor deve incluir o bem-estar dos espectadores, além da pessoa que está na pior situação)¹³. Se o contratualismo for verdadeiro, não existe, nem mesmo em princípio, um número de frustrações dos espectadores da Copa do Mundo que ultrapasse a força da razão de Jones para rejeitar um princípio que envolva manter a transmissão. É por causa de sua rejeição de Continuidade que o contratualismo é anti-agregativo.

Naturalmente, a rejeição de continuidade não significa que o contratualista acredite, necessariamente, que *todas* as razões são descontínuas entre si. É possível, por exemplo, que ao comparar fardos não exatamente idênticos, mas razoavelmente próximos, seja possível para o contratualista aceitar a continuidade destas razões, desde que os fardos sejam suficientemente próximos. Nesta questão, existe uma disputa interna entre os contratualistas. Basta aqui dizer que os contratualistas estão de acordo quando os fardos são tão diferentes que razões envolvendo os princípios relevantes são descontínuas, isto é, não podem ser convertidas em valores numéricos da mesma escala. Aceitar a condição de Continuidade significa aceitar que todo fator moralmente relevante pode ser convertido em valores de uma mesma escala — em teorias consequencialistas, por exemplo, é uma escala que representa a “bondade” dos estados de coisas trazidos pelas consequências de um ato ou distribuição. Qualquer consideração pode entrar nessa escala, o que significa que um número suficientemente alto de qualquer tipo de consideração pode ser mais forte que qualquer outra consideração.

Na seção sobre a teoria contratualista de Rawls, comentei que seu procedimento para construção dos princípios da justiça pode ser encarado como uma combinação entre a posição original e um princípio de decisão racional sob ignorância, conhecido como *maximin*, segundo o qual devemos escolher atos que possuam o maior valor mínimo. Junto do véu da ignorância, a regra *maximin* parece nos levar a resultados muito parecidos com os do contratualismo contemporâneo desenvolvido por Scanlon. *Maximin* é também anti-agregativo: não importa quão alto seja o valor máximo de algum resultado se o ato relevante tem um valor mínimo menor que alguma alternativa. Para enxergar isto, considere o caso da escolha entre pizzas da seção 2.1. Não

¹³ Faço essa distinção porque *Maximin*, que rejeita Pareto, pode aceitar Continuidade, isto é, a conversão dos fatores moralmente relevantes em um valor numérico, sem se complicar com o caso Jones, já que o valor numérico que representa o valor de um estado será apenas o do indivíduo com o menor bem-estar (neste caso, Jones) e as melhoras para outros indivíduos (nesse caso, os espectadores) serão ignoradas.

importa o quão alto o valor máximo de pedir pizza de pesto seja, maximin sempre recomendará a pizza marguerita, já que o valor mínimo de pedir a pizza de pesto é menor. Isto se aplica também ao caso Jones: em situação de ignorância, isto é, sem saber qual a probabilidade de ser um espectador ou Jones, a escolha exigida pelo princípio é salvar Jones independentemente da quantidade de espectadores, porque o valor mínimo de ignorar o sofrimento de Jones é baixíssimo no estado do mundo em que *você* é Jones.

Dito isso, o contratualismo de Scanlon difere de maximin de duas formas importantes. Em primeiro lugar, Maximin rejeita Pareto. Definimos a regra maximin como recomendando a maximização do valor mínimo de cada ato. A regra, portanto, ignora todos os valores não-mínimos. Em casos como o da escolha da pizza este fato pode não parecer tão relevante, mas considere uma escolha entre as distribuições $R = (2, 3, 3, 500)$ e $S = (2, 3, 3, 4)$. Como Maximin concentra-se apenas no valor mínimo, ele é indiferente entre as duas. Este não é o caso do contratualismo, já que, como foi mostrado acima, aqueles que estão em pior situação em S que em R podem sempre rejeitar S razoavelmente, já que esta rejeição não coloca um fardo maior em outros.

Além disso, embora Maximin seja não-agregativo, ele não rejeita continuidade. Continuidade é uma propriedade que permite que o valor de distribuições seja representado por um número real. Como Iwao Hirose aponta em *Moral Aggregation*, “[...] Maximin ranqueia estados de coisas, e a bondade relativa é representada de tal maneira que um estado de coisas $x = (w_1, w_2, \dots, w_n)$ é pelo menos tão bom quanto outro estado de coisas $y = (w'_1, w'_2, \dots, w'_n)$ se e somente se $\min(w_1, w_2, \dots, w_n) \geq \min(w'_1, w'_2, \dots, w'_n)$.” (2015, p. 41. Tradução minha.). Isto quer dizer que o valor de uma distribuição (3, 10, 20) para Maximin pode ser representado como $v(3,10,20) = 3$, já que 3 é o valor mínimo da distribuição — uma teoria agregativa como o utilitarismo, por exemplo, diria que $v(3,10,20) = 33$, a soma total. Isto funciona porque o foco no valor mínimo que Maximin exige é preservado por este tipo de representação: Maximin prefere (4, 8, 10) a (2, 3000, 4000) já que, segundo Maximin, $v(4, 8, 10) = 4$ e $v(2, 3000, 4000) = 2$. A capacidade que Maximin possui de atribuir um valor numérico real a distribuições mostra que é um princípio contínuo. Repare que o mesmo não pode ser dito do contratualismo. O contratualismo implica, por exemplo, que uma distribuição $x = (3, 10, 20)$ pode ser razoavelmente rejeitada se comparada a uma distribuição $y = (5, 8, 9)$. No entanto, isso não quer dizer que o contratualismo

pode representar o valor de x como $v(3, 10, 20) = 3$, mesmo que 3 seja o valor mínimo. Para ver isso, basta comparar duas distribuições com o mesmo valor mínimo. Para o contratualismo, o valor de $(3, 10, 20)$ não é igual ao valor de $(3, 50, 100)$. Nem, tampouco, é algum tipo de soma dos valores. O contratualismo pode apenas ordenar as distribuições, sem atribuir valores numéricos reais.

Existe um princípio que estende Maximin aos outros valores de uma distribuição nesses casos em que o valor mínimo é o mesmo, chamado Leximin. Segundo esta regra, devemos maximizar o valor mínimo obtível, mas “se os piores resultados foram iguais, deve-se escolher uma alternativa em que o *segundo pior* resultado é tão bom quanto possível [...] e assim por diante” (PETERSON, 2017, p. 45. Tradução minha.). Leximin exige a escolha de R em vez de S porque 500, o quarto pior valor de R , é superior a 4, o quarto pior valor de S , e os outros piores valores estão empatados. Naturalmente, isto significa que Leximin aceita Pareto, já que qualquer aumento de valor que não diminua os outros em relação às alternativas faz com que a distribuição com valor aumentado seja preferida. Leximin rejeita continuidade, já que dá prioridade aos piores resultados de maneira que não exista um número ou tamanho de melhores resultados que faça diferença na escolha a não ser que os piores resultados estejam todos empatados. Então, assim como o contratualismo, Leximin aceita Imparcialidade, Comparabilidade Interpessoal e Pareto; mas é não-agregativo, pois rejeita Continuidade.

Isso não significa que o contratualismo possa, necessariamente, ser reduzido a uma versão moral de Leximin. Como reparei na comparação entre contratualismo e a regra maximin, ressalto uma diferença importante: embora o contratualismo se assemelhe a Leximin porque confere algum tipo de consideração especial daqueles que estão em pior situação, o contratualismo é uma teoria concentrada na maneira em que as ações afetam cada pessoa; Leximin, por sua vez, concentra-se no nível de bem-estar geral dos menos favorecidos. Recorde a modificação do caso Jones, que discuti no primeiro capítulo, na qual os espectadores estão em um nível de bem-estar baixíssimo, menor que o de Jones. Como Leximin confere significado apenas àqueles que estão na pior situação (salvo empates, que requerem consideração do segundo pior, e assim por diante), este princípio recomenda ignorar Jones nesse caso, já que os benefícios aos espectadores, desta vez em pior situação que Jones, são o que importa. O contratualismo, no entanto, não acredito que leva ao mesmo resultado, já que a reivindicação dos espectadores não pode ser mais relevante que a de

Jones no que se refere *a este ato específico*; mesmo que estejam em níveis mais baixos de bem-estar *no geral*, suas reivindicações não são mais fortes que a de Jones em relação à escolha entre parar ou continuar a transmissão, já que esta decisão afeta o bem-estar de Jones muito mais que o de cada espectador individualmente. O contratualismo moral não é uma teoria que simplesmente privilegia aqueles que estão em nível pior no geral. É, mais especificamente, uma teoria que confere maior peso às razões daqueles que serão *mais afetados* por uma ação ou princípio. É claro que, em muitos casos, as duas tendências coincidem, mas situações como o caso Jones modificado podem apontar para a diferença. Volto a citar uma passagem em que essa característica do contratualismo fica evidente:

No contratualismo, quando consideramos um princípio nossa atenção é naturalmente dirigida primeiro àqueles que estariam em pior situação sob este princípio. Isso porque se alguém possui fundamentos razoáveis para objetar ao princípio é *provável* que sejam estas pessoas. Não se segue, contudo, que o contratualismo sempre nos obriga a selecionar o princípio em que as expectativas dos que estão em pior situação são maiores. (SCANLON, 1982, p. 123. Grifo do autor. Tradução minha.)

Nesta questão, o contratualismo é distinto de Leximin.¹⁴

3.6 Teorias agregativas, princípios agregativos e a separação entre as pessoas

No capítulo anterior, argumentei contra a noção moral de agregação, com base na separação normativa entre as pessoas. Neste capítulo, apresentei duas versões de contratualismo, defendendo que resolvem o problema da separação ao rejeitar a agregação. O contratualismo é, como disse, *não-agregativo* ou *anti-agregativo*, em oposição a teorias como o utilitarismo clássico, que satisfazem a agregação. Gostaria de examinar uma última distinção que diz respeito à maneira em que o contratualismo evita a agregação.

O contratualismo moral é uma teoria que define a moralidade de ações referindo-se à rejeitabilidade (ou, no caso de Rawls, aceitabilidade) de princípios. Contratualistas recorrem ao resultado de um procedimento de escolha hipotético (seja a posição original de Rawls ou o teste de rejeitabilidade de princípios de Scanlon) para definir o status destes princípios. Mas repare que não existe nada intrínseco ao procedimento que exclua a escolha de princípios agregativos. Os

¹⁴ As mesmas considerações deste parágrafo valem, é claro, para a diferença entre Contratualismo e Maximin, que foi comentada na seção sobre as diferenças entre o contratualismo de Rawls e o de Scanlon.

princípios adotados por Rawls são resultado da posição original e da regra maximin para decisões sob ignorância; os de Scanlon são resultado do Princípio do Fardo Maior e das Restrições Individualista e Personalista. É verdade que estes princípios não são agregativos. Entretanto, alguém poderia manter a ideia geral do contratualismo, isto é, a ideia de que a moralidade de ações e princípios é determinada por sua aceitação ou rejeição em determinadas circunstâncias, e mesmo assim acabar com princípios agregativos.

De fato, o economista John Harsanyi oferece uma teoria deste tipo. Seu modelo de equiprobabilidade, muito semelhante ao de Rawls, deve respeitar a separação entre as pessoas tanto quanto qualquer outra forma de contratualismo, já que é baseado na escolha oferecida a cada pessoa individualmente. Suas consequências morais, contudo, envolvem um princípio agregativo (o princípio da utilidade média). Nesse ponto, o contratualista parece estar em um dilema: deve assumir ou que a separação das pessoas não é uma objeção à agregação ou que o contratualismo, por permitir princípios agregativos, não respeita a separação entre as pessoas. Iwao Hirose sugere uma objeção deste tipo. Segundo ele, como fica evidente pela teoria de Harsanyi, se procedimentos contratualistas respeitam a separação, então a separação entre as pessoas não nos leva necessariamente a princípios não-agregativos, o que sugere que não seja um bom motivo para rejeitar a agregação moral (HIROSE, 2014, p. 83-84). Por outro lado, se a separação for um motivo deste tipo, então o contratualismo não a respeita, já que pode levar a princípios agregativos.

Acredito que a separação das pessoas como razão para rejeitar a agregação, bem como o contratualismo como uma teoria não-agregativa no sentido relevante, podem ser salvos se feita uma distinção entre teorias agregativas e princípios agregativos. Considere, por exemplo, o utilitarismo clássico, a teoria de que ações moralmente corretas são as que maximizam o saldo geral de bem-estar. Mostrei, no primeiro capítulo, que o utilitarismo clássico satisfaz cada uma das quatro características estruturais da agregação. A maneira em que a teoria *define* ações moralmente corretas é agregativa: no utilitarismo clássico, a moralidade das ações *depende diretamente* do resultado do processo agregativo. Esse uso da combinação do bem-estar de diferentes pessoas em um único valor numérico faz com que a teoria fira a separação das pessoas, como argumentei no primeiro capítulo. O utilitarismo clássico é uma *teoria* agregativa da moralidade, nesse sentido, porque sua definição da moralidade envolve a agregação.

Essa classificação não parece se aplicar à teoria de Harsanyi. Como o processo de justificação de princípios e, por consequência, ações é muito semelhante ao de Rawls, a diferença entre a teoria de Harsanyi e o utilitarismo clássico deve ficar clara facilmente. A ideia da teoria de Harsanyi é que os princípios morais adequados são aqueles que os sujeitos racionais escolheriam na situação de equiprobabilidade. A teoria está preocupada, portanto, com a maneira em que estes princípios podem ser justificados a *cada indivíduo*, e é o fato de que poderiam ser justificados que fundamenta sua correção moral. Esta definição da moralidade é bastante diferente da definição utilitarista, por exemplo, e de qualquer definição de uma teoria agregativa, já que para teorias agregativas o que importa é o valor numérico agregado que representa um estado de coisas em que a ação ou princípio foi adotado. A teoria da moralidade de Harsanyi respeita a separação das pessoas porque é capaz de oferecer uma justificação a cada indivíduo. É claro que, apesar disso, Harsanyi acaba com um princípio agregativo, o utilitarismo médio. Mas isso não faz da teoria de Harsanyi uma teoria agregativa no sentido relevante desta discussão. Harsanyi chega a um princípio agregativo por causa da aplicação de um método de decisão racional duvidoso para decisões deste tipo, e essa má formulação da racionalidade de cada indivíduo distorce seu contratualismo fortemente, ao ponto de que ele acaba com uma versão utilitarista de contratualismo. Contudo, não acredito que ele possua uma visão agregativa da moralidade, porque pensa este domínio como um espaço daquilo que pode ser justificado a cada um. Como comentei acima, existem razões para rejeitar esta teoria, mas a separação das pessoas, em meu ver, não é uma delas.

O mesmo vale para os contratualismos de Rawls e Scanlon. Sua definição da moralidade não é agregativa. São teorias baseadas na justificação de princípios a cada pessoa (tanto a posição original quanto as restrições de Scanlon são artifícios que impedem um processo agregativo de justificação). Mesmo se as teorias recomendarem algum princípio agregativo, sua justificação não será agregativa e, portanto, respeita a separação das pessoas no sentido relevante. O próprio Scanlon, discutindo a relação entre o caso Jones e os casos em que a agregação parece plausível¹⁵, sugere que o contratualismo é capaz de considerar a separação mesmo sendo possível que aceite alguns princípios agregativos (SCANLON, 1998, p. 235). Isto não faz do contratualismo uma

¹⁵ Como, por exemplo, o caso mencionado no capítulo 1, trazido por Taurek (1977), que envolve a escolha entre salvar uma ou cinco pessoas.

teoria agregativa, porque não define a moralidade em termos de resultados da agregação; é, simplesmente, uma teoria moral não-agregativa que permite alguns tipos de princípio agregativo.

Após argumentar contra a noção moral de agregação, no segundo capítulo, e sugerir uma teoria moral não-agregativa, o contratualismo moral, no terceiro capítulo, discutirei, nos capítulos seguintes, dois problemas associados ao contratualismo moral na literatura. O capítulo a seguir é dedicado ao problema dos números. O objetivo será oferecer uma defesa contratualista para este problema.

4 O problema dos números

No segundo capítulo, foi apresentada uma discussão sobre os problemas da noção moral de agregação. No terceiro, introduzi o contratualismo, uma teoria moral não-agregativa que se pretende superior a suas contrapartes agregativas. Eliminar a agregação de sua análise da moralidade permitiu ao contratualismo resolver as dificuldades associadas à noção. Existem, no entanto, problemas sérios em teorias morais não-agregativas. Neste capítulo, discuto um dos principais: o problema dos números. Começo com uma exposição deste problema para depois apresentar possíveis saídas para o contratualismo. Concluo que, ainda que existam dificuldades em algumas das estratégias oferecidas na literatura, existe esperança em diferentes tipos de resposta não-agregativa ao problema dos números que são compatíveis com o contratualismo.

4.1 Números e agregação

Teorias não-agregativas da moralidade, como é o caso do contratualismo moral, têm dificuldades para explicar a relevância moral dos números em casos que envolvem a escolha entre salvar menos ou mais vidas. Considere um caso como esse:

Dois barcos: Dois barcos estão afundando rapidamente. Em um deles está uma pessoa A, e no outro, estão duas, B e C. Você tem tempo para salvar apenas um dos barcos.

A primeira reação da grande maioria das pessoas a um caso como esse é a de que, salvo alguma consideração especial, somos moralmente obrigados a salvar o barco com duas pessoas. O raciocínio mais simples para justificar essa opinião é o de que é melhor salvar duas pessoas em vez de uma, já que o resultado de duas mortes é pior que o de uma morte. Teorias agregativas como diferentes tipos de consequencialismo conseguem explicar muito bem essa intuição: devemos salvar o barco com duas pessoas porque, *ceteris paribus*, o saldo de bem é maior se duas pessoas forem salvas. O fato de que podemos somar o bem trazido por salvar as duas pessoas para avaliar as consequências é o que permite que teorias deste tipo obtenham o resultado correto em casos como *Dois barcos*.

O mesmo não pode ser dito, no entanto, no caso de teorias não-agregativas como o contratualismo. A preocupação está baseada no seguinte raciocínio: segundo o Contratualismo, aparentemente, você tem as mesmas razões para salvar qualquer um dos barcos. Pelo Princípio do

Fardo Maior, as três pessoas possuem a mesma consideração (o salvamento de sua vida). No entanto, como a Restrição Individualista limita o levantamento de razões a razões individuais, ou seja, tanto B como C podem apenas referir-se a si mesmos para justificar princípios, não existem motivos claros para escolher um barco ou outro. O problema é criado pela exclusão da agregação porque, se contratualistas aceitassem a agregação, poderiam apelar para uma espécie de “soma” das razões de B às razões de C, que derrotariam as de A. No entanto, com a Restrição Individualista em ação, este movimento não pode ser realizado de maneira óbvia.

O problema dos números pode, então, ser entendido da seguinte forma: teorias não-agregativas não conseguem o resultado desejado em casos que envolvem a escolha entre grupos menores e maiores de pessoas, porque são incapazes de “somar os números”. Existem casos em que os números importam, e teorias como o contratualismo possuem dificuldades para explicar por que é certo escolher o maior número em situações deste tipo. No que se segue, exploro algumas possibilidades de resposta de parte dos contratualistas.

4.2 *Os números contam?*

A primeira reação disponível a um contratualista é rejeitar o problema dos números, argumentando que, em casos como *Dois barcos*, não temos obrigação moral de salvar o barco com duas pessoas; em termos morais, os números simplesmente *não importam*. Se este resultado contraintuitivo puder ser defendido, o contratualista não precisa se preocupar com o problema dos números. Esta não é uma estratégia muito popular entre aqueles contratualistas e aqueles que rejeitam a agregação de modo geral, mas merece ser examinada, porque acredito que é mais difícil de rejeitar do que parece ao primeiro olhar. A posição de que, em casos deste tipo, os números não importam, foi famosamente defendida por John Taurek em seu artigo de 1977 “Should the numbers count?”.

Taurek começa nos pedindo para considerar uma situação bastante semelhante a *Dois barcos*.

A situação é a de que eu tenho um estoque de algum remédio que salva vidas. Seis pessoas vão certamente morrer se não forem tratadas com o remédio. Mas uma das seis pessoas precisa de todo o remédio para sobreviver. Cada um dos outros cinco precisa apenas de um quinto do remédio. O que devo fazer? (TAUREK, 1977, p. 294. Tradução minha).

O caso é estruturalmente similar a *Dois barcos* porque envolve a decisão entre salvar a vida de um grupo com mais pessoas ou salvar a vida de um grupo menor. Assim como em *Dois Barcos*, a decisão intuitivamente correta é salvar o grupo maior. De acordo com Taurek, no entanto, o que um agente moral deveria fazer é jogar uma moeda justa e decidir qual grupo salvar baseando-se no resultado aleatório desse lance de moeda. Taurek chega a esta conclusão a partir de três considerações: em primeiro lugar, ele acredita que é permissível salvar o grupo menor em casos deste tipo; em segundo, defende que não seria pior que cinco pessoas morressem em vez de apenas uma; e, finalmente, Taurek crê que um procedimento como o lance de moeda é a melhor maneira de demonstrar igual respeito a cada indivíduo, já que confere a cada um a mesma chance de ser salvo (HIROSE, 2014, p. 111-112). Examinamos as três considerações.

A maioria das pessoas analisa o caso de Taurek e conclui que salvar o maior grupo é a única atitude permissível e, portanto, é moralmente obrigatória. Taurek concorda que salvar o maior grupo é permissível, mas discorda que seja a única atitude permitida em casos deste tipo. Segundo ele, salvar o menor grupo é igualmente permissível. Para defender essa afirmação, ele nos pede, inicialmente, para considerar uma variação do caso, na qual David é a pessoa no grupo menor que precisa de todo o remédio (TAUREK, 1977, p. 295). Suponha que David é o dono do remédio. Neste caso, muitas pessoas concordam que é permissível que David se salve, em vez de se sacrificar pelo grupo de cinco pessoas. Se isso é verdade, então porque não é permitido que um agente qualquer salve David? Como Taurek coloca,

Eu não consigo imaginar que pudesse dar a David alguma razão para que *ele* devesse achar melhor que os cinco estranhos continuem a viver do que ele continue a viver [...] Além disso, e eu gostaria de enfatizar este fato, ao negar seu remédio às cinco pessoas [...] ele não viola os direitos de ninguém. Nenhum destes cinco possui uma reivindicação legítima do remédio de David nessa situação [...] Mas se é moralmente permissível que David dê o remédio a si mesmo nessa situação, porque seria moralmente proibido que eu fizesse o mesmo? É o meu remédio [...] Eu não violo os direitos de ninguém quando uso meu remédio para salvar a vida de David. (TAUREK, 1977, p. 300-301. Grifo do autor. Tradução minha.)

O primeiro movimento de Taurek é argumentar que David pode salvar a si mesmo se possuir o remédio. Acredito que essa afirmação é verdadeira. A base desse raciocínio está no fato de que possuímos permissões especiais para conferir maior peso a nossas considerações — naturalmente, essas permissões não são absolutas, mas em alguns casos existem, nos termos de Scheffler (1992),

prerrogativas centradas em agentes.¹⁶ Hirose explica a permissibilidade da ação de David ao se salvar de maneira parecida, assumindo que é injustificável colocar uma demanda de sacrifício tão grande (nesse caso, a própria vida) em David (HIROSE, 2014, p. 114). Depois de estabelecer que seria permitido para David se salvar, Taurek argumenta que, se David pode se salvar, então qualquer pessoa poderia salvar David. Esta conclusão está apoiada na ideia de que o que podemos moralmente fazer a alguém é o que esta pessoa pode fazer a si mesma — o que é chamado por Frick em seu vídeo “5. Should the Numbers Count?” (JDAFRI2, 2020) de *Princípio da Parcialidade Legítima*. O argumento de Taurek em favor da permissibilidade de salvar uma pessoa pode ser entendido da seguinte maneira:

1. Podemos salvar a nós mesmos em vez dos cinco estranhos.
2. O que podemos fazer a alguém é o mesmo que essa pessoa pode fazer a si mesma (princípio da parcialidade legítima).
3. Portanto, podemos salvar a pessoa singular em vez dos cinco estranhos.

Repare que este argumento estabelece apenas a *permissibilidade* de salvar a pessoa singular, e não sua obrigatoriedade. Mesmo Taurek acredita que salvar o grupo maior é permitido. Esta observação é importante, já que a conclusão de Taurek não é que devemos salvar o menor grupo, mas que devemos *jogar uma moeda* para decidir qual dos grupos salvar.

O argumento é válido, mas possui premissas questionáveis. Em especial, a premissa 2, o Princípio da Parcialidade Legítima, conflita com as razões que poderiam nos levar a aceitar a premissa 1. Como foi mencionado acima, a motivação para pensar que podemos salvar nós mesmos em vez de cinco estranhos é a ideia de que possuímos permissões especiais que dizem respeito ao fato de que a decisão afeta *a nós mesmos*. Mas, se estas permissões existem, então o Princípio da Parcialidade Legítima é falso, já que algumas coisas que a pessoa poderia fazer a si mesma não podem ser realizadas por mim — o único motivo pela qual aquela pessoa pode realizar é uma permissão especial, que tem a ver com sua agência. Como Hirose (2014, p. 114) aponta, no caso David, pedir para que ele se sacrificasse para salvar os cinco estranhos seria injustificável,

¹⁶É claro que um utilitarista convicto, por exemplo, rejeita a existência dessas razões morais, mas tanto pior para o utilitarista, já que elas se encontram na moralidade do senso comum.

mas o mesmo não vale para um agente moral separado do caso, já que nenhum sacrifício está sendo pedido a esta pessoa. A permissibilidade de atos deste tipo não é transferível, porque estão em jogo permissões que são ligadas a agentes específicos. Acredito que o Princípio da Parcialidade Legítima é falso, e por isso é difícil ver como o argumento de Taurek pela permissibilidade seria bem-sucedido.

A segunda consideração de Taurek é uma espécie de rejeição da noção impessoal de bem. Segundo ele, a morte de cinco pessoas não é pior *simpliciter* do que a morte de uma pessoa, porque não existe *alguém* para quem essas mortes sejam piores. Isto é, para Taurek, algo só pode ser bom ou ruim *para alguém*, o que significa que a moralidade não se importa com “a perda de indivíduos”, mas com “a perda *para* indivíduos” (HIROSE, 2014, p.116).

A afirmação de que devemos salvar o maior grupo em vez do menor foi fundamentada na afirmação de que, todo o resto constante, é pior que essas cinco pessoas morram [...] é um juízo avaliativo que não posso aceitar [...] Eu concedo que para cada uma das cinco pessoas, seria pior se David sobrevivesse e essa pessoa morresse do que se David morresse e os cinco sobrevivessem. Mas, é claro, da perspectiva de David, é o contrário [...] [Alguns dirão que] é uma coisa pior, não necessariamente para alguém em particular, ou relativa aos fins de alguém, mas apenas uma coisa pior em si [...] Eu não sou capaz de oferecer uma explicação satisfatória do significado de juízos deste tipo (TAUREK, 1977, p. 303-304. Tradução minha.).

Concordo em linhas gerais com a ideia de que a moralidade interpessoal exige a justificação *para* indivíduos (existindo ou não uma noção impessoal de bem, acredito que é de pouca relevância para a moralidade). No entanto, ainda que não exista um sentido para afirmação de que a morte de cinco pessoas é pior que a de uma, não acho que isto implica que o fato de que cinco pessoas morrerão em vez de apenas uma não pode ser usado para justificar o salvamento de muitas vidas *para* David. De fato, este tipo de raciocínio é o que será explorado por contratualistas em seções futuras deste capítulo. Portanto, não acredito que a rejeição de uma noção impessoal de bem direcione uma teoria a aceitar a conclusão de Taurek de que os números não importam.

A terceira consideração de Taurek é a ideia de que um lance de moeda é a melhor estratégia para conferir igual respeito aos indivíduos porque dá a cada um a mesma chance de sobreviver — neste caso, 50%. No entanto, como Hirose aponta, esta não parece ser a única maneira de conferir respeito a todos indivíduos (HIROSE, 2014, p. 120). De fato, praticamente todas as teorias morais explicam de sua própria maneira no que o igual respeito consiste. Um utilitarista poderia dizer que

respeito igual é conferido quando todos são incluídos no cálculo da utilidade; um prioritarista pode apelar para a ideia de quantidade de bem-estar de cada um, pesados pelo seu nível de bem-estar; um contratualista entende o respeito a cada pessoa como a justificabilidade dos princípios que afetam aquela pessoa. Chances iguais é uma maneira de demonstrar respeito, mas não parece ser a única, e nem aplicável a todas as situações. Acredito, portanto, que nenhuma das considerações levantadas por Taurek em favor da conclusão de que devemos jogar uma moeda em casos como *Dois barcos* ou seu caso do remédio é convincente para este propósito. A primeira opção para o contratualista, que seria adotar a visão de Taurek sobre a importância dos números (a saber, a de que, em termos de moralidade, eles não contam), deve ser descartada. Os números importam, e o problema dos números deve ser respondido.

4.3 O argumento do desempate

A resposta mais popular para casos como *Dois barcos* é que devemos salvar o maior grupo. Acredito, junto da maioria dos filósofos, que esta é a resposta correta. Como pode o contratualismo motivar esta resposta? Um dos argumentos promissores é oferecido pelo próprio Scanlon (1998, p. 232), e é conhecido como “argumento do desempate”. A ideia é a seguinte: quando existem duas razões conflitantes de mesmo tamanho, a adição de uma razão de um dos lados pode servir para *desempatar* a decisão, tornando salvar o grupo com mais pessoas a decisão correta.

Este raciocínio está baseado em dois requerimentos propostos por Scanlon em casos que envolvem decisões deste tipo. Em primeiro lugar, a ideia de que é preciso dar relevância à vida de cada pessoa; em segundo lugar, a ideia de que é preciso dar importância igual à vida de cada pessoa (SCANLON, 1998, p. 233). Essas duas regras explicam por que o procedimento de decisão proposto por Taurek é inadequado: lembre-se que, segundo o contratualismo, princípios morais corretos são aqueles aos quais ninguém poderia levantar uma objeção razoável. Scanlon aponta que se adotarmos um princípio como o lance de moeda de Taurek, pelo menos um dos membros do grupo com maior quantidade de pessoas poderia reclamar que o fato de que sua vida deve ser salva não está sendo levado em consideração, já que o princípio adotado “permite que o agente decida o que fazer da mesma maneira que teria permitido se este membro não estivesse presente, e existisse apenas uma pessoa em cada grupo” (SCANLON, 1998, p. 232. Tradução minha.).

Um princípio de salvar mais vidas, no entanto, é compatível com as duas regras propostas por Scanlon: as razões de cada uma das três pessoas possuem o mesmo peso, e todas são levadas em consideração na decisão final. Supomos que existe razão para salvar A e B. As duas razões se equilibram, já que estão baseadas nas mesmas considerações. Nesse caso, segundo Scanlon, “a razão apresentada pelas necessidades da segunda pessoa em um dos dois grupos deve ter o poder de pelo menos desempatar” (SCANLON, 1998, p. 232. Tradução minha.). Para ilustrar o desempate, imagine, primeiro, Dois Barcos apenas com as pessoas A e B. As razões que o agente possui para salvar cada um dos dois barcos encontram-se em equilíbrio: uma para cada lado. Adicionar a pessoa C ao caso tem que fazer alguma diferença, porque agora temos uma razão para salvar C. Mas se princípios como o de Taurek fossem verdadeiros, a adição de C ao caso não mudaria em nada a maneira em que decidimos qual barco salvar. Nesse caso, o princípio de lance da moeda pode ser razoavelmente rejeitado pela pessoa C, já que esta tem direito a consideração na decisão; mas a pessoa A não pode rejeitar o princípio de salvar mais vidas com base na mesma consideração, já que sua vida entra de maneira relevante na decisão, mesmo que o barco com mais pessoas seja o escolhido no final.

Essa estratégia funciona? O argumento não convence críticos como Michael Otsuka, que acreditam que o contratualismo não consegue realmente evitar a agregação. Para ele, o argumento do desempate envolve um tipo de agregação — o que derrotaria um dos atrativos do contratualismo como teoria moral. Uma maneira de interpretar o argumento é entender que as considerações de C, somadas à de B, são suficientes para derrotar as considerações de A (OTSUKA, 2000). Otsuka nos pede para imaginar uma balança: após medirmos A e B, a balança encontra-se em equilíbrio. O “desempate” acontece quando adicionamos a consideração de C. Mas “adicionar” razões é algo que apenas teorias agregativas podem fazer. As razões que temos para salvar o barco de B e C são razões relativas a um grupo de pessoas, a saber, o grupo composto por B e C. Se o argumento do desempate funcionar assim, ele fere a *Restrição Individualista*. Para Otsuka, o contratualista não pode ter o melhor de dois mundos, evitando a agregação apenas quando lhe convém.

Existem, no entanto, maneiras alternativas de entender o argumento do desempate segundo as quais o argumento não envolve apelos a grupos de pessoas. Rahul Kumar (2001) sugere que a melhor maneira de entender o argumento é que as razões de A e B se neutralizam, sobrando apenas as razões de C. A decisão acontece em dois estágios: primeiro neutralizamos as razões de A e B,

e depois consideramos a razão de C. Podemos chamar essa estratégia de “desempate por neutralização”. Desta maneira, não existe apelo a um grupo de pessoas: pense que em nenhum dos estágios envolve considerar as razões de mais de uma pessoa de cada vez.

Ainda assim, a interpretação da neutralização levanta algumas dúvidas. Como apontado por Otsuka (2006, p. 119), a estratégia de Kumar parece sugerir que em casos como *Dois barcos*, sem a pessoa C, o agente não tem razão para salvar qualquer um dos barcos, já que as razões de A e B se neutralizam. No entanto, não é isso que desejamos: imaginamos que em casos onde temos dois barcos com uma pessoa cada um, temos razão para salvar algum dos barcos.

Acredito que existem outras maneiras de mostrar porque o argumento do desempate não envolve agregação. Para tal, precisamos de uma caracterização mais precisa da agregação como a que foi discutida no capítulo sobre o assunto. A agregação interpessoal é composta de quatro características (HIROSE, 2014, p. 31-41):

Comparabilidade interpessoal: é possível comparar os fatores moralmente relevantes de uma pessoa com os de outra.

Imparcialidade: duas alternativas são igualmente boas se elas diferem apenas no que diz respeito à identidade das pessoas.

Pareto: se uma opção x é melhor (mais justa/mais razoável etc.) para alguém que outra opção y e é pelo menos tão boa quanto y para outras pessoas, x é melhor que y.

Continuidade: ordenamentos podem ser representados por um valor numérico real. Isto implica que, para quaisquer dois fatores moralmente relevantes a e b, existe uma quantidade n de b que possui valor objetivo maior que a.

Como expliquei no capítulo anterior, o contratualismo moral é compatível com três dessas características: Comparabilidade Interpessoal, Imparcialidade e Pareto. De fato, os dois requerimentos de Scanlon suportam essas três condições. O segundo requerimento, segundo o qual cada vida deve ter igual importância, está na ideia de Imparcialidade, enquanto o primeiro, segundo o qual cada vida deve ter importância positiva nas decisões, incorpora Pareto. As duas só podem ser pensadas se comparações de importância forem possíveis, o que vindica Comparabilidade Interpessoal. O que torna o contratualismo não-agregativo é sua rejeição da Continuidade. Otsuka

acusa o argumento do desempate de envolver uma forma implícita de agregação interpessoal porque faz referência a “salvar o grupo maior”, realizando, de certa forma, uma “soma razões”. Não é claro, no entanto, que isso seja suficiente para chamar o argumento de “agregativo”, pelo menos não no sentido em que o contratualismo moral entende a agregação. Se a mera soma de razões é agregativa, então não se trata do mesmo tipo de agregação que os contratualistas desejam rejeitar.

Para que o argumento do desempate pudesse ser criticado por ser agregativo de maneira efetiva, Otsuka deveria mostrar que ele se compromete com a condição de Continuidade. Mas essa condição não aparece no argumento, mesmo que exista referência ao tamanho de grupos e uma soma de razões para desempatar a decisão. O motivo pelo qual a soma de razões é permitida para o desempate *é justamente porque as razões são, nesse caso contínuas, isto é, de mesmo grau de relevância moral*. Em *Dois Barcos*, cada pessoa possui uma razão idêntica. Somar razões para desempatar é diferente de somar razões de ordem distinta. Essa rejeição da Continuidade pode ser vista em Scanlon:

[O argumento do desempate] dependia do fato de que se um dos membros do grupo com duas pessoas estivesse ausente então a razão positiva para salvar uma pessoa seria equilibrada por uma razão idêntica [...], criando um empate [...] Mas quando os danos em questão são desiguais, não podemos criar um empate [...] então não podemos usar esse argumento do ‘desempate’. (SCANLON, 1998, p. 235. Tradução minha.)

O motivo pelo qual os números importam em casos como *Dois barcos*, mas não importam em casos como o de Jones e a transmissão da Copa do Mundo, por exemplo, é o seguinte: em *Dois barcos* os danos discutidos são iguais para todas as partes, no caso de Jones o sofrimento de uma das partes é muito maior que a de outra. Desde que os danos em questão sejam iguais, a soma de razões pode ser feita sem aderir à Continuidade, então o argumento do desempate não apela à agregação, pelo menos não o tipo de agregação interpessoal que é objeto de discussão até agora.

Existe outra reclamação em relação ao argumento do desempate. Scanlon pretende que este argumento estabeleça um princípio de salvar mais vidas para casos como *Dois Barcos*. Para isso, estabelece dois requerimentos que são cumpridos pelo princípio de salvar mais vidas, mas não por princípios que permitam salvar o menor grupo, como o de Taurek. Os requerimentos de Scanlon, no entanto, podem ser compatíveis com outros princípios. Segundo Hirose (2015, p. 155-156), o argumento do desempate não estabelece necessariamente um princípio de salvar mais vidas,

porque é consistente com uma loteria ponderada. Imagine que o agente organiza uma loteria ponderada, dando ao barco de A $\frac{1}{3}$ de chance de ser salvo e ao barco de B e C $\frac{2}{3}$ de chance de ser salvo. Neste método de decisão, a presença de C faz uma diferença positiva (aumenta a chance de salvação do barco de 50% para 66% em relação à decisão apenas entre A e B), o que satisfaz a primeira exigência do argumento do desempate. Além disso, as razões de cada pessoa possuem o mesmo peso, já que é como se cada pessoa trouxesse $\frac{1}{3}$ de chance de ser salvo para a decisão final.

Acredito que Hirose está basicamente certo ao dizer que os dois requerimentos do argumento do desempate são cumpridos por uma loteria ponderada e que, portanto, o argumento até agora descrito é compatível com tal procedimento de decisão. Antes de discutir se existem motivos internos ao contratualismo para preferir o princípio de salvar mais vidas a uma loteria ponderada, gostaria de ressaltar que o problema dos números já perdeu bastante força neste ponto do debate. Isso acontece porque, diferente de princípios que permitem salvar menos pessoas, como é o caso do princípio de Taurek, a loteria ponderada não é tão implausível. Isto é, se fosse o caso que o contratualismo moral não conseguisse motivar um princípio de salvar mais vidas em vez de uma loteria ponderada, não acredito que seria um grande baque para a plausibilidade da teoria. De fato, a loteria ponderada está de acordo com muitos dos princípios que acreditamos serem relevantes para decisões morais, e pode ser visto como a melhor maneira de conferir peso a cada pessoa. Lembre-se, também, que este tipo de loteria só poderia ser usado, segundo o contratualismo, em casos em que as razões conflitantes possuem peso idêntico, o que bloqueia a maior parte dos casos absurdos de loteria ponderada que podemos imaginar. Se o argumento do desempate nos fornece razões para rejeitar lances de moeda, mas é indiferente entre loterias ponderadas e salvar mais vidas, acredito que o argumento serve como uma resposta ao problema dos números, de certa forma. De fato, alguns filósofos *preferem* raciocínios morais que deixem aberta a possibilidade de uso dos dois métodos, adotando um approach misto para casos deste tipo (PETERSON, 2009; PETERSON, 2010).

Ainda assim, a maioria dos contratualistas, Scanlon incluso, acredita que uma teoria moral ideal chega no princípio de salvar mais vidas para casos como *Dois Barcos* (SCANLON, 1998, p. 234). Se os requerimentos do argumento do desempate forem expandidos, é possível excluir a loteria ponderada e ficar apenas com o princípio desejado. Rasmussen (2011, p. 213) sugere um terceiro requerimento, que é uma espécie de desenrolar do primeiro: a ideia de que as razões de

cada pessoa precisam não apenas ter *algum* peso, mas *o maior peso possível*. Este requerimento faz sentido, já que é plausível que poderíamos rejeitar razoavelmente princípios que não estão conferindo o peso máximo às razões de todos envolvidos (a noção de peso máximo, aqui, é claro, é limitada por como o peso adicional conferido a cada razão afeta outros indivíduos).

O princípio de salvar mais vidas atende este requerimento, mas a loteria ponderada não atende. Para ver como isso acontece, considere primeiro o caso *Dois Barcos* resolvido pelo princípio de salvar mais vidas. Existem três pessoas no caso: A, B e C. A está em um barco e B e C estão no outro. Antes de sabermos em qual barco estaremos, este princípio nos dá uma chance de $\frac{2}{3}$ de ser salvo e $\frac{1}{3}$ de não ser salvo, já que temos uma chance de $\frac{2}{3}$ de estar no barco com maior número de pessoas. A loteria ponderada, no entanto, confere uma chance de salvamento menor, porque multiplica as probabilidades de estar em um barco pela probabilidade deste barco ser salvo. Então, antes de sabermos qual barco estaremos, a loteria nos dá apenas $\frac{5}{9}$ de chance de sobrevivência. Para entender isso, considere que a chance total de sobrevivência em *Dois Barcos* sob uma loteria ponderada é: $\frac{2}{3}$ (estar no barco com mais pessoas) \times $\frac{2}{3}$ (este é o barco escolhido na loteria) $+ \frac{1}{3}$ (estar no barco com menos pessoas) \times $\frac{1}{3}$ (este é o barco escolhido na loteria) $= \frac{4}{9} + \frac{1}{9} = \frac{5}{9}$. A chance de $\frac{2}{3}$ oferecida pelo princípio de salvar mais vidas equivale a $\frac{6}{9}$, o que mostra que a loteria ponderada falha no terceiro requerimento conferindo *a todos* uma chance menor de sobrevivência.¹⁷

Este argumento é interessante e motiva a escolha do princípio de salvar mais vidas para os contratualistas através do argumento do desempate, resolvendo de maneira adequada o problema dos números. Ele envolve, no entanto, um movimento que é discutido internamente entre contratualistas, que é a adoção de uma espécie de véu da ignorância limitado. O truque que permite a exclusão da loteria ponderada é o uso das probabilidades de sobrevivência *antes de sabermos em qual barco estaremos*. Perceba que o princípio de salvar mais vidas confere a cada pessoa uma chance maior de ser salvo que a loteria ponderada apenas quando não sabemos quem é a pessoa que está no barco sozinha, já que, do ponto de vista de alguém que conhece a posição de cada pessoa nos barcos, o princípio de salvar mais vidas confere à pessoa A 0% de chance de salvamento, porque segundo este princípio, devemos sempre salvar mais vidas em situações como

¹⁷ O mesmo raciocínio, aplicado a outra versão do caso, é apresentado por Rasmussen (2011, p. 213).

Dois barcos. A ideia que está por trás do argumento inspirado por Rasmussen (2011), então, é a de que contratualistas devem analisar a razoabilidade de princípios a partir da maneira em que podem ser rejeitados com base em razões individuais sem que os indivíduos saibam exatamente como os princípios os afetarão, ou seja, “de um ponto no tempo antes de serem adotados” (SUIKKANEN, 2019, p. 79. Tradução minha.). A aceitação desta ideia geral é uma das maneiras contratualistas de resolver o problema dos números, e por isso deve ser explicitada.

4.4 Véu da ignorância

Segundo o contratualismo moral, ações corretas moralmente são aquelas que estão baseadas em princípios que não poderiam ser razoavelmente rejeitados. Uma qualificação que pode ser feita sobre esta análise da correção moral diz respeito a quanta informação sobre sua relação com os princípios os indivíduos possuem na hora de levantar objeções. Quando Scanlon formulou o contratualismo moral, ainda que tenha se baseado na teoria da justiça de Rawls, ele decidiu excluir o “véu da ignorância” como mecanismo de imparcialidade para as decisões do contrato. É possível, no entanto, argumentar que este recurso a decisões em situação de ignorância deve fazer parte do contratualismo moral quando estamos tentando decidir entre casos em que os números importam. Esta visão, aplicada ao caso *Dois barcos*, é base de um raciocínio que demonstrou porque, para os contratualistas defensores do argumento do desempate como resposta ao problema dos números, o princípio de salvar mais vidas é superior a uma loteria ponderada. Argumento, no entanto, que a adoção da ideia de que os princípios morais relevantes para o caso devem ser justificados para cada pessoa sem que essa pessoa saiba qual posição ocupa na distribuição dos lugares nos barcos torna o argumento do desempate desnecessário, porque esta ideia em si já fundamenta a escolha de um princípio de salvar mais vidas, resolvendo o problema dos números — pelo menos se for aceita de maneira forte.

Para enxergar como a adoção de um véu limitado resolve a questão, basta considerar que $\frac{1}{3}$ das pessoas envolvidas no caso está no barco solitário e $\frac{2}{3}$ estão no barco com duas pessoas. Sem saber qual das pessoas somos, faz sentido pensar que existe uma chance maior de estarmos no barco com duas pessoas. Se este é o caso, o princípio de salvar mais vidas não pode ser razoavelmente rejeitado nem mesmo pela pessoa A, já que, antes da pessoa A saber que está no barco solitário, ela não possui fundamentos razoáveis para rejeitar um princípio que nos obrigue a salvar o maior grupo, já que este é o princípio que confere à pessoa A a melhor chance de

sobrevivência (da perspectiva de um agente que desconhece seu lugar na distribuição). Um argumento parecido foi oferecido por Schelling (2006). Segundo ele, quando não sabemos em qual grupo estamos, “é simplesmente uma questão de probabilidade que [estejamos] no grupo maior” (SCHELLING, 2006, p. 141. Tradução minha.). Isso significa que é racional para cada agente escolher o princípio de salvar mais vidas. Hirose chama este argumento de *argumento probabilístico* (HIROSE, 2014, p. 170).¹⁸

O argumento probabilístico possui uma estrutura semelhante a uma teoria já discutida no capítulo 2: a teoria da justiça de John Rawls (e, de certa forma, a teoria de Harsanyi, também). A ideia por trás da teoria de Rawls é que princípios morais corretos são aqueles que seriam escolhidos por agentes racionais sob um véu da ignorância. O argumento probabilístico também transforma a decisão do caso *Dois barcos* em uma decisão individual sob incerteza ignorante e recomendam como princípio moralmente aceitável aquele que agentes racionais escolheriam.

Existem, então, dois princípios para decisão sob ignorância que podem ser aplicados para chegar ao princípio de salvar mais vidas. O argumento probabilístico de Schelling parece aplicar o princípio da razão insuficiente, segundo o qual em casos de ignorância, isto é, incerteza sem que possamos conferir probabilidades a cada estado, devemos assumir que existe probabilidade igual de estarmos em qualquer uma das três posições, fazendo que a decisão racional seja salvar o barco com mais pessoas. Embora o princípio da razão insuficiente possua seus problemas quando aplicado a casos que envolvem fardos de tamanhos diferentes (como foi discutido no caso da aplicação de Harsanyi deste princípio à decisão na posição original, onde podem existir diferenças de resultado enormes), neste caso acredito que a escolha do princípio é aceitável para um contratualista, já que o fardo envolvido é o mesmo. Repare que o mesmo princípio para decisões sob ignorância não pode ser usado para casos como o caso Jones na Copa do Mundo; a motivação teórica para isto é semelhante à discutida no caso do argumento do desempate. Existe, ainda, a possibilidade de motivar o princípio de salvar mais vidas através de Leximin, já que este princípio de decisão sob ignorância é sensível ao fato de que devemos maximizar o valor mínimo obtido no segundo pior resultado quando os piores resultados possíveis empatam. Este raciocínio não é

¹⁸ O argumento probabilístico de Schelling não é pensado especificamente para casos como dois barcos, mas possui características estruturais que permitem sua aplicação a esse tipo de caso.

suportado por um agente que se importe unicamente com seu próprio bem-estar, já que nesse caso salvar qualquer um dos barcos seria racional. Mas é relativamente pacífico supor que é melhor salvar uma pessoa a mais quando não existe custo para fazê-lo, então acredito que Leximin aplicado sob véu da ignorância fundamenta a racionalidade do princípio de salvar mais vidas. Se estes dois princípios servem para mostrar que agentes racionais sem saber em qual barco estão escolheriam o princípio de salvar mais vidas, então eles fundamentam uma objeção razoável que pode ser feita a outros princípios, estabelecendo, para os contratualistas, a moralidade do princípio de salvar mais vidas. Acredito que é uma das saídas mais promissoras disponíveis ao contratualismo para este clássico problema das teorias não-agregativas. Ela está, no entanto, apoiada no mesmo insight moral do argumento do desempate, já que só funciona quando os fardos distribuídos são os mesmos.

4.5 O argumento do princípio mais razoável

Gostaria de discutir outro argumento disponível ao contratualista para escolher o princípio de salvar mais vidas. Chamo esse argumento de “argumento do princípio mais razoável”, e é largamente inspirado em um argumento de Kamm (1993), chamado pela autora de “argumento da agregação”. Kamm descreve o argumento da seguinte forma:

Se (1) é pior que B e C morram que se apenas B morrer [...]; e (2) é igualmente ruim se A morrer ou se B morrer [...]; então (3) por substituição, deve ser pior se B e C morrerem do que se apenas A morrer. Isto é, se “<” = “pior que”, se $B + C < B$ e $A = B$, então $B + C < A$ (KAMM, 1993, p. 85. Tradução minha.)

Este argumento estabelece de maneira convincente que, em casos como *Dois barcos*, é melhor salvar mais pessoas. Além disso, como Hirose (2014, p. 163-166) aponta, este argumento não é agregativo, pelo menos não no sentido técnico de agregação moral que estamos discutindo nesta dissertação. O raciocínio de Kamm, a rigor, depende de duas características da agregação que não são, por si só, nem em conjunto, agregativas. Hirose (2014, p. 163; 2001) sugere que o princípio de salvar mais vidas pode ser motivado apenas com Pareto e Imparcialidade. Lembre-se que *Imparcialidade* é a propriedade segundo a qual duas alternativas são igualmente boas se elas diferem apenas no que diz respeito à identidade das pessoas e *Pareto* diz que se uma opção x é melhor (mais justa/mais razoável etc.) para alguém que outra opção y e é pelo menos tão boa quanto y para outras pessoas, x é melhor que y. *Pareto* e *imparcialidade* podem ser usados para

defender o princípio de salvar mais vidas. Imagine três situações possíveis resultantes de uma decisão no caso *Dois barcos*:

1: salvamos A, B e C morrem;

2: salvamos B, A e C morrem;

3: salvamos B e C, A morre;

Por Imparcialidade, 1 e 2 são igualmente ruins, já que neste caso estamos desconsiderando qualquer diferença moral entre os indivíduos e a única diferença inicial entre A, B e C é a identidade. Por Pareto, 3 é melhor que 2, já que 3 é melhor que 2 para a pessoa C e não é pior para nenhuma outra pessoa (A morre tanto no resultado 2 quanto no 3). Se isso é verdade, então 3 deve ser melhor que 1, já que 1 e 2 possuem valor igual. Imparcialidade e Pareto são propriedades estruturais da agregação moral, mas não são suficientes para esta. Nos capítulos anteriores mencionei como alguns princípios e teorias são compatíveis com Pareto e Imparcialidade, mas não são agregativos. O princípio Leximin e o contratualismo moral são exemplos de princípios e teorias não-agregativas que aceitam Pareto e Imparcialidade. O argumento da agregação de Kamm, portanto, não é realmente agregativo.

Existe uma suposição feita neste argumento, no entanto, que muitos contratualistas gostariam de rejeitar. Como Hirose (2014, p. 166) ressalta, o argumento estabelece que salvar mais pessoas é *melhor*, mas para estabelecer a *correção moral* do princípio de salvar mais vidas, seria necessário assumir uma espécie de consequencialismo, isto é, a ideia de que princípios moralmente corretos são os que levam aos *melhores resultados*. Esta característica do argumento está perigosamente próxima do utilitarismo. Repare que os argumentos anteriores discutidos não envolviam referências a quanto os resultados dos princípios discutidos eram bons, mas sim à maneira em que os princípios podem ser justificados para cada pessoa, movimento esse que é um dos principais insights morais do contratualismo. O argumento de Kamm, por sua vez, apela à ideia de que salvar mais vidas é *melhor*. Contratualistas não gostariam de assumir esse raciocínio.

Ainda assim, acho que podemos adaptar o argumento, através da explicação de Hirose (2001; 2014, p. 163) de como Pareto e Imparcialidade possibilitam a escolha de salvar o maior número, para uma análise contratualista. Recorde as três situações possíveis:

1: salvamos A, B e C morrem;

2: salvamos B, A e C morrem;

3: salvamos B e C, A morre;

Se como fator moralmente relevante usarmos a ideia de que um resultado é *melhor* que outro, temos o argumento de Kamm. Mas se os fatores moralmente relevantes envolvidos em Pareto e Imparcialidade forem traduzidos para a razoabilidade ou justificabilidade de sua aplicação, o consequencialismo desaparece do raciocínio. Considerando que A, B e C possuem o mesmo tipo de razão para rejeitar princípios, por *Imparcialidade*, 2 é tão razoável e justificável quanto 1. Por *Pareto*, 3 é mais razoável e justificável que 2. Portanto, 3 é mais razoável e justificável que 1. Isto significa dizer que existe alguma vantagem normativa nas razões que suportam princípios que levam a 3 em relação às razões que suportam princípios que não o fazem. Se isso é verdade, então princípios que levam a 3, como o de salvar mais vidas, não podem ser razoavelmente rejeitados com base nas razões da pessoa A, por exemplo, o que estabelece o princípio de salvar mais vidas como um princípio moral defendido pelo contratualismo. Esta versão modificada do argumento é o que chamo de argumento do princípio mais razoável. É possível que existam reclamações em relação a uma característica deste argumento: a suposição de que a razoabilidade das razões oferecidas pode depender do resultado da aplicação de um princípio. Contudo, o fato de que o contratualismo é uma teoria de tendência deontológica não significa que os resultados de ações e princípios nunca são levados em conta; significa, sim, que existem outros fatores dignos de consideração moral, mas os resultados da aplicação de um princípio são fornecedores de força normativa para razões por excelência. Portanto, não acredito que a adaptação do argumento seja problemática para contratualistas. A insistência na razoabilidade dos princípios em vez do fato de que possuem *melhores* resultados é o que faz o argumento modificado compatível com o contratualismo, enquanto o de Kamm não era.

Assim como os outros argumentos discutidos, parte do que faz o argumento funcionar é o fato de que em casos como *Dois barcos* os fardos envolvidos são os mesmos (todas as pessoas estão em posição de oferecer razões individuais ligadas à própria vida). Além disso, o argumento do princípio mais razoável funciona assumindo propriedades estruturais (Pareto e Imparcialidade) que são compatíveis com o contratualismo e com os outros argumentos discutidos. Por isso, não

acredito que seja tão diferente do argumento do desempate, por exemplo, mas é outra forma de explicar como o contratualismo pode adotar o princípio de salvar mais vidas em casos como *Dois barcos*.

4.6 Reivindicações relevantes agregadas

A ideia de agregação de reivindicações relevantes é a última proposta que irei sugerir como um caminho para o contratualista que deseja defender um princípio de salvar mais vidas. Esta visão não é completamente não-agregativa porque envolve, em certo sentido, uma forma limitada de agregação. Contudo, essa “forma limitada de agregação” fere uma das características estruturais da agregação moral e, portanto, não é agregativa no sentido técnico que estamos discutindo nesta dissertação. RRA (“Reivindicações relevantes agregadas”) foi criada por Alex Voorhoeve (2014) com o objetivo de unir as vantagens normativas dos approaches agregativos com as vantagens dos approaches não-agregativos na ética distributiva. Outras propostas similares de agregação limitada podem ser encontradas em Scanlon (1998), Kamm (1996), Steuwer (2021) e Temkin (2011). Acredito que a parte “agregativa” deste tipo de teoria pode ser traduzida em termos não-agregativos, e com alguma reformulação, a ideia de reivindicações relevantes pode ser incorporada ao contratualismo. A teoria RRA é descrita por Voorhoeve da seguinte maneira:

1. Cada indivíduo cujo bem-estar está em questão tem uma reivindicação para que você o ajude [...]
2. As reivindicações de diferentes indivíduos competem apenas caso elas não possam ser conjuntamente satisfeitas
3. As reivindicações de um indivíduo são tanto mais fortes:
 - a) Quanto mais seu bem-estar aumentaria se fosse ajudado; e
 - b) Quanto menor for o nível de bem-estar a partir do qual esse aumento ocorreria.
4. Uma reivindicação é relevante se e somente se é suficientemente forte em relação à reivindicação mais forte com a qual compete.
5. Você deveria escolher uma alternativa que satisfaz a maior soma de reivindicações relevantes ponderando suas forças. (VOORHOEVE, 2014, p. 66. Tradução minha.)

Esta teoria tem algumas características semelhantes ao contratualismo moral contemporâneo. Repare, por exemplo, que a condição 1 possui a mesma função da *Restrição Individualista*, que é

garantir que as razões usadas para decidir princípios sejam individuais. A condição 4, por sua vez, é uma versão modificada do *Princípio do Fardo Maior*, que introduz a noção de relevância. A parte agregativa da teoria está na condição 5, que permite a agregação de reivindicações, mas apenas quando elas são relevantes entre si. Está claro que uma caracterização precisa das consequências morais de RRA depende, então, de uma explicação desta noção de relevância.

RRA usa a ideia de relevância relativa para limitar a agregação de reivindicações. Segundo a condição 4, a reivindicação mais forte não é necessariamente vencedora — como é no caso do Princípio do Fardo Maior ou em princípios de decisão como Leximin e Maximin — mas ela compete apenas com reivindicações de um nível de força parecido. Voorhoeve propõe um critério para estabelecer a relevância de reivindicações baseado na ideia de “perspectiva pessoal permissível” (VOORHOEVE, 2014, p.72. Tradução minha.). Segundo a moralidade do senso comum, em alguns casos temos permissão moral para priorizar nosso bem-estar em vez do de outros, mesmo quando adotar esta perspectiva envolve um fardo maior aos outros que o fardo que nos seria conferido caso não adotássemos a perspectiva pessoal permissível. Por exemplo, é permissível priorizar nossa própria vida em vez da vida de um estranho, e é também permissível nos salvar da paralisia cerebral completa em vez da vida de um estranho. Não seria permissível, no entanto, nos salvar de uma dor de cabeça forte em vez da vida de um estranho: esse é um caso em que a diferença na força das reivindicações é tão grande que a perspectiva pessoal não pode ser adotada. Esta forma permissível de parcialidade é chamada por Voorhoeve de “perspectiva pessoal permissível” e é base para estabelecer a relevância relativa de reivindicações. Segundo Voorhoeve, uma reivindicação A é relevante relativa a outra B com a qual compete somente se seria permissível, do ponto de vista pessoal, satisfazer A em vez de B:

A reivindicação da pessoa A é relevante relativa à reivindicação conflitante de B somente se, em uma comparação um-a-um de suas reivindicações da perspectiva pessoa permissível de A, a reivindicação de A é pelo menos tão importante quanto a de B. Ou, para colocar este teste em termos de ação permissível: a reivindicação de A é relevante relativa à reivindicação conflitante de B somente se, em uma situação em que A deve escolher se satisfaz sua reivindicação ou a de B e em que nenhum fator moralmente relevante além da preocupação mínima requerida por B está no caminho de A agir em seu autointeresse, seria permissível que A satisfizesse sua própria reivindicação. (VOORHOEVE, 2014, p. 72. Tradução minha.).

Este critério para relevância ainda inclui algum grau de vagueza, já que está baseado em distinções morais intuitivas do senso comum, mas é um bom guia e está relacionado com a ideia de que razões

individuais só podem competir quando estiverem em grau semelhante de força, que é uma ideia atrativa dos approaches não-agregativos como o contratualismo. Dito isso, dado um grupo de razões relevantes, a agregação destas razões de nível parecido pode ser feita. Novamente vemos que esse procedimento é semelhante ao argumento do desempate, embora envolva agregação, porque está centrado na ideia de que os números importam apenas quando os fardos envolvidos estão na mesma ordem de grandeza. A principal diferença é que enquanto o argumento do desempate e o Princípio do Fardo Maior exigem que as razões estejam realmente empatadas, a noção de relevância de RRA permite um pouco mais de flexibilidade. Essa flexibilidade traz vantagens e desvantagens que estão além do escopo deste capítulo (por exemplo, uma vantagem é a capacidade de lidar com casos limítrofes de comparação de fardos, com os quais o contratualismo tradicional tem problema; uma desvantagem é a possibilidade de *slippery slopes*, isto é, que a relevância de uma reivindicação para outra sirva como ponte para que reivindicações muito pequenas se tornem relevantes para reivindicações grandes, reduzindo RRA a uma teoria agregativa¹⁹).

RRA possui os resultados intuitivos nos casos que estou discutindo. Em *Dois barcos*, como as reivindicações de A, B e C são idênticas, elas são relativas entre si. Isso significa que a agregação pode ser feita nesse caso, e então devemos salvar o barco com mais pessoas. RRA, então, suporta o princípio de salvar mais vidas neste caso. No entanto, RRA não encontra os mesmos problemas de agregação que encontramos em teorias agregativas tradicionais. No caso Jones na Copa do Mundo, por exemplo, a agregação é bloqueada, porque a reivindicação de cada um dos espectadores é irrelevante em relação à de Jones.

Gostaria de ressaltar que embora Voorhoeve use a palavra “agregação” para o procedimento de soma usado quando as reivindicações são relevantes, RRA não adota o que estamos chamando de “agregação moral”. Isso porque RRA, assim como o contratualismo, fere a propriedade de Continuidade, segundo a qual ordenamentos podem ser representados por um valor numérico real. Aceitar a condição Continuidade significaria aceitar que, em termos de fatores moralmente relevantes, *todos* são contínuos entre si. O contratualismo claramente rejeita essa afirmação, já que a união entre Restrição Individualista e Princípio do Fardo Maior implica que

¹⁹ Este tipo de objeção é oferecido por Parfit (2003).

apenas razões idênticas são contínuas. RRA rejeita Continuidade de outra maneira: a inserção da noção de relevância relativa significa que apenas reivindicações relevantes são contínuas entre si. O que Voorhoeve chama de “agregação limitada” é, na definição técnica de agregação moral, uma forma não-agregativa de teoria.

RRA é estruturalmente similar ao contratualismo e sua maneira de evitar o problema dos números através da noção de agregação de reivindicações relevantes espelha, de certa forma, o procedimento do argumento do desempate. Acredito que a introdução dessa noção de relevância no lugar do Princípio do Fardo Maior é uma opção interessante para contratualistas, e mais uma forma de solucionar o problema dos números mantendo-se fiel à ideia original da teoria, que é a construção de uma teoria moral não-agregativa. Neste capítulo, quatro opções diferentes (embora estruturalmente similares) de defesa do contratualismo contra o problema dos números foram exploradas. Acredito que a ideia principal de que os números importam para decisões morais somente quando os fardos são parecidos é o que está por trás de cada uma dessas respostas, e é um insight moral que pode ser formulado de diferentes maneiras. O problema dos números, portanto, não é uma sentença de morte para o contratualismo. No próximo capítulo, um outro tipo de objeção a teorias não-agregativas como o contratualismo será analisado: as questões que envolvem imposições de risco.

5 Contratualismo e Risco

O capítulo anterior realizou uma discussão de um tipo de objeção comumente feita ao contratualismo moral que está baseado em sua rejeição da noção de agregação moral, o problema dos números. Este capítulo será dedicado à discussão de outro tipo de objeção relacionado à exclusão da agregação. Segundo alguns críticos, o contratualismo moral é uma teoria inadequada para lidar com casos que envolvem imposição de riscos. A discussão sobre a moralidade do risco é relativamente recente na filosofia moral e ainda está em aberto: muitos tipos de teoria encontram dificuldades para fornecer princípios intuitivos da moralidade dos riscos. Neste capítulo, explorarei algumas alternativas para que o contratualismo seja capaz de fornecer uma explicação adequada desse fenômeno moral. Começo expondo o problema moral do risco, em seguida salientando como este se aplica ao contratualismo moral. Depois, considero algumas alternativas de resposta, retomando a distinção entre contratualismo *ex post* e contratualismo *ex ante*. Argumento que uma versão *ex ante* de contratualismo ou uma versão *ex post* sensível ao risco são capazes de obter resultados interessantes na explicação da moralidade de imposição de riscos.

5.1 Contratualismo e o problema moral do risco

Porque existe pouca certeza e segurança absoluta na vida, boa parte das decisões que tomamos possuem potenciais resultados indesejados. Chamaremos este fenômeno geral de risco. A presença do fenômeno do risco impacta situações moralmente de diversas maneiras. Algumas ações arriscadas são imorais por causa do risco que impõem a outros. Considere, por exemplo, o caso do motorista embriagado que não causou nenhum acidente. Julgamos que dirigir nestas condições foi moralmente errado, porque existia um risco considerável de causar acidentes de trânsito com resultados terríveis. Impor risco a outros, no entanto, nem sempre é imoral. Basta lembrar que existe risco de acidente quando dirigimos sem embriaguez também, sem que esta ação seja moralmente condenável. Existem casos, ainda, em que impor riscos a outros é moralmente correto ou, ainda, obrigatório. Um caso relevante é a aplicação de vacinas: mesmo quando uma vacina envolve riscos graves (podendo ser até letais), acreditamos que é obrigação do Estado e dos cidadãos oferecer e tomar vacinas, respectivamente, quando esta ajuda a prevenir doenças perigosas.

As teorias morais contemporâneas mais populares têm dificuldades no que diz respeito à explicação das nuances da moralidade do risco. Tomemos, a título de ilustração, o caso das teorias agregativas, como o utilitarismo, o igualitarismo teleológico e o prioritarismo. No capítulo 1, ao discutir a noção moral de agregação e as teorias nela baseadas, levantei uma objeção originalmente trazida por Voorhoeve e Fleurbaey (2012) relacionada à ideia de separação entre as pessoas. Recorde-se que a objeção envolvia um caso de risco, já que as operações realizadas por não curavam *com certeza* Diego ou Rafael, mas sim conferiam *uma chance* de cura. Este exemplo nos mostra que um princípio da moralidade de impor riscos a outras pessoas é o seguinte: a distribuição dos riscos importa para a justificação das imposições de risco. Nos casos discutidos no capítulo 1, a falha central das teorias agregativas é o fato de que elas eram incapazes de diferenciar o caso Intrapessoal, onde o risco de perda e a chance de ganho eram dados à mesma pessoa, e o caso Interpessoal, onde um risco era imposto a Diego sem que ele recebesse alguma chance de ganho. Teorias agregativas, portanto, não conseguem lidar adequadamente com essa característica da moralidade do risco.

Outras características deste fenômeno moral são problemáticas para teorias não-agregativas. Graus de risco importam na avaliação moral de ações arriscadas, bem como seus potenciais resultados negativos, o que significa que quando julgamos ações sob risco, estamos atentos pelo menos a (i) qual é o grau de risco dos resultados negativos; (ii) quão ruins são os possíveis resultados negativos; e (iii) quão benéficos são os possíveis resultados positivos. Dirigir embriagado, por exemplo, é errado porque as chances de gerar acidentes são consideravelmente altas e os resultados dos acidentes são bastante ruins. Além disso, os resultados positivos potenciais geralmente podem ser obtidos por outros meios, como pedir carona a pessoas sóbrias. Dirigir em condições normais, no entanto, embora possua resultados possíveis tão ruins quanto dirigir embriagado (por exemplo, um acidente fatal), não é considerado moralmente errado porque o grau de risco é consideravelmente baixo e os resultados positivos de permitir a prática são enormes. Daí obtemos um princípio plausível da moralidade do risco, segundo o qual ações arriscadas, mesmo as que possuem potenciais resultados muito indesejados, podem ser moralmente justificadas quando seu grau de risco é suficientemente baixo e seus potenciais resultados benéficos são consideráveis.

O problema é o seguinte: as teorias não-agregativas parecem proibir ações intuitivamente permitidas pela moralidade quando estas envolvem risco de resultados muito negativos, mesmo quando o grau de risco é baixo. Este problema vem do fato de que é difícil avaliar a conformidade de ações às normas morais quando elas envolvem uma mistura de resultados positivos e negativos possíveis, já que alguns dos resultados podem nos levar a permitir a ação, enquanto outros nos fazem querer proibir a ação (HANSSON, 2013, p. 29). Em uma teoria agregativa, este problema não existe, já que a agregação dos riscos em um valor único permite que uma ação que afeta diversas pessoas seja tomada de um ponto de vista “único”. No caso do contratualismo, no entanto, este procedimento não pode ser feito sem qualificações. Para ilustrar o problema, considere este caso, trazido por Frick (2015):

Vacinação em Massa (Vítimas desconhecidas): Um milhão de crianças pequenas são ameaçadas por um terrível vírus, que com certeza vai matar todas se não fizermos nada. Devemos escolher entre a produção de massa de uma das duas vacinas (limites de capacidade nos impedem de produzir as duas):

- A Vacina 1 com certeza salva a vida de *todas* as crianças. No entanto, a vacina não dá proteção completa contra o vírus. Se uma criança recebe a Vacina 1, o vírus vai com certeza paralisar uma de suas pernas, e assim ele ou ela vai ter que caminhar de muletas para o resto da vida.
- A Vacina 2 é arriscada. Ela dá a cada criança uma chance de 999/1000 de sobreviver ao vírus completamente intocada. Entretanto, para cada criança existe uma chance de 1/1000 de que a Vacina 2 vá ser completamente ineficaz e a criança vai ser morta pelo vírus (FRICK, 2015, p. 181-182. Tradução minha.).

A escolha moral intuitiva é a Vacina 2. O raciocínio por trás desta escolha da moralidade de senso comum é simples: ainda que exista um pequeno risco de morte, a Vacina 2 é superior porque o fardo da paralisia *certa* imposto a *todas* as crianças é muito grande. Como o risco de morte é pequeno, a aplicação da Vacina 2 não é apenas permissível, mas moralmente obrigatória neste caso. Teorias agregativas como o utilitarismo conseguem obter este resultado sem problemas, já que apenas somando o bem-estar esperado que seria produzido por cada uma das vacinas, fica claro que a Vacina 2 é superior neste caso. O mesmo não pode ser dito com facilidade do contratualismo. Relembre que, segundo o contratualismo, a moralidade de ações é definida pela razoabilidade dos princípios que a permitem, e esta razoabilidade depende, primariamente, de duas características da teoria: a *Restrição Individualista* e o *Princípio do Fardo Maior*. Um argumento disponível aos críticos do contratualismo passa pelo seguinte caminho: no caso descrito acima, a pessoa com a razão mais forte é a criança que morreria por causa da Vacina 2. Por causa da

Restrição Individualista, as crianças que seriam paralisadas caso a Vacina 1 seja escolhida não podem “unir” suas reivindicações. Pelo *Princípio do Fardo Maior*, a reivindicação da criança que morreria é mais forte que a de uma criança que seria paralisada, o que faz com que a criança que morreria pela Vacina 2 seja capaz de rejeitar razoavelmente um princípio que exija a distribuição da Vacina 2. Assim, se este raciocínio está correto, o contratualismo incorre no problema de proibir ações arriscadas mesmo quando o baixo grau de risco justifica ou exige sua performance, do ponto de vista moral. No que se segue, exploro alternativas de resposta para contratualistas.

5.2 *Contratualismo ex post e ex ante*

A análise de casos que envolvem imposição de riscos motivou uma distinção teórica entre contratualistas que diz respeito à perspectiva temporal ou informacional adotada pelos agentes idealizados que realizam rejeições aos princípios candidatos da moralidade. Uma opção é exigir que as razões levantadas incluam informação total sobre como a adoção de princípios afetará cada indivíduo, isto é, os indivíduos *sabem* quais fardos que os afetarão serão resultado da adoção de um princípio moral. Esta versão é conhecida como contratualismo *ex post* e conta com defensores como o próprio Scanlon em sua formulação inicial do contratualismo (1998), Rürger (2018) e Suikkanen (2019). De acordo com Suikkanen, no contratualismo *ex post*:

Nós imaginamos um conjunto de mundos que são quase idênticos ao mundo atual, exceto pelo fato de que princípios morais diferentes foram adotados neles. Como resultado disso, indivíduos agem de maneiras diferentes nesses mundos e isso afeta as vidas que eles acabam levando. Indivíduos podem então fazer objeções a tais princípios com base nos fardos pessoais pelos quais estes princípios foram causalmente responsáveis (SUIKKANEN, 2019, p.78. Tradução minha.).

Em contraste, o contratualismo *ex ante* postula que as razões relevantes devem estar baseadas em uma perspectiva temporal anterior à adoção dos princípios, ou seja, sem que os indivíduos saibam como cada princípio os afetará, mas levando em consideração as chances que o princípio confere a cada pessoa (SUIKKANEN, 2019, p. 79). O contratualismo *ex post*, portanto, está preocupado principalmente com os resultados, enquanto o *ex ante* concentra-se nas chances oferecidas a cada pessoa (STEUWER, 2021b, p. 112).

A objeção da seção anterior parece assumir uma forma *ex post* de contratualismo, já que envolve trazer uma objeção que só seria possível caso cada pessoa possuísse informação completa sobre como a vacinação afetará os vacinados. No momento da aplicação, não sabemos quem são

as crianças que sofrerão os efeitos colaterais letais da vacina. Mas, segundo o contratualismo *ex post*, a decisão deve ser tomada do ponto de vista informado sobre os resultados. Isso significa que as crianças (ou os guardiões) que sofrerão os fardos da Vacina 2 possuem fundamento para rejeitar razoavelmente ao princípio. Esta é uma consequência contraintuitiva do contratualismo *ex post*. Defensores desta visão precisam rejeitar o contratualismo *ex post*, ou mostrar como ele não implica a escolha da Vacina 1 ou defender a escolha da Vacina 1. Na próxima seção, investigo a última alternativa.

5.3 O argumento da informação irrelevante e o argumento do caso individual

Para ver como defensores da visão *ex post* defendem este resultado incômodo da teoria, considere um tipo de argumento encontrado na literatura, formalizado por Frick (2015, p. 194) e por ele chamado de *Argumento da informação irrelevante*. O contratualista *ex post* começa sugerindo que existe certo espelhamento entre o caso de vacinação apresentado e um caso similar no qual sabemos que a aplicação da Vacina 2 causará a morte de mil crianças e sabemos quais serão as vítimas. Administrar a vacina arriscada quando sabemos quais são as crianças que serão afetadas seria moralmente proibido. Mas, argumenta o contratualista *ex post*, qual é a diferença moral entre *Vacinação em Massa (Vítimas desconhecidas)* e *Vacinação em Massa (Vítimas conhecidas)*? Como temos boa razão para acreditar que o número de mortes será o mesmo — a probabilidade de que mil crianças irão morrer se a Vacina 2 for aplicada em *Vacinação em Massa (Vítimas desconhecidas)* é, como lembra Frick (2015, p. 183), superior a 99% — a diferença mais óbvia é que no primeiro caso não conhecemos as identidades dos afetados. Esta informação, segundo alguns defensores do contratualismo *ex post*, é irrelevante do ponto de vista moral: o que importa é sabermos que *alguém sofrerá um fardo*. Observe, por exemplo, o que Reibetanz escreve:

Desde que saibamos que a aceitação de um princípio vai afetar *alguém* de uma certa maneira, devemos conferir a esta pessoa uma reivindicação baseada na magnitude completa daquele fardo ou benefício, mesmo se não pudermos identificar a pessoa de antemão. Apenas se não soubermos se a aceitação de um princípio vai afetar alguém de uma certa maneira que deveríamos alocar a cada indivíduo uma reivindicação baseada em seus fardos e benefícios esperados sob aquele princípio (REIBETANZ, 1998, p. 304. Grifo da autora. Tradução minha.)

Acompanhe esta formalização do argumento, adaptada de Frick (2015, p. 194):

- (1) Uma ação é permissível se e somente se for justificável para cada pessoa (Contratualismo moral).
- (2) Em casos como Vacinação em Massa (Vítimas conhecidas), a Vacina 2 não pode ser justificada para as crianças que serão negativamente afetadas, porque suas mortes são fardos maiores que a paralisia causada pela Vacina 1.
- (3) Casos como *Vacinação em Massa (Vítimas desconhecidas)* possuem praticamente o mesmo número de mortes esperadas que *Vacinação em Massa (Vítimas conhecidas)*, ainda que não possamos saber as identidades das pessoas que morrerão.
- (4) Saber ou não a identidade das pessoas que morrem é irrelevante, do ponto de vista da justificabilidade da escolha de vacina.
- (5) Se 2, 3 e 4 são verdadeiros, então existem crianças em *Vacinação em Massa (Vítimas desconhecidas)* para as quais não podemos justificar a escolha da Vacina 2.

Conclusão: Se 1 e 5 são verdadeiros, a escolha da Vacina 2 em *Vacinação em Massa (Vítimas desconhecidas)* é moralmente proibida.

Assim, o contratualista *ex post* tenta defender sua escolha da Vacina 1. Esta é uma escolha que contradiz a moralidade do senso comum, o que coloca esta forma de contratualismo em péssima situação quando se trata de avaliar riscos. Entretanto, o argumento é válido. Isso significa que sua correção depende da verdade das premissas. A premissa essencial para que o argumento funcione é a premissa 4. Exceto essa, acredito que todas as outras premissas são verdadeiras.

Além de ser base para uma conclusão moral contraintuitiva, a premissa 4 também contradiz a maneira em que o contratualismo *ex ante* sugere que tomemos decisões em *Vacinação em Massa (Vítimas desconhecidas)*. Antes de seguir discutindo o argumento, gostaria de mostrar como o contratualismo *ex ante* lida com este caso. A versão *ex ante* do contratualismo moral não encontra dificuldade neste caso do ponto de vista da comparação com a moralidade do senso comum. Segundo o contratualismo *ex ante*, o que importa do ponto de vista moral em casos de imposição de risco não são os resultados do princípio, mas as chances que o princípio oferece a cada pessoa. Como não sabemos quais crianças serão afetadas letalmente pela Vacina 2, esta opção oferece a mesma chance de sobrevivência para cada pessoa. Resta saber se esta escolha pode ser justificada

para cada um. Frick (2015, p. 187) sugere que devemos pensar neste caso, portanto, como um caso de decisão individual sob risco. A decisão entre vacinas do ponto de vista individual pode ser vista neste caso:

Vacinação individual: uma única criança (chamemos de Clara) está ameaçada por um terrível vírus, que com certeza a matará se não fizermos nada. Como guardiões de Clara, devemos decidir qual das duas vacinas disponíveis lhe dar. Como antes,

- A Vacina 1 com certeza vai salvar a vida de Clara, mas com o custo de que uma de suas pernas ficará paralisada pelo vírus.
- A Vacina 2 vai dar a Clara uma chance de 999/1000 de sobrevivência ao vírus completamente ílesa; no entanto, existe uma chance de 1/1000 de que a Vacina 2 vá ser completamente ineficaz para Clara e ela será morta pelo vírus (FRICK, p.186. Tradução minha.).

A teoria da decisão nos fornece uma maneira de avaliar este problema através da teoria da utilidade esperada, que pode ser aplicada quando conhecemos as probabilidades envolvidas em uma decisão sob incerteza. A ideia é que devemos descontar o peso de fardos e benefícios na decisão de acordo com as probabilidades associadas a suas ocorrências. O raciocínio que leva à escolha da Vacina 2 é bastante intuitivo: como a paralisia de uma perna é um fardo considerável, é razoável assumir que a paralisia certa da Vacina 1 possui utilidade esperada menor que a chance de 99,9% de cura completa oferecida pela Vacina 2. O fardo da morte possível é descontado pelo fato de que a probabilidade de sua ocorrência é muito baixa.

Como cada pessoa está em situação igual em Vacinação em Massa (Vítimas desconhecidas) — já que não sabemos quem serão os afetados negativamente — se uma escolha é justificada para uma criança individual, deve ser justificada para todas. Frick chama este movimento de *Argumento do caso individual*, segundo o qual:

[...] ao avaliar a correção de uma ação arriscada, devemos nos concentrar em como ela afeta cada indivíduo, comparado com as alternativas possíveis [...] o caso pode ser analiticamente decomposto em um milhão de apostas individuais que ocorrem em paralelo [...] As probabilidades oferecidas a cada criança quando administramos a Vacina 2 são exatamente as mesmas que as oferecidas para Clara no caso Vacinação Individual, e a alternativa é a mesma também. Portanto, ao produzir e administrar em massa a Vacina 2, estamos fazendo para cada criança o que um guardião, preocupado somente com os interesses *daquela* criança deveria fazer para ela [...] Dado isso, como pode qualquer criança em Massa (Vítimas desconhecidas) rejeitar razoavelmente um princípio que nos permite escolher a Vacina 2? (FRICK, 2015, p. 187-188. Tradução minha.)

Este argumento, então, ilustra a maneira em que contratualistas *ex ante* pensam sobre casos de risco social deste tipo. Para o contratualismo *ex ante*, a justificção das imposições de risco leva em conta não o resultado, isto é, como um princípio afeta uma pessoa, mas “a qualidade das chances que ação deu a esta pessoa *ex ante*” (FRICK, 2015, p. 188. Grifo do autor. Tradução minha.). Desta maneira, as pessoas “azaradas” não possuem fundamento para rejeitar razoavelmente princípios se as probabilidades oferecidas forem boas o suficiente para justificar a aposta no princípio do ponto de vista individual.

É importante, como ressalta Frick (2015, p. 189-190), distinguir o contratualismo *ex ante* das teorias contratualistas de Rawls e Harsanyi, que envolvem a adoção de um véu da ignorância. Embora possam parecer similares por trabalharem com a ideia de que princípios justificáveis são os que agentes racionais e razoáveis escolheriam se não soubessem alguns fatos sobre como os princípios os afetariam, existe uma diferença crucial. Segundo as teorias contratualistas de Rawls e Harsanyi, princípios corretos da moralidade são aqueles que seriam escolhidos por agentes racionais sob um véu da ignorância. Este modelo é aplicado a diferentes tipos de princípio moral, desde decisões sobre a estrutura básica da sociedade como no caso de Rawls, até princípios mais gerais sobre regulação do comportamento. É preciso lembrar, no entanto, que este véu da ignorância é *artificial*. Em nenhum momento a ignorância sobre nossa posição social, preferências, crenças realmente existiu; a posição original é um procedimento artificial, uma espécie de filtro que introduz imparcialidade nos modelos de Rawls e Harsanyi. Existe, de certa forma, um véu da ignorância no contratualismo *ex ante*, mas é um véu da ignorância *natural*, isto é, um véu da ignorância criado pelo fato de que em casos de risco social, as pessoas *realmente* não sabem em que posição estão e como serão afetados pela adoção dos princípios disponíveis.

Se este tipo de raciocínio sobre casos de risco social estiver correto, então a premissa 4 do *Argumento da informação irrelevante* é falsa, porque o fato de que não sabemos as identidades das pessoas que morrerão faz diferença moral: a saber, faz com que a decisão deva ser encarada de maneira análoga ao ponto de vista de uma aposta individual. O argumento do caso individual oferecido por Frick tem a vantagem normativa de evitar o problema moral do risco, já que reflete o fato de que graus suficientemente baixos de risco podem justificar ou até exigir ações arriscadas. Como contratualistas *ex post* podem defender a premissa 4 do *Argumento da informação irrelevante* da força intuitiva do argumento do caso individual?

Uma primeira sugestão é a ideia de que a certeza de que alguém sofrerá um fardo é suficiente para rejeitar princípios arriscados como um que permite a Vacina 2. Frick (2015, p. 195) chama esta ideia de *Argumento da perda certa*. A fundamentação do argumento está presente na passagem anteriormente citada de Reibetanz (1998, p. 304). Seguindo essa linha de raciocínio, o que importa é que alguém será altamente prejudicado; e o fato de que temos certeza disso pode ser usado como base para rejeitar princípios que impõem este tipo de risco. Acredito que este argumento é pouco persuasivo. Em primeiro lugar, deve existir uma distinção entre rejeições a um princípio baseadas no fato de que este princípio impõe um fardo a mim e rejeições baseadas no fato de que ele impõe um fardo a *alguém*. O propósito da Restrição Individualista é justamente diferenciar entre diferentes tipos de reivindicação de acordo com sua relação com o agente. No contratualismo, a força de uma rejeição depende de como a adoção de um princípio está relacionada com minhas razões individuais. Assim, quando rejeito um princípio porque ele poderia impor fardo a alguém, toda a força dessa rejeição deve vir do fato de que esse alguém *poderia* ser eu. Mas isso não pode ser força suficiente nos casos que estamos discutindo, já que é implausível que esta reivindicação seja mais forte que a das pessoas que serão paralisadas, já que a probabilidade de morte é tão pequena. Até porque, no fim das contas, a própria pessoa que está objetando à pequena chance de morte é uma pessoa que será com certeza paralisada se a Vacina 1 for escolhida. Como Frick ressalta, este argumento “implica uma estranha assimetria estranha entre apostas individuais e casos de imposição de risco social”, já que seria muito estranho, do ponto de vista individual, que os guardiões de Clara julgassem que a Vacina 1 é melhor que a 2 para Clara com base na ideia de que *alguém* morrerá se a Vacina 2 for escolhida, já que no caso individual a probabilidade de alguma morte é inferior a 0,1% (FRICK, 2015, p. 196. Tradução minha.). É difícil imaginar, então, porque no caso *Vacinação em Massa (Vítimas desconhecidas)* “uma criança azarada deveria ter uma reivindicação maior que a de Clara, só porque, mesmo que *ele* não tivesse sido azarado, é estatisticamente certo que alguma *outra* criança teria morrido” (FRICK, 2015, p. 196. Tradução minha.). A ideia de “perda certa” não parece motivar a premissa 4 do argumento da informação irrelevante.

Outra forma de defender a premissa em questão vem do determinismo (FRICK, 2015, p. 197). O determinismo é, grosso modo, a ideia de que todos os eventos futuros são determinados pelos eventos que já ocorreram juntos das leis da natureza. A suposição de que o determinismo é verdadeiro, pelo menos em níveis macrofísicos, é amplamente aceita. Se este é o caso, as crianças

azaradas, isto é, as que sofrem os efeitos negativos da Vacina 2, nunca tiveram chance real de se curar, porque o evento “morte” já estava determinado desde o começo. Este fato está relacionado à ideia de que o conhecimento sobre as identidades das vítimas é irrelevante.

Recorde que nossos juízos sobre os dois casos de Vacinação em Massa eram de que no caso em que não conhecemos as vítimas a Vacina 2 é moralmente exigida, enquanto no caso em que conhecemos as vítimas esta Vacina é proibida. Estas intuições estão apoiadas na distinção de significância moral entre crianças “azaradas”, isto é, aquelas que receberam uma chance justa de sobrevivência, mas acabaram morrendo, e crianças “arruinadas”, aquelas que foram, essencialmente, mandadas para morrer quando a Vacina 2 foi escolhida, porque não possuíam chance nenhuma de sobrevivência.²⁰ Se o determinismo for verdadeiro, então é plausível supor que, do ponto de vista objetivo, todas as crianças são “arruinadas”, já que é comum supor que o determinismo é incompatível com a noção objetiva de probabilidade e, portanto, com a noção objetiva de sorte. Se as crianças “azaradas” do caso *Vacinação em Massa (Vítimas desconhecidas)* forem na verdade iguais às crianças “arruinadas” de *Vacinação em Massa (Vítimas conhecidas)*, então o conhecimento das identidades deve ser irrelevante do ponto de vista moral, e o argumento *ex post* da informação irrelevante funciona.

O argumento do determinismo para defender a premissa 4 depende de afirmações questionáveis, no entanto. A primeira é a própria ideia de que o determinismo é incompatível com probabilidades objetivas. Como Steuwer (2021b, p. 135) aponta, existe na literatura da metafísica contemporânea muitas abordagens de probabilidade objetiva compatíveis com o determinismo. A discussão detalhada destas abordagens está fora do escopo desta dissertação, mas esboçarei uma reconstrução do argumento de Steuwer, largamente inspirado em Glynn (2010) e List e Pivato (2015). A existência de leis probabilísticas nas chamadas “ciências especiais”, isto é, ciências de níveis de explicação superior, como as ciências sociais, a psicologia, teoria evolutiva, meteorologia etc., nos fornecem alguma razão para supor que chances objetivas existam, já que seria difícil explicar os mesmos fenômenos puramente com probabilidades epistêmicas. Para reconciliar esta ideia com o determinismo, Steuwer sugere que as ideias de “determinismo” e “chance” se aplicam não ao mundo em geral, mas a diferentes *níveis específicos*. Isso significa que probabilidades

²⁰ Este tipo de distinção é feito por Frick (2015, p. 182-183).

objetivas podem não existir no nível físico mas existirem em níveis de ordem mais alta de descrição. Considere o nível relevante para o raciocínio moral, que Steuwer chama de *nível agencial*. Este é o nível de descrição referente a agentes e suas ações (STEUWER, 2021b, p. 136). Temos boa razão para acreditar que o mundo não é determinístico neste nível, já que nos referimos apenas às decisões disponíveis a agentes, e não às propriedades microfísicas detalhadas do mundo. Steuwer oferece como exemplo um lance de moeda. Mesmo que seja possível, em princípio, descrever fisicamente um lance de moeda de maneira determinística, com condicionais do tipo “Se eu jogar a moeda com uma força x , ela cairá coroa”, “Se eu jogar a moeda com força y , cairá cara”, “Se eu jogar a moeda com uma força z , cairá coroa” e assim por diante, estas proposições não são opções disponíveis ao agente no nível agencial, o que significa que quando estamos preocupados com a avaliação de decisões morais, as proposições relevantes devem ser do tipo “Jogar a moeda possui probabilidade 50% de sair cara e 50% coroa”. Estas probabilidades são objetivas, isto é, características do mundo; mas são características do mundo no nível agencial, não no nível microfísico. Se isto é verdade, o determinismo não elimina completamente a noção objetiva de probabilidade, o que mantém a distinção entre crianças “azaradas” e crianças “arruinadas” que sustenta a escolha da Vacina 2. Esta resposta ao argumento do determinismo depende da questão metafísica sobre a natureza das probabilidades objetivas, mas é uma maneira de manter a intuição *ex ante*.

Frick (2015, p. 199) sugere uma resposta que precisa de menos suposições. O argumento do determinismo tenta solapar a distinção entre crianças azaradas e crianças arruinadas a partir da constatação de que probabilidades objetivas não existem em um mundo determinístico. No entanto, a distinção não depende da existência de probabilidades objetivas. As probabilidades relevantes discutidas em casos como *Vacinação em Massa (Vítimas desconhecidas)* são probabilidades *epistêmicas*, isto é, referentes ao grau de crença racional que um agente teria em determinadas proposições. Do ponto de vista da justificação de nossas ações, as probabilidades epistêmicas são as que importam, não as objetivas, porque é através delas que somos capazes de fornecer razões para decisões. Como Frick ressalta, é difícil imaginar um motivo para que, no caso individual, alguém preferisse estar em uma loteria indeterminística com probabilidades objetivas a uma loteria determinística com probabilidades meramente epistêmicas, se as probabilidades forem as mesmas. Fatos puramente epistêmicos, como não sabermos as identidades das pessoas que morrerão com a Vacina 2, ou não sabermos qual é o bilhete vencedor, podem influenciar a

moralidade e a racionalidade da decisão. Contrário ao que sugere o *Argumento da informação irrelevante*, a decisão moral em *Vacinação em Massa (Vítimas desconhecidas)* depende crucialmente do fato de que não sabemos as identidades das vítimas — a melhor maneira de justificar nossas ações a cada pessoa é afetada pelo nosso conhecimento na hora de tomar a decisão. Se lembrarmos que fatos epistêmicos são, assim como fatos sobre moralidade e racionalidade, um tipo de fato normativo, esta conclusão não é assim tão surpreendente. Não é o caso, portanto, que o determinismo sustenta a premissa 4 do *Argumento da informação irrelevante*.

Até aqui, o contratualismo *ex ante* se encontra em vantagem teórica em relação ao contratualismo *ex post*. Isso acontece porque a versão *ex ante* obtém o resultado intuitivo no caso de risco social discutido, resolvendo assim o problema moral do risco do qual o contratualismo *ex post* sofre. As próximas duas seções são dedicadas a duas objeções comumente levantadas contra o contratualismo *ex ante*.

5.4 O problema das regras *ex ante*

A ideia principal do contratualismo *ex ante* é que a justificação de ações envolvendo risco depende das chances oferecidas a cada pessoa, em vez dos resultados dos princípios relevantes. Alguns contratualistas *ex post* rejeitam este modelo de justificação porque, segundo eles, ele permite que ações injustificáveis passem no teste da rejeição razoável. Um exemplo clássico, que chamarei de *Experimento Médico*, foi trazido por Scanlon:

Considere qualquer princípio que nos permita impor fardos muito severos a uma pequena minoria de pessoas, escolhidas aleatoriamente (fazendo deles cobaias involuntárias de experimentos médicos dolorosos e perigosos, por exemplo) para beneficiar uma larga maioria. Um contratualista gostaria de manter aberta a possibilidade de que tal princípio fosse razoavelmente rejeitado por causa dos fardos severos que envolve (SCANLON, 1998, p. 208-209. Tradução minha.).

Se os críticos estão corretos, o contratualismo *ex ante* é incapaz de rejeitar princípios como o do Experimento Médico, já que é plausível que sejam do interesse de cada pessoa do ponto de vista *ex ante*, isto é, sem que os indivíduos conheçam seu lugar na distribuição de fardos e benefícios. Este é um problema similar aos problemas de distribuição de risco clássicos de teorias agregativas, mas que parece afetar a versão *ex ante* do contratualismo. Frick chama esta objeção de *Problema das regras ex ante* (FRICK, 2015, p. 202).

Rahul Kumar (2015, p. 36-39) respondeu a este tipo de objeção apontando para o fato de que o uso do corpo de cada indivíduo confere uma razão poderosa o suficiente para a rejeição razoável de princípios como o do caso Experimento Médico. Acredito que este não deve ser o fator relevante, no entanto, porque é possível pensar em casos estruturalmente similares a Experimento Médico que não envolvem o *uso* do corpo e que parecem também objeccionáveis. Isto é, o uso do corpo é parte do que dá razões para rejeição de princípios deste tipo; mas é possível rejeitá-los de outras maneiras também.

A sugestão de Frick de um Teste de Decomposição parece mais adequada. Segundo o teste proposto por Frick:

[...] Se uma regra ou procedimento pode ser decomposto em uma sequência de estágios causais distintos, cada um envolvendo a ação voluntária de algum agente [...], então é permissível adotar e agir de acordo com essa regra ou procedimento somente se as ações por ele requeridas em cada estágio são justificáveis para cada pessoa *naquele momento* (FRICK, 2015, p. 205. Grifo do autor. Tradução minha.).

A ideia, portanto, é que o caso *Experimento Médico* pode ser dividido em dois estágios: em primeiro lugar, escolhemos, aleatoriamente, quais pessoas serão cobaias; depois, conduzimos os experimentos nas pessoas selecionadas. Embora o primeiro estágio seja do interesse *ex ante* de todos, o segundo estágio não pode ser justificado para as pessoas que foram escolhidas, justamente porque as chances a elas oferecidas não são justas. Aplicar o teste a Vacinação em Massa ajuda a entender seu apelo: no caso das Vítimas desconhecidas, como não sabemos quem será afetado pela vacina, não faz sentido dividir a escolha em dois estágios; a decisão possui um estágio apenas que é justificável. Acredito que o teste proposto por Frick é uma boa maneira de defender o contratualismo *ex ante* do problema das regras *ex ante*, porque capta exatamente qual é o problema que torna princípios como o do *Experimento Médico* injustificáveis.

5.5 *Vidas estatísticas e vidas identificadas*

A última objeção a ser discutida é a de que o contratualismo *ex ante* privilegia o salvamento de vidas identificadas em detrimento das chamadas “vidas estatísticas”, mesmo em casos em que princípios deste tipo pareçam ser injustificáveis. Este é o principal problema do contratualismo *ex ante* e o mais difícil de ser resolvido. Considere este caso, adaptado de Frick (2015, p. 212):

Mineiros (1 vs. 100): P.F, um mineiro, está preso em um poço. Se não o salvarmos, ele morrerá em alguns dias. Entretanto, salvá-lo terá altos custos. Temos que escolher entre duas opções: Ou *Resgate*, na qual gastamos toda nossa verba para salvar P.F; ou *Prevenção*, na qual gastamos nossos fundos para melhorar a segurança na mina. *Prevenção* diminui o risco de morte de cada funcionário de 3% para 1%; como a mina possui 100 mineiros, espera-se que isso salve duas vidas. Se escolhermos esta opção, P.F morrerá.

Este é um tipo interessante de caso. Na maioria dos casos discutidos na filosofia moral, a opinião da moralidade de senso comum está de acordo com a da maioria dos filósofos, e o objetivo da argumentação é modificar teorias para que possam se adequar a estes juízos — tudo isso faz parte do equilíbrio reflexivo como método na filosofia moral. Casos como *Mineiros (1 vs. 100)*, no entanto, são difíceis de avaliar do ponto de vista do senso comum. Por um lado, é comum que as pessoas confirmem certo privilégio a salvar vidas identificadas. Por outro, esta característica do raciocínio moral é costumeiramente enxergada como meramente psicológica e sem base racional ou normativa que justifique moralmente essa ideia. O fato de que *Prevenção* salva mais vidas a longo prazo parece dar forte suporte moral para esta opção.

O contratualismo *ex ante*, no entanto, implica que no caso acima, devemos salvar P.F porque a teoria privilegia, de certa forma, o salvamento de vidas identificadas. Isto acontece porque as vítimas identificadas possuem uma razão forte para rejeitar *Prevenção*, enquanto cada uma das potenciais vítimas estatísticas possui uma razão descontada pela baixa probabilidade. A mesma característica que tornava o contratualismo *ex ante* atrativo no caso *Vacinação em Massa*, no qual estamos lidando com as chances oferecidas a um grupo só, torna-se problemática quando lidamos com casos em que um grupo compete com outro. Esta conclusão pode ser encarada como um problema para o contratualismo *ex ante*, mas sigo Frick (2015) em argumentar que *Resgate* pode ser a resposta moralmente correta.

No caso *Mineiros (1 vs. 100)*, a reivindicação de P.F é mais forte que a reivindicação de qualquer outro mineiro individual. Para ver isto, compare *Mineiros (1 vs. 100)* com *Mineiros (1 vs. 1)*. Neste caso, temos a escolha entre salvar P.F ou diminuir o risco de morte de um outro mineiro de 3% para 1%. Como a morte de P.F é *certa* se não escolhermos *Resgate*, parece que sua reivindicação é mais forte. Assim como em *Vacinação em Massa (Vítimas desconhecidas)*, a diferença entre arruinados (P.F se escolhermos *Prevenção*) e azarados (Os mineiros que morrerão

no futuro se escolhermos *Resgate*) é moralmente significativa. Esta consideração é base do que Frick chama de *Argumento Pró-vida identificada*.

Este raciocínio é interessante e acredito que ilumine boa parte das decisões morais que costumemente precisamos tomar sobre distribuições de riscos. No entanto, o viés pró-vida identificada do contratualismo *ex ante* é insensível ao número de vidas estatísticas do outro lado. Isso significa que, segundo a versão *ex ante* de contratualismo, devemos escolher *Resgate* em um caso *Mineiros (1 vs. 1000)*; o mesmo vale para *Mineiros (1 vs. 10000)* (FRICK, 2015, p. 218-219). Este resultado é bastante implausível, porque o número esperado de mortes nesses casos, se diminui de 3% para 1%, é reduzido em 20 e 200, respectivamente. Salvar P.F em *Mineiros (1 vs. 100)* pode ser justificado em parte porque possuímos razão moral para preferir salvar uma vida identificada a salvar duas vidas estatísticas, mas o argumento por trás dessa conclusão abre a porta para o sacrifício de inúmeras vidas estatísticas.

Responder a este desafio é complicado para contratualistas *ex ante*. Como o problema das vidas identificadas está ligado intimamente aos principais insights morais do contratualismo *ex ante*, talvez seja impossível construir uma versão plausível da teoria que obtenha outro resultado que não o salvamento de vidas identificadas em casos como *Mineiros (1 vs. 1000)*. Uma alternativa adotada por Frick é desistir da ambição de que o contratualismo seja uma teoria completa da moralidade. Ele propõe uma visão *pluralista* da moralidade interpessoal, segundo a qual nem todas as propriedades que fazem uma ação ser imoral estão relacionadas ao fato de que não podem ser razoavelmente rejeitadas (FRICK, 2015, p. 220). Isto significa que o contratualismo não seria capaz de oferecer uma análise da correção moral; o que a teoria faz, na verdade, segundo esta sugestão, é propor uma propriedade de nível superior (a injustificabilidade) que faz com que diferentes ações sejam erradas do ponto de vista moral. Acredito que o contratualismo perde bastante apelo se adotarmos esta postura, já que se torna, na verdade, outra coisa: deixa de ser uma teoria que fornece uma explicação completa da moralidade interpessoal.

Steuwer (2021b, p. 141-142) oferece outra explicação de porque o contratualismo *ex ante* não precisa se preocupar com o problema das vidas identificadas. Segundo o autor, o contratualismo é uma teoria bastante similar às teorias deontológicas e baseadas em direitos. Como resaltei no capítulo 2, o contratualismo moral é uma teoria baseada, de certa forma, na ideia de que a moralidade é um conjunto de permissões e proibições (as obrigações existem quando apenas

uma opção é permitida). As proibições do contratualismo podem ser entendidas como restrições deontológicas. É normal que filósofos de inclinação deontológica aceitem que estas restrições podem ser feridas em casos específicos, para evitar catástrofes morais. Steuwer sugere, então, que existe um limiar de consequências negativas a partir do qual devemos abandonar a moralidade contratualista, já que é uma teoria limitada ao entendimento da moralidade interpessoal. Esta opção tem o resultado, no entanto, de excluir o contratualismo como uma teoria moral útil para muitos casos de distribuição e de risco social.

Nenhuma das duas opções acima, que envolvem desistir do contratualismo como uma teoria completa da moralidade interpessoal, me interessa muito. Na verdade, acho que a opção de aceitar a conclusão do contratualismo *ex ante* sobre as vidas identificadas é a melhor solução se quisermos manter esta teoria. Embora inicialmente implausível, gostaria de elencar algumas razões para defender esta ideia. Em primeiro lugar, a moralidade do senso comum não é muito clara sobre esta questão: como mencionei anteriormente, existe um certo privilégio que agentes conferem a salvar vidas identificadas. Reflexão acurada sobre casos como *Mineiros (1 vs 1000)* pode fazer com que mudemos de ideia sobre esta questão. Não acho tão óbvio que escolher *Resgate* seja moralmente injustificável. Evidentemente, é uma opção *pior* no quesito consequências. Mas se a moralidade fosse apenas produto das consequências de ações, não haveria motivos para sermos contratualistas para começo de conversa. Então não acho que o contratualismo *ex ante* necessariamente esteja em conflito com nossas intuições morais sobre os casos discutidos; estas intuições são bastante complexas.

Em segundo lugar, gostaria de ressaltar que a urgência imposta no caso pode fornecer razão moral para preferir *Resgate*. Da maneira em que o caso é descrito, possuímos verba *apenas* para escolher uma das opções. Mas como *Prevenção* é uma opção que envolve o futuro, não escolher *Prevenção* não significa necessariamente que as mortes irão ocorrer: alguma coisa poderia mudar para que a mina adquirisse mais verba antes que as mortes estatísticas ocorram, por exemplo. É claro que, no contexto de um experimento mental, devemos encarar a decisão como uma só. Mas a base racional para escolher *Resgate* pode ser o fato de que as vidas estatísticas poderiam ser salvas, no futuro, por outros meios, enquanto a vida identificada de P.F só pode ser salva agora.

Minha proposta é a seguinte: um princípio moral com algum apelo pode ser incorporado ao contratualismo para explicar a razão pela qual escolher *Resgate* é justificável. Chamo esse

princípio moral de *Urgência*. Segundo este princípio, quando precisamos escolher entre salvar uma vida *agora* e salvar vidas que poderiam ser salvas no futuro por outros meios, é permitido salvar a vida em perigo urgente. É claro que *Urgência* não se aplica a todos os casos em que precisamos escolher entre salvar uma vida agora ou mais vidas no futuro; esta permissão especial só é acionada quando estamos lidando com vidas futuras que poderiam ser salvar por outras decisões futuras. Embora o caso dos mineiros seja subdescrito, é plausível que suas vidas possam ser salvas no futuro por outros meios, o que justifica a escolha de *Resgate* mesmo em casos como *Mineiros (1 vs 1000)*, salvando o contratualismo *ex ante* desta objeção. O problema da minha proposta é que o contratualismo *ex ante* continua com dificuldades em versões deste caso em que *Urgência* não se aplica; casos em que não é possível salvar as vidas estatísticas a não ser que escolhamos *Prevenção* agora. Em casos assim, as vidas estatísticas também estão em posição de urgência. E o contratualismo *ex ante* não consegue fornecer um raciocínio para a escolha de *Prevenção*; se a intuição de que esta é a escolha correta for muito forte, então o contratualismo *ex ante* perde força.

Minha conclusão sobre a plausibilidade do contratualismo moral levando em consideração o problema moral do risco é que a teoria precisa de reformas, mas é promissora em algum sentido. O problema das vidas estatísticas é real e precisa ser resolvido ou rejeitado. Não acredito, no entanto, que o contratualismo esteja em sérios problemas por não oferecer uma boa explicação do risco social neste caso específico, já que nenhuma das teorias morais tradicionais é capaz de oferecer uma explicação boa de todos os tipos de casos de risco moral. A moralidade do risco é um dos tópicos mais espinhosos da filosofia moral contemporânea. Além disso, acho que o contratualismo *ex ante*, apesar da consequência estranha nos casos entre vidas identificadas e vidas estatísticas, possui artifícios teóricos interessantes para abarcar tanto a ideia de que a distribuição de riscos deve ser justa quanto a ideia de que ações arriscadas podem ser justificadas dependendo do grau de risco envolvido. Este já é um grande avanço no caminho de desenvolver uma teoria moral que adequadamente analisa decisões morais sob risco.

6 Conclusões

A discussão até aqui avançou na justificação da escolha do contratualismo moral como teoria não-agregativa da moralidade. O primeiro capítulo, centrado na noção moral de agregação, trouxe um poderoso argumento contra as teorias agregativas, inspirado pelo argumento de John Rawls contra o utilitarismo: o de que teorias agregativas não respeitam a separação entre as pessoas. Algumas tentativas de resposta do lado dos agregacionistas foram consideradas e rejeitadas. Além disso, foi demonstrado que mesmo teorias agregativas que incluem cláusulas de distribuição igualitária encontram problemas com a separação das pessoas. A partir dessas considerações, filósofos morais têm três opções: reformar as teorias agregativas de maneira radical, rejeitar que a noção de separação entre as pessoas tenha importância moral ou abandonar a agregação e buscar teorias não-agregativas da moralidade. A primeira opção é interessante, embora seja difícil de imaginar reformas capazes de resolver o problema sem abandonar a ideia de agregação, já que existe uma conexão íntima entre a combinação dos fatores morais em um valor único e o apagamento das distinções entre as pessoas. A segunda alternativa, como foi brevemente discutido, é pouco atrativa, dada a plausibilidade da relevância moral da separação entre pessoas e das distinções morais que se baseiam nessa ideia. A decisão entre casos interpessoais e intrapessoais, inspirada no caso de Voorhoeve e Fleurbaey (2012), explora um tipo de distinção moral (sobre atribuições de risco) que teorias agregativas são incapazes de realizar.

A opção mais atrativa, e que é desenvolvida no capítulo 3, é a procura de uma teoria anti-agregativa da moralidade. A teoria candidata analisada nesta dissertação é o contratualismo moral, formulado de maneira influente por T.M Scanlon (1998). Segundo esta visão, a correção moral de ações depende de como os princípios que as permitem podem ser justificados a cada pessoa individualmente; se um princípio não pode ser razoavelmente rejeitado por alguém com base em considerações individuais, é um princípio aceitável. Durante este capítulo, noções centrais do contratualismo, como a ideia de rejeição razoável, entendida através da *Restrição Individualista* e do *Princípio do Fardo Maior*, foram apresentadas a fim de conferir maior clareza às análises morais contratualistas. As origens desta versão de moral de contratualismo, encontradas na teoria da justiça de Rawls, bem como as diferenças entre as duas devem ter sido clarificadas. Um ponto de contato especialmente importante entre as duas formas de contratualismo é a centralidade que conferem à noção de *justificação moral*. Tanto para Rawls como para o contratualismo de Scanlon, o que torna ações moralmente erradas ou corretas é uma propriedade superior sobre as ações que está ligada à sua justificabilidade; isto quer dizer que ações são erradas *porque* são injustificáveis, e não o contrário. Também foram exploradas as relações entre contratualismo moral, diferentes

princípios não-agregativos de decisão como Maximin e Leximin, e as características estruturais da agregação. O motivo pelo qual o contratualismo não é agregativo é sua rejeição da propriedade Continuidade: esta observação será relevante no resto da discussão.

Os capítulos 4 e 5 lidaram com objeções sofridas pelo contratualismo em função de ser uma teoria anti-agregativa. A primeira dificuldade encontrada é como lidar com casos em que os números parecem fazer diferença moral — como, por exemplo, a decisão entre salvar 1 ou 5 vidas. Este é o assunto do capítulo 4. O principal argumento contratualista para motivar um princípio de salvar mais vidas, conhecido como argumento do desempate, pode ser defendido através do uso de uma forma limitada de véu da ignorância. Existem ainda outras propostas para motivar tal princípio: examinei o argumento probabilístico de Schelling (2006), o argumento do princípio mais razoável, inspirado em Kamm (1993) e Hirose (2001; 2014) e as formas limitadas de agregação como o princípio RRA (Reivindicações agregadas relevantes), proposto por Voorhoeve (2014). Todas estas propostas são atrativas de alguma forma, porque estão baseadas no insight original por trás do argumento do desempate, que é a ideia de que podemos usar os números para decidir dilemas morais quando os fardos em questão são de níveis similares.

No último capítulo, me concentrei nos casos de risco social. A exclusão do contratualista da agregação cria uma dificuldade na avaliação de casos em que uma imposição de risco é justificada porque a atividade possui baixo grau de risco e altos benefícios esperados — um exemplo pungente é o caso *Vacinação em Massa* discutido em diferentes versões. A distinção entre contratualismo *ex post* e contratualismo *ex ante* iluminou duas maneiras diferentes de pensar sobre a justificação moral de princípios envolvendo decisões sob risco: para contratualistas *ex post*, o que importa são os resultados efetivos de uma ação arriscada; para contratualistas *ex ante*, importam as chances que um princípio oferece a cada pessoa. Foi demonstrado como o contratualismo *ex ante* não sofre do problema anteriormente discutido; além disso, não existe motivação forte o suficiente para aceitar o principal argumento por trás do contratualismo *ex post*, conhecido como Argumento da Informação Irrelevante.

No restante do capítulo, duas possíveis objeções ao contratualismo *ex ante* são consideradas. O *Problema das regras ex ante*, segundo o qual o contratualismo *ex ante* acaba sancionando ações injustificáveis quando estão no interesse *ex ante* de todos, foi resolvido com a sugestão de Frick (2015) de um teste de decomposição que exige que ações arriscadas sejam justificadas em cada estágio de efetivação. A segunda objeção é a que considero a maior dificuldade para o contratualismo moral, o problema das vidas estatísticas. O contratualismo privilegia o salvamento de vidas identificadas em vez de um número maior de vidas estatísticas,

isto é, vidas que temos razão para esperar que serão perdidas, mesmo sem saber a quem as vidas perdidas pertencem. A característica teórica que permitia ao contratualismo *ex ante* sucesso no caso *Vacinação em Massa (Vítimas Desconhecidas)* cria um problema sério. Defendi o contratualismo *ex ante* de alguns casos problemáticos como *Mineiros (1 vs 100)*, mas reconheço, junto de Frick (2015) e Steuwer (2021b), que a extensão do problema para casos como *Mineiros (1 vs 10000)* é difícil de resolver sem abandonar as ambições analíticas com as quais o contratualismo é proposto — que envolvem fornecer uma análise completa da moralidade interpessoal. Um princípio de urgência, proposto por mim como forma de explicar por que, em casos hipotéticos como os discutidos, existe certa força moral na reivindicação do salvamento de uma vida identificada agora contra múltiplas vidas perdidas no futuro, é uma forma de defender o contratualismo da implausibilidade inicial de suas consequências em casos como este. O fato de que o contratualismo encontra dificuldade quando se trata da moralidade de ações arriscadas não é surpresa, como deixo claro no início do capítulo. De fato, existe um problema moral do risco que afeta boa parte das principais teorias da moralidade. O contratualismo, na verdade, é uma teoria promissora neste quesito, se comparada, por exemplo, ao utilitarismo ou a teorias deontológicas tradicionais, porque consegue, pelo menos na versão *ex ante*, acomodar duas intuições importantes sobre o risco moral: (i) algumas ações arriscadas são justificadas ou até exigidas, dependendo da relação entre o grau de risco e os benefícios da prática; (ii) a maneira em que os riscos são distribuídos importa do ponto de vista moral, e a imposição de riscos deve respeitar a separação entre as pessoas. Ainda existe, no entanto, trabalho a ser feito por contratualistas neste campo.

De modo geral, o contratualismo moral se mostra uma teoria promissora no campo anti-agregativo das teorias morais. Fundamentada na ideia de justificação moral, é capaz de captar muitos dos juízos normativos intuitivos que regem a discussão da filosofia moral contemporânea, bem como obter sucesso onde outras teorias não conseguem, como na ética distributiva, em boa parte dos casos de risco social e na explicação da motivação moral. O movimento contratualista de rejeitar a agregação cria problemas para a teoria que podem ser resolvidos quando prestamos atenção cuidadosa ao conceito de agregação moral e o que realmente implica. A investigação filosófica sobre as implicações normativas e metaéticas desta teoria é justificada dados os resultados positivamente interessantes que o contratualismo inspirado por Scanlon trouxe à literatura da filosofia moral.

REFERÊNCIAS

- ARRHENIUS, Gustaf; BYKVIST, Krister. Future Generations and Interpersonal Compensations: Moral Aspects of Energy Use. Manuscrito.
- BRICKER, Adam Michael. Do Judgements about Risk Track Modal Ordering? **Thought: A Journal of Philosophy**, v. 7, n. 3, p. 200–208, 2018.
- BUCHAK, Lara. **Risk and Rationality**. [s.l.]: Oxford University Press, 2013.
- BUCHAK, Lara. Risk and Tradeoffs. **Erkenntnis**, v. 79, n. S6, p. 1091–1117, 2014.
- DE OLIVEIRA, Gustavo Oliva. Contratualismo como alternativa na filosofia moral. Manuscrito, 2019a.
- DE OLIVEIRA, Gustavo Oliva. Contratualismo ex post e ex ante: como evitar a agregação. *In*: ALVES, Eduardo; GABOARDI, Gregory; DA COSTA, Claiton Silva; *et al* (Orgs.). **XIX Semana Acadêmica do PPG em Filosofia da PUCRS - Volume 1**. [s.l.: s.n.], 2019b, p. 131–139.
- EBERT, Philip A.; SMITH, Martin; DURBACH, Ian. Varieties of Risk. **Philosophy and Phenomenological Research**, v. 101, n. 2, p. 432–455, 2020.
- FLEURBAEY, Marc; VOORHOEVE, Alex. Decide As You Would With Full Information! An Argument Against Ex Ante Pareto. *In*: NORHEIM, Ole; HURST, Samia; EYAL, Nir; *et al* (Orgs.). **Inequalities in Health: Concepts, Measures, and Ethics**. [s.l.]: Oxford University Press, 2013.
- FRICK, Johann. Contractualism and Social Risk. **Philosophy and Public Affairs**, v. 43, n. 3, p. 175–223, 2015.
- FRIED, Barbara H. Can Contractualism Save Us from Aggregation? **The Journal of Ethics**, v. 16, n. 1, p. 39–66, 2012.
- GLYNN, Luke. Deterministic chance. **British Journal for the Philosophy of Science**, v. 61, n. 1, p. 51–80, 2010.
- HARSANYI, John. Can the Maximin Principle serve as a basis for morality? A Critique of John Rawls' Theory. **American Political Science Review**, v. 69, p. 594–606, 1975.
- HARSANYI, John. Morality and the Theory of Rational Behavior. **Social Research**, v. 44, n. 4, p. 623–656, 1977.

HARSANYI, John. Morality and the theory of rational behaviour. *In*: SEN, Amartya; WILLIAMS, Bernard (Eds.). **Utilitarianism and beyond**. Cambridge: Cambridge University Press, 1982.

HIROSE, Iwao. **Moral Aggregation**. [s.l.]: Oup Usa, 2014.

HOLM, Sune. The Luckless and the Doomed. Contractualism on Justified Risk-Imposition. **Ethical Theory and Moral Practice**, v. 21, n. 2, p. 231–244, 2018.

JDAFRI2. **5. Should the Numbers Count?** [s.l.: s.n.], 2020. Disponível em: <<https://www.youtube.com/watch?v=j6LQfgM4YPA>>. Acesso em: 19 out. 2021.

KAVKA, Gregory S. The numbers should count. **Philosophical Studies**, v. 36, n. 3, p. 285–294, 1979.

KUMAR, Rahul. Risking and Wronging. **Philosophy and Public Affairs**, v. 43, n. 1, p. 27–51, 2015.

LENMAN, James. Contractualism and risk imposition. **Politics, Philosophy and Economics**, v. 7, n. 1, p. 99–122, 2008.

LIST, Christian; PIVATO, Marcus. Emergent Chance. **Philosophical Review**, v. 124, n. 1, p. 119–152, 2015.

M. KAMM, F. **Morality, Mortality: Rights, Duties, and Status**. [s.l.]: Oxford University Press USA, 1996.

MCGINN, Colin. Reasons and unreasons, **New republic**, May 24, 1999, p.35.

ORD, Toby. **Why I'm Not a Negative Utilitarian**. Disponível em: <<http://www.amirrorclear.net/academic/ideas/negative-utilitarianism/>>. Acesso em: 5 abr. 2021.

OTSUKA, Michael. Prioritarianism and the Separateness of Persons. **Utilitas**, v. 24, n. 3, p. 365–380, 2012.

OTSUKA, Michael; VOORHOEVE, Alex. Why it matters that some are worse off than others: An argument against the priority view. **Philosophy and Public Affairs**, v. 37, n. 2, p. 171–199, 2009.

PARFIT, Derek. Equality and priority. **Ratio**, v. 10, n. 3, p. 202–221, 1997.

PARFIT, Derek. Justifiability to each person. **Ratio**, v. 16, n. 4, p. 368–390, 2003.

PETERSON, Martin. **An introduction to decision theory**. 2. ed. [s.l.]: Cambridge University Press, 2017.

PETERSON, Martin. Some Versions of the Number Problem Have No Solution. **Ethical Theory and Moral Practice**, v. 13, n. 4, p. 439–451, 2010.

PETERSON, Martin. The Mixed Solution to the Number Problem. **Journal of Moral Philosophy**, v. 6, n. 2, p. 166–177, 2009.

PETTIT, P. Can Contract Theory Ground Morality?. In: DREIER, J. (org.). **Contemporary debates in moral theory**. Oxford: Blackwell Publishing, 2006.

_____. **The Common Mind**. Nova Iorque: Oxford University Press, 1993.

PRITCHARD, Duncan. Risk: Risk. **Metaphilosophy**, v. 46, n. 3, p. 436–461, 2015.

RASMUSSEN, Katharina Berndt. Should the probabilities count? **Philosophical Studies**, v. 159, n. 2, p. 205–218, 2012.

RAWLS, John. **Justiça como equidade**: uma reformulação. São Paulo: Martins Fontes, 2003.

RAWLS, John. **O liberalismo político**. São Paulo: Martins Fontes, 2011.

RAWLS, John. **Uma teoria da justiça**. 4 ed. São Paulo: Martins Fontes, 2016.

RAWLS, John. **Political Liberalism**. New York: Columbia University Press, 2005.

RAWLS, John. **Lectures on the history of moral philosophy**, org. Barbara Herman. Cambridge, Massachusetts: Harvard University Press, 2000.

REIBETANZ, Sophia. Contractualism and aggregation. **Ethics**, v. 108, n. 2, p. 296–311, 1998.

RESNIK, Michael D. **Choices: An Introduction to Decision Theory**. [s.l.]: Univ of Minnesota Press, 1987.

RIDGE, M. Contractualism and the new and improved redundancy objection. **Analysis**, v. 63, n. 4, p. 337–342, 2003.

RÜGER, Korbinian. On Ex Ante Contractualism. **Journal of Ethics and Social Philosophy**, v. 13, n. 3, 2018.

SANDBERG, Joakim; JUTH, Niklas. Ethics and Intuitions: A Reply to Singer. **The Journal of Ethics**, v. 15, n. 3, p. 209–226, 2011.

SANDERS, John T. Why the numbers should sometimes count. **Philosophy and Public Affairs**, v. 17, n. 1, p. 3–14, 1988.

SCANLON, T. M. **The Difficulty of Tolerance: Essays in Political Philosophy**. [s.l.]: Cambridge University Press, 2003.

SCANLON, T. M. **Being Realistic About Reasons**. [s.l.]: Oxford University Press, 2014.

SCANLON, Thomas. **What We Owe to Each Other**. [s.l.]: Belknap Press of Harvard University Press, 1998.

SCANLON, Thomas M. Contractualism and utilitarianism. *In*: SEN, Amartya Kumar; WILLIAMS, Bernard Arthur Owen (Orgs.). **Utilitarianism and Beyond**. [s.l.]: Cambridge University Press, 1982, p. 103--128.

SCHEFFLER, Samuel. **Human Morality**. [s.l.]: Oxford University Press, 1992.

SEN, Amartya; WILLIAMS, Bernard. **Utilitarianism and Beyond**. [s.l.]: Cambridge University Press, 1982.

SINGER, Peter. Ethics and Intuitions. **The Journal of Ethics**, v. 9, n. 3–4, p. 331–352, 2005.

SMART, J. J. C.; WILLIAMS, Bernard. **Utilitarianism: For and Against**. [s.l.]: Cambridge University Press, 1973.

SMART, J. J. C.; WILLIAMS, Bernard. **Utilitarianism: For and Against**. [s.l.]: Cambridge University Press, 1973.

STEUWER, Bastian. Aggregation, Balancing, and Respect for the Claims of Individuals. **Utilitas**, v. 33, n. 1, p. 17–34, 2021.

STEUWER, Bastian. Contractualism, Complaints, and Risk. **Journal of Ethics and Social Philosophy**, v. 19, n. 2, 2021.

SUIKKANEN, Jussi. Ex Ante and Ex Post Contractualism: A Synthesis. **The Journal of Ethics**, v. 23, n. 1, p. 77–98, 2019.

TAUREK, John. Should the numbers count? **Philosophy and Public Affairs**, v. 6, n. 4, p. 293–316, 1977.

TEMKIN, Larry S. Inequality: A Complex, Individualistic, and Comparative Notion. **Noûs**, v. 35, n. s1, p. 327–353, 2001.

TEMKIN, Larry S. **Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning**. [s.l.]: Oxford University Press, 2012.

VOORHOEVE, Alex; FLEURBAEY, Marc. Egalitarianism and the Separateness of Persons. **Utilitas**, v. 24, n. 3, p. 381–398, 2012.



Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Graduação
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: prograd@pucrs.br
Site: www.pucrs.br