



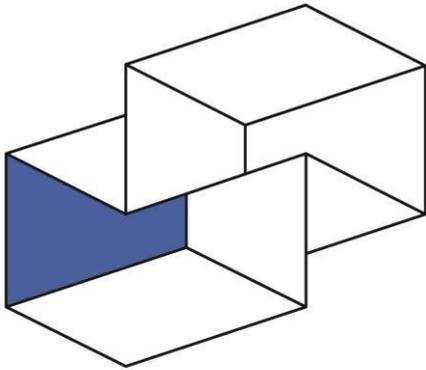
# **Investigação Filosófica**

**Revista de Filosofia**

ISSN: 2179-6742

**Investigação Filosófica, v. E4, Rio de Janeiro, 2016, 211 p.**

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
INSTITUTO DE FILOSOFIA E CIÊNCIAS SOCIAIS  
PROGRAMA DE PÓS-GRADUAÇÃO LÓGICA E METAFÍSICA



**PPGLM**

**Programa de Pós-Graduação  
Lógica e Metafísica**

Pós-Graduação em Filosofia - UFRJ

**Coordenador**

Rodrigo Guerizoli

**Vice Coordenador**

Carolina de Melo Bomfim Araújo

Revista desenvolvida em parceria com o Programa de Pós-Graduação Lógica e Metafísica (PPGLM) da Universidade Federal do Rio de Janeiro (UFRJ).

INVESTIGAÇÃO FILOSÓFICA  
<http://periodicoinvestigacaofilosofica.blogspot.com.br/>  
[ifilosofica@gmail.com](mailto:ifilosofica@gmail.com)

**Editores Responsáveis**

Rodrigo Reis Lastra Cid

**Coordenadores Editoriais**

Luiz Helvécio Marques Segundo  
Mayra Moreira da Costa  
Pedro Vasconcelos Junqueira Gomlevsky

**Conselho Editorial**

Danillo de Jesus Ferreira Leite  
Guilherme da Costa Assunção Cecílio  
Luis Fernando Munaretti da Rosa  
Luiz Helvécio Marques Segundo  
Luiz Maurício Bentim da Rocha Menezes  
Mário Augusto Queiroz Carvalho  
Mayra Moreira da Costa  
Pedro Vasconcelos Junqueira Gomlevsky  
Renata Ramos da Silva  
Rodrigo Alexandre de Figueiredo  
Rodrigo Reis Lastra Cid  
Sagid Salles Ferreira  
Tiago Luís Teixeira de Oliveira

**Conselho Consultivo**

Alexandre Meyer Luz  
Alexandre Noronha Machado  
Carlos Eduardo Evangelisti Mauro  
Desidério Orlando Figueiredo Murcho  
Guido Imaguire  
Mário Nogueira de Oliveira  
Michel Ghins  
Roberto Horácio de Sá Pereira  
Rodrigo Guerizoli Teixeira  
Rogério Passos Severo  
Sérgio Ricardo Neves de Miranda  
Ulysses Pinheiro

**Equipe Técnica**

Logotipo: Thiago Reis

INVESTIGAÇÃO FILOSÓFICA

Revista de Filosofia Semestral

Volume E4, 2016, 211p.

Publicação digital

ISSN: 2179-6742

1. Filosofia – Periódicos. 2. Universidade Federal do Rio de Janeiro. Instituto de Filosofia e Ciências Sociais. Programa de Pós-Graduação Lógica e Metafísica. 3. Blog Investigação Filosófica.

## Sumário / Contents

<b>Editorial</b> .....	05
------------------------	----

### Artigos/Articles

<b>Artigo 1:</b> Nomes próprios: por que a teoria causal de Perry não dá conta da retenção de crenças?.....	06
<b>Autor:</b> Sagid Salles	

### Traduções/Translations

<b>Tradução 1:</b> Nomes, Essência e Possibilidade.....	30
<b>Autor:</b> Scott Soames	
<b>Tradutor:</b> L. H. Marques Segundo	

<b>Tradução 2:</b> Metáfora.....	64
<b>Autor:</b> Max Black	
<b>Tradutor:</b> Carlos André Fernandes	
<b>Revisores:</b> L. H. Marques Segundo e Aluízio Couto	

<b>Tradução 3:</b> Conhecimento <i>a priori</i> .....	86
<b>Autor:</b> Albert Casullo	
<b>Tradutor:</b> Luiz Helvécio Marques Segundo e Alexandre Meyer Luz	
<b>Revisor:</b> Eduardo Benkendorf	

<b>Tradução 4:</b> O analítico e o sintético, o necessário e o possível, o <i>a priori</i> e o <i>a posteriori</i> .....	134
<b>Autor:</b> Scott Soames	
<b>Tradutor:</b> Luiz Helvécio Marques Segundo	

<b>Tradução 5:</b> A descoberta do mundo: O Ser Intemporal.....	164
<b>Autor:</b> Reinhardt Grossman	
<b>Tradutor:</b> L. H. Marques Segundo	
<b>Revisor:</b> Renato Mendes Rocha	

<b>Tradução 6:</b> O direito a um eleitorado competente.....	179
<b>Autor:</b> Jason Brennan	

**Tradutor:** Aluízio Couto  
**Revisor:** L. H. Marques Segundo

**Tradução das notas de “Conhecimento *a priori*”, de Albert Casullo.....212**

Investigação Filosófica, v. 6, n. 2, 2015. (ISSN: 2179-6742)

Editorial

## **Investigação Editorial**

Convidamos os nossos leitores a investigar nosso volume especial. Nele, apresentamos um artigo sobre filosofia da linguagem e epistemologia, de Sagid Salles, e diversas traduções de importantes filósofos contemporâneos, como Scott Soames, Max Black, Albert Casullo, Reinhardt Grossman e Jason Brennan. Esses textos vão desde a filosofia da linguagem até a metafísica e a filosofia política. Esperamos que vocês tenham uma boa e agradável leitura filosófica.

*Rodrigo Cid*  
Editor

# NOMES PRÓPRIOS: POR QUE A TEORIA CAUSAL DE PERRY NÃO DÁ CONTA DA RETENÇÃO DE CRENÇAS

Sagid Salles<sup>1</sup>

**RESUMO:** Tenho dois objetivos principais neste texto. Primeiro, defenderei que a versão de Perry da teoria causal da referência não dá conta de casos de retenção de crença. Faço isto por meio de um contraexemplo à teoria de Perry. Segundo, defenderei que uma versão específica de teoria da identificação é capaz de lidar com o contraexemplo apresentado e, neste sentido, é superior à teoria de Perry.

**PALAVRAS-CHAVE:** Nomes próprios, retenção de crença, teoria causal de Perry, teoria da identificação, teoria da referência.

**ABSTRACT:** I have two goals in this paper. First, I will claim that Perry's version of causal theory of reference is not able to account cases of belief retention. I shall do this by presenting a counterexample to Perry's theory. Second, I will claim that a specific version of theory of identification is able to deal with the example and, in this sense, must be preferred to Perry's theory.

**KEYWORDS:** Proper names, belief retention, Perry's causal theory, theory of identification, theory of reference.

## Introdução

Acredito que uma versão da Teoria da Identificação, que tem seus antecedentes em Evans, está correta em relação à referência

---

<sup>1</sup> Doutorando pelo Programa de Pós-Graduação em Lógica e Metafísica da Universidade Federal de Ouro Preto e bolsista pelo programa de Bolsas de Doutorado da CAPES.

dos nomes próprios.<sup>2</sup> Sustentei que esta versão é superior às diferentes versões da Teoria Causal, de Kripke, Donnellan e Devitt. Mas nada disse sobre a versão desenvolvida por Perry. A versão da Teoria Causal desenvolvida por ele é muito próxima da Teoria da Identificação que defendi. A principal diferença consiste no papel que ambos damos à causação na referência. Perry aceita que um elo causal de um tipo apropriado seja condição suficiente para o sucesso de um ato de referência através de um nome, eu recuso isto.

A principal motivação de Perry para a sua teoria é que ela dá conta de explicar casos de retenção de crenças. Meu principal objetivo neste artigo é defender que Perry não consegue realmente resolver problemas envolvendo retenção de crenças. Para isto repetirei um exemplo dado anteriormente contra outras versões da Teoria Causal da referência, e sustentarei que ele também pode ser usado para mostrar falhas na perspectiva de Perry. Penso que o erro de Perry é considerar um elo causal de um tipo apropriado condição suficiente para a referência bem-sucedida.

O artigo é dividido em cinco seções. Na primeira apresento um esboço da teoria que defendo. Na segunda, apresento a versão da Teoria Causal defendida por Perry. Ambas serão apresentadas apenas em caráter de esboço, sem mencionar detalhes que poderiam ser interessantes. Na terceira e quarta apresento e discuto um contraexemplo a Perry, sustentando que a sua teoria falha em explicar casos de retenção de crença. Embora não seja meu principal objetivo, sugeri, na quinta seção, que a minha perspectiva sobre a referência é mais promissora que a de Perry.

I

---

<sup>2</sup> Não apenas Evans (1982), mas também Campbell (2002), Jackson (2010), Sainsbury (2005) defendem perspectivas muito próximas da apresentada na parte I deste artigo. Mas não entrarei em quaisquer questões interpretativas aqui. Uma apresentação detalhada de minha versão de teoria da identificação para a referência, relacionando-a com as versões destes autores e respondendo a um conjunto de possíveis objeções pode ser encontrada em Salles (2013), principalmente a parte III. Minhas críticas à teoria causal nas versões de Kripke e Devitt podem ser encontradas na parte II do referido trabalho.

Antes de iniciar a discussão, é preciso fazer algumas observações preliminares. Alguma independência entre *referir-se* por um nome e *entender* o que foi dito através de um nome pode lhe parecer plausível. Evans deu alguma importância a isto. Pense no seguinte exemplo. Imagine que você esteja recebendo a visita de um amigo inglês que não fala ou entende português. Ambos estão numa festa, aqui no Brasil, e seu amigo demonstra nítido interesse em conhecer uma garota. Então seu amigo lhe pergunta (em inglês) como pode dizer, em português, que está interessado nela. Você então lhe explica que ele deve proferir a seguinte frase para ela.

1) Hitler foi uma boa pessoa.

Após treinar a pronuncia da frase, seu amigo vai até a garota e, com um sorriso no rosto, profere 1. A garota, chocada com o que ouviu, vai embora. Pergunte-se: terá seu amigo se referido a Hitler através do proferimento de 1? Quem aceita que o inglês teve sucesso em se referir a Hitler e dizer algo sobre ele, tem de aceitar alguma distinção entre *referir-se* e *entender*. Pois está claro pelo exemplo que o inglês não foi capaz de entender o que fez. Sua intenção era dizer que estava interessado na garota, e ele pensou estar dizendo isto quando proferiu 1. Assim, mesmo que ele tenha se referido a Hitler por 1, ele não entendeu o que foi dito pelo seu proferimento. Ele se referiu através de um nome sem ser capaz de entender o que o nome diz.

Eu penso que casos como este se baseiam em uma confusão entre um falante *ter de fato sucesso em se referir por um nome* e os ouvintes *terem razões para pressupor que o falante teve sucesso*. Nosso inglês não teve realmente sucesso em se referir a qualquer pessoa pelo proferimento de 1. Mas a garota tinha razões para pressupor que ele teve. (Ela não sabia que ele não conhecia o idioma, não sabia que estava sendo enganado, etc.). Seja como for, não vou argumentar a favor disto aqui. Menciono o ponto apenas para ressaltar que, daqui por diante, quando falar em referência, estarei arbitrariamente me restringindo aos casos nos quais os usuários do

nome são capazes não apenas de se referirem, mas também de entender o que dizem com o nome.

Alguém que aceite a distinção acima entre referir-se e entender pode pensar que um falante que é meramente capaz de se referir por um nome (mas não de entender o que o nome diz) não é capaz de usar o nome para expressar e sustentar crenças sobre o referente. Neste artigo, contudo, estou interessado justamente nos casos envolvendo crenças. Assim, quando dizer que um falante é capaz de se referir, também estarei aceitando que é capaz de sustentar e ter crenças sobre o referente do nome.

O problema da referência dos nomes próprios envolve, como nota Devitt (1981: 26-32, 1999: 66), pelo menos dois fenômenos. O primeiro é o da *fixação da referência*. Casos de fixação da referência são aqueles em que atribuímos um nome a um objeto. Geralmente isto é feito com um batismo, que pode ser cerimonial ou informal. O segundo é o fenômeno da *referência por empréstimo* ou *referência parasitária*. Após um nome ser atribuído a um objeto, ele pode ser difundido de falante para falante formando uma longa cadeia que vai muito além da circunstância de atribuição. Alguém usa um determinado nome, você ouve, e torna-se capaz de usar o mesmo nome para se referir ao mesmo objeto. Dizemos que você *tomou de empréstimo* a capacidade que o outro tinha de se referir pelo nome. O foco deste artigo será o segundo tipo de caso, isto é, a referência por empréstimo.

Uma *Teoria da Identificação* é aquela que aceita que a identificação é uma condição necessária para o sucesso da referência. De modo mais preciso, uma condição necessária para que um sujeito S tenha sucesso em se referir a um objeto *x*, através de um nome N, é que ele seja capaz de identificar *x*. Entende-se *ser capaz de identificar x* como ser capaz de diferenciar *x* de todos os outros objetos do mundo.

Nesta concepção, portanto, o Descritivismo Clássico, normalmente atribuído a Frege e Russell, é uma teoria da identificação. O mesmo vale para o Descritivismo Causal de Kroom (1987), Lewis (1984) e Jackson (2010). Estas perspectivas têm em comum o fato de aceitarem que pelo menos uma condição necessária para o sucesso da referência através de um nome é que o falante saiba que o referente do nome é o único a satisfazer alguma condição.<sup>3</sup> Mas saber que um objeto é o único a satisfazer uma determinada condição é um modo de diferenciar este objeto de todos os outros do mundo. Ou seja, de identificar este objeto. Mas nem toda versão de Teoria da Identificação precisa ser Descritivista neste sentido. A perspectiva que apresento a seguir é um exemplo de teoria da identificação, mas não é um caso de teoria descritivista.<sup>4</sup>

Para facilitar a discussão, chamemos de *consumidor* a cada usuário de um nome que se refere por empréstimo pelo nome.<sup>5</sup> Agora, a ideia é que consumidores usam nomes como mecanismos de identificação. Isto é, eles usam os nomes para identificar objetos. Você pode objetar que isto seria extremamente implausível, dado que diferentes objetos podem possuir o mesmo nome. O nome “Sócrates”, por exemplo, é usado para falar de dois indivíduos distintos, o jogador do Corinthians e o filósofo. Sendo assim, a mera posse deste nome não me permitirá diferenciar um do outro. Consequentemente, não me permitirá identificar qualquer um deles. Aliás, dado que é uma prática comum de nossa comunidade usar um mesmo nome para diferentes indivíduos, seria implausível sugerir que nomes servissem como mecanismos de identificação.

---

<sup>3</sup> Mas nem toda forma de Descritivismo precisa ser uma Teoria da Identificação. Pode-se imaginar, por exemplo, que a condição determinante da referência não precisa ser conhecida por cada usuário do nome, mas pela comunidade como um todo.

<sup>4</sup> Quando digo que minha perspectiva não é descritivista, quero dizer que não está comprometida com a tese de que uma condição necessária para um falante S se referir a um objeto O, através de um nome N, é que S tenha conhecimento proposicional de que o referente do nome é o único a possuir uma propriedade determinada. Conforme veremos, minha teoria não nos compromete com isto.

<sup>5</sup> Tomo o termo *consumidor* de empréstimo de Evans (1982) e McCulloch (1989: 282-283).

A objeção está correta, e revela outro aspecto importante da teoria. Nomes são apenas parte de um material maior que usamos para identificar seus referentes. Chamei este material anteriormente de *material identificador* (Salles, 2013: 62). Chamei à outra parte deste material, seguindo Jackson, de *marcadores*. Assim, o material identificador de alguém pode, por exemplo, ser formado por um nome + marcadores como “é um filósofo”, “é um jogador de futebol”, etc. Portanto, embora nomes sirvam como mecanismos de identificação, eles não têm de fazer seu trabalho sozinho.

Voltemos ao caso do nome “Sócrates”. Existem pelo menos dois objetos com este nome, de modo que a mera posse do nome não nos permitirá identificar qualquer um em particular. Neste contexto, podemos usar os marcadores que associamos ao nome para tornar nossa busca mais precisa. Marcadores como “é um filósofo”, “é grego”, etc. nos permitem diferenciar um indivíduo particular *dentre todos aqueles chamados “Sócrates”*.

Com isto, não estou sugerindo que os marcadores têm de ser *verdadeiros* dos referentes dos nomes para que a identificação ocorra. Um modo natural de interpretar a perspectiva acima seria pensar que o indivíduo identificado é aquele chamado “Sócrates” que *de fato* for um filósofo, for grego, etc. Não é esta a minha sugestão. O que importa não é que os marcadores sejam verdadeiros do referente do nome, mas que sejam informações difundidas acerca do referente. Assim, o objeto identificado será aquele chamado “Sócrates” acerca de quem os marcadores “é um filósofo”, “é grego”, etc. forem informações difundidas. Com algumas exceções, a verdade dos marcadores não importa.

Além disso, não estou sugerindo que o usuário do nome tenha de ter conhecimento proposicional de alguma descrição metalinguística como “o referente de N é quem quer que seja o objeto chamado N acerca de quem os marcadores tais e tais são difundidos”. Muito menos que os falantes comuns tenham que ser capazes de

especificar quaisquer condições deste tipo. Minha sugestão é apenas que os falantes *sabem como* usar seu material identificador para fins de identificação. Em outras palavras, um consumidor que usa um nome N com sucesso em se referir tem a habilidade prática de identificar o seu referente como aquele chamado N acerca de quem seus marcadores são difundidos.

Para tornar um pouco mais claro como a coisa toda funciona, pense no seguinte exemplo. Imagine que um grande historiador da filosofia descubra que Sócrates não foi realmente um filósofo, e sequer era grego. Sócrates foi apenas um amigo íntimo de Platão, e este difundiu um conjunto de estórias sobre ele. Suponha que nosso historiador publique um artigo em uma renomada revista de filosofia. Agora, imagine que todo o material identificador que um falante chamado “João” possui seja o nome “Sócrates” + o marcador “foi um filósofo da Grécia antiga”. Neste caso, embora João não saiba, nenhum de seus marcadores é uma informação verdadeira acerca de Sócrates. Mas isto não o impede de identificar Sócrates através de seu material identificador. Suponha que ele tente fazer uma busca no Google, a fim de aumentar seus conhecimentos sobre Sócrates. Coincidentemente, o primeiro artigo que ele clica é o de nosso historiador. Este artigo está repleto de frases como “ao contrário do que todos pensam, *Sócrates* não foi realmente um *filósofo*, sequer era *grego*”. João estará em perfeitas condições de perceber que o artigo trata da mesma pessoa que ele procurava. Não porque seu marcador é verdadeiro (ele não é!), mas porque ele é difundido. O indivíduo que ele encontrará será aquele chamado “Sócrates” acerca de quem seus marcadores são informações (verdadeiras ou não) difundidas. O que é realmente importante para a determinação do referente do nome “Sócrates”, como usado por João, é que ele encara seus marcadores como informações difundidas acerca do referente.

Por fim, repare que uma condição necessária para o sucesso de João em se referir pelo nome “Sócrates” é que exista ou tenha existido uma prática anterior de usar este nome na qual seus

marcadores foram difundidos acerca do referente. Isto não é uma desvantagem, dado que o que queremos explicar é o empréstimo da referência. Os casos de empréstimo da referência são justamente aqueles nos quais o sucesso de um falante em se referir é explicado em termos do sucesso de outros falantes.

Isto sugere a seguinte imagem geral sobre como um falante S pode identificar o referente de um nome N. S deve dispor de um material identificador, formado pelo nome N + um marcador - ou um conjunto de marcadores - £. S pode usar N para procurar alguém chamado N. Mas dado que há vários indivíduos chamados N, isto não será suficiente.<sup>6</sup> Assim, S pode usar seu marcador £ para distinguir um indivíduo específico dentre aqueles chamados N: aquele acerca de quem £ é uma informação difundida. Para que a identificação de S tenha sucesso, é necessário que exista uma prática de usar o nome N para se referir a um objeto x, na qual £ seja uma informação difundida acerca de x.

## II

Sustentei anteriormente que a perspectiva esboçada acima é superior a algumas versões da Teoria Causal da referência. Nomeadamente, as de Kripke (1972:91), Devitt (1981, 1999) e Donnellan (1974). O caso do último é um pouco mais complicado porque ele mesmo parece ser reticente quanto ao papel da causalidade na referência. Do modo como o interpreto, o essencial de sua teoria é apenas que a referência é um fenômeno histórico, e não que é causal.

Mas o que seria uma Teoria Causal da referência parasitária dos nomes? Neste texto, trato por Teoria Causal qualquer teoria que aceite que um elo causal de um tipo apropriado é, no mínimo, uma condição suficiente para a referência por nomes. De modo mais preciso, uma condição suficiente para um sujeito S se referir a um objeto x, através de um nome N, é que ele esteja em uma relação

---

<sup>6</sup> Na verdade, o próprio conhecimento do falante de que é permissível que diferentes objetos possuam o mesmo nome seria suficiente para que o falante, sendo competente, não usasse nomes isolados de seus marcadores como mecanismos de identificação.

causal apropriada com  $x$ . Eu ignoro se tal elo é uma condição necessária ou não. Seja como for, segue-se disso que a identificação não é uma condição necessária para o sucesso da referência. Nesta formulação, a Teoria Causal é inconsistente com a Teoria da Identificação.

Tanto Kripke como Devitt forneceram versões sobre como a referência dos consumidores funciona. Ambos sustentaram que um elo causal de um tipo apropriado é, pelo menos, uma condição suficiente para o sucesso do empréstimo da referência. Ambos aceitam, por exemplo, que um consumidor pode ter sucesso em se referir a algo por um nome mesmo que (i) não conheça qualquer condição que somente o referente do nome satisfaz, (ii) não seja capaz de identificar o referente, (iii) não se lembre que lhe introduziu o nome, etc. Mas foi Devitt quem tornou a teoria mais precisa.

Resumidamente, a perspectiva de Devitt sobre a referência parasitária é a seguinte:

*Ao ouvir  $S$  proferir o nome  $N$  para se referir a  $x$ , um sujeito  $S'$  pode, devido a uma relação causal com  $S$  (ele ouviu o proferimento de  $S$ ), adquirir a habilidade de usar  $N$  para se referir a  $x$ .*

A ideia é que um consumidor adquire a habilidade de usar um nome  $N$  para se referir a um objeto  $x$  ao entrar em contato com o exercício que outro sujeito faz de sua habilidade. Imagine que você seja capaz de se referir a Sócrates, mas eu não. Então, você me diz que Sócrates foi um grande filósofo. Neste momento, você exercitou a sua habilidade de usar o nome "Sócrates". Ao ouvir o que disse, eu adquiro a habilidade de usar o mesmo nome para me referir ao mesmo indivíduo ao qual você se referiu. O mesmo poderia acontecer se eu lesse uma frase proferida por você contendo o nome "Sócrates". Mas onde está o elo causal? Resposta: Devitt está interpretando a percepção (ouvir, ver, etc.) como um tipo de relação causal. Uma condição suficiente para o sucesso da referência de um

consumidor é que tudo corra bem com estes elos que, no fim das contas, recuarão até a situação do batismo do referente.

Perry aceita que a existência de um elo causal apropriado é uma condição suficiente para o sucesso da referência através de nomes. O diferencial de sua versão é que o importante para o elo causal é a informação associada ao nome, e não o nome em si. Na perspectiva de Devitt, quando vamos contar a história do elo causal, temos de nos focar nos diferentes proferimentos do nome, passando de falante para falante. Na perspectiva de Perry, o que é realmente importante é olhar para a história da difusão das informações que o consumidor associa ao nome. Em suas palavras:

(...) that it is not the name itself (ersatz or proper) that is of interest, but the grouping of predicates. It is the predicates grouped that lie at the end of a causal chain originating with the person I am thinking about.<sup>7</sup>

Para tornar a ideia mais clara, comecemos do começo. Primeiro, Perry apela a uma noção teórica muito próxima daquilo que chamei anteriormente de *material identificador* e outra próxima daquilo que chamei de *marcadores*. À primeira, ele chama de *arquivo*, à segunda, de *feixes ou agrupamentos de predicados*. Um arquivo pode ser formado por um nome + um agrupamento de predicados, exatamente como o material identificador o pode. Como na perspectiva apresentada na parte anterior, este agrupamento de predicados é, ou representa, informações acerca do referente do nome. Mas Perry pensa que um consumidor não precisa usar este material para identificar o referente. O referente do nome será quem quer que seja a origem causal das informações contidas no arquivo. Portanto, enquanto na minha perspectiva o material tem também uma função identificadora, na perspectiva de Perry ele tem apenas

---

<sup>7</sup> Perry, John (1980). A Problem About Continue Belief. In: *The Problem of the Essential Indexical*. Oxford University Press, 1993, pp. 69-90. p.85.

uma função causal. Mais uma vez, elo causal é uma condição suficiente para o sucesso da referência.

Não é minha intenção entrar em detalhes sobre a ontologia dos arquivos ou do material identificador. O próprio Perry parece ter a intenção de que sejam entidades mentais ou estejam de algum modo localizados no cérebro. Mas ele também não vai muito além. Para os nossos propósitos, a natureza dos arquivos não importa.

Uma consequência interessante deste tipo de perspectiva é que os nomes são, em certo sentido, dispensáveis. Podemos ter um arquivo formado apenas por predicados. Não é difícil pensar num exemplo em que isto ocorra. Imagine que uma testemunha ocular de um crime vá depor na delegacia. Ela diz ao delegado coisas como “ele era gordo, alto, e muito violento”. O delegado lhe pergunta se sabe o nome do suspeito. A testemunha responde que “não”. Nossa testemunha havia adquirido várias informações sobre o suspeito no dia anterior, formando um arquivo contendo predicados como “x é gordo”, “x é alto”, etc. Mas o arquivo não contém qualquer nome. Isto não impede a testemunha de ser capaz de se referir ao sujeito, muito menos de formar e reter crenças sobre ele.

O mesmo pode ser dito de minha perspectiva. Os marcadores de um indivíduo podem ser suficientemente ricos para permitir-lhe identificar um objeto independentemente de ele possuir um nome. Para pensar em outro exemplo, imagine que um especialista em Platão, devido a um caso raro de amnésia, esqueça o nome do indivíduo acerca de quem ele é especialista. Ele continua lembrando do que aprendeu, apenas esquece o nome. Sem dúvida, isto não o impede de fazer referência a este indivíduo, nem de manter suas crenças sobre ele.

A possibilidade destes casos coloca problemas interessantes. Se nomes são dispensáveis, então qual é, afinal, a sua utilidade? Eles têm alguma utilidade? Problemas como estes serão ignorados aqui. Me concentrarei apenas em casos envolvendo nomes próprios.

É preciso reconhecer que nem a minha perspectiva nem a de Perry são totalmente claras quanto a alguns pontos. Está Perry exigindo que cada informação presente no arquivo tenha de ser causalmente derivada do referente do nome? Ou pode ser o caso que algumas poucas informações falhem em satisfazer esta condição? A coisa pode ser ilustrada por um exemplo. Eu acredito que Sócrates é um grande filósofo. João acredita que Sócrates é o maior jogador do Corinthians. Obviamente, nós temos pessoas diferentes em mente. Mas não sabemos disso. Ao contar a João que Sócrates foi um grande filósofo grego, ele pensa que estou falando do jogador. Então, acrescenta o predicado “x é um grande filósofo grego” ao seu arquivo sobre o jogador. Agora, uma das informações que seu arquivo contém é causalmente derivada do filósofo, mas todas as outras são derivadas do jogador. João ainda será capaz de se referir e ter crenças sobre algum deles?

Novamente, o mesmo problema pode ser colocado para a minha perspectiva, embora fosse apropriado falar em outros termos. O material identificador de João, após o processo, conterá tanto informações difundidas sobre o filósofo quanto sobre o jogador. João ainda será capaz de se referir e ter crenças sobre algum deles? Problemas como este também serão ignorados ao longo do texto. As teorias acima serão propositalmente mantidas vagas e em caráter de um esboço.

Por fim, deixe-me dizer alguma coisa sobre a motivação de Perry para a sua teoria. A principal motivação é a sua aparente capacidade de lidar com problemas envolvendo aquisição e retenção de crenças. Casos de retenção de crença são aqueles em que um indivíduo mantém uma mesma crença ao longo do tempo. Suponha que ontem você tenha se encontrado com Maria e adquirido a crença de que ela está rica. Hoje você conta a novidade para um amigo: Maria está rica, sabia? Neste caso, você claramente reteve uma crença por um período de tempo. Ontem você *adquiriu* a crença de que Maria está rica, hoje você *continua* acreditando nisto. Este é um

caso simples, e não esperamos que alguma teoria plausível tenha dificuldades em lidar com ele. Mas existem casos mais complicados.

Para se ter uma leve ideia, voltemos ao suposto fato de nomes serem, no sentido antes mencionado, dispensáveis. Vimos dois tipos de casos. No primeiro, o indivíduo adquire e retém uma crença sobre o suspeito sem nunca ter possuído um nome para ele. No segundo, o indivíduo perde o nome que possuía, mas ainda retém crenças sobre o seu referente (Platão). Perry pode tratar os dois casos facilmente, pois o que é relevante para a determinação do referente e retenção de crenças não é o nome, mas as informações contidas no arquivo. Em ambos os casos, os indivíduos continuam possuindo um arquivo com determinado número de informações causalmente derivadas de um determinado objeto. O que faz com que o objeto da crença do primeiro seja o suspeito é que estas informações são causalmente derivadas do suspeito. O que faz com que o objeto da crença do segundo seja Platão é que suas informações são causalmente derivadas de Platão. Em suma, o que garante a retenção é o fato de, em cada caso, um mesmo indivíduo estar na origem causal das informações o tempo todo.

O tipo de perspectiva da identificação apresentado anteriormente também dá conta destes casos. A diferença é que a explicação seria dada em termos de identificação. Em ambos os casos, o sujeito manteria a sua capacidade de identificar um objeto independentemente de nunca ter tido ou ter perdido o nome.

Seja como for, penso que a perspectiva de Perry não fornece um tratamento adequado seja da referência, seja da retenção de crenças.

### III

Nas próximas seções, defendo que a teoria causal de Perry não dá conta realmente dos casos de retenção de crença. A estrutura de meu argumento será a seguinte. Primeiro, apresento um caso envolvendo a habilidade de um falante de usar um mesmo nome para

se referir e sustentar crenças sobre duas pessoas diferentes. Mais especificamente, nosso falante será capaz de usar o nome para sustentar uma crença C1 e outra crença C2, em que C1 e C2 são crenças sobre indivíduos diferentes. Em seguida, imagino um momento posterior específico do tempo, e coloco a seguinte pergunta: terá nosso falante retido as crenças C1 e C2? O objetivo é investigar o que a teoria de Perry nos diz sobre este caso. Sustentarei que existem três possibilidades de resposta, e defenderei que nenhuma das três está disponível a ele. Então concluo que a sua versão de teoria causal não dá conta deste caso. Por fim, explico como a perspectiva da identificação apresentada anteriormente pode tratar o mesmo caso, e sustento que ela é superior à de Perry. Passemos então ao exemplo.

Imagine uma conversa entre Marcos e seu amigo. Marcos não está interessado no que seu amigo diz, mas vez em outra finge estar prestando atenção. Em uma destas vezes, ele ouve o seguinte:

2) Maria é legal.

Depois disto, Marcos se distrai novamente. Ao longo do tempo, ele percebe que seu amigo começa a falar de outra pessoa, também chamada “Maria”, mas não presta atenção no que diz. Pouco depois ele ouve o seguinte:

3) Maria é legal.

Marcos sabe que 3 foi proferida para falar de outra pessoa, diferente da pessoa referida em 2. Suponha que tudo tenha corrido bem com ambos os proferimentos de seu amigo. Em outras palavras, seu amigo teve sucesso em se referir tanto por 2 quanto por 3.

Antes de começarmos a discutir este exemplo, façamos algumas estipulações. Sabemos que a pessoa referida pelo proferimento de 2 é diferente da pessoa referida pelo proferimento de 3. Usarei “\*” para indicar quando estou falando da segunda. Assim,

quando disser que Maria\* é tal e tal, estarei falando da pessoa referida por 3. Quando quiser falar do nome usado em 3, apenas acrescento aspas.

Penso que os seguintes pontos são intuitivos sobre o exemplo até agora.

- Marcos adquiriu a habilidade de se referir a Maria através do nome proferido em 2 (“Maria”).
- Marcos adquiriu a capacidade de se referir a Maria\* através do nome proferido em 3 (“Maria\*”).
- Marcos adquiriu, podemos supor, tanto uma crença sobre Maria (a crença de que Maria é legal) quanto uma crença sobre Maria\* (a crença de que Maria\* é legal). A primeira (C1) é uma crença diferente da segunda (C2).

Considero que nenhum dos pontos acima requer maior argumentação. Qualquer um que queira recusar algum (se houver quem o faça) terá o ônus da prova.

Mas agora continuemos com o exemplo. Imagine que Marcos finalmente se liberta de seu amigo e vai para casa. Cansado, ele só quer uma noite de sono. Ao acordar, lembra que foi introduzido a duas (e somente duas) pessoas chamadas “Maria”, que ambas eram legais e que se tratava de pessoas diferentes.<sup>8</sup> Ele não se lembra de qualquer outra coisa. Não sabe quem lhe introduziu os nomes, nem quando foi, etc. Em outras palavras, Marcos tem a seguinte crença:

C3: Existem duas, e somente duas, pessoas diferentes chamadas “Maria” que me foram introduzidas.

Repare que C3 é uma crença claramente diferente tanto de C1 quanto de C2. Se você ainda não está convencido disto, talvez a formalização

---

<sup>8</sup> O termo “introduzido” não deve sugerir que Marcos pense que foi apresentado às duas garotas pessoalmente. De fato, podemos supor que ele não sabe se foi introduzido às garotas pessoalmente ou se apenas ouviu, como realmente aconteceu, alguém lhe proferir os nomes.

ajude. Deixe “F” representar o predicado “chama-se “Maria””, e “G” representar o predicado “foi introduzido a Marcos”. A crença C3 pode ser formalizada como segue:

$$\exists x \wedge \exists y: [(x \neq y) \wedge (Fx \wedge Fy) \wedge (Gx \wedge Gy) \wedge \forall z ((Fz \wedge Gz) \rightarrow (z=x \vee z=y))]$$

Obviamente, nem a crença C1 nem a crença C2 possui esta forma lógica. Agora, pergunte-se: terá Marcos retido a sua crença de que Maria é legal ou a crença de que Maria\* é legal? Em outras palavras: Marcos reteve as crenças C1 e C2?

Há três possibilidades de resposta, que cito abaixo:

Opção 1: Marcos reteve tanto C1 quanto C2.

Opção 2: Marcos reteve somente uma das duas crenças acima (C1 ou C2).

Opção 3: Marcos não reteve nem C1 nem C2.

No que segue, considero todas as possibilidades acima, e sustento que nenhuma das opções está disponível a Perry. A conclusão é que a teoria não prediz qualquer resultado satisfatório para o exemplo acima.

#### **IV**

Antes de mais nada, repare que Marcos poderia usar certos artifícios para se referir e sustentar crenças sobre uma das duas garotas em particular. Ele poderia estipular, por exemplo, que deseja falar daquela pessoa, dentre as duas que lhe foram apresentadas, que nasceu primeiro. Não há dúvidas de que, neste caso, ele poderia se referir e sustentar crenças sobre esta pessoa. Por exemplo, se Maria\* foi a primeira a nascer, então ele poderia sustentar uma crença sobre Maria\*.

Este caso, contudo, não nos ajuda muito. Primeiro, porque nós podemos simplesmente restringir nossa pergunta ao momento em que ele acordou, antes de ele fazer qualquer estipulação deste tipo.

Segundo, e mais importante, supor que qualquer estipulação deste tipo é necessária seria um ponto a favor de teorias da identificação e contra perspectivas causais como a de Perry. Ora, ao fazer esta estipulação, Marcos está identificando o objeto de sua crença como a primeira a nascer, dentre as duas que lhe foram apresentadas. Perspectivas causais recusam que identificação seja uma condição necessária para referência. Terceiro, sequer seria claro se, após a estipulação, ainda temos um caso de retenção de crença, que é o que nos interessa.<sup>9</sup> Assim, devemos imaginar que Marcos não fez qualquer estipulação deste tipo. Dito isto, passemos à discussão do exemplo, considerando cada opção de resposta da seção anterior.<sup>10</sup>

*Opção 1: Marcos reteve tanto a crença C1 quanto a crença C2.* Para aceitar esta opção, Perry teria de sustentar que Marcos possui dois arquivos diferentes. Ambos os arquivos teriam o nome “Maria” + a informação de que o seu referente é legal e diferente do referente do nome contido no outro arquivo. A informação contida em um arquivo seria causalmente derivada de Maria, e a informação contida no outro seria causalmente derivada de Maria\*. Somente este estado de coisas permitiria a Perry sustentar que Marcos reteve ambas as crenças.

Penso que isto é implausível. Marcos não reteve as duas crenças. Para notar isto, imagine que ele tente contar a José uma das duas crenças que reteve. Tanto faz se você imagina que conta que Maria é legal ou se conta que Maria\* é legal. Como ele poderia fazer isto? Ele poderia apelar a artifícios como *aquela que nasceu primeiro*.

---

<sup>9</sup> Afinal, poderia ser o caso que ele simplesmente perdeu a crença e, após a estipulação, passou a acreditar naquilo novamente. Mas isto é diferente de reter uma crença. Além disso, não é óbvio que este caso envolvendo a estipulação envolva uma crença do mesmo tipo que Marcos tinha antes. Mas isto envolve problemas que não nos dizem respeito aqui.

<sup>10</sup> Em discussões deste texto, eu sou repetidamente apresentado a esta pergunta: mas Marcos não poderia falar daquela que nasceu primeiro? Ou não poderia usar artifícios que o permitiria distinguir a pessoa de quem ele fala de todos os outros objetos? Não poderia estipular uma propriedade, por exemplo, que somente uma delas satisfaz? A resposta, ad nauseum, é “sim, ele poderia fazer isto tudo”. Mas isto não afeta a questão principal: tais estipulações ou artifícios são necessários para o sucesso da referência de Marcos? Se sim, então nós já temos um argumento contra a teoria causal de Perry, que alega que um elo causal é condição suficiente. Por outro lado, se a resposta é “não”, então podemos prosseguir com o exemplo apenas assumindo que elas não ocorrem.

Poderia dizer coisas como: conheci duas pessoas chamadas “Maria”, e me refiro à mais nova. Mas, como vimos, tais artifícios estão vetados. Tudo que ele possui é a crença de que foi introduzido a duas pessoas com este nome e que ambas eram legais. Neste contexto, seria implausível supor que ele fosse capaz de contar ao José uma das crenças que adquiriu. A suposição de que ele possui os dois arquivos nos leva à conclusão de que Marcos tem uma crença que não pode expressar.

Suponha que Marcos tente expressar uma de suas crenças através de 4.

4) Maria é legal.

Qual crença ele está expressando? Está ele expressando C1 ou C2? Qualquer escolha aqui seria arbitrária. A razão de ser arbitrária é que ele não é capaz de expressar qualquer uma das crenças.<sup>11</sup>

A primeira objeção que alguém pode fazer a isto é que estou confundindo ter uma crença com ser capaz de expressá-la. Marcos reteve ambas as crenças, mas acontece que não pode expressá-las. Isto não é absurdo. Para pensar num exemplo radical, há alguma plausibilidade na ideia de que cachorros possuem crenças, mas não são capazes de expressá-las linguisticamente. Mas não se trata de cometer esta confusão. O problema com o caso de Marcos é que estamos assumindo que ele possui um arquivo contendo um nome mais uma informação causalmente derivada de um objeto. Além disto, assumimos que isto é condição suficiente para a referência e retenção de crenças. Mas, em seguida, negamos que ele possa expressar a sua crença. O que é implausível é a ideia de que ele

---

<sup>11</sup> Repare que a tese de que Marcos não é capaz de expressar qualquer uma das duas crenças em particular é consistente com a tese de que ele expressa *alguma* crença ao proferir 4. Uma hipótese, por exemplo, é que ele refira conjuntamente às duas mulheres através de 4. Quando eu digo “meus pais estão em casa”, refiro simultaneamente a meu pai e a minha mãe. Do mesmo modo, talvez Marcos refira simultaneamente a Maria e a Maria\* quando profere 4. Deve-se notar, contudo, que isto não afeta o argumento que apresento logo em seguida. O argumento só depende da suposição de que ele é incapaz de expressar tanto a crença específica de que Maria é legal, quanto a crença específica de que Maria\* é legal. Nem a possibilidade de que ele seja incapaz de expressar qualquer crença que seja, nem a possibilidade de que seja capaz de expressar alguma outra crença (como a de que ambas são legais) afeta o ponto do meu argumento.

possa usar este material para reter uma crença, mas não para expressá-la.

Outra objeção consiste em recusar que Marcos não possa expressar a sua crença. Ao proferir 4, ele expressou alguma das duas crenças que possui. Apenas acontece que nem nós nem Marcos sabemos qual é. Quando proferiu 4, Marcos *acessou* um de seus arquivos. Se o arquivo que ele acessou contém informações causalmente derivadas da Maria, então ele expressou a crença C1. Se contém informações causalmente derivadas de Maria\*, expressou uma a crença C2.

A razão pela qual acho isto implausível é que retira totalmente a responsabilidade de Marcos pelo que expressa. Lembre-se que Marcos não saberá explicar de qual das duas pessoas chamadas “Maria” ele fala. Tudo que ele sabe é que foi introduzido a duas e que ambas eram legais. Nada disso lhe permite sustentar que fala de uma delas em particular pelo proferimento de 4. Ele simplesmente profere uma frase, e seu cérebro faz todo o trabalho por ele. É como se ele pudesse decidir “vou proferir 4 para falar de uma das duas pessoas que me foram apresentadas” e depois, seu cérebro decidisse de qual delas ele fala. Parafraseando Armstrong sobre Quine, Marcos tem poder sem ter responsabilidade.

Resumindo. Começamos com a suposição de que Marcos reteve os dois arquivos. Neste caso, teria retido ambas as crenças C1 e C2. Sustentei que, se ele reteve alguma das duas crenças, seria capaz de expressá-la. Mas ele não é capaz de expressar qualquer uma das duas. Logo, não reteve qualquer uma. Talvez ele tenha retido alguma crença, diferente destas, ou talvez não tenha retido qualquer crença, mas isto em nada afeta o ponto do argumento.

*Opção 2: Marcos reteve somente uma das duas crenças.* Neste caso, ele terá retido uma crença acerca de uma das duas garotas. Para isto, ele tem de ter preservado pelo menos um dos dois arquivos. Se o arquivo contém informação causalmente derivada de Maria, então ele reteve a crença de que Maria é legal (C1), mas se o

arquivo contém informação causalmente derivada de Maria\*, então reteve a crença de que Maria\* é legal (C2).

Talvez o que tenha tornado implausível a ideia de que Marcos fosse capaz de expressar suas crenças no caso anterior foi a suposição de que ele possuísse os dois arquivos. Se supormos que ele possui apenas um, tudo correria bem. Ao proferir 4, Marcos acessa aquele arquivo que possui.

Mesmo que isto fosse plausível (e não penso que seja), teríamos o problema de explicar por que apenas um arquivo foi retido. O que há de especial em relação a um arquivo que o outro não tem? O exemplo não parece dar margem para sustentar que um deles tem algo de especial, que fez com que fosse retido. Nem mesmo para sustentar que um deles tem algo de problemático, que fez com que não fosse retido. O mais plausível seria sustentar ou que Marcos tem os dois ou que não tem qualquer um dos dois arquivos. Conclusão, a opção 2 não está disponível à Perry.

*Opção 3: Marcos não reteve nem C1 nem C2.* Para sustentar isto, Perry teria de afirmar que Marcos não reteve qualquer um dos arquivos, nem aquele com informação causalmente derivada de Maria, nem aquele com informação causalmente derivada de Maria\*.

Que razão Perry poderia fornecer a favor de que Marcos perdeu os dois arquivos? Marcos lembra que foi apresentado a duas pessoas diferentes, cada uma delas chamada “Maria”. Ele também sabe que recebeu o predicado “x é legal” de cada uma delas. Nada parece ter prejudicado a permanência dos arquivos. Nós temos o predicado “x é legal” ocorrendo ora como causalmente derivado de um indivíduo, ora como causalmente derivado de outro. O que impediria o nome “Maria” de estar associado a um destes predicados formando um arquivo e ao outro deles formando outro arquivo? Não vejo um elemento na teoria de Perry que nos permita responder a isto.

Talvez o problema esteja no fato de haver dois arquivos contendo o mesmo nome. Mas isto não pode ser um problema. Nós frequentemente somos capazes de nos referirmos a indivíduos

diferentes através do mesmo nome. Sou, por exemplo, capaz de me referir a dois indivíduos diferentes chamados “Sócrates”. Também o fato de os arquivos conterem o mesmo predicado não deve ser importante. Afinal, somos capazes de nos referirmos a vários indivíduos que satisfazem o predicado “x é bonito”.

Outra opção seria alegar que o problema está no fato de os arquivos conterem o mesmo nome e o mesmo predicado ao mesmo tempo. Em outras palavras, não poderia haver dois arquivos com *as mesmas coisas dentro*. Para que a coisa funcione bem, o material de cada arquivo teria de ser diferente. Penso que isto está na direção correta, mas esta resposta também não está disponível a Perry. Ele teria de explicar por que o material é importante numa perspectiva na qual o que determina o objeto de uma crença é a origem causal das informações. Seria necessário uma justificativa para aceitarmos uma restrição geral à formação de arquivos com o mesmo material. Esta justificativa terá de ser consistente com uma perspectiva causal, que aceita um elo causal apropriado como *condição suficiente* para a referência.

A situação é a seguinte. Antes de ir dormir, Marcos possuía duas crenças diferentes. A crença de que Maria é legal e a crença de que Maria\* é legal. Perry explica isto alegando que ele tinha dois arquivos diferentes, cada um contendo pelo menos o nome “Maria” mais o predicado “x é legal”. Estamos supondo que, ao acordar, ele não possuía mais qualquer arquivo. O que Perry tem dificuldades de explicar é o que deu errado neste processo.

Uma última alternativa seria alegar que o elo causal foi perdido porque os arquivos se misturaram. Antes, Marcos possuía dois arquivos contendo informações causalmente derivadas de dois indivíduos diferentes. Após acordar, de algum modo as informações se misturaram e formaram um único arquivo contendo informações sobre dois indivíduos diferentes.<sup>12</sup> Como os elos causais se misturaram, ele perde a capacidade de se referir e de sustentar

---

<sup>12</sup> Este terceiro arquivo seria diferente dos dois anteriores. De modo que, neste sentido, Marcos perdeu os dois anteriores.

crenças sobre os indivíduos isoladamente. Isto explica também porque ele não consegue, como vimos antes, expressar qualquer crença sobre um indivíduo particular através do proferimento de 4. Por fim, explica porque Marcos não reteve suas crenças.

Mas há razão para supor a existência de um único arquivo misturado? Marcos não fez nenhuma confusão deste tipo ao acordar. Ele não mistura informações sobre indivíduos diferentes. Ele lembra perfeitamente que foi introduzido aos nomes de *duas pessoas diferentes* e que recebeu informações sobre cada uma delas em particular. É difícil ver, neste contexto, como uma explicação causal poderia apontar para uma *mistura* dos arquivos ou da informação contida neles.

Em conclusão, a teoria de Perry apresenta problemas nos três casos possíveis. Ele não consegue explicar o que acontece com Marcos, se ele retém ou não suas crenças e a razão disto acontecer ou não. Lembremos que a motivação de Perry para a sua teoria era justamente a sua facilidade em lidar com problemas de retenção de crenças. Contudo, eu acabei de apresentar um caso de retenção de crenças do qual a teoria não dá conta. Meu diagnóstico é que há algo errado com esta teoria. Acredito que a teoria da identificação apresentada anteriormente pode ser usada para explicar o que está errado com a perspectiva de Perry.

## V

Termino considerando o que uma Teoria da Identificação, como aquela esboçada na parte I, nos diria sobre este exemplo. Primeiro, consideremos a situação de Marcos depois de acordar. Tudo que ele possuía naquela situação era a informação de que foi apresentado a duas pessoas chamadas “Maria” e que cada uma delas era legal. Assim, todo o material identificador que possui é o nome “Maria” mais o marcador “x é legal”. Mas este material identificador, obviamente, é insuficiente para identificar qualquer indivíduo particular. Dado que Marcos é incapaz de identificar qualquer indivíduo particular através deste material, ele também é incapaz de

se referir e sustentar crenças sobre qualquer indivíduo particular. Portanto, ele não retém qualquer uma das duas crenças que possuía antes. Marcos não reteve nem C1 nem C2.

Repare que isto não é o mesmo que afirmar que ele fez qualquer confusão ou que misturou arquivos. Nós podemos, se desejarmos, supor que ele possuía duas daquelas entidades que Perry chama de *arquivos*. Cada uma contendo o nome “Maria” mais o marcador “é legal”. Ainda assim, contudo, ele não terá retido qualquer uma das duas crenças. Isto porque o que importa não é apenas a origem causal dos arquivos ou da informação contida neles, mas o que fazemos com o material contido neles. Este material é usado para identificação de objetos. Mas neste caso, devido ao fato de os dois arquivos conterem um material muito pobre, eles não servirão para este propósito. Marcos sequer é capaz de diferenciar Maria de Maria\* com o material que possui. Portanto, não é capaz de identificar qualquer uma das duas.

Dado o que Marcos se lembra, nós estamos intitulados a afirmar que ele possui, *pelo menos*, a crença de que *existe um x e existe um y, tal que: x e y são chamados “Maria”, x e y são legais, e x é diferente de y...* Isto é, Marcos possui a crença C3. Mas esta é uma crença diferente tanto de C1 quanto de C2. Cada uma das duas últimas é uma crença sobre um objeto particular, C3 não. Ele está ciente de que foi introduzido duas vezes ao nome “Maria”, cada uma fazendo referência a um indivíduo diferente, e que recebeu a informação de que ambas eram legais. Mas isto é insuficiente para permitir-lhe identificar cada uma das pessoas em particular. Ele não é capaz de diferenciar Maria de todos os outros objetos do mundo, e nem é capaz de diferenciar Maria\* de todos os outros objetos do mundo. Mais uma vez, ele sequer é capaz de diferenciar uma da outra. Por isto, não possui nenhuma das duas últimas crenças.

Conforme dito antes, ele seria capaz se enriquecesse seu material com mais marcadores.<sup>13</sup> Ele pode fazer isto estipulando que

---

<sup>13</sup> Embora isto não signifique que ele reteve a crença.

deseja falar daquela, dentre as duas, que nasceu primeiro. Em todo caso, estipulações como esta envolvem uma peculiaridade. Seu sucesso em se referir dependeria de a estipulação feita ser verdadeira de um e somente um indivíduo. Se ele estipula que deseja falar daquela, dentre as duas chamadas “Maria” que lhe foram apresentadas, que nasceu primeiro, então seu sucesso em se referir dependerá de existir um e somente um indivíduo, dentre as duas, que nasceu primeiro. Este seria uma das exceções em que os marcadores teriam de ser verdadeiros do referente.

O mesmo pode acontecer em relação aos momentos antes de dormir. Marcos possuía tanto a crença de que Maria é legal como a de que Maria\* é legal. Mas era capaz de identificar cada uma de um modo próximo ao que foi dito acima. Ele poderia, por exemplo, identificar Maria como *a pessoa a quem seu amigo se referiu da primeira vez*, e Maria\* como *a pessoa a quem seu amigo se referiu da segunda vez*. A diferença é que no momento da conversação (e talvez tempo depois) ele não precisou fazer qualquer estipulação. Isto estava dado pelo contexto.

Portanto, a resposta ao exemplo reside no quão rico é o material identificador de Marcos. No momento depois de acordar, parece claro que o material é pobre e não lhe permite identificar ninguém em particular. Assim, ele não reteve as suas crenças. No momento da conversação parece claro que o material era suficientemente rico, de modo que Marcos era capaz de identificar ambas as mulheres, e sustentar crenças sobre cada uma delas em particular.

O mesmo valerá para os momentos após a conversação e, no entanto, antes de Marcos ir dormir. Suponha que Marcos conversou com seu amigo há duas horas, mas que ainda não foi dormir. Marcos reteve as duas crenças que tinha no momento da conversação? Isto dependerá da riqueza do material retido por ele. Se, neste momento, tudo que ele lembra é que foi introduzido duas vezes ao nome “Maria” e que o referente de ambos era legal, então não terá retido

nem C1 nem C2. A sua retenção destas crenças dependerá de seu material identificador permanecer rico o suficiente para identificar cada uma das mulheres em particular.

Assim, conseguimos explicar tanto como Marcos pode possuir duas crenças diferentes em alguns momentos antes de ir dormir quanto por que não as possui depois de acordar. A perspectiva da identificação tem mais facilidade que a de Perry para lidar com o exemplo acima. O erro de Perry, penso, foi não ver o papel identificador do material do falante.

Em conclusão. Perry pensa que uma vantagem de sua Teoria Causal é explicar a retenção de crenças. O que garante a retenção de uma crença sobre  $x$  é a existência de um arquivo contendo informações causalmente derivadas de  $x$ . Sustentei que esta perspectiva não tem realmente sucesso em explicar a retenção de crenças. Fiz isto através do exemplo de Marcos. Defendi que nenhuma das respostas disponíveis a Perry é satisfatória quanto ao caso. Por sua vez, sugeri que a perspectiva da identificação apresentada na seção I é capaz de explicar o que acontece no caso de Marcos. A perspectiva de Perry é bem próxima da minha. A principal diferença é que acrescento um papel identificador aos marcadores, recusando que alguma espécie de elo causal seja condição suficiente para referência e retenção de crenças. Deve-se manter em mente, contudo, que nada disse sobre a necessidade de um elo causal.

Nem de longe isto é algo decisivo a favor da perspectiva da identificação. Isto principalmente porque existe um conjunto de exemplos que supostamente mostram que é demasiado exigente postular identificação como condição necessária para a referência. Há uma crença razoavelmente difundida de que falantes completamente ignorantes acerca do referente do nome podem, ainda assim, ter sucesso em se referir pelo nome. E o mesmo vale para a capacidade de sustentar e reter crenças sobre o referente do nome. O próprio Perry pensa isto (1980:83). Por outro lado, acredito já ter mostrado

em outro lugar que estes exemplos não são realmente um problema (Salles, 2013:74-81). Mas esta é outra história.

### **Observações Finais**

Não acho que a solução de Perry para a referência por empréstimo esteja totalmente errada. Aliás, a perspectiva de Perry é muito próxima daquela que eu mesmo defendo. De fato, acho que Perry acertou em quase tudo. Seu erro foi não reconhecer um papel adicional dos marcadores, o papel identificador que eles têm.

Repare, portanto, que não recuso que elos causais são uma parte importante da explicação correta da referência. De fato, a perspectiva da identificação apresentada na seção I aceita tais elos. Eu sequer recuso que a existência de um elo causal de um tipo apropriado seja uma condição necessária para o sucesso da referência por empréstimo. Novamente, a única coisa que defendi é que a teoria de Perry erra ao supor que certo tipo de elo causal é *suficiente* para o sucesso da referência.

Por fim, é importante entender que, pelo menos em um sentido, eu **não** sou um descritivista. O descritivismo é muitas vezes interpretado como afirmando que: uma condição necessária para o sucesso da referência é que o usuário do nome tenha conhecimento proposicional de que o referente do nome é o único a possuir certa propriedade. Se interpretarmos o descritivismo como comprometido com esta tese, então não serei um descritivista. A teoria da identificação que defendi para o fenômeno da referência por empréstimo não requer qualquer conhecimento proposicional deste tipo. De fato, tudo que ela exige é que falantes tenham, pelo menos, a *habilidade prática* de identificar o referente do nome, que os falantes pelo menos *saibam como* usar o seu material para tal fim. Aparentemente, nada nos comprometeria a interpretar esta habilidade prática como conhecimento proposicional. Seja como for, a relação entre o tipo de tese que defendo e teorias descritivistas foi

investigada, em maior detalhe, em outro lugar (Salles: 2013: seção 3.16).

### **Referências**

CAMPBELL, John (2002). *Reference and Consciousness*. Oxford Cognitive Science Series. Clarendon press. Oxford.

DEVITT, Michael (1981) "A Causal theory of designation (1 e 2)". In: *Designation*. Columbia University Press.

DEVITT, Michael. Sterelny, Kim (1999) *Language and Reality: an introduction to the philosophy of language*. Oxford, Blackwell.

DONNELLAN, Keith (1974) "Speaking of Nothing". In: *The Philosophical Review*. Vol. 83, No 1, pp. 3-31. Duke University Press.

EVANS, G. (1982) *The Varieties of Reference*. Clarendon Press - Oxford University Press - New York. 2002.

JACKSON, Frank (2010). *Language, Names, and Information*. Blackwell.

KRIPKE, Saul A. (1972/1980) *Naming and Necessity*. Oxford, Blackwell, 1990.

KROON, Frederick W. (1987). "Causal descriptivism", *Australasian Journal of Philosophy*, 65:1, 1-17.

LEWIS, David (1984) "Putnam's Paradox" In: *Papers in metaphysics and epistemology*. Cambridge University Press, 1999. pp. 56-77

McCULLOCK, Gregory (1989). *The Game of the Name: Introducing Logic, Language and Mind*. Oxford University Press. 2003.

PERRY, John (1980) "A Problem About Continued Belief". In: *The Problem of Essential Indexical and Other Essays*. Oxford University Press, 1993.

SALLES, Sagid (2013). *Nomes Próprios: Referência e Identificação*. Disponível em: <http://ppglm.files.wordpress.com/2008/12/dissertacao-ppglm-sagid-ferreira.pdf>.

## NOMES, ESSÊNCIA E POSSIBILIDADE\*

Scott Soames

Tradução de L. H. Marques Segundo<sup>14</sup>

### PRÉVIA DO CAPÍTULO

1. *A significância do Naming and Necessity*
2. *Por que as descrições não fornecem o significado dos nomes*

O argumento modal de que os significados dos nomes não são dados pelas descrições que os falantes associam a eles.

3. *Designação rígida*

Designação Rígida e o Argumento Modal

Definição de designação rígida; o argumento de Kripke de que os nomes são rígidos ao passo que as descrições tipicamente associadas a eles não; a tentativa de evitar esse argumento apelando-se para as descrições rigidificadas; a extensão do argumento modal para rebater essa proposta.

Uma confusão a ser evitada

Como é que a afirmação de que *n* designa *o* pode não ser necessária, muito embora *n* designe *o* no que diz respeito a todos os estados do mundo possíveis.

Designação rígida e essencialismo

O uso da designação rígida para rebater a objeção de Quine de que não faz sentido dizer de *o*, independentemente de como ele é descrito, que ele tem uma propriedade essencialmente.

Designação rígida, mundos possíveis e “identificação transmundo”

Mundos possíveis não são universos alternativos que requerem critérios para a identificação das contrapartes dos objetos efetivamente existentes; eles são estados possíveis nos quais o universo poderia ter estado; o sentido no qual os designadores rígidos são usados para “estipular” esses estados.

---

\* “Names, Essence, and Possibility”. *Philosophical Analysis in Twentieth Century*, vol. 2: *The Age of Meaning*. Princeton University Press, 2002. Capítulo 14.

<sup>14</sup> Bolsista CAPES, doutorando pela UFSC.

4. *Por que as descrições normalmente não fixam os referentes dos nomes*

Os argumentos contra a versão fraca fixadora da referência da teoria das descrições. / A concepção da cadeia causal quanto à determinação da referência de Kripke. / As cadeias históricas fornecem descrições fixadoras de referência? Por que elas não fornecem; resolução de uma falta de clareza sobre a fixação da referência.

**A significância do *Naming and Necessity***

Neste capítulo começamos nossa discussão do livro de Saul Kripke, *Naming and Necessity*, que foi originalmente apresentado como três longas palestras públicas na Universidade de Princeton em janeiro de 1970, quando Kripke tinha 29 anos.<sup>15</sup> Na época foi feita uma gravação das palestras, e dois professores do departamento de filosofia de Princeton, Gilbert Harman e Thomas Nagel, transcreveram as gravações. Kripke adicionou as notas de rodapé, e mais tarde escreveu um prefácio para a versão em livro. O impacto das palestras foi profundo e imediato, e a sua influência aumentou com os anos. Na filosofia da linguagem, o *Naming and Necessity* está entre as obras mais importantes de sempre, competindo com o trabalho clássico de Frege no final do século dezanove, e de Russell e Tarski na primeira metade do século vinte. Além da filosofia da linguagem, mudou fundamentalmente o modo no qual se fazia filosofia. Os aspectos mais importantes da obra são (i) um conjunto de teses sobre o significado e a referência de nomes próprios, (ii) um conjunto correspondente de teses sobre o significado e a referência dos termos para categorias naturais como *calor*, *luz*, *ouro*, *água*, e *tigre*, (iii) uma defesa convincente dos conceitos metafísicos de necessidade e possibilidade, (iv) uma nítida distinção entre a noção metafísica de necessidade e a noção epistemológica de aprioricidade, (v) poderosos argumentos de que há verdades necessárias que são conhecíveis apenas *a posteriori*, e verdades *a priori* que são contingentes e, por conseguinte, não necessárias, e (vi) a defesa persuasiva da inteligibilidade do essencialismo – *i.e.*, a tese de que faz sentido caracterizar objetos como tendo algumas de suas propriedades essencialmente, e outras acidentalmente. Além desses aspectos explícitos da obra, a discussão em *Naming and Necessity* teve amplas implicações para o que veio a ser conhecido como *externismo* sobre o significado e a crença – grosso modo, a ideia de que os significados das palavras de alguém, assim como os conteúdos de suas crenças,

<sup>15</sup> Saul Kripke, *Naming and Necessity*, Cambridge, MA: Harvard, 1980 (Trad. Portuguesa: *O Nomear e a Necessidade*. Gradiva 2012), originalmente publicado em D. Davidson e G. Harman, eds., *Semantics of Natural Languages* (Dordrecht: Reidel, 1972); as citações serão do texto 1980.

são parcialmente constituídos por fatos inteiramente fora dele. Finalmente, o *Naming and Necessity* desempenhou um grande papel na rejeição implícita, embora difundida, da perspectiva – bastante popular entre os filósofos da linguagem comum – de que a filosofia nada mais é do que análise da linguagem.

### **Por que descrições não dão os significados dos nomes**

Começaremos a nossa investigação do trabalho de Kripke examinando a sua discussão da teoria descritivista dos nomes próprios, da qual ele distingue duas versões. De acordo com a primeira, os nomes próprios têm os mesmos significados que as descrições que os falantes associam a eles. De acordo com a segunda, embora os nomes possam não ser sinônimos das descrições, o referente de um nome próprio  $n$ , quando usado por um falante num dado instante, é determinado, como questão de regra linguística, como sendo um e único objeto que satisfaz as descrições associadas a  $n$  pelo falante naquele instante. Uma vez que o significado de um termo supostamente determina a sua referência, a primeira versão da teoria descritivista é entendida como implicando a segunda. A conversa, no entanto, não se dá; pode acontecer que o referente de um nome seja, como questão de regra semântica, determinado por uma descrição, muito embora o nome não seja sinônimo da descrição. Essas duas versões da teoria descritivista são expressas pelas teses 1 e 2.<sup>16</sup>

### **DUAS TESES SOBRE OS NOMES**

Tese 1: O significado de um nome  $n$  (para um falante num dado momento) é dado por uma descrição, uma conjunção de descrições, ou um agregado de descrições  $D$  que o falante associa a  $n$  naquele instante. Se  $D$  dá o significado de  $n$ , então a substituição de um pelo outro preserva tanto o significado quanto a proposição expressa. Assim, se  $S'$  resulta de  $S$  pela substituição de  $D$  por uma ou mais ocorrências de  $n$  em  $S$ , então  $S$  e  $S'$  significam a mesma coisa e expressam a mesma proposição.

Tese 2: O referente de um nome  $n$  (para um falante num dado momento) é semanticamente fixado (determinado) por uma descrição, uma conjunção de descrições, ou um agregado de descrições  $D$  que o

---

<sup>16</sup> Essas são reconstruções das duas teses da maneira como Kripke as entende. Para as próprias formulações de Kripke, veja pp. 71-80 do *Naming and Necessity*.

falante associa a  $n$  naquele instante. Se  $D$  fixa o referente de  $n$ , então:

- (i) o falante acredita que  $D$  se aplica a um único indivíduo;
- (ii) se  $D$  se aplica a um único indivíduo  $o$ , então  $o$  é o referente de  $n$ ;
- (iii) se  $D$  não se aplica a um único indivíduo, então  $n$  não tem referente.
- (iv) o falante sabe (é capaz de saber) *a priori* que **se  $n$  existe, então  $n$  é  $D$**  expressa uma verdade.

Na palestra 1, Kripke oferece um argumento contra a tese 1, que veio a ser conhecido como *o argumento modal*. Ilustremos o argumento com um exemplo. Seja  $n$  o nome *Aristóteles*. Sejam (1-8) candidatos à descrição  $D$  que dá o significado do nome.

1. O fundador da lógica formal
2. O maior aluno de Platão
3. O professor de Alexandre
4. O famoso filósofo grego nomeado "Aristóteles"
5. O último grande filósofo da antiguidade
6. A conjunção de 1-5
7. A conjunção de todas as descrições que o falante associa a  $n$
8. Um agregado de descrições, incluindo 1-5, que o falante associa a  $n$

Para os nossos propósitos, a afirmação de que o significado de um nome é dado por um agregado de descrições  $D_1 \dots D_n$  será tomado com equivalente à afirmação de que o significado do nome é dado pela descrição ***a coisa da qual a maior parte (ou um número suficiente) das seguintes são verdadeiras: é  $D_1$ , é  $D_2$ , ..., é  $D_n$*** . Podemos agora testar a afirmação de que o significado de *Aristóteles* é dado por uma ou mais das descrições (1-8), aplicando o seguinte teste modal.

### **O TESTE MODAL**

Se  $D$  dá o significado de  $n$ , então a proposição (verdade/declaração/afirmação) expressa pela frase

***Se  $n$  existiu, então  $n$  era  $D$ .***

é uma verdade necessária.

A razão por trás desse teste é que se D tem o mesmo significado de *n*, então a substituição de uma pela outra numa frase não mudará a proposição expressa (ou a afirmação feita). Mas isso significa que a frase **Se *n* existiu, então *n* era D** expressa a mesma proposição (diz a mesma coisa) que a frase **Se D existiu, então D era D**. Uma vez que a última frase expressa uma verdade necessária, a primeira tem também de expressar. Ao usar a terminologia *verdade necessária* diremos o seguinte: uma proposição é uma verdade necessária sse (i) é verdadeira dado o modo como o mundo efetivamente é e (ii) teria sido verdadeira caso o mundo estivesse em algum outro estado possível no qual poderia ter estado.

Se a tese 1 estiver correta, então tem de haver alguma descrição D que os falantes associam ao nome “Aristóteles” tal que a proposição expressa por **Se Aristóteles existiu, então Aristóteles foi D** é uma verdade necessária. De fato, uma vez que D dá o significado de “Aristóteles”, a proposição expressa por essa frase deveria ser necessária e conhecível *a priori*. Mas, argumenta Kripke, não há tal descrição D. Por exemplo, considere a descrição *o fundador da lógica formal* como uma possível candidata. Para aplicar o teste modal, perguntamos se a afirmação *Se Aristóteles existiu, então Aristóteles foi o fundador da lógica formal* é uma verdade necessária. Dizer que é uma verdade necessária é dizer que não há maneira possível na qual o mundo poderia ter sido que tornaria a antecedente, *Aristóteles existiu*, verdadeira, e a conseqüente, *Aristóteles foi o fundador da lógica formal*, falsa. Mas isso não parece correto. Pelo contrário, parece que o mundo poderia ter estado numa circunstância na qual Aristóteles existiu, mas não fundou qualquer lógica. Uma vez que Aristóteles poderia ter existido sem ser o fundador da lógica formal, então a afirmação condicional *Se Aristóteles existiu, então Aristóteles foi o fundador da lógica formal* não é uma verdade necessária. Assim, o nome *Aristóteles* não significa a mesma coisa que a descrição *o fundador da lógica formal*.

Esse resultado não é isolado. O mesmo argumento poderia ser fornecido a outras descrições candidatas (1-8), ou a outras descrições com as quais naturalmente poder-se-ia pensar como substitutas do nome *Aristóteles*. A razão para isso é que aproximadamente todas as descrições que as pessoas hoje associam ao nome têm a ver com os famosos feitos de Aristóteles, especialmente aqueles envolvendo a sua filosofia. No entanto, como Kripke corretamente mostrou, nenhuns desses feitos eram condições necessárias para Aristóteles existir.<sup>17</sup> Aristóteles poderia ter existido ainda que não tivesse ido

<sup>17</sup> Mesmo ser nomeado “Aristóteles” não era uma condição necessária para Aristóteles existir.

para a filosofia, ou feito algo importante. Desse modo, Kripke está pronto para avançar o argumento que acabamos de oferecer usando qualquer descrição ou agregado de descrições baseadas nos famosos feitos ou nas características mais conhecidas de Aristóteles. Mas uma vez que elas fornecem o principal conteúdo descritivo que a maioria de nós associa ao nome, ele conclui que a teoria descritivista do significado é incorreta enquanto teoria de como a maioria de nós usa o nome.

A força desse argumento contra a teoria descritivista do significado dos nomes próprios pode ser apreciada perguntando-se que condições uma descrição teria de satisfazer a fim de bloquear o argumento. No mínimo, a descrição *D* teria de ser tal que a proposição expressa por ***Se *n* existiu, então *n* foi *D**** seria necessária e *a priori*. O problema é que para aproximadamente todos os nomes que se possa pensar, tais descrições não podem ser obtidas. Se isso estiver correto, então a tese 1 é falsa.

O argumento que acabamos de fornecer contra a tese 1 **não** vai contra a tese 2, que expressa a perspectiva de que as descrições fixam semanticamente os referentes dos nomes. A razão disso é que a tese 2 **não** alega que nomes e descrições são sinônimos. Foi justamente a afirmação de que os nomes têm os mesmos significados que certas descrições que nos permitiu ver a tese 1 como comprometida com a ideia de que quando substituimos nomes e descrições uns pelos outros numa frase, não mudamos a proposição expressa, e, por conseguinte, não mudamos o estatuto modal ou epistemológico da frase. Suponha, no entanto, que tivéssemos uma teoria que nos dissesse não que *D* dá o significado de *n*, mas apenas que *D* fixa semanticamente o referente de *n*. Tal teoria sobre a fixação do referente **não** se comprometeria com a ideia de que a substituição de um pelo outro tem de preservar a proposição expressa. Mas caso a substituição pudesse mudar a proposição expressa, não teríamos razão para pensar que também não poderia mudar o estatuto modal ou epistêmico da frase. Assim, embora o argumento modal possa ser visto como refutando a tese de que nomes são sinônimos das descrições que os falantes a eles associam, não refuta por si próprio a perspectiva de que essas descrições determinam semanticamente os referentes dos nomes. Assim, se quisermos criticar a versão fixadora da referência da teoria descritivista, é preciso de argumentos adicionais. Nas palestras 2 e 3 Kripke faz isso. Mas antes de considerarmos esses argumentos, precisamos examinar alguns dos conceitos centrais que ele emprega.

## Designação Rígida

### *A designação rígida e o argumento modal*

O primeiro desses conceitos é o de um designador rígido.<sup>18</sup>

#### DESIGNAÇÃO RÍGIDA

Um termo singular  $t$  é um designador rígido de um objeto  $o$  sse designa  $o$  no que diz respeito a todos os estados possíveis do mundo (em que  $o$  existe); e, além do mais,  $t$  nunca designa algo além de  $o$  (no que diz respeito a qualquer estado possível do mundo).

Se um termo singular é um designador rígido de um objeto  $o$ , então as frases que contêm  $t$  são verdadeiras quando tomadas como descrições de estados possíveis alternativos do mundo,  $w_1, w_2, w_3$ , sse um e o mesmo objeto  $o$  tem as propriedades relevantes naqueles estados alternativos. Por exemplo, se  $t$  é um designador rígido de um objeto  $o$ ,  $F$  expressa a propriedade  $\varphi$ , e  $p$  é a proposição expressa por  **$Ft$** , então (i)  $p$  é verdadeira relativo ao estado efetivo do mundo sse como as coisas são efetivamente no mundo,  $o$  tem a propriedade  $\varphi$ , e (ii)  $p$  é verdadeira no que diz respeito a qualquer estado possível  $w$  do mundo sse relativo a  $w$ ,  $o$  tem a propriedade  $\varphi$  (i.e., se o mundo estivesse no estado  $w$ , então  $o$  teria a propriedade  $\varphi$ ). Se  $t$  é um designador não-rígido de  $o$ , então embora (i) permaneça como acima, (ii) não. Se  $t$  é não-rígido, então há frases  **$Ft$** , propriedades  $\varphi$ , proposições  $p$  e estados do mundo  $w$  e  $w^*$  tal que  $p$  é expressa por  **$Ft$** , e ou  $p$  é verdadeira no que diz respeito a  $w$  muito embora  $o$  não tenha  $\varphi$  no que diz respeito a  $w$ , ou  $p$  é falsa no que diz respeito a  $w^*$  muito embora tenha  $\varphi$  no que diz respeito a  $w^*$  (ou ambos).

Por exemplo, considere o par (9) e (10).

9. O vencedor das eleições presidenciais dos Estados Unidos de 1996 era um democrata.
10. Bill Clinton era um democrata.

A descrição *o vencedor das eleições presidenciais dos Estados Unidos de 1996* e o nome *Bill Clinton* designam o mesmo indivíduo  $o$ . Uma vez que  $o$  era um democrata, tanto (9) quanto (10) são verdadeiras

---

<sup>18</sup> Essa é uma versão simplificada que evita várias complicações. Por exemplo, deixamos de fora indexicais e variáveis de modo que não precisamos relativizar a referência a contextos e atribuições de valores a variáveis. Para uma definição que leve em conta essas complicações, veja o meu *Reference and Description: The Case Against Two-Dimensionalist*, Princeton University Press, 2005. Para os nossos propósitos consideremos como termos singulares nomes e descrições definidas singulares como *o tal e tal*.

no que diz respeito ao mundo como efetivamente é (foi). Isso não é assim no que diz respeito a um estado possível  $w$  do mundo em que Clinton concorreu como um democrata mas o republicano Bob Dole vence as eleições de 1996. O indivíduo cuja filiação partidária determina o valor de verdade de (10) relativo a estado do mundo  $w$  é Bill Clinton, o mesmo indivíduo que é relevante para determinar o valor de verdade de (10) no estado efetivo do mundo. Assim, (10) é verdadeira no que diz respeito a  $w$ . No entanto, o indivíduo cuja filiação partidária determina o valor de verdade de (9) relativo a  $w$  não é Bill Clinton, mas antes, Bob Dole. Assim, (9) é falsa no que diz respeito a  $w$ . Esse exemplo mostra que a descrição *o vencedor das eleições presidenciais dos Estados Unidos de 1996* é não-rígida, ao passo que o nome *Bill Clinton* é rígido.

Por que para um designador rígido o mesmo objeto  $o$  é relevante no que diz respeito a todos os estados possíveis do mundo, ao passo que para um designador não-rígido os objetos relevantes diferem de estado do mundo para estado do mundo? A resposta pode ser dada como se segue: se  $t$  é rígido, então qualquer objeto que seja designado por  $t$  é designado por  $t$  no que diz respeito a todos os estados possíveis do mundo em que esse objeto exista, e nada além desse objeto é designado por  $t$  no que diz respeito a qualquer estado do mundo; mas se  $t$  é não-rígido, então ou  $o$  que é efetivamente designado por  $t$  não é designado por  $t$  no que diz respeito a outros estados possíveis do mundo (em que esse objeto existe), ou algo além desse objeto é designado por  $t$  no que diz respeito a algum estado do mundo.<sup>19</sup> Isso sugere um teste linguístico para determinar se um termo singular arbitrário em Português é um designador rígido.

### **UM TESTE LINGUÍSTICO**

$t$  é um designador rígido sse a frase ***O indivíduo que é (foi) efetivamente  $t$  não poderia ter existido sem ser  $t$ , e nada além do indivíduo que é (foi) efetivamente  $t$  poderia ter sido  $t$***  expressa uma verdade.

Posto de outro modo: um termo singular  $t$  do Português é um designador rígido sse as frases relevantes da forma (11) e (12) são

---

<sup>19</sup> Essa ideia é uma mera aproximação que ignora certas complicações abstrusas que surgem em casos especiais quando os indexicais são considerados e a noção de designação rígida é relativizada ao contexto. (O mesmo vale para o teste linguístico que se segue).

falsas. Um termo singular é não-rígido sse ou (11) ou (12) são verdadeiras.

11. O indivíduo que é (foi) efetivamente  $t$  poderia ter existido sem ser  $t$ .
12. Poderia ter sido o caso que alguém além do indivíduo que efetivamente é (foi)  $t$  fosse  $t$ .

Kripke sustenta que se aplicarmos esse teste descobriremos que os nomes próprios são designadores rígidos, ao passo que a maioria das descrições comuns não são. Isso não quer dizer que ele pensa que nenhuma descrição seja rígida; por exemplo, ele consideraria *a raiz quadrada positiva de 25* e *o indivíduo que é idêntico a Saul Kripke* como sendo rígidas. No entanto, ele acredita que no que diz respeito à maioria dos nomes comuns, descrições como essas não são boas candidatas para serem usadas pelos falantes para dar os significados ou estabelecer a referência.<sup>20</sup> Assim, no caso da vasta maioria dos nomes comuns, ele sustentaria que as descrições associadas a eles são não-rígidas. Em particular, as descrições (1-8) associadas ao nome *Aristóteles* em nosso exemplo são não-rígidas.

Com isso em mente, podemos apresentar o argumento modal de Kripke. Em essência, o argumento é o seguinte:

- (P1) Os nomes são designadores rígidos.  
(P2) Comumente, as descrições associadas aos nomes pelos falantes não são designadores rígidos.
- 

(C) Portanto, os nomes comumente não são sinônimos das descrições associadas a eles pelos falantes.

Esse argumento foi imediatamente reconhecido como um desafio poderoso ao descritivismo sobre os significados dos nomes próprios, e continua a ser aceito por muitos filósofos hoje em dia.

Recentemente, no entanto, uma resposta ao argumento tem tido certa aceitação entre alguns teóricos que continuam a ser atraídos pelo descritivismo. A resposta é baseada na observação de que para qualquer descrição não-rígida ***o F***, podemos formar uma descrição rígida ***o F efetivo*** que designa, quando estamos a falar de qualquer

---

<sup>20</sup> Deixamos de lado aqui casos especiais como o numeral “2”, que concebivelmente poderia ser definido como *o sucessor de 1*.

estado possível do mundo, o indivíduo que é *F* no estado efetivo do mundo.<sup>21</sup> Considere, por exemplo, a frase (13).

13. Poderia ter sido o caso que o vencedor efetivo das eleições presidenciais de 1996 não venceu as eleições presidenciais de 1996.

Intuitivamente, o que essa frase diz é verdadeiro. Isso significa que tem de haver um estado possível do mundo, *w*, tal que a frase (14), como usada por nós, é verdadeira quando tomada como uma descrição de *w*.

14. O vencedor efetivo das eleições presidenciais de 1996 não venceu a eleição presidencial de 1996.

Isso, por sua vez, significa que o indivíduo que conta como referente do nosso uso do termo *o vencedor efetivo das eleições presidenciais de 1996*, quando usada para dizer algo sobre *w*, é o indivíduo que venceu a eleição, não com respeito a *w*, mas no que diz respeito ao mundo como efetivamente é. A frase (14) é verdadeira no que diz respeito a *w*, pois esse indivíduo – a pessoa que foi a vencedora no que diz respeito a estado efetivo do mundo – não venceu no que diz respeito a *w*. A lição aqui é que o resultado de se adicionar o operador de efetividade a uma descrição ***o F*** resulta numa nova descrição, ***o F efetivo***, que designa rigidamente o objeto que a primeira descrição designa no estado efetivo do mundo (se tal objeto é unicamente designado por ***o F*** relativo ao estado efetivo do mundo). Essa ideia tem sido usada por alguns descritivistas pós-kripkianos para sugerir que os nomes são sinônimos, não das descrições comuns, mas das descrições rigidificadas pelo uso do operador de efetividade. Essa tese, com efeito, ataca a premissa P2 acima, e **não** é refutada pelo argumento modal que Kripke oferece.

Não obstante, a proposta falha por outras razões. Primeiro, se a proposta fosse correta, então a proposição expressa por ***Se n existiu, então n foi D*** seria o mesmo que a proposição expressa por ***Se o D efetivo existiu, então o D efetivo foi D***. Essa última proposição é algo conhecível *a priori*, independentemente de qualquer evidência empírica. No entanto, quando *n* é um nome próprio comum, a proposição expressa por ***SE n existiu, então n foi D*** tipicamente não é conhecível *a priori*. Assim, as duas proposições

<sup>21</sup> Essa afirmação é de algum modo um exagero, e ignora as complicações mencionadas nas notas 4 e 5. Os interessados num caso mais completo deveriam ver a discussão do uso do operador de efetividade para rigidificar descrições na obra citada na nota 4.

são diferentes, e a proposta de salvar a tese 1 vai a pique. (Esse ponto está implícito no *Naming and Necessity*. Retornaremos a ele mais tarde quando discutirmos o conteúdo da palestra 2 de Kripke.)

Há também um segundo problema com a proposta que Kripke não tratou, mas que é discutido detalhadamente no capítulo 2 do meu *Beyond Rigidity*.<sup>22</sup> Mencionarei apenas o seu ponto principal. A proposição de que o F efetivo é G é uma proposição que diz sobre o estado efetivo do mundo que o indivíduo que é unicamente F no que diz respeito a ele é também G.<sup>23</sup> Consequentemente, é possível acreditar que o F efetivo é G apenas se estivermos em algum tipo de contato epistêmico com o estado efetivo do mundo, e, em virtude disso, acreditar em certa proposição sobre ele. Suponhamos, para fins da argumentação, (i) que todos nós, vivendo no mundo como efetivamente é, temos tal contato com esse estado do mundo, e (ii) que, por causa disso, quando acreditamos que o F é G, automaticamente qualificamos como **também** acreditando, do estado efetivo do mundo, que o indivíduo que é unicamente F no que diz respeito a ele é também G. No entanto, ainda que garantamos isso, temos também de reconhecer que as coisas mudam quando consideramos, não o que os agentes efetivamente acreditam, mas o que é possível para os agentes acreditarem. Certamente, é possível acreditar que Aristóteles foi um filósofo sem estar em qualquer contato epistêmico com o estado efetivo do mundo e, por conseguinte, sem acreditar em qualquer coisa sobre ele. Por exemplo, se vários fatos irrelevantes sobre Aristóteles tivessem sido diferentes, eu poderia ainda ter acreditado que Aristóteles foi um filósofo. Dizer isso é apenas dizer que há algum estado possível do mundo, *w*, diferente do estado efetivo do mundo, tal que, no que diz respeito a *w*, acredito que Aristóteles foi um filósofo, muito embora no que diz respeito a esse estado do mundo não posso acreditar em qualquer coisa acerca do estado efetivo do mundo. Não posso acreditar em algo sobre o estado efetivo do mundo uma vez que o estado efetivo do mundo é uma propriedade total ou máxima que representa o modo pelo qual as coisas efetivamente são. Tivesse o mundo estado num estado *w*, eu poderia ter sido familiar com o modo como as coisas são no que diz respeito a *w*, e, por isso, ter epistemicamente estado em contato com *w*, mas não preciso estar familiarizado com outros estados do mundo maximamente possíveis, e, por isso, não preciso estar familiarizado com o estado do mundo efetivo. Se isso estiver correto, mostra que para qualquer descrição **o D efetivo**, é

<sup>22</sup> Scott Soames, *Beyond Rigidity* (Nova York: Oxford University Press, 2002).

<sup>23</sup> Nesse parágrafo uso “F” e “G” como letras esquemáticas. “D” é usado como uma variável metalinguística.

possível acreditar na proposição expressa por *Aristóteles foi um filósofo* sem acreditar na proposição expressa por **O D efetivo foi um filósofo**, e, por conseguinte, que as duas proposições são diferentes.<sup>24</sup> Assim, a proposta de que nomes são sinônimos de descrições rigidificadas pelo uso do operador de efetividade é falsa.

### **A confusão a ser evitada**

Antes de prosseguir, paremos para afastar uma confusão muito fácil de se fazer, e muito importante de ser evitada. A confusão surge de um enigma gerado pelas duas afirmações seguintes.

- (i) O nome “Aristóteles” é um designador rígido. Assim, para todos os estados possíveis  $w$  do mundo, refere-se ao mesmo indivíduo – o homem Aristóteles – no que diz respeito a  $w$ .
- (ii) Não é uma verdade necessária que Aristóteles se chamasse “Aristóteles”; poderia ter sido o caso de o nome “Aristóteles” não se referir a Aristóteles. Assim, tem de haver algum estado do mundo,  $w$ , tal que a afirmação de que “Aristóteles” não refere Aristóteles é verdadeira no que diz respeito a  $w$ .

Ambas as afirmações são verdadeiras, e ambas seriam endossadas por Kripke. No entanto, isso poderia parecer embaraçoso, uma vez que (i) e (ii) podem parecer inconsistentes. O que faz (i) e (ii) parecerem inconsistentes é a tendência de tacitamente se aceitar (iii) como algo demasiadamente óbvio.

- (iii) A relação ternária *\_\_\_ refere \_\_\_ no que diz respeito a \_\_\_* (tacitamente invocada em (i)) estabelece-se entre o nome “Aristóteles”, o homem Aristóteles e um estado  $w$  do mundo sse é verdadeiro no que diz respeito a  $w$  que a relação binária *\_\_\_ refere \_\_\_* (invocada em (ii)) estabelece-se entre “Aristóteles” e Aristóteles – *i.e.*, sse a afirmação de que “Aristóteles” refere Aristóteles é verdadeira quando tomada como uma descrição de  $w$ .

---

<sup>24</sup> Esse passo no argumento supõe tacitamente, assim como o argumento anterior contra a proposta, que  $x$  **acredita que S** reporta uma relação entre o crente e a proposição expressa por S. Alguns proponentes da ideia de se analisar os nomes como descrições rigidificadas usando o operador de efetividade disputa essa suposição. Para respostas à tentativa mais sistemática de desenvolver tal ponto de vista, veja o meu “Saul Kripke, the Necessary A Posteriori, and the Two-Dimensionalism Heresy”, em M. Garcia-Carpintero e J. Maciá, eds., *The Two-Dimensionalism Framework* (Oxford: Oxford University Press, 2006), e o meu *Reference and Description*.

Embora o princípio (iii) pudesse parecer inegável à primeira vista, é na verdade falso. No decorrer do capítulo seguimos Kripke ao tomar a relação ternária *\_\_\_ refere \_\_\_ no que diz respeito a \_\_\_* estabelecida entre um nome *n*, um objeto *o* e o estado do mundo *w* sse ***n*, quando usado por nós aqui e agora no mundo como efetivamente é, refere o objeto o, quando as nossas palavras são tomadas como descrições de w**. Por causa disso, *n* pode referir *o* no que diz respeito a *w* ainda que (a) o nome *n* não exista no que diz respeito a *w*, ou (b) no que diz respeito a *w* o nome *n* não é usado pelos falantes para referir algo, ou (c) no que diz respeito a *w* o nome *n* é usado pelos falantes para referir algo que não *o*. Aquilo a que os falantes teriam referido ao usar nome *n* tivesse o mundo estado em *w* é irrelevante quanto a se *n* refere *o* no que diz respeito a *w*. No entanto, ao que os falantes teriam referido ao usar *n* tivesse o mundo estado na circunstância *w* é crucial para determinar o par de nomes e objetos aos quais a relação binária *\_\_\_ refere \_\_\_* se aplica no que diz respeito a *w*. **É verdade no que diz respeito a *w* que o nome *n* se refere ao objeto *o* sse tivesse o mundo estado no circunstância *w*, os falantes teriam usado *n* para referir *o***. Assim, o que (ii) diz é que há estados do mundo tais que, tivesse o mundo estado nessas circunstâncias, os falantes não teriam usado “Aristóteles” para referir Aristóteles. Isso é compatível com a afirmação feita por (i) – nomeadamente que, aqui e agora no mundo como efetivamente é, usamos o nome “Aristóteles” para referir ao homem Aristóteles quando as nossas palavras são tomadas como descrições de qualquer estado do mundo que seja.

### ***Designação Rígida e Essencialismo***

Por toda a discussão de Kripke sobre nomes na palestra 1 de *Naming and Necessity* ele toma por garantido que a distinção entre as propriedades essenciais de um objeto e suas propriedades contingentes é uma distinção legítima. Uma propriedade essencial de um objeto é uma propriedade que o objeto não poderia deixar de ter em qualquer circunstância na qual ele exista. Uma propriedade contingente, ou acidental, é aquela que o objeto tem, mas poderia ter existido sem ela. Exemplos de propriedades contingentes minhas são a propriedade de viver em Princeton, a propriedade de ser um pai e a propriedade de ser um filósofo. Exemplos incontestados de propriedades essenciais minha são mais raros, mas as seguintes parecem ser boas candidatas: a propriedade de ser humano, a propriedade de ter um cérebro, a propriedades de ter um corpo constituído de moléculas, a propriedade de ser mortal, a propriedade de não ser idêntico a Saul Kripke.

Há uma conexão íntima entre a noção de um designador rígido e a alegação de que um objeto possui uma propriedade essencialmente. Essa conexão é expressa por (i).<sup>25</sup>

- (i) Se  $n$  é um designador rígido de  $o$  e  $F$  é um predicado que expressa a propriedade  $P$ , então a afirmação de que  $P$  é uma propriedade essencial de  $o$  é equivalente à afirmação de que **é necessário que se  $n$  existe, então  $n$  é  $F$ .**

A equivalência mencionada em (i) é relevante a um enigma usado por Quine há mais de duas décadas antes de *Naming and Necessity* a fim de lançar dúvida sobre a inteligibilidade do essencialismo.<sup>26</sup> Quine afirmou primeiro que a noção de uma propriedade essencial de um objeto é definida em termos da noção de necessidade, e segundo, que seja qual for a compreensão de necessidade que tenhamos, ela é expressa pelo nosso uso do predicado *é uma verdade necessária*, que se aplica a frases, ou pelo uso do operador *é necessário que...*, que se liga a frases. Assim, pensou ele, se para darmos sentido à ideia de que **é essencialmente  $F$**  se aplica a certo objeto  $o$ , não temos escolha a não ser ver essa afirmação como baseada na suposição de que a frase **é necessário que se  $t$  existe, então  $t$  é  $F$**  é verdadeira para alguma escolha designada de um termo  $t$  que se refere a, ou descreve,  $o$ .

No entanto, Quine também observou que para qualquer objeto  $o$ , haverá alguns termos  $t$  que referem  $o$  que tornarão a frase **É**

---

<sup>25</sup> Alguma explicação da minha terminologia. Os predicados **expressam** propriedades ao serem **verdadeiros dos** objetos no que diz respeito aos estados do mundo. Note que a propriedade que o predicado expressa não é relativizada a diferentes estados do mundo. Se  $F$  expressa  $P$ , então para qualquer estado do mundo,  $w$ ,  $F$  é verdadeiro de um objeto  $o$  no que diz respeito  $w$  sse  $o$  tem  $P$  no que diz respeito a  $w$ .

<sup>26</sup> Veja Quine “Notes on Existence and Necessity”, *Journal of Philosophy* 40 (1943): 113-27; “The Problem of Interpreting Modal Logic”, *Journal of Symbolic Logic* 12 (1947): 43-48; “Reference and Modality” em *From a Logical Point of View* (Cambridge: Cambridge University Press, 1953, 1961, 1980/ Trad. Brasileira: *De um Ponto de Vista Lógico*, São Paulo: Editora Unesp, 2011); e “Three Grades of Modal Involvement”, originalmente publicado em 1953, reimpresso em Quine, *The Way of Paradox* (Nova York: Random House, 1966). Para explanação e crítica a Quine, cada qual excelente a seu modo, veja David Klapan “Opacity” em E. Hahn e P. A. Schilpp, eds., *The Philosophy of W. V. Quine* (La Salle, IL: Open Court, 1986), e John Burgess “Quinus ab omni naevo vindicatus”, em Ali A. Kazmi, ed., *Meaning and Reference* (Calgary: University of Calgary Press, 1998). Burgess assinala que se alguém (erradamente) considera a necessidade com sendo simplesmente analiticidade, e considera a analiticidade como sendo uma propriedade de frases – o que fizeram tanto Quine quanto aqueles que ele criticava na altura – então o seu ceticismo sobre as “propriedades analiticamente essenciais de um objeto” faz muito mais sentido (embora já não tenha algo a ver com a necessidade metafísica genuína ou com as propriedades metafisicamente essenciais). Kaplan argumenta (i) que a posição de Quine sobre o assunto estava ligada a uma oposição amplamente errada à “quantificação” em construção não-extensionais, (ii) que alguém que aceite a verdade lógica pode aceitar versões benignas de essencialismo, e (iii) que embora a inteligibilidade de teses essencialistas mais robustas não devessem estar em dúvida, a verdade ou falsidade de tais teses não levanta questões metafísicas substanciais.

**necessário que se  $t$  existe, então  $t$  é  $F$**  falsa, ainda que haja outros termos  $t$  que refiram  $o$  que a torne verdadeira. Assim, pensou ele, relativo a um modo de descrever  $o$  pode ser que a propriedade expressa por  $F$  seja uma propriedade essencial de  $o$ , ao passo que relativo a um modo diferente de descrever  $o$ , a propriedade expressa por  $F$  não será uma propriedade essencial de  $o$ . Mas e se considerarmos  $o$  por si próprio, independentemente de qualquer descrição? É a propriedade expressa por  $F$  uma das propriedades essenciais de  $o$  ou não? Pareceria que nada há a dizer.

Eis um exemplo representativo da apresentação de Quine do suposto problema embaraçoso.

Talvez eu possa chamar a atenção ao sentido apropriado da confusão como se segue. Concebivelmente pode-se dizer dos matemáticos que são necessariamente racionais e não necessariamente bípedes; e dos ciclistas que são necessariamente bípedes e não necessariamente racionais. Mas e de um indivíduo que dentre suas excentricidades está ser matemático e ciclista? É esse indivíduo concreto necessariamente racional e contingentemente bípede, ou vice-versa? Apenas na medida em que estamos a falar referencialmente do objeto, sem qualquer preconceito especial contra o pano de fundo de agrupar matemáticos em oposição a ciclistas ou vice-versa, não há semelhança de sentido agrupar alguns de seus atributos como necessários e outros como contingentes. Alguns de seus atributos contam como importantes e outros sem importância, sim; alguns como duradouros e outros como transitórios; mas nenhum como necessário ou contingente.<sup>27</sup>

Seja  $i$  algum indivíduo que é tanto um matemático brilhante quanto um ciclista campeão, e suponhamos que *o maior matemático do mundo* e *o maior ciclista do mundo* ambas designam  $i$ . Assim, uma vez que (15a) é, defensavelmente, verdadeira, ao passo que (15b) não é,

- 15a. É necessário que: se o maior matemático do mundo existe (*i.e.*, se há tal indivíduo como o maior matemático do mundo), então o maior ciclista do mundo é racional.

---

<sup>27</sup> Quine, *Word and Object*, p. 199.

- b. É necessário que se o maior matemático do mundo existe (*i.e.*, se há tal indivíduo como o maior matemático do mundo), então o maior matemático do mundo é bípede.

segue-se, da perspectiva de Quine, que relativo a escolha de descrever *i* como *o maior matemático do mundo*, ser racional é uma das propriedades essenciais de *i*, mas ser bípede não. No entanto, se escolhermos descrever *i* como *o maior ciclista do mundo*, obtermos o resultado oposto. Uma vez que (16a) é defensavelmente verdadeira, ao passo que (16b) não,

- 16a. É necessário que: se o maior ciclista do mundo existe (*i.e.*, se há tal indivíduo como o maior ciclista do mundo), então o maior ciclista do mundo é bípede.
- b. É necessário que: se o maior ciclista do mundo existe (*i.e.*, se há tal indivíduo como o maior ciclista do mundo), então o maior ciclista do mundo é racional.

segue-se, da perspectiva de Quine, que relativo à escolha de descrever *i* como *o maior ciclista do mundo*, ser bípede é uma das propriedades essenciais de *i*, mas ser racional não. De acordo com isso, pensa Quine, não faz sentido perguntar acerca de *i* em si mesmo, independentemente de qualquer modo de descrevê-lo, quais de suas propriedades são essenciais e quais não são.

Em geral, Quine supunha que não havia maneira fundamentada, não arbitrária, de selecionar, para um objeto arbitrário *o* e uma propriedade *P*, que tipo de termo *t* deveria ser usado pra subscrever as afirmações a fim de que *o* possuisse ou não possuisse *P* essencialmente. Assim, a sua doutrina era a de que não faz sentido perguntar se um objeto tem uma propriedade essencialmente independentemente de como é descrito. Ao invés, os objetos têm ou não propriedades essencialmente apenas **relativo** aos modos de descrevê-los.

A relevância de tudo isso para Kripke é que se, como ele sustenta, há uma distinção genuína entre designadores rígidos e não-rígidos, então os designadores rígidos fornecem uma conexão fundamentada entre as afirmações sobre as propriedades essenciais dos objetos e as afirmações sobre que frases expressam, e quais não expressam, verdades necessárias. Apenas frases do tipo mencionado em (i) contêm um designador **rígido** do objeto *o* é que são relevantes para o problema de se *o* possui a propriedade *P* essencialmente ou não.

Quando consideramos se um objeto possui uma propriedade essencialmente, usamos um designador rígido para falar de um e o mesmo objeto no que diz respeito a todos os estados do mundo possíveis. Porque o designador é rígido, o problema de o objeto ter a propriedade no que diz respeito a todos os estados do mundo é equivalente ao problema da frase ou fórmula que atribui a propriedade ao referente do designador rígido ser verdadeira em todos aqueles estados do mundo. Os valores de verdade das outras frases que contêm designadores não-rígidos do objeto são simplesmente irrelevantes.<sup>28</sup>

Desse modo, Kripke rebate a objeção de Quine à inteligibilidade do essencialismo. A situação dialética é esta: começamos com uma distinção intuitiva. Embora eu seja corretamente descrito como o filósofo de Princeton que foi educado em Seattle, ser um filósofo, trabalhar em Princeton e ter sido educado em Seattle são propriedades contingentes minhas - eu poderia ter existido ainda que não tivesse sido educado em Seattle, ido para a filosofia, ou trabalhado em Princeton. Em contraste, ser um ser senciente e não ser idêntico a Saul Kripke parecerem ser propriedades essenciais minhas - parece não haver cenários possíveis no qual eu exista mas não seja um ser senciente, ou no qual sou Saul Kripke. Todos entendem essas afirmações. Embora possa haver desacordos e incertezas sobre que propriedades se enquadram em quais categorias, todos nós reconhecemos a inteligibilidade das afirmações desse tipo antes de recebermos qualquer instrução em filosofia. Quine então chega com uma objeção. Ele oferece um argumento que se destina a mostrar que todos nós estávamos sem perceber a falar besteira. No entanto, a sua objeção assenta-se uma premissa falsa - nomeadamente, que não há maneira não arbitrária de selecionar, para um dado objeto  $o$  e a propriedade  $P$  expressa pelo predicado  $F$ , que tipo de termo  $t$  para designar  $o$  deveria ser usado para construir as afirmações, ***É necessário que se  $t$  existe, então  $t$  é  $F$*** , de acordo com as quais a verdade ou falsidade das afirmações essencialistas sobre  $o$  dependem. Kripke refuta essa premissa mostrando que os designadores rígidos, e apenas os designadores rígidos, fornecem a conexão entre as alegações sobre a necessidade das afirmações, por um lado, e as propriedades essenciais dos objetos, por outro. Com a objeção de Quine fora do caminho, a nossa convicção pré-filosófica de que as afirmações essencialistas são inteligíveis continuam incontestadas.

---

<sup>28</sup> Note que uma vez que *o maior matemático do mundo* e *o maior ciclista do mundo* são ambas não rígidas, as frases em (15) e (16) são irrelevantes à questão de se o indivíduo denotado por elas é ou não essencialmente racional, ou essencialmente bípede.

A essa altura é preciso estar de guarda contra certa resposta bastante familiar dos quínianos não convencidos. Com certeza, dirão, que se a designação rígida faz sentido, então o essencialismo também faz sentido. Mas a designação rígida faz sentido? Tome o nome *Aristóteles*, por exemplo. Dizer que é um designador rígido é dizer que a afirmação feita pelo nosso uso de uma frase como *Aristóteles foi um filósofo* é verdadeira quando avaliada no que diz respeito a um estado possível do mundo, *w*, (e.g., um estado em que o homem *h* ao qual efetivamente chamamos “Aristóteles” nunca conheceu ou estudou com Platão) sse *h* foi um filósofo (i.e., teve a propriedade de ser um filósofo) no que diz respeito a *w*. Mas isso faz só sentido se fizer sentido perguntar acerca do indivíduo *h*, independentemente de qualquer descrição, se *h* teve certas propriedades relativo a algum estado meramente possível do mundo. Certamente, esse é um tipo de coisa que Quine questionou. Assim, sustenta o cético quíniano, o apelo de Kripke à designação rígida é circular, e a poderosa objeção de Quine permanece intacta.

Na minha opinião, e na de Kripke, essa manobra comum do cético quíniano é inteiramente inadequada. Começamos com uma distinção pré-filosófica intuitiva cuja inteligibilidade é reconhecida por praticamente todos. Quine oferece uma objeção. Ele pretende mostrar que há algo incoerente no modo pelo qual todos nós olhamos para as coisas. O ônus da prova de demonstrar que há alguma incoerência interna em nosso pensamento, algum conflito entre os diferentes aspectos de nossas ideias que não notamos é dele. A resposta de Kripke à objeção de Quine mostra que ele não faz isso. A essa altura não adianta responder que o legítimo cético quíniano – aquele determinado a não aceitar, a qualquer custo, a inteligibilidade das afirmações essencialistas – não aceitaria as pressuposições da resposta de Kripke. Kripke não precisa provar a inteligibilidade do essencialismo a partir de premissas que fossem aceitas até mesmo pelo cético mais determinado, não mais do que do oponente do ceticismo radical sobre o mundo externo precisa provar a existência de outros objetos além dele próprio e suas ideias para a satisfação de um filósofo determinado a tomar uma posição solipsista. É suficiente rebater algumas objeções razoáveis que tais cétricos pudessem levantar na tentativa de nos persuadir que, pelos padrões que até mesmo nós reconhecemos, nossas perspectivas cotidianas do senso comum estão erradas.

À luz disso, não é surpreendente que a resposta final de Kripke a Quine nessa questão tenha um sabor distintamente mooriano. Imediatamente ao parágrafo que sumariza a objeção de Quine à

inteligibilidade da afirmação de que objetos têm propriedades essenciais e acidentais independente de como são descritos, Kripke diz o seguinte:

Sugeriu-se mesmo na literatura que, embora pudesse haver algum tipo de intuição por trás de uma noção de necessidade (pensamos que algumas coisas poderiam ter sido de outro modo; e que outras não poderiam ter sido de outro modo), essa noção [de uma distinção entre propriedades necessárias e contingentes] é apenas uma doutrina criada por algum filósofo ruim que (acho) não percebeu que há diversas maneiras de se referir à mesma coisa. Não sei se alguns filósofos não perceberam isso; de qualquer forma está longe de ser verdade que essa ideia [de que se pode atribuir de maneira inteligível uma propriedade como essencial ou acidental a um objeto independentemente de sua descrição] é uma noção que não tem conteúdo intuitivo, que nada significa ao homem comum. Suponha que alguém dissesse, apontando Nixon, “Este é o cara que poderia ter perdido”. Outra pessoa diz “Não, se você o descreve como ‘Nixon’, então ele poderia ter perdido; mas certamente se descrevê-lo como o vencedor, então não é verdade que ele poderia ter perdido”. Ora, quem é que está sendo o filósofo aqui, o homem intuitivo? Parece-me obviamente que é o segundo. O primeiro homem diria, e com grande convicção, “Bem, com certeza, o vencedor da eleição *poderia ter sido um outro alguém*. O vencedor efetivo, tivesse o curso da campanha sido diferente, poderia ter sido o perdedor, e outra pessoa o vencedor; ou poderia não ter tido qualquer eleição. Assim, tais termos como ‘o vencedor’ e ‘o perdedor’ não designam o mesmo objeto em todos os mundos possíveis. Por outro lado, o termo ‘Nixon’ é apenas um *nome deste homem*”. Ao perguntar se é necessário ou contingente *Nixon* ter vencido a eleição, você está a fazer a pergunta intuitiva se em alguma situação contrafactual *este homem* teria de fato perdido a eleição. Se alguém pensa que a noção de uma propriedade necessária ou contingente (esqueça se há quaisquer propriedades necessárias não triviais [e considere] apenas inteligibilidade da noção) é uma noção de filósofo sem conteúdo intuitivo, está errado. Certamente que alguns filósofos pensam que algo ter conteúdo intuitivo é um indício bastante inconclusivo em seu favor. Penso que seja um indício bastante forte em favor de algo. Realmente não sei que indício mais conclusivo se poderia ter a favor de algo, no fim das contas.

Mas, em todo caso, as pessoas que pensam que a noção de propriedade accidental não é intuitiva têm a intuição contrária, penso.<sup>29</sup>

No fundo, a posição de Kripke frente ao cético quínio acerca da inteligibilidade do essencialismo é muito parecida com a posição de Moore frente aos céticos que ele enfrenta.<sup>30</sup> Para Kripke, (i) há uma pressuposição inicial forte de que tanto o nosso discurso contrafactual quanto a distinção entre propriedades essenciais e accidentais que o acompanha são inteligíveis, e (ii) a fim de sustentar a distinção, é suficiente rebater os argumentos céticos destinados a demonstrar que tal distinção não poderia ser coerente.

Se, como acredito, esse for o modo correto de se ver a situação, por que o ceticismo de Quine sobre essa questão foi tão influente por tanto tempo? Em minha opinião, três fatores desempenharam papéis importantes. Primeiro, por muitos anos o aparato técnico da semântica dos mundos possíveis de Kripke, incluindo o seu conceito de designação rígida, ou era não existente, ou não amplamente compreendido, ou imperfeitamente apreendido (e às vezes confundido com perspectivas estranhas e implausíveis). Sem um conceito útil e facilmente aplicável de designação rígida não era inteiramente claro qual deveria ser a resposta à objeção cética de Quine.<sup>31</sup> Segundo, como em muitas discussões sobre o ceticismo, as discussões de quem tem o ônus da prova se tornaram completamente confusas desde o início, com os defensores de Quine recusando-se (perversamente) a aceitar a pressuposição inicial da inteligibilidade do discurso contrafactual comum ao qual recorreram Kripke e outros proponentes da distinção entre propriedades essenciais e accidentais. Terceiro, a já familiar confusão da necessidade com a analiticidade desempenhou um grande papel obscurecendo as características centrais em questão.

---

<sup>29</sup> *Naming and Necessity*, pp. 41-42. Os colchetes são de Kripke.

<sup>30</sup> Veja os capítulos 1 e 2 do volume 1 para uma discussão da resposta de Moore ao ceticismo.

<sup>31</sup> Para falar a verdade, houve antecipações históricas das doutrinas de Kripke sobre os designadores rígidos e seu potencial uso na resposta às objeções *a la* Quine, incluindo discussões importante de Bertrand Russell, Raymond Smullyan, Frederick Fitch, Ruth Barcan Marcus, Paul Ziff, Arthur Prior, Dagfinn Follesdal, Keith Donnellan, Peter Geach, e outros. Para um pano de fundo histórico veja o meu “Revisionism about Reference” e “More Revisionism about Reference”, junto de “Marcus, Kripke, and Names” e “How not to Write History of Philosophy” de John Burgess, todos dos quais podem ser encontrados em Paul W. Humphreys e James H. Fetzer, eds., *The New Theory of Reference* (Dordrecht, Boston e Londres: Kluwer, 1998). Embora escritos em resposta a uma controversa torpe envolvendo crédito histórico apropriado, esses artigos contêm informação significativa sobre os precursores de Kripke e sobre as primeiras tentativas de se responder a Quine. (Um aspecto diferente da controvérsia é coberto por Burgess em seu manuscrito não publicado “Geach, Donnellan, Kripke, and Names”). Para um levantamento útil e extenso de tópicos relevante, veja também Stephen Neale “On a Milestone of Empiricism”, em Alex Orenstein e Petr Kotatko, eds., *Knowledge, Language, and Logic* (Dordrecht e Londres: Kluwer, 2000).

As discussões de Quine deixam claro que tanto ele identificava a analiticidade com a necessidade, quanto tomava a analiticidade como sendo uma propriedade de frases. Dado essa perspectiva de que as afirmações sobre que propriedades um objeto possui essencialmente ou acidentalmente são afirmações sobre que frases sobre o objeto são necessárias, ele naturalmente concluiu que as afirmações sobre as propriedades essenciais ou acidentais dos objetos têm, no fim das contas, de ser afirmações sobre que frases contendo termos que designam esses objetos são verdadeiras em virtude do significado. Uma vez que não há uma conexão direta e natural ligando explicitamente afirmações linguísticas sobre os significados das frases com as afirmações equivalentes sobre as propriedades essenciais ou acidentais dos objetos, ele naturalmente conclui que tais afirmações sobre objetos têm de ter sido confundidas.<sup>32</sup> Antes da articulação clara e da defesa de uma concepção metafísica de necessidade de Kripke – não ligada ou dependente de conceitos linguísticos como o de analiticidade – não havia modo com que os filósofos interessados nessas questões pudessem tê-las visto com clareza. Uma vez que a distinção foi feita, a clareza e a utilidade da noção de designação rígida e da inteligibilidade das afirmações essencialistas se tornaram inseparáveis e completamente irresistíveis.

### ***Designação Rígida, Mundos Possíveis e o Critério de “Identificação Transmundo”***

Conectaremos a noção de rigidez com a natureza dos mundos possíveis e com outro problema, ou pseudoproblema, que Kripke levanta na palestra 1 – a necessidade de um critério de “identificação transmundo”. Diz-se às vezes que antes de podermos avaliar a verdade ou falsidade da afirmação de que poderia ter sido o caso que Nixon, digamos, fosse tal e tal, temos de resolver o problema de quem conta como Nixon em mundos possíveis diferentes. Há diversas ideias diferentes conectadas a essa que Kripke rejeita. Primeiro, se – por alguma razão (talvez simplesmente por causa da terminologia *mundo*) – você pensa que os mundos são grandes objetos concretos – universos alternativos que existem realmente, mas numa parte do espaço e do tempo inacessíveis a nós – então pode lhe parecer óbvio que cada um de nós habita apenas um mundo (assim como cada um de nós tem apenas uma localização num determinado tempo na Terra). De acordo com essa imagem, nenhuns dos indivíduos que

---

<sup>32</sup> O relato dessa linha de raciocínio é muito bem contado por John Burgess em “Quinus ab omni naevus vindicatus”. Como Burgess aponta, a situação dialética foi de pouca ajuda que os principais oponentes pré-kripkianos de Quine – que defendiam a modalidade *de re*, o essencialismo, e a quantificação em construções modais – explicavam comumente a sua concepção de necessidade identificando-a com a analiticidade ou com a verdade lógica.

existem em outros mundos podem ser idênticos a qualquer um de nós, incluindo Nixon. De acordo com essa imagem, o máximo que se pode fazer é estabelecer critérios para aqueles que naqueles mundos sejam similares o bastante ao nosso Nixon de modo que possam ser apropriadamente descritos como desempenhado o “papel de Nixon”. (Como se sempre pudéssemos aprender sobre tais mundos afinal.) Essa não é a concepção de mundos possíveis de Kripke.

Para Kripke, um mundo possível é um estado possível do mundo – um modo como tudo poderia ter sido. É, com efeito, uma propriedade máxima que o universo poderia ter tido.<sup>33</sup> Dizer que há mundos possíveis em que Nixon perdeu as eleições é apenas dizer que há propriedades que o universo poderia ter tido que são tais que, se o universo as tivesse tido, Nixon teria perdido as eleições. Ao especificar essas propriedades – *i.e.*, esses estados do mundo – podemos referir diretamente ao próprio Nixon. Não temos de obter um critério descritivo que tenha de ser satisfeito no caso de alguém desempenhar o “papel de Nixon”.

Podemos por isso de outro modo. Ao especificar os estados possíveis do mundo não nos restringimos ao uso de termos descritivos gerais. Não estamos restritos a dizer que as coisas como *os estados do mundo são aqueles que são de tal modo que tivesse o universo estado em qualquer um deles, então alguém que se formou numa pequena faculdade no estado mais populoso no país e que foi primeiro vice-presidente, que mais tarde se tornou presidente embora forçado a renunciar, teria sido tal e tal*. Se esses fossem os únicos tipos de descrições que pudéssemos dar dos estados possíveis do mundo, então precisaríamos de critérios de identidade para descobrir qual indivíduo, se algum, poderia ser tomado como Nixon no que diz respeito a um dado estado do mundo. No entanto, não estamos restritos desse modo. Não há razão pela qual as descrições puras, não contaminadas por nomes ou outros designadores rígidos, deveriam ter um estatuto privilegiado na especificação dos estados do mundo. Podemos, se quisermos, especificar uma classe de estados do mundo como aqueles em que Nixon tem certa propriedade P. Uma vez que “Nixon” é um designador rígido, ao fazer isso estamos a especificar os estados do mundo como aqueles em que certo indivíduo possui P. Dados esses estados do mundo, não temos, então, de especificar quem Nixon é.

---

<sup>33</sup> Para discussões úteis dessa concepção de mundos possíveis, veja Robert Stalnaker, “Possible Worlds”, *Nous* 10 (1976): 65-75, reimpresso em Michael Loux, ed., *The Actual and Possible* (Ithaca, NY: Cornell University Press, 1979); e Nathan Salmon, “On the Logic of What Have Might Have Been”, *The Philosophical Review* 98 (1989): 3-34.

Kripke trata disso na seguinte passagem:

A tendência de exigir descrições puramente qualitativas de situações contrafactuais tem muitas fontes. Uma, talvez, seja a confusão do epistemológico e do metafísico, entre aprioricidade e necessidade. Se alguém identifica a necessidade com a aprioricidade, e pensa que os objetos são nomeados por meio de propriedades unicamente identificadoras, pode pensar que as propriedades usadas para identificar o objeto é que, sendo conhecidas *a priori*, têm de ser usadas para identificá-lo em todos os mundos possíveis, para descobrir qual objeto é Nixon. Contra isso, repito: (1) Geralmente, as coisas não são “descobertas” na situação contrafactual, são estipuladas; (2) os mundos possíveis não precisam ser dados puramente de maneira qualitativa como se estivéssemos os vendo através de um telescópio.<sup>34</sup>

Embora a principal tese dessa passagem pareça claramente estar correta, é preciso de uma última clarificação. Quando Kripke fala de *estipular um mundo possível*, ele não quer dizer que aquilo que é possível é uma questão de estipulação. Ele quer dizer que as possibilidades que optamos por selecionar, ou falar delas, é uma questão que depende de nós. Depende de nós estipular, ou especificar, quais os estados possíveis em que o mundo genuinamente poderia estar nos interessam, e que desejamos fazer afirmações sobre eles. Além disso, o fato de que podemos usar o nome “Nixon” quando especificamos uma classe de estados possíveis do mundo não significa que nenhuma especificação possam falhar. Poderíamos tentar especificar estados possíveis do mundo nos quais Nixon tem certa propriedade P, mas que de fato não a tem, pois, Nixon não poderia ter tido essa propriedade. Por exemplo, não poderíamos estipular de maneira bem-sucedida uma situação possível na qual Nixon é um objeto inanimado. Nesse tipo de caso não há tal estado possível do mundo correspondente à nossa especificação. Isso é apenas dizer que nossas especificações são falíveis; de fato, a observação de que são falíveis mostra que os fatos sobre a possibilidade não são criados ou determinados pelas nossas estipulações. O mesmo se aplica para as especificações feitas envolvendo descrições. Em geral, as descrições não têm prioridade sobre os nomes ou outros designadores rígidos na especificação dos estados possíveis do mundo.

---

<sup>34</sup> *Naming and Necessity*, pp. 49-50.

## Porque as descrições normalmente não fixam os referentes dos nomes

*Argumentos contra a versão fraca da fixação da referência da teoria descritivista*

Revisemos onde estamos. Até agora em nossa discussão do *Naming and Necessity* fizemos as seguintes coisas:

- (i) Definimos as noções de designadores rígidos e não-rígidos, e argumentamos que nomes próprios são rígidos, ao passo que a maioria das descrições que os falantes associam aos nomes não são.
- (ii) Discutimos o argumento modal de Kripke de que nomes não significam a mesma coisa que as descrições não-rígidas que os falantes associam a eles. Além disso, estendemos esse argumento para mostrar que nomes não significam o mesmo que descrições rigidificadas pelo uso do operador de efetividade.
- (iii) Clarificamos a distinção entre propriedades essenciais e não essenciais, e tentamos explicar o que Kripke quis dizer com o discurso dos mundos possíveis.

Tendo feito isso, examinaremos então uma teoria mais fraca sobre a relação entre nomes e descrições. Na palestra 2, Kripke considera a possibilidade de que descrições possam, como ele diz, *fixar os referentes dos nomes* sem fornecer os seus significados. A ideia é algo como isto: muito embora os nomes próprios não tenham significados no sentido usual, algo tem de ser responsável pelo estabelecimento e preservação da conexão entre um nome e aquilo que ele refere. Algo tem de determinar a que se refere a elocução “Aristóteles” feita por alguém. Talvez as descrições semanticamente associadas aos nomes façam isso afinal. Isto é, pode ser que cada nome próprio seja associado a certas descrições que forneçam o critério para determinar aquilo a que eles se referem – descrições que são parte de seu significado, e dominadas pelos falantes competentes, muito embora não forneçam sinônimos para o nome.<sup>35</sup>

<sup>35</sup> Ao dizer que as condições descritivas que fixam a referência são semanticamente associadas ao nome, que são parte do significado do nome, e que são dominadas pelos falantes competentes, distinguimos a tese interessante, porém controversa, de que o referente de *n* é semanticamente por uma descrição da tese trivial e desinteressante de que é possível descrever o processo pelo qual a palavra adquiriu o seu significado, ou a sua referência. Contudo, isso não significa que todas as palavras têm os seus significados, ou referentes, semanticamente fixados por descrições. Uma coisa é descrever como as

Uma vez que as descrições fixam o que o nome designa no que diz respeito ao estado efetivo do mundo, a conexão entre o nome e o referente se torna rígida. Assim, no que diz respeito a qualquer estado possível do mundo,  $w$ , o nome designa o indivíduo que satisfaz as descrições no que diz respeito ao estado efetivo do mundo – satisfaça ou não essas descrições no que diz respeito a  $w$ . Por exemplo, suponha que temos uma frase  **$F_n$**  na qual o referente de  $n$  é semanticamente fixado pelo conjunto  $D$  de descrições. De acordo com essa teoria, pode-se pensar nas condições de verdade da proposição expressa por essa frase com sendo determinadas com se segue. Primeiro, determina-se que objeto  $o$  satisfaz unicamente  $D$  no que diz respeito ao estado efetivo do mundo. Então, dado esse objeto, podemos ver que a proposição expressa pela frase é verdadeira no que diz respeito a um estado possível do mundo arbitrário  $w$  sse  **$F_x$**  é verdadeiro de  $o$  relativo a  $w$ .

No processo de testar essa teoria, Kripke isola os corolários (i-iv) da teoria.

O referente do nome  $n$  (para um falante num determinado tempo) é semanticamente fixado por uma descrição, um conjunto de descrições, ou um agregado de descrições  $D$ . Se  $D$  fixa o referente de  $n$ , então:

- (i) o falante acredita que  $D$  se aplica a um único indivíduo;
- (ii) se  $D$  se aplica a um único indivíduo  $o$ , então  $o$  é o referente de  $n$ ;
- (iii) se  $D$  não se aplica a um único indivíduo, então  $n$  não tem referente;
- (iv) o falante sabe (ou é capaz de saber) *a priori* que a frase ***Se  $n$  existe, então  $D$  é  $n$***  expressa uma verdade.

Na palestra 2 Kripke examina essas teses uma por uma e oferece o que considera como contraexemplos a elas. Não demorarei nisso, mas tentarei dizer o suficiente para indicar a ideia principal. Ao fazer isso, empregarei uma estratégia que está implícita na discussão de Kripke. Considerarei que as descrições que são candidatas a fixar a referência de um nome  $n$  para um falante são mais ou menos aquelas que o falante ofereceria, seja inicialmente ou sob algumas idealizações razoáveis, se perguntado ***A quem ou ao que você se refere com  $n$ ?*** Assim, ao oferecer esses contraexemplos, presumirei que se um falante usa descrições para fixar semanticamente o

---

palavras obtiveram os seus significados e referentes que têm; outra é dizer que as descrições são partes dos significados das palavras da maneira contemplada pela teoria descritivista.

referente de um nome, então ele normalmente já estaria ciente, ou facilmente conseguiria ficar ciente, dessas descrições, e, conseguiria fornecê-las se pedido.<sup>36</sup> Mais tarde, quando chegarmos à própria teoria positiva de Kripke sobre como a referência é determinada, consideraremos a questão de se, ao relaxar essa condição, sua própria teoria positiva poderia ser posta na forma de uma descrição associada ao nome por um falante.

Consideraremos agora os corolários. Primeiro o corolário (i). Kripke observa que no caso de muitos nomes a informação descritiva que os falantes associam ao nome é demasiado empobrecida para apanhar um indivíduo unicamente. Um exemplo disso é fornecido pelo nome *Cícero*. O que a maioria de nós sabe sobre Cícero? Muitos de nós não sabem muito mais além de que ele foi um famoso romano, talvez um estadista e orador de algum tipo. Presumivelmente, no entanto, havia mais do que um famoso estadista e orador romano. A maioria de nós reconhece isso. Assim, diria Kripke, nem mesmo acreditamos que a descrição que associamos ao nome apanha um indivíduo unicamente. Portanto, esse é um contraexemplo ao corolário (i). Não obstante, o nosso uso do nome *Cícero* refere a uma única pessoa. Kripke conclui disso que o mecanismo linguístico que determina o referente do nosso uso do nome tem de ser algo além daquilo sustentado pela teoria descritivista.

Há um ponto aqui que merece ser notado, e que mostra que esse tipo de exemplo é mais comum do que se poderia inicialmente pensar. Imagine que temos um falante que sabe mais sobre Cícero além dele ser um famoso estadista e orador romano. Suponha que ele sabe certo fato que Quine gostava de mencionar – nomeadamente que Cícero foi o famoso estadista romano que primeiro denunciou Catilina publicamente. Agora essa descrição de fato apanha o homem Cícero unicamente. Assim, você poderia pensar que, para esse falante particular, a objeção de Kripke ao corolário (i) não funciona. No entanto, é preciso ser cuidadoso ao tirar essa conclusão. Pois a descrição em questão contém um nome próprio – *Catilina*. E poder-se-ia pedir ao falante para dar sua descrição de Catilina. Se o falante for como muito de nós, o melhor que ele poderia fazer seria dizer que Catilina é o líder romano denunciado publicamente pela primeira vez

---

<sup>36</sup> Alguns descritivistas pós-kripkianos não aceitam essa limitação aos candidatos a descrições que fixam semanticamente os referentes. A minha própria opinião é que uma vez essa limitação seja abandonada, torna-se difícil fazer a importante distinção indicada na nota anterior entre as descrições do processo causal pré-semântico pelo qual as palavras adquirem significado e referência, e as descrições que são parte dos significados dos termos que são dominados pelos falantes competentes. Para uma crítica das versões recentes de descritivismo que, em minha opinião, se atrapalham com essa distinção, veja as obras citadas na nota 10.

por Cícero. Pelo que temos um par de nomes, *Cícero* e *Catilina*, cada qual associado a uma descrição que determina um único indivíduo, mas apenas se o nome contido na descrição já tem uma referência independentemente. Se a informação do falante é esgotada por essas descrições, então a teoria descritivista não conseguirá explicar como o referente de cada nome é determinado.

Esse exemplo introduz uma forte exigência imposta pela forma pura da teoria descritivista quando tomada como uma teoria sobre como os referentes de todos os nomes são fixados. O que ela exige é que cada nome seja associado a **propriedades puramente descritivas** que são suficientes para determinar unicamente seu referente. Mas isso é altamente contra-intuitivo. Se para cada nome que você usa você procurar por tais propriedades associadas a esses nomes, estou certo de que terminará com a conclusão de que não as tem. Se isso estiver correto, então os falantes sequer consideram-se como tendo o tipo de informação descritiva associada a cada nome que a forma pura da teoria requer. Tanto pior para o corolário (i).

Vamos ao corolário (ii) e (iii). Os exemplos do tipo *Cícero* que discutimos são casos nos quais referimos um indivíduo, muito embora a informação puramente descritiva à nossa disposição não apanhe unicamente um indivíduo. Outro tipo de caso que Kripke considera não tem como problema a falta de informação, mas antes a existência de informação errada. Considere o nome *Tales*. Tudo o que sei sobre Tales é que ele foi um filósofo pré-socrático que sustentou que tudo é água. Suponha, no entanto, que havia certo indivíduo chamado de *Tales* por seus contemporâneos, ou em qualquer caso chamado por algum nome que quando traduzido e passado a nós veio como *Tales*. Suponha, além disso, que seus contemporâneos o atribuíram um ponto de vista que ele nunca sustentara. Suponha que ele nunca sustentou que tudo fosse água, mas antes que acreditava em algo mais sensato. Não obstante, o relato propagado e alterado sobre ele, e que chegou até nós, é o de que Tales sustentava que tudo é água. Nesse caso imaginário, a descrição que associamos ao nome não designa a pessoa que realmente refere. Poderia mesmo haver um outro filósofo pré-socrático que fosse ermitão e que ninguém o conhecesse. Ainda que por acaso ele de fato sustentasse que tudo fosse água, e, por conseguinte, satisfizesse a descrição que associamos ao nome, isso não o tornaria Tales. O nome que usamos não o referiria, mas antes referiria ao filósofo originalmente confundido. Esses pontos são indicações *prima facie* de que os corolários (ii) e (iii) da teoria descritivista são falsos.

Outro caso desse tipo que Kripke cita é o de Peano. A principal coisa que a maioria das pessoas que ouviu falar de Peano acredita sobre ele é que ele foi o criador dos agora axiomas padrão da aritmética elementar – os chamados *axiomas de Peano*. De fato, ele publicou esses axiomas, e as pessoas se lembram dele por isso. No entanto, numa nota de rodapé ele creditou os axiomas a outro matemático – Dedekind. A nota foi praticamente esquecida, e Peano terminou sendo creditado pela maioria das pessoas pelos axiomas. Aceitemos que os axiomas realmente foram criados por Dedekind. Antes de ler *Naming and Necessity* eu não pensava que eu soubesse disso, embora eu já tivesse ouvido falar dos *axiomas de Peano*. Se eu tivesse alguma descrição associada ao nome *Peano* naquela época, seria, penso, *o descobridor dos axiomas da aritmética*. Contudo, o meu uso do nome *Peano* não refere Dedekind.

Objetar-se-á, talvez, que eu tinha alguma outra descrição associada ao nome que de fato referia Peano. Consideremos algumas candidatas a tal descrição. A primeira candidata é a descrição parasitária:

*a pessoa a qual a maioria das pessoas referem quando usam o nome  
“Peano”*

A ideia por trás dessa sugestão é que a pessoa que não sabe o bastante para descrever Peano unicamente pode fixar o referente de seu uso do nome via essa descrição parasitária, contanto que a maioria das outras pessoas tenha os recursos para referir correta e unicamente a Peano. O problema com essa ideia é que ela corre o risco ou de malogro ou de circularidade. Se a maioria das pessoas não tem outras descrições que são independentemente bem-sucedidas em referir Peano, então o teórico descritivista está preso ao resultado que aqueles que usam a nossa descrição parasitária falham em referir. Mas podemos imaginar casos em que a maioria das pessoas carecem de tais descrições não parasitárias; de fato, isso pode ainda ser verdadeiro no caso do nome *Peano*. Não obstante, as pessoas nessas situações se referem a alguém, e a pessoa a que elas se referem é Peano.

E as outras candidatas à descrição fixadora da referência?

Considere:

*a pessoa a quem a maioria dos especialistas referem quando usam o  
nome “Peano”*

Essa não é boa, uma vez que não especifica de que tipo de especialistas – em ópera italiana? Pintura renascentista? etc. – estamos a falar. Estaríamos melhor com *a pessoa de quem a maioria dos matemáticos referem quando usam o nome “Peano”*? Ora, a matemática é um campo amplo, e poderia facilmente ocorrer de a maioria dos matemáticos terem apenas a descrição de Dedekind associada ao nome. Não obstante, ainda usam o nome para se referir a Peano, e não a Dedekind. E a descrição *a pessoa a qual a maioria dos especialistas em Peano referem quando usam o nome “Peano”*? Mas isso nos deixa num círculo, uma vez que a fim de descobrir quem Peano é, temos primeiro de localizar os especialistas em Peano, ao passo que para descobrir quem é um especialista em Peano – i.e., que tem conhecimento versado do homem Peano – temos primeiro de descobrir quem Peano é.

Uma circularidade similar afeta

*a pessoa a qual os axiomas da aritmética são comumente atribuídos*

O que é atribuir os axiomas da aritmética a alguém? Suponho que seja dizer que pessoa os descobriu. Ora, o que diriam as pessoas se as perguntássemos *Quem descobriu os axiomas da aritmética*? Provavelmente muitas responderiam *Peano*, dizendo com isso que a quem quer que seja que o nome se refira, descobriu os axiomas da aritmética. Assim, a quem o nome refere? De acordo com a teoria, refere-se à pessoa que satisfaz a descrição *a pessoa a qual os axiomas da aritmética são comumente atribuídos*. Desse modo, a fim de determinar a pessoa que satisfaz a descrição, temos primeiro de ter o referente do nome *Peano*, mas, se essa versão da teoria descritivista estiver correta, não podemos fazer isso até que primeiro determinemos o que satisfaz a descrição. Portanto, caímos num círculo novamente. O desfecho disso tudo é que é difícil obter uma descrição fixadora da referência não-circular que não esteja sujeita a contraexemplos claros e óbvios. Com base nisso, Kripke conclui que os corolários (ii) e (iii) da versão fixadora da referência da teoria descritivista são falsos.

Isso nos deixa com o corolário final da teoria descritivista.

- (iv) o falante sabe (ou é capaz de saber) *a priori* que a frase ***Se n existe, então D é n*** expressa uma verdade.

Por que esse corolário é parte da teoria? Ora, se a regra linguística pela qual o referente de um nome é fixado é a de que o nome tem de referir a quem quer que seja designado por certa descrição específica

D, então alguém sabe (ou é capaz de saber), simplesmente em virtude de conhecer essa regra, que a frase **Se *n* existe (existiu), então *n* é (foi) D** não pode não ser verdadeira. Por quê? Porque se D não designa algo, então *n* também não, e a frase é trivialmente verdadeira em virtude da falsidade de seu antecedente. Por outro lado, se D designa algo, então *n* designa a mesma coisa, e a consequente é garantida como verdadeira. De qualquer modo, a frase condicional como um todo é garantida como verdadeira.<sup>37</sup> Tudo isso alguém sabe (ou é capaz de saber) apenas por conhecer as regras da linguagem, sem qualquer investigação empírica. Embora se pudesse certamente levantar a questão de se é a isso que o conhecimento *a priori* genuíno equivale, no momento seguiremos a indicação de Kripke na discussão de seus exemplos e designaremos por “*a priori*” nada mais do que aquilo que pode ser conhecido apenas com base no entendimento da linguagem.<sup>38</sup> Portanto, se a versão fixadora da referência da teoria descritivista estiver correta, o corolário (iv) deveria sempre se aplicar.

No entanto, quando consideramos nomes particulares, vemos que ele tipicamente falha. Considere o nome *Colombo*, por exemplo. Aqui, a descrição mais saliente que muitas pessoas associam ao nome é algo como *o europeu que velejou da Espanha em 1492 procurando por uma nova rota para a Ásia, mas que, ao invés, acabou descobrindo a América*. No entanto, a regra linguística que determina o referente do nome certamente **não** é que, por definição, ele é seja quem for que satisfaça essa descrição. Pois se nos perguntarmos como sabemos que se Colombo realmente existiu, foi o europeu que velejou da Espanha em 1492 procurando por uma nova rota para a Ásia, mas que, ao invés, acabou descobrindo a América, certamente não diremos que sabemos isso *a priori*, simplesmente pelo entendimento da linguagem. Ao invés, sabemos disso porque lemos nos livros didáticos, e porque os nossos professores nos disseram. E como as nossas fontes sabiam disso? Presumivelmente por meios similares envolvendo vários tipos de especialistas, testemunho, documentos remanescentes, artefatos e similares. Mas se assim for, então o nosso conhecimento repousa em evidência empírica, e, por isso, é *a posteriori*. Além do mais, a nossa crença sobre Colombo continua sujeita à revisão caso alguns dos indícios históricos nos quais nos fiamos se mostrem falsos, errados, inexatos, ou simplesmente

---

<sup>37</sup> Se alguém está preocupado que uma frase contendo um nome que falha em referir pode não expressar uma proposição, e por isso falha em ser verdadeira, então pode mudar a afirmação do corolário para ser lido como *O falante sabe (ou é capaz de saber) a priori que se “n existe” expressa uma verdade, então “n é (foi) D”*. Uma vez que a questão não afeta o resultado da presente discussão, deixá-la-ei de lado.

<sup>38</sup> Essa concepção do *a priori* será revisada e criticamente avaliada no capítulo 16.

incompletos. Embora não esperamos que aconteça, é certamente concebível que novos indícios pudessem ser descobertos e mostrassem que Colombo nunca deixou a Espanha, mas enviou alguém em seu lugar. Isso mostra que nem a proposição expressa por (15), nem a afirmação de que a frase (15) expressa uma verdade e é conhecível *a priori* no sentido em que indicamos.

15. Se Colombo existiu, então Colombo foi o europeu que velejou da Espanha em 1492 procurando por uma nova rota para a Ásia, mas, ao invés, acabou descobrindo a América.

Parece, portanto, que todos os principais corolários da versão fixadora da referência da teoria descritivista são falsos – quando tomados como expressando generalizações universais que cobrem todos os nomes próprios. Com base nisso, Kripke conclui que **não há regra semântica geral** especificando que o referente de um nome é o indivíduo designado pelas descrições associadas a ele pelos falantes. No entanto, isso não significa que nunca fixamos o referente de um nome descritivamente. Pelo contrário, Kripke pensa que podemos e, em casos raros, fixar semanticamente o referente de um nome estipulando-se que tem de ser o que quer que seja que satisfaça certa descrição. Ele dá o exemplo do nome *Netuno*, que ele sugere poder ter sido originalmente introduzido como um nome próprio para seja o que for que causou certas perturbações na órbita esperada do planeta Urano. Kripke insiste que ainda que o nome seja introduzido por uma estipulação para referir ao que quer que venha a satisfazer certa descrição, ainda assim o nome **não** é sinônimo da descrição, mas antes um designador rígido. Além do mais, ainda que o nome seja primeiramente introduzido por uma descrição fixadora da referência, mais tarde, quando é passado de falante para falante, a descrição pode ser perdida e o nome pode vir a ser entendido sem referência à descrição. Se Kripke estiver correto, tal pode ter sido a história do nome *Netuno*. Certamente, no entanto, já não é *a priori* para nós que a frase *Se Netuno existiu, então Netuno causou as perturbações na órbita esperada de Urano* expressa uma verdade (se alguma vez expressou).

### ***A concepção histórico-causal de Kripke da determinação da referência***

De acordo com Kripke, então, a referência de um nome é às vezes semanticamente fixada por uma descrição. Mas na maioria dos casos não é assim que a referência é inicialmente determinada; e mesmo quando é, a associação semântica do nome com a descrição fixadora

da referência provavelmente tem vida curta. Isso levanta a questão de como, na vasta maioria dos casos, a referência de um nome é determinada. Kripke apresenta uma imagem positiva de determinação da referência que tenta responder essa questão. A sua ideia é extremamente simples e de senso comum. Um nome particular é introduzido para um objeto ou pessoa. Depois de o nome ter sido introduzido por algum tipo de batismo linguístico, aqueles que introduzem o nome começam a usá-lo na conversação para se referir a seu portador. Novas pessoas ouvem o nome e começam a usá-los pretendendo referir o mesmo indivíduo que suas fontes costumavam referir. Esse processo continua com o nome sendo passado de um usuário ao próximo, com cada usuário formando um elo na cadeia de transmissão ou herança da referência. Geralmente, algum conteúdo descritivo acompanhará a transmissão do nome, mas tipicamente esse conteúdo variará de falante para falante, e como a cadeia de uso se torna maior, pode acabar que para muitos falantes o nome tenha mais informação errônea do que informação exata sobre seu referente associada a ele. Não importa. A informação que uma pessoa acaba associando ao nome **não** é o que determina o seu referente. Ao invés, o referente de um uso de um nome por um falante *x* é determinado pela cadeia histórica que conecta o uso de *x* aos falantes a partir dos quais *x* adquiriu o nome, conectando esses falantes às suas fontes e remontando em última instância ao indivíduo batizado com o nome. De acordo com essa imagem, se uma pessoa que obtém um nome pretende usá-lo para se referir a seja qual indivíduo for que foi referido por suas fontes desse nome, então habitualmente não importa tanto que outras crenças ele possa ter sobre o referente. O que determina a referência não são as crenças dos falantes, mas a cadeia do uso na qual o falante está inserido. Para Kripke, referir tipicamente não é algo que se faça isolado; é uma atividade comunitária.

Essa é uma imagem geral. Dever-se-ia notar, no entanto, que Kripke não é explícito sobre a natureza exata da cadeia de uso, ou sobre que fatos têm de ocorrer a fim de que um uso de um nome conte com um elo na cadeia que determina a referência posterior. Suponha, por exemplo, que a pessoa da qual eu primeiro obtive o nome *Platão* estivesse a falar de seu vizinho, que ele acreditava ser muito sábio. Suponha que após falar dessas pessoas eu tivesse muitas conversas nas quais o nome *Platão* foi usado para descrever o famoso biógrafo de Sócrates. A certa altura, li sobre Platão e, por isso, descobri o nome impresso. Sequer tinha lido traduções da obra de Platão. Tudo isso poderia ter sido verdade, ainda que eu erradamente presumisse

que a pessoa da qual primeiro ouvi o nome estivesse falando do mesmo indivíduo que o restante das pessoas. Nesse tipo de caso, a quem eu agora refiro quando uso o nome *Platão*? Refiro-me, assim como minha fonte primária, ao seu sábio vizinho? Ou refiro-me, como a maioria de minhas outras fontes desse nome, ao filósofo antigo? Certamente que a última resposta é a correta. No entanto, em lugar algum Kripke apresenta uma teoria precisa e explícita que dá um veredito claro aos casos potencialmente problemáticos como esse. Seu objetivo, diz ele, não é fornecer tal teoria, mas ao invés esboçar uma imagem alternativa de como a referência comumente funciona – uma imagem na qual uma cadeia histórica de usos **de algum modo** conecta o falante ao referente, ao invés de uma imagem na qual a conexão entre o falante e o referente é fornecida pelas descrições que o falante associa ao nome.

A incompletude da imagem de Kripke é também ilustrada pelo fenômeno da mudança de referência. Às vezes, um nome pode começar se referindo a uma coisa, e então passar por um período no qual essa coisa é confundida com outra coisa, e ainda mais tarde contará como referindo apenas à segunda coisa – ainda que ninguém jamais tenha tentado conscientemente redefinir o nome ou mudar a sua referência. Tal caso é apresentado por Gareth Evans num artigo chamado “The Causal Theory of Names”.<sup>39</sup> Ele dá o exemplo do nome *Madagascar*. Ele sustenta que uma versão do nome foi originalmente usada para designar alguma parte do continente africano. No entanto, quando os árabes e os europeus chegaram, pegaram o nome dos habitantes pensando erradamente que fosse um nome da grande ilha destacada a sudeste da costa da África. Como Evans relata, eles usaram esse nome com a intenção tanto de referir a ilha quanto de referir a área que os habitantes locais designavam com esse nome pensando que os dois eram um e o mesmo. Após algum período de tempo o referente do nome se tornou, sem ambiguidade, a ilha. Se genuínos, os exemplos desse tipo não desacreditam a concepção histórica de Kripke de como o referente de um nome é fixado. No entanto, mostram que há um problema substancial não trivial sobre o que entra no estabelecimento dos elos da cadeia histórica.

### ***A teoria histórico-causal fornece descrições fixadoras de referência aos descritivistas?***

Com isso em mente, podemos passar à sugestão às vezes feita pelos descritivistas persistentes que o que Kripke fez foi simplesmente

---

<sup>39</sup> *Proceedings of the Aristotelian Society*, supplementary volume 47 (1973), 187-208; reimpresso em seu *Collected Paper* (Oxford: Clarendon Press, 1985).

fornecer um tipo específico da versão fixadora da referência da teoria descritivista dos nomes. A ideia deles, em termos mais simples, é que a referência de um nome *n* para um falante particular é semanticamente determinada por alguma descrição extraível da teoria histórico-causal de Kripke da transmissão e da herança da referência. David Lewis apresenta a ideia em seu artigo de 1997 “Naming the Colours”, com se segue:

Kripke e seus aliados não refutaram a teoria descritivista da referência, pelo menos para nomes de pessoas e lugares? Por que, então, deveríamos esperar que o descritivismo funcione melhor para nomes e experiência de cores? [...] Discordo. O que foi satisfatório e verdadeiramente refutado foi uma versão do descritivismo na qual os sentidos descritivos supostamente eram uma questão de proezas famosas e outras peculiaridades distintas. Uma versão melhor sobrevive ao ataque: o *descriptivismo causal*. O sentido descritivo associado ao nome poderia, por exemplo, ser “o lugar que ouvi falar com o nome de ‘Taromeo’” ou “a fonte causal desse espécime: Taromeo”, e para uma explicação da relação que é invocada aqui, consultamos os escritos dos teóricos causais da referência.<sup>40</sup>

Para manter as coisas simples, podemos ilustrar essa ideia usando a descrição *o indivíduo ao qual a pessoa ou pessoas das quais adquiri o nome referido quando o usaram*. Uma das coisas que faz a ideia de que o uso de uma pessoa de um nome é fixado por tal descrição parecer plausível é a exigência, reconhecida por Kripke, de que para uma cadeia de transmissão da referência ser criada pela transmissão de um nome de um falante a outro, a pessoa que adquire o nome tem de pretender que sua referência seja parasitária da referência de suas fontes. O descritivista pode ser visto como propondo pôr essa exigência na forma de uma descrição que fixa semanticamente o referente do nome.

Embora a ideia possa parecer razoável, há, penso, diversas razões para se preocupar. Primeiro, não é claro que os falantes tenham invariavelmente implícito em mente, dentre todas as descrições diferentes que associam a dado nome, alguma descrição fixadora da referência precisa para ele. Sabemos que a descrição *o indivíduo ao qual a pessoa ou pessoas das quais primeiro adquiri o nome*

---

<sup>40</sup> P. 353, nota 22 de “Naming the Colours”, em seu *Papers in Metaphysics and Epistemology* (Cambridge University Press, 1999), originalmente publicado em *Australasian Journal of Philosophy* 75 (1997). Outros descritivistas contemporâneos expressam ideias similares. Por exemplo, veja David Chalmers, “On Sense and Intension”, *Philosophical Perspectives* 16, *Language and Mind* (Oxford: Blackwell, 2002), 135-82.

*referiram quando o usaram* nem sempre apanha o meu referente para *n*. Além do mais, nem mesmo é completamente claro aos teóricos precisamente que descrição parasitária desse tipo será suficiente para tratar todos os casos problemáticos diferentes. No momento, simplesmente não há teoria histórico-causal precisa e explícita que possa vir a ser uma descrição que seja adequada para todos os casos. Além do mais, ainda que os teóricos viessem a obter tal descrição, está longe de ser óbvio que os falantes comuns sempre a devessem ter implicitamente à disposição ao usar um nome próprio.

Segundo, ainda que aceitemos que os falantes sempre tenham uma descrição parasitária apropriada associada a um nome, ter-se-ia de mostrar que eles de algum modo concordam com essa descrição antes de determinar o referente do nome sobre todas as outras descrições que associam a ele. Isso pode não ser fácil. Se pedíssemos a falantes comuns que nos fornecessem as descrições que especificam mais fiavelmente os referentes dos diferentes nomes que usam, é claro que **não** apresentariam espontaneamente as descrições parasitárias não circulares relevantes.

Talvez, se os guiássemos o bastante nos experimentos mentais do estilo de Kripke sobre a referência, teríamos algum sucesso em extrair as descrições parasitárias que aproximam os resultados corretos a muitos casos. O descritivista poderia então postular que, assim como o escravo no *Menon* de Platão, esses falantes comuns têm sempre de ter tido, e dadas inconscientemente anterior a, as descrições fixadoras de referência relevantes. No entanto, esse relato certamente tem de ser considerado como especulativo ao extremo.

Terceiro, levando-se tudo isso em conta, é importante ter em mente a distinção mencionada nas notas 21 e 22 entre descrições dos processos causais pré-semânticos pelos quais as palavras adquirem e conservam seus significados e/ou referentes, e as descrições que são partes dos significados por meio disso adquiridos, e, por conseguinte, apreendidos pelos falantes competentes que dominam as palavras. Embora claramente haja descrições do primeiro tipo para toda palavra em nossa linguagem – incluindo *se*, *e*, e *mas* – isso não mostra que haja descrições do segundo tipo para todas elas. A doutrina de que os referentes dos nomes são fixados semanticamente pelas descrições associadas a eles pelos falantes é uma tese ambiciosa e problemática que vindica a existência de descrições do segundo tipo para todos os nomes.

Finalmente, não se deveria pensar que a versão fixadora da referência da teoria descritivista – concebida como uma teoria

semântica implicitamente dominada pelos falantes competentes – de algum modo **tem** de estar correta. Poder-se-ia haver certo processo pelo qual a referência fosse passada de falante para falante. Poderia ser que, comumente, quando se adquire um nome, pretende-se usá-lo para referir seja qual for a pessoa apanhada dos usos para referi-la, ou seja, ao que for que é apropriadamente usado para referir na comunidade linguística como um todo. Uma vez que já se tenha adquirido o nome, pode-se começar a usá-lo para expressar crenças sobre o objeto que ele representa. Depois, pode-se esquecer completamente quase todos os aspectos de como se adquiriu o nome, embora preservando as crenças relevantes expressas ao se usar o nome. Se, nessa última vez, perguntar-se o que determina o referente do uso que se faz do nome, então a resposta poderia ser que a sua referência simplesmente é herdada das crenças usada para expressá-la – possa-se ou não obter alguma descrição correta de como originalmente se adquiriu o nome, ou alguma descrição correta dos últimos usos do nome que possam ter mudado a sua referência original. (Pense no caso de Platão). Se isso estiver correto, então pode haver um processo natural de herança da referência pelo qual os últimos usos de um nome herdam a sua referência dos primeiros usos, ainda que os próprios falantes não tenham de ter qualquer compreensão completa e acurada do processo. Nada há de incoerente com essa ideia. Portanto, nada há de estranho com a ideia de que poderia não haver versão fixadora da referência da teoria descritivista correta que seja parte daquilo que se aprende quando se domina uma linguagem.<sup>41</sup>

Finalmente, penso que há uma falta de clareza no *Naming and Necessity* quanto ao tipo de pergunta que fazemos quando perguntamos como a referência de um nome é fixada para um falante num dado momento. Kripke escreve frequentemente como se essa fosse uma pergunta semântica sobre uma regra que os falantes dominam quando aprendem um nome – uma regra do tipo **“n” quando usado num contexto C refere (a) seja o que for que satisfaça tais e tais descrições, ou (b) seja o que for que esteja no fim de tal e tal cadeia de transmissão de referência.**

---

<sup>41</sup> Dever-se-ia também lembrar que ainda que alguma versão fraca da teoria descritivista estivesse correta, de acordo com a qual os referentes dos nomes fossem semanticamente fixados por descrições *a la* Lewis, a teoria não nos ajudaria resolver o problema que os descritivistas tradicionalmente têm enfrentado – a saber, especificar o conteúdo das proposições expressas pelas frases contendo nomes, e cridas e asseridas pelos falantes que aceitam, ou profere assertivamente, tais frases. Quando digo que os babilônios antigos acreditavam que Vênus era uma estrela, não estou dizendo que elas acreditavam que o objeto que está conectado por uma cadeia histórica de uso apropriada à minha elocução da palavra *Vênus* era uma estrela. Assim, ainda que, parecendo extremamente improvável, alguma versão dessa de descritivismo pudesse funcionar como uma abordagem semântica da fixação de referência, não **resolveria** o problema do conteúdo que tem sido de grande interesse aos descritivistas.

Nesse modo de ver as coisas, a pergunta *Como a referência de um nome é fixada?* é usada para expressar uma exigência de uma regra semântica que seja implicitamente apreendida pelos falantes, e que pudesse ser usada por eles na especificação das condições de verdade das frases que contêm o nome. Mas esse não é o único tipo de pergunta que poderia ser feita com essa frase.

Às vezes, por exemplo, poderia ser usada para fazer uma pergunta pragmática sobre qual dos vários conteúdos de um termo é usado numa ocasião particular. Por exemplo, poder-se-ia perguntar o que determina se, numa certa ocasião, “David” é usado para referir David Kaplan ou David Lewis, assim como poder-se-ia perguntar o que determina se, numa ocasião particular, a palavra “banco” é usada para falar de um assento ou de uma instituição financeira. As fontes causais desses usos particulares das palavras podem muito bem ser relevantes para responder essas perguntas. No entanto, as perguntas não são perguntas semânticas sobre quais significados são atribuídos às palavras; ao invés, são perguntas sobre como, quando uma palavra comporta múltiplos significados, determina-se qual significado um falante está usando.

Pode-se ainda usar *Como o referente de um nome é fixado?* para fazer um outro tipo de pergunta. Enquanto teorizadores, podemos querer saber (i) como um nome, ou alguma outra expressão, inicialmente veio a ter o significado e/ou referência que tem e (ii) em virtude de quais aspectos de seu uso pelos falantes esse nome conserva esse significado e referência na linguagem comum da comunidade. Essas são perguntas teóricas que podem surgir mesmo após ter-se estabelecido precisamente qual o significado e a referência de uma dada expressão, seja na linguagem da comunidade como um todo, seja como usada por um falante numa ocasião particular. Essas perguntas são perguntas fundacionais sobre os processos causais que originalmente conferiram às expressões suas propriedades semânticas, e que as preservam na linguagem com as propriedades que têm.

Quando Kripke discute a versão fixadora da referência da **teoria descritivista**, ele está claramente a discutir a proposta acerca da semântica dos nomes. Por causa disso, o seu modo de conceber a discussão geral – *O que fixa a referência?* – pode parecer sugerir que as duas respostas que ele dá a essa pergunta – *Descrições em certos casos incomuns* e *Cadeias históricas de transmissão de referência para a maior parte dos nomes próprios* – estão a par. Uma vez que a afirmação acerca das descrições claramente pode ser entendida

como semântica, e, conseqüentemente, acerca das regras linguísticas que os falantes têm de dominar ao aprender vários nomes, isso encoraja o leitor incauto a pensar erradamente a mesma coisa sobre as cadeias históricas. Acredito que, na discussão de Kripke, seja essa a fonte do último suspiro dos descritivistas na tentativa de interpretar sua teoria histórico-causal da transmissão da referência como fornecendo as descrições necessárias para uma teoria descritivista da semântica dos nomes correta. Isso é o que os leva à perspectiva de que, detalhes à parte, a referência de *n* para um falante particular seja determinado semanticamente por uma descrição como *o indivíduo ao qual as minhas fontes se referiram quando usaram n*.

Em minha opinião, todo esse modo pensar está errado. Nada há de especial aqui sobre a semântica dos nomes. Comumente, quando se usa qualquer palavra na linguagem de uma comunidade, faz-se isso com a intenção de que ela transmita seja qual for o significado e a referência que já adquiriu. Esse é um fato sobre o uso de todas as expressões, não sobre a semântica de algumas. Na medida em que haja perguntas adicionais sobre os nomes, elas serão pragmáticas e fundacionais. Os comentários de Kripke sobre os batismos iniciais deveriam ser vistos como respostas à pergunta fundacional *Em virtude de que esses termos originalmente vêm a referir aquilo que referem na linguagem?* Seus comentários sobre as cadeias causais de transmissão da referência deveriam ser entendidos como fornecendo informação relevante à resposta à pergunta pragmática *Como determinar a qual dos diversos portadores do nome uma elocução particular sua se refere?* e à pergunta pragmática *Em virtude de que o nome continua a referir a esse objeto na linguagem?*<sup>42</sup> E se nos perguntarem, *E o que uma teoria **semântica** de tais nomes deveria nos dizer?*, a resposta mais razoável, creio, é que deveria nos dizer ao que esses termos referem na linguagem comum da comunidade, e nada mais. No entanto, isso nos leva para além daquilo que Kripke explicitamente está comprometido, e à beira de um dos problemas fundamentais não resolvidos levantados pelo *Naming and Necessity*. Esse problema, que é relevante de maneira central à discussão revolucionária da distinção entre verdade necessária e verdade *a priori*, será tratado no próximo capítulo.

---

<sup>42</sup> Esse modo de olhar para a discussão de Kripke da fixação da referência via batismo e cadeias históricas de transmissão de referência é explicado vigorosamente defendido por meu orientando, Jonathan McKeown-Green, em sua dissertação de doutorado em Princeton não publicada *The Primacy of Public Language*.

## Metáfora\*

Max Black

Tradução de Carlos André Fernandes

Revisão de L. H. Marques Segundo e Aluizio Couto<sup>43</sup>

“Metaphor are no arguments, my pretty maiden.”

(*The Fortunes of Nigel*, Book 2, Ch. 2.)

Dar atenção às metáforas de um filósofo é depreciá-lo – é algo como elogiar um lógico por sua bela caligrafia. O vício em metáforas é considerado ilícito de acordo com o princípio segundo o qual daquilo que só se pode falar metaforicamente, não se deve falar o que quer que seja. No entanto, a natureza da ofensa não é clara. Gostaria de fazer algo para afastar o mistério que envolve o tópico; mas uma vez que os filósofos (apesar de seu notório interesse na linguagem) têm negligenciado tanto o assunto, terei de contar com a ajuda dos críticos literários. Eles, pelo menos, não aceitam o mandamento “não cometerás metáfora”, e nem presumem que a metáfora seja incompatível com o pensamento sério.

### I

As perguntas as quais eu gostaria de ver respondidas dizem respeito à “gramática lógica” da “metáfora” e às palavras que tenham significados relacionados. Seria satisfatório ter respostas convincentes às perguntas: “Como reconhecemos um caso de

---

Artigo publicado originalmente em *Proceedings of the Aristotelian Society*, New Series, Vol. 55 (1954 - 1955), pp. 273-294.

<sup>43</sup> Fernandes é graduado em Filosofia pela UFOP. Marques Segundo é Bolsista CAPES, doutorando em Filosofia pela UFSC. Aluizio Couto é Mestrando pela UFMG, com bolsa CNPq.

metáfora?”, “Existem quaisquer critérios para detectar metáforas?”, “Metáforas podem ser traduzidas em expressões literais?”, “A metáfora é considerada apropriadamente como uma decoração do ‘senso comum’?”, “Quais são as relações entre metáfora e símile?”, “Em que sentido, se houver algum, uma metáfora é ‘criativa’?”, “Qual é a razão de se usar uma metáfora?”. (Ou, mais resumidamente, “O que queremos *dizer* por ‘metáfora’?”. As perguntas exprimem tentativas de tornar mais claros alguns dos usos da palavra “metáfora” – ou, se alguém preferir o modo material, a análise da noção de metáfora.)

A lista não é completa, e várias das perguntas se sobrepõem de maneiras bastante óbvias. Mas espero que elas ilustrem suficientemente o tipo de investigação que se pretende.

Seria útil se pudéssemos partir de alguma lista consensual de “casos claros” de metáfora. Uma vez que a palavra “metáfora” tem alguns usos inteligíveis, embora vagos ou oscilantes, tem de ser possível construir tal lista. Presumivelmente, deveria ser mais fácil concordar se algum item deveria ser incluído do que sobre qualquer análise proposta da noção de metáfora.

Talvez a seguinte lista de exemplos, selecionada não completamente ao acaso, possa servir:

- (i) “O diretor cortou o direito a discussão.”
- (ii) “Uma cortina de fumaça de testemunhas”.
- (iii) “Uma melodia persuasiva”.
- (iv) “Vozes de mata-borrão” (Henry James).
- (v) “Os pobres são os negros da Europa” (Baudelaire).
- (vi) “A luz é nada mais que a sombra de Deus” (Sir Thomas Browne).

- (vii) “Oh querida criança branca, fortuita como os pássaros, brincando em meio a linguagens arruinadas.” (Auden).

Espero que todas essas sejam aceitas como *instâncias* inconfundíveis de metáfora, quaisquer que sejam os juízos que possam em último caso ser feitos sobre o significado de “metáfora”. Os exemplos são oferecidos como casos claros de metáfora, mas, com a possível exceção do primeiro, seriam impróprios como “paradigmas”. Se quiséssemos ensinar o significado de “metáfora” a uma criança, precisaríamos de exemplos mais simples, como “As nuvens estão chorando”, ou “As raízes estão lutando entre si”. (É importante que se encontre exemplos de personificação?) Mas tenho tentado incluir algumas advertências de possíveis complexidades que mesmo metáforas relativamente fáceis possam gerar.

Considere o primeiro exemplo - “O diretor cortou o direito a discussão”. Um ponto inicial óbvio é o contraste entre a palavra “cortou” e as palavras restantes pelas quais ela é acompanhada. Isso seria comumente expresso dizendo-se que “cortou” tem aqui um sentido metafórico enquanto as outras palavras têm sentido literal. Apesar de apontarmos a frase inteira como uma instância (um “caso claro”) de metáfora, a nossa atenção se volta rapidamente a uma única palavra, cuja presença é a razão aproximada para a atribuição. E observações similares podem ser feitas sobre os próximos quatro exemplos da lista, sendo as palavras cruciais, respectivamente, “cortina de fumaça”, “argumentativo”, “mata-borrão” e “negros”.

(Mas a situação é mais complicada nos dois últimos exemplos da lista. Na citação de Sir Thomas Browne, “Luz” supostamente deve ter um sentido simbólico, e certamente significar muito mais do que significaria num contexto de um manual de ótica. Aqui o sentido metafórico da expressão “a sombra de Deus” impõe um significado mais rico do que o usual ao conteúdo da frase. Efeitos similares podem ser notados na passagem de Auden (considere, por exemplo,

o significado de “branco” na primeira linha). Terei de negligenciar tais complexidades neste artigo.)

Em geral, quando falamos de uma metáfora relativamente simples, estamos nos referindo a uma frase ou a uma outra expressão na qual *algumas* palavras são usadas metaforicamente embora as restantes sejam usadas não-metaforicamente. Uma tentativa de construir uma frase inteira de palavras que sejam usadas metaforicamente resulta num provérbio, numa alegoria ou num enigma. Nenhuma análise preliminar da metáfora cobrirá satisfatoriamente mesmo um exemplo tão trivial como “à noite todas as vacas são pretas”. E casos de simbolismo (no sentido em que o castelo de Kafka é um “símbolo”) também necessitam de tratamento separado.

## II

“O diretor cortou o direito à discussão”. Ao dizer que essa frase é um caso de metáfora, estamos dizendo que pelo menos uma palavra (aqui, a palavra “cortou”) está sendo usada metaforicamente na frase, e que pelo menos uma das palavras restantes está sendo usada literalmente. Chamemos à palavra “Cortou” o *foco* da metáfora, e ao restante da frase na qual essa palavra ocorre, *enquadramento*. (*Estamos agora usando metáforas – e misturando-as de alguma maneira? Isso importa?*) Uma noção que precisa ser clarificada é a de “uso metafórico” do foco de uma metáfora. Entre outras coisas, seria bom compreender como a presença de um único enquadramento pode resultar no uso metafórico da palavra complementar, enquanto que a presença de um enquadramento diferente para a mesma palavra não resulta em metáfora.

Se a frase sobre o comportamento do diretor fosse traduzida palavra por palavra em alguma língua estrangeira na qual isso fosse possível, diríamos que a frase traduzida é um caso *exatamente da mesma* metáfora. Assim, chamar uma frase de um caso de metáfora

é dizer algo acerca do seu *significado*, não sobre sua ortografia, seu padrão fonético ou sua forma gramatical<sup>44</sup>. (Para usar uma distinção bem conhecida, a “metáfora” tem de ser classificada como um termo pertencente à “semântica” e não à “sintaxe” – ou a qualquer investigação *física* sobre a linguagem.)

Suponha que alguém diga “Gosto de cortar minhas memórias regularmente”. Deveríamos ou não dizer que ele está usando a mesma metáfora como no caso já discutido? Nossa resposta dependerá do grau de similaridade que estamos dispostos a atribuir na comparação dos dois “enquadramentos” (pois temos o mesmo “foco” em cada uma das vezes). Diferenças nos dois enquadramentos produzirão *algumas* diferenças na interação<sup>45</sup> entre foco e enquadramento nos dois casos. Se considerarmos as diferenças como suficientemente notáveis para nos justificar a dizer que as frases são *duas* metáforas é uma questão de decisão arbitrária. “Metáfora” é, na melhor das hipóteses, uma palavra vaga, e temos de ser cautelosos ao atribuir a ela regras de uso mais estritas do que aquelas que são efetivamente encontradas na prática.

Até agora, tenho tratado “metáfora” como um predicado aplicável apropriadamente a certas expressões, sem dar atenção a quaisquer ocasiões nas quais as expressões são usadas, ou aos pensamentos, ações, sentimentos e intenções dos falantes em tais ocasiões. E isso é seguramente correto para *algumas* expressões. Reconhecemos, sem a necessidade de saber quem usou a expressão, ou em quais ocasiões ou com que intenção, que chamar um homem de “latrina” é usar uma metáfora. As regras da nossa linguagem determinam que algumas expressões têm de contar como metáforas; e um falante não pode mais mudar isso do mesmo modo que pode legislar que “vaca” deva significar o mesmo que “ovelha”. Mas temos

---

<sup>44</sup> Qualquer parte do discurso pode ser usada metaforicamente (ainda que os resultados sejam escassos e desinteressantes no caso das conjunções), qualquer forma de expressão verbal pode conter um foco metafórico.

<sup>45</sup> Aqui estou usando a linguagem apropriada para a “perspectiva interacionista” da metáfora que é discutida no final deste artigo.

também de reconhecer que as regras estabelecidas da linguagem deixam uma ampla margem para variação, iniciativa e criação individuais. Há indefinidamente muitos contextos (incluindo aproximadamente todos os contextos interessantes) onde o significado de uma expressão metafórica tem de ser reconstruído a partir das intenções do falante (e outras pistas), pois as regras amplas do uso padrão são demasiado gerais para fornecer a informação necessária. Quando Churchill, em uma famosa expressão, chamou Mussolini de “aquele *utensílio*”, o tom de voz, a colocação verbal e o pano de fundo histórico ajudaram a deixar claro *qual* metáfora estava sendo usada. (Contudo, mesmo aqui, é difícil ver como a expressão “aquele *utensílio*” poderia ser ainda aplicada a um homem a não ser como um insulto. Aqui, como em outra parte, as regras gerais de uso funcionam como limitações na liberdade do falante de dizer tudo o que ele quiser. Esse é um exemplo, embora ainda simples, do quanto o reconhecimento e a interpretação de uma metáfora podem exigir atenção às *circunstâncias particulares* de sua elocução.

É especialmente notável que, em geral, não há qualquer regra padrão para o grau de *importância* ou *ênfase* a ser atribuído a um uso particular de uma expressão. Para saber o que o usuário de uma metáfora quer dizer, necessitamos saber o quão “seriamente” ele trata o foco metafórico. (Ele ficaria igualmente satisfeito em ter um sinônimo aproximado ou somente *aquela* palavra serviria? Devemos considerar apressadamente a palavra, observando somente suas implicações mais óbvias – ou deveríamos dar importância às suas associações menos imediatas?) No discurso podemos usar ênfase e fraseologia como dicas. Mas no discurso impresso ou escrito, mesmo estes auxílios rudimentares estão ausentes. No entanto, esse “peso” de certa forma esquivo de uma metáfora (suposta ou detectada<sup>46</sup>) é de grande importância prática na exegese.

---

<sup>46</sup> Quero aqui que se dê, tanto quanto possível, o menor “peso” possível a essas palavras!

Tomemos um exemplo filosófico. Se a expressão “forma lógica” deveria ser tratada em um enquadramento particular como tendo um sentido metafórico, dependerá da medida em que o usuário é considerado consciente de alguma suposta analogia entre argumentos e outras coisas (vasos, nuvens, batalha, piadas) “forma”. Dependerá ainda mais se o escritor deseja que a analogia seja ativada na mente de seus leitores; e quanto o seu próprio pensamento depende e está nutrido pela suposta analogia. Não temos de esperar que as “regras da linguagem” sejam de muita ajuda em tal investigação. (Há, de acordo com isso, um sentido de “metáfora” que pertence à “pragmática” ao invés de pertencer à “semântica” – e esse sentido pode ser o que merece mais atenção).

### III

Tentemos a explicação mais simples possível que possa ser dada do significado de “O diretor cortou o direito à discussão” e vejamos o quanto longe ela nos levará. Um comentário plausível (para aqueles que presumivelmente têm a mentalidade por demais literal para compreender o original) pode ser alguma coisa como o que se segue:

“Um falante que usa a frase em questão é entendido como alguém que quer dizer *alguma coisa* sobre um diretor e seu comportamento em alguma reunião. Em vez de dizer clara ou *diretamente* que o diretor lidou sumariamente com objeções, ou impiedosamente suprimiu comentários irrelevantes, ou algo do tipo, o falante escolheu usar a palavra (‘cortou’) que, estritamente falando, significa algo diferente. Mas um ouvinte inteligente pode facilmente imaginar o que o falante tinha em mente.”<sup>47</sup>

Essa explicação trata a expressão metafórica (chamemo-la “M”) como uma substituta para alguma outra expressão literal (“L”, digamos) que teria expressado o mesmo significado, tivesse ela sido

---

<sup>47</sup> Note como esse tipo de paráfrase naturalmente carrega alguma implicação de *falha* por parte do autor da metáfora. Há uma forte sugestão que ele deve ter tirado da cabeça aquilo que ele realmente quis dizer – a metáfora é descrita como um modo de encobrir falta de clareza e incertezas.

usada. De acordo com essa perspectiva, o significado de *M*, em sua ocorrência metafórica, é apenas o significado *literal* de *L*. O uso metafórico de uma expressão consiste, de acordo com essa perspectiva, no uso dessa expressão em outro sentido que não seja o seu sentido apropriado ou normal, em algum contexto que permita o sentido inapropriado ou anormal a ser detectado e apropriadamente transformado. (As razões oferecidas para uma performance tão notável serão discutidas mais tarde.)

A qualquer perspectiva que sustente que uma expressão metafórica é usada no lugar de alguma expressão *literal* equivalente, chamarei *perspectiva substitutivista da metáfora*. (Gostaria que essa etiqueta cobrisse também qualquer análise que considere a frase inteira, que é o *locus* da metáfora, como substituindo um conjunto de frases literais.) Até recentemente, uma ou outra forma de perspectiva substitutivista tem sido aceita por muitos autores (comumente críticos literários ou autores de livros de retórica) que tiveram qualquer coisa a dizer sobre as metáforas.

Tomemos alguns exemplos. Whetely define uma metáfora como “uma palavra substituída por outra em razão da Semelhança ou Analogia entre suas significações.”<sup>48</sup> O verbete do dicionário Oxford (para saltar aos tempos modernos) também não é muito diferente disso: “Metáfora: A figura de linguagem na qual um nome ou termo descritivo é transferido para algum objeto diferente, mas análogo àquele que o termo é apropriadamente aplicável; um exemplo disso é uma expressão metafórica.”<sup>49</sup> Tão solidamente fortificada está a perspectiva expressa por essas definições que um autor

<sup>48</sup> Richard Whately, *Elements of Rhetoric* (7th revised. ed., London, 1846), p. 280.

<sup>49</sup> Em “Figura” temos: “Qualquer das várias “formas” de expressão, desviando do arranjo ou uso normal das palavras, que são adotadas a fim de dar beleza, variedade ou força a uma composição; e.g., Aposiopese, Hipérbole, Metáfora, etc.” Se tomarmos isso estritamente podemos ser conduzidos a dizer que a transferência de uma palavra não adotada com o interesse de introduzir “beleza, variedade, ou força” tem de necessariamente falhar em ser um caso de metáfora. Ou a “variedade” automaticamente cobrirá *toda* a transferência? Notar-se-á que a definição do D.O.I. não é uma melhoria de Whately. Onde ele fala de uma “palavra” sendo substituída, o D.O.I. prefere “nome ou termos descritivos.” Se isso quer dizer restringir as metáforas a substantivos (e adjetivos?), está demonstravelmente errado. Mas, se não, qual é o “termo descritivo” que supostamente significa? E por que a referência de Whately à “Semelhança ou Analogia” foi aplicada a uma referência à analogia apenas?

contemporâneo que esteja explicitamente argumentando a favor de uma perspectiva da metáfora diferente e mais sofisticada, todavia, escorrega no velho modelo ao definir metáfora como “dizendo uma coisa e querendo dizer outra.”<sup>50</sup>

Segundo a perspectiva substitutivista, o foco de uma metáfora, a palavra ou expressão que tem um uso distintamente metafórico no interior de um enquadramento literal, é usada para comunicar um significado que poderia ter sido expresso literalmente. O autor substitui *M* por *L*; é a tarefa do leitor inverter a substituição, usando o significado literal de *M* como um indício para o pretendido significado literal de *L*. Compreender uma metáfora é como decifrar um código ou decifrar um enigma.

Se agora perguntarmos por que, de acordo com essa perspectiva, o autor deveria dar a seus leitores a tarefa de resolver um quebra-cabeça, dois tipos de resposta nos serão oferecidas. A primeira é que de fato pode não haver equivalente literal, *L*, disponível na linguagem em questão. Os matemáticos falaram da “perna” de um ângulo porque não havia expressão literal concisa para denotar uma linha delimitadora; dizemos “lábios de cereja” porque não há modo que seja tão conveniente para dizer rapidamente como os lábios são. As metáforas preenchem as lacunas no vocabulário literal (ou, ao menos, supre a falta de abreviações convenientes). Assim vista, a metáfora é uma espécie de *catacrese*, que definirei como o uso de uma palavra em algum novo sentido a fim de reparar a lacuna no vocabulário. Catacrese é dar novos sentidos a velhas palavras.<sup>51</sup> Mas se uma catacrese serve como uma necessidade genuína, o novo sentido introduzido rapidamente tornará

---

<sup>50</sup> Owen Barfield, “Poetic Diction and Legal Fiction,” in *Essays Presented to Charles Williams* (Oxford, 1947), pp. 106–127. A definição de metáfora ocorre na p.111, onde a metáfora é tratada como um caso especial do que Barfield chama “ato de apontar”. Todo o ensaio merece ser lido.

<sup>51</sup> O D.O.I. define catacrese como: O uso impróprio de palavras; aplicação de um termo a uma coisa que ela não denota apropriadamente, abuso ou deturpação de um tropo ou metáfora.” Quero excluir as sugestões pejorativas. Nada há de deturpado ou abusivo em alongar velhas palavras para ajustá-las a novas situações. Catacrese é simplesmente um caso de surpreendente transformação de significado que está constantemente ocorrendo em qualquer linguagem viva.

parte do sentido *literal*. “Laranja” pode originalmente ter sido aplicado à cor por catacrese; mas agora a palavra é aplicada “apropriadamente” (e não metaforicamente) tanto à cor quanto à fruta. Curvas “osculadoras” não se beijam por muito tempo e passam rapidamente a um contato matemático mais prosaico. E o mesmo ocorre em outros casos. Quando bem-sucedida, o destino da catacrese é desaparecer.

Há, contudo, muitas metáforas sobre as quais as virtudes atribuídas à catacrese não podem se aplicar, porque há, ou tem de haver, algum equivalente literal facilmente disponível e igualmente conciso. Portanto, no exemplo um pouco infeliz<sup>52</sup> “Richard é um leão”, que escritores contemporâneos têm discutido com maçante insistência, o significado literal é tomado como sendo o mesmo que o da frase “Richard é bravo”.<sup>53</sup> Aqui, a metáfora não tem a pretensão de enriquecer o vocabulário.

Quando a catacrese não pode ser invocada, as razões para substituir uma expressão indireta, metafórica, são tidas como estilísticas. É-nos dito que a expressão metafórica pode (em seu uso literal) se referir a mais objetos concretos que poderia sua equivalente literal; e isso supostamente dá prazer ao leitor (o prazer de ter os próprios pensamentos afastados de Richard e aproximados do irrelevante leão). Novamente, o leitor é tomado como alguém que aprecia a resolução de problemas – ou tomado como alguém que se deleita com a habilidade do autor que ora parcialmente oculta e ora parcialmente revela seu significado. Ou as metáforas dão um choque de “surpresa agradável” – e assim por diante. O princípio por trás dessas “explicações” parece ser: Quando em dúvida sobre alguma

---

<sup>52</sup> Podemos imaginar um indivíduo qualquer dizendo isso hoje em dia e seriamente querendo dizer qualquer coisa? Acho que é difícil fazê-lo. Mas na falta de um contexto autêntico de uso, qualquer análise está sujeita a ser diluída, óbvia e não proveitosa.

<sup>53</sup> Uma discussão desse exemplo, completa de diagramas, será encontrada em *Meaning and Change of Meaning* de Gustaf Stern (Goteborgs Hogskolas Arsskrift, vol. 38, 1932, part 1), pp.300 ff. A abordagem de Stern tenta apresentar como o leitor é conduzido pelo contexto a *selecionar* da conotação de “leão” o atributo (bravura) que se aplicará ao homem Richard. Entendo-o como defendendo uma forma de perspectiva substitutivista.

peculiaridade da linguagem, atribua sua existência ao prazer que ela dá ao leitor. Um princípio que tem o mérito de funcionar bem na falta de quaisquer indícios.<sup>54</sup>

Quaisquer que sejam os méritos de tais especulações sobre a resposta do leitor, elas concordam em fazer da metáfora uma *decoreção*. Exceto em casos em que uma metáfora é uma catacrese que remedia algumas imperfeições temporárias da linguagem literal, o propósito da metáfora é entreter e divertir. Seu uso, de acordo com essa perspectiva, sempre constitui um desvio do “comum e do estilo estritamente apropriado” (Whately).<sup>55</sup> Desse modo, se os filósofos têm algo mais importante para fazer do que dar prazer aos seus leitores, à metáfora pode não ter um lugar sério na discussão filosófica.

#### IV

A perspectiva de que uma expressão metafórica tem um significado que é uma transformação de seu significado literal normal é um caso especial de uma perspectiva mais geral acerca da linguagem “figurativa”. Essa sustenta que qualquer figura de linguagem envolvendo mudança semântica (e não meramente mudança sintática, como a inversão da ordem normal da palavra) consiste em alguma transformação de um significado *literal*. O autor fornece não seu significado pretendido,  $m$ , mas uma função sua,  $f(m)$ ; a tarefa do leitor é aplicar a função inversa,  $f^{-1}$ , e assim obter  $f^{-1}(f(m))$ , *i.e.*,  $m$ , o significado original. Quando funções diferentes são usadas, diferentes tropos resultam. Portanto, na ironia, o autor diz o *oposto* do que ele

---

<sup>54</sup> Aristóteles atribui o uso da metáfora ao prazer na aprendizagem; Cícero liga o prazer na metáfora ao gozo da ingenuidade do autor na passagem do imediato, ou na vívida apresentação do objeto principal. Para a referência dessas e outras perspectivas tradicionais veja E. M. Cope, *An Introduction to Aristotle's Rhetoric* (London, 1867), “Appendix B to Book III, Ch. II: On Metaphor”.

<sup>55</sup> Assim Stern (*op. Cit.*) diz de todas as figuras de linguagem que “elas visam servir à função expressiva e útil do discurso melhor que ‘afirmações simples’” (p.296). Uma metáfora produz uma “intensificação” (*steigerung*) do objeto, mas o fator que conduz a seu uso “envolve a função expressiva e efetiva (útil) do discurso, não a função simbólica e comunicativa” (p. 290). Isso é para dizer que metáforas podem revelar sentimentos e outras predisposições para agir e perceber de vários modos– mas elas tipicamente não *dizem* qualquer coisa.

pretende, na hipérbole, ele *exagera* seu significado, e assim por diante.

Qual é, então, a função transformadora característica envolvida na metáfora? A resposta a isso tem sido: *analogia* ou *similaridade*. *M* é ou similar ou análoga em significado a sua equivalente literal *L*. Uma vez que o leitor tenha detectado a razão da analogia ou símile pretendida (com a ajuda do enquadramento, ou pistas extraídas de um contexto mais geral), ele pode refazer o caminho do autor e assim alcançar o significado literal original (o significado de *L*).

Se o autor sustenta que uma metáfora consiste na *apresentação* da analogia ou similaridade subjacente, ele estaria adotando o que eu chamarei uma “*perspectiva comparativista*” da metáfora. Quando Schopenhauer chamou uma prova geométrica de ratoeira, ela estava, segundo tal perspectiva, *dizendo* (ainda que não explicitamente): “Uma prova geométrica é *como* uma ratoeira, uma vez que ambas oferecem uma recompensa ilusória, atraem suas vítimas gradualmente, levam a uma surpresa desagradável, etc.” Esta é uma perspectiva da metáfora como um *símile* condensado ou elíptico. Notar-se-á que uma “*perspectiva comparativista*” é um caso especial de uma “*perspectiva substitutivista*.” Pois sustenta que o enunciado metafórico pode ser substituído por uma *comparação* equivalente literal.

Whately disse: “O Símile ou a Comparação podem ser considerados como diferindo de uma metáfora apenas na forma; a semelhança, que está implícita na metáfora, nesse caso é *explicitada*”.<sup>56</sup> Bain disse que “A metáfora é uma comparação implícita no mero uso do termo” e acrescenta que “é na circunstância de estar confinada a uma palavra, ou quando muito a uma expressão, que temos de procurar pelas peculiaridades da metáfora – suas

<sup>56</sup> Whately, *loc. Cit.* Ele prossegue traçando uma distinção entre “Semelhança propriamente dita, *i.e.*, semelhança *direta* entre os próprios objetos em questão, (como quando falamos de ‘planalto’, ou comparamos grandes ondas a *montanhas*)” e “Analogia, que é a semelhança de Proporção – uma similaridade das relações que elas mantêm com outros objetos; como quando falamos de ‘*luz* da razão’, ou de ‘*revelação*’, ou comparamos um guerreiro ferido e aprisionado a um navio encalhado”.

vantagens por um lado, e perigos e abuso por outro”.<sup>57</sup> Essa perspectiva da metáfora, como símile ou comparação condensada, tem sido muito popular.

A principal diferença entre uma perspectiva substitutivista (do tipo anteriormente considerado) e a forma especial disso a que tenho chamado de perspectiva comparativista pode ser ilustrada pelo exemplo simples de “Richard é um leão”. De acordo com a primeira perspectiva, a frase significa aproximadamente o mesmo que “Richard é corajoso”; de acordo com a segunda, aproximadamente o mesmo que “Richard é *como* um leão (em ser corajoso)”, sendo as palavras adicionadas nos parênteses entendidas, embora não explicitamente enunciadas. Na segunda tradução, como na primeira, o enunciado metafórico é tomado como ocupando o lugar de um enunciado *literal* equivalente. Mas a perspectiva da comparação fornece uma paráfrase mais elaborada, visto que o enunciado original é interpretado tanto como sendo sobre leões quanto sobre Richard.<sup>58</sup>

A principal objeção contra uma perspectiva comparativista é que ela sofre de uma vagueza que beira a vacuidade. É de se esperar que estejamos perplexos sobre como alguma expressão (*M*) usada metaforicamente pode funcionar no lugar de alguma expressão literal (*L*) considerada como um sinônimo aproximado; e a resposta oferecida é que aquilo que *M* representa (em seu uso literal) é *similar* ao que *L* representa. Mas, o quão informativo isso é? Há uma tentação em pensar nas similaridades como “objetivamente dadas”, de modo que uma pergunta da forma “no que diz respeito a *P*, *A* é como *B*?” tenha uma resposta definida e predeterminada. Se assim fosse, símiles seriam regidos por regras tão rígidas quanto aquelas que controlam as afirmações da física. Mas semelhança sempre

---

<sup>57</sup> Alexander Bain, *English Composition and Rhetoric* (Enlarged edition, London, 1887), p.159.

<sup>58</sup> A perspectiva comparativista provavelmente deriva de uma breve declaração de Aristóteles manifesta na *Poética*: “A metáfora consiste em dar à coisa um nome que pertence à outra; sendo a transferência ou do gênero para a espécie, ou da espécie para o gênero, ou da espécie para espécie, ou sobre a base de analogia” (1457b). Não tenho espaço para uma discussão detalhada de Aristóteles como ela merece. Uma defesa competente de uma perspectiva baseada em Aristóteles será encontrada em *The World of Imagery* (London, 1927, especialmente. pp. 67 ff.) de S. J. Brown.

admite graus, de modo que uma pergunta verdadeiramente “objetiva” necessitaria de tomar alguma forma tal como “É A mais parecido com B do que com C no que diz respeito a P?” – ou, talvez, “É A mais próximo de B do que de C em tal e tal escala de graus de P?”. Contudo, na proporção em que abordamos tais formas, os enunciados metafóricos perdem sua efetividade e seu objetivo. Precisamos das metáforas apenas nos casos em que não pode haver qualquer questão tão precisa quanto os enunciados científicos. Um enunciado metafórico não é um substituto para uma comparação formal ou qualquer outro tipo de enunciado literal, mas tem suas próprias capacidades e realizações *distintas*. Frequentemente dizemos “X é M”, evocando alguma conexão atribuída entre M e uma atribuída a L (ou, ao invés, a um sistema indefinido  $L_1, L_2, L_3, \dots$ ) em casos em que, antes da construção da metáfora, teria sido difícil para nós expressá-la a fim de encontrar uma semelhança *literal* entre M e L. Seria mais iluminante em alguns desses casos dizer que a metáfora *cria* a similaridade do que dizer que ela formula alguma similaridade anteriormente existente.<sup>59</sup>

## V

Volto agora a considerar um tipo de análise que chamo de *perspectiva interacionista* da metáfora. Ela me parece ser livre dos principais defeitos da perspectiva da substituição e da perspectiva da comparação e oferece um importante *insight* sobre os usos e as limitações da metáfora.<sup>60</sup>

---

<sup>59</sup> Muito mais necessitaria ser dito em um exame completo da perspectiva comparativista. Seria revelador, por exemplo, considerar tipos contrastantes de casos nos quais uma comparação formal é preferida a uma metáfora. Uma comparação é muitas vezes um prelúdio a uma apresentação explícita das bases da semelhança contanto que não esperemos que uma metáfora explique a si própria. (Cf. A diferença entre *comparar* a face de um homem com uma máscara de lobo ao procurar por pontos semelhantes– e ver a face humana *como* vulpina.) Mas, sem dúvida, o limite entre *algumas* metáforas e *alguns* símiles não é preciso.

<sup>60</sup> As melhores fontes são os escritos de I. A. Richards, especialmente o Capítulo 5 (“Metaphor”) e o capítulo 6 (“Command of Metaphor”) de seu *The Philosophy of Rhetoric* (Oxford, 1936). Capítulos 7 e 8 de seu *Interpretation in Teaching* (London, 1938) abrange muito do mesmo assunto. Em *Greek Metaphor* (Oxford, 1936), W. Bedell Stanford defende, com erudição e destreza, o que ele chama de uma “teoria integracionista” (veja especialmente pp. 101 ff.). Infelizmente, ambos os autores têm grandes problemas em tornar clara a natureza da posição que estão defendendo. O capítulo 18 de *The Structure of Complex Words* (London, 1951) de W. Empson é uma discussão útil da perspectiva de Richard sobre a metáfora.

Começemos com o seguinte enunciado: “Na formulação mais simples, quando usamos uma metáfora temos dois pensamentos sobre coisas diferentes juntamente ativos e suportados por uma única palavra ou frase cujo significado é um resultante de sua interação.”<sup>61</sup>

Podemos descobrir o que aqui se pretende ao empregar a observação de Richard ao nosso exemplo anterior: “Os pobres são os negros da Europa”. A perspectiva da substituição, em sua forma mais crua, diz-nos que alguma coisa está sendo *indiretamente* dita sobre os pobres da Europa. (Mas o que? Que eles são uma classe oprimida, uma vergonha permanente aos ideais oficiais da comunidade, que pobreza é herdada e indelével?) A perspectiva da comparação diz que a epigrama *apresenta* alguma comparação entre os pobres e os negros. Em oposição a ambas, Richard diz que nossos “pensamentos” sobre os europeus pobres e os negros (americanos) são “ativados juntamente” e “interagem” para produzir um significado que é um resultante dessa interação.

Penso que isto tem de significar que em dado contexto a palavra foco “negros” obtém um *novo* significado, que *não* é exatamente seu significado nos usos literais, nem exatamente o significado que qualquer substituta literal teria. O novo contexto (o “enquadramento” da metáfora, em minha terminologia) impõe *extensão* de significado sobre a palavra foco. E considero que Richard está a dizer que para a metáfora funcionar o leitor tem de manter-se ciente da extensão do significado – tem de prestar atenção ao significado velho e novo juntos.<sup>62</sup>

Mas como se dá essa extensão ou mudança de significado? Em um ponto, Richards fala de “características comuns” dos dois termos (os pobres e negros) como “o fundamento da metáfora” (*op. cit.*, p. 117), de modo que em seu uso metafórico uma palavra ou expressão

<sup>61</sup> *The Philosophy of Rhetoric*, p.93. Richards também diz que a metáfora é “fundamentalmente um empréstimo e um intercurso entre *pensamentos*, uma transação entre contextos” (p.94). A metáfora, diz ele, requer duas ideias “que cooperam em um significado inclusivo” (p.119).

<sup>62</sup> Talvez seja isso que conduz Richards a dizer que “falar da identificação ou fusão que uma metáfora realiza é quase sempre enganoso e pernicioso” (*op.cit.*, p. 127).

tem de conotar somente uma *seleção* de características conotadas em seus usos literais. Isso, contudo, parece um retorno à análise mais antiga e menos sofisticada que ele está tentando substituir.<sup>63</sup> Ele está em terra mais firme quando diz que o leitor é forçado a “conectar” as duas ideias (p.125). E nessa “conexão” reside o segredo e o mistério da metáfora. Falar de “interação” de dois pensamentos “ativados conjuntamente” (ou, uma vez mais, de sua “inter-iluminação” ou “cooperação”) é *usar* uma metáfora enfatizando os aspectos dinâmicos de uma boa resposta do leitor a uma metáfora não trivial. Não tenho problemas com o *uso* de metáforas (caso sejam boas) para falar sobre a metáfora. Mas seria bom que usássemos várias metáforas para evitar sermos enganados pelo casual encanto das nossas preferidas.

Tentemos, por exemplo, pensar em uma metáfora como um *filtro*. Considere o enunciado “O homem é um lobo”. Aqui, podemos dizer que estão *dois* objetos - o *objeto principal*, Homem (ou: homens) e o *objeto auxiliar*, Lobo (ou: lobos). Ora, o significado pretendido da frase metafórica em questão não será transmitido ao leitor suficientemente ignorante sobre lobos. Não é preciso que o leitor saiba o significado padrão de “lobo” - ou seja capaz de usar aquela palavra em sentido literal. É preciso apenas que saiba aquilo a que chamarei de o *sistema de lugares-comuns associados*. Imagine que se peça a um leigo para dizer, de muito especial, aquilo que ele pensa ser verdadeiro acerca dos lobos; o conjunto de afirmações resultante seria aproximado ao que estou chamando de o sistema de lugares-comuns associados à palavra “lobo”. Presumo que em qualquer cultura as respostas dadas por diferentes pessoas ao teste sugerido concordariam aproximadamente, e que mesmo um eventual especialista, que pode ter um notável conhecimento do assunto, ainda saberia “aquilo que o homem na rua pensa sobre o assunto”.

---

<sup>63</sup> Comumente, Richard tenta mostrar que a similaridade entre os dois termos é, na melhor das hipóteses, *parte* da base para a interação dos significados em uma metáfora.

Do ponto de vista do especialista, o sistema de lugares-comuns associados pode incluir meias verdades ou erros claros (como quando uma baleia é classificada como um peixe); mas o importante para a efetividade da metáfora não é que o lugar comum seja verdadeiro, mas antes que ele seja pronta e livremente despertado. (Por ser assim, uma metáfora que funciona em uma sociedade pode parecer absurda em outra. Homens que consideram que lobos são reencarnações de humanos mortos dariam ao enunciado “O homem é um lobo” uma interpretação diferente daquela que tenho assumido).

Para formular a questão de outro modo: usos literais da palavra “lobo” são regidos por regras sintáticas e semânticas, regras que quando violadas produzem falta de sentido ou autocontradições. Em acréscimo, sugiro, usos literais das palavras normalmente comprometem o locutor com a aceitação de um conjunto padrão de crenças sobre lobos (trivialidades correntes) que são de domínio comum dos membros de alguma comunidade falante. Negar qualquer parte do lugar comum aceito (e.g., dizendo que os lobos são vegetarianos - ou facilmente domesticados) produz um efeito paradoxal e provoca uma demanda por justificação. Um locutor que diz “lobo” é normalmente tomado como alguém que pressupõe estar se referindo, em algum sentido da palavra, a algo feroz, carnívoro, traiçoeiro, e assim por diante. A ideia de um lobo é parte de um sistema de ideias não nitidamente delineado mas que, no entanto, é suficientemente definido a ponto de admitir enumeração detalhada.

O efeito, então, de (metaforicamente) chamar um homem de “lobo” é evocar o sistema-lobo de lugares comuns relatados. Se o homem é um lobo, ele se alimenta de outros animais, é feroz, faminto, envolvido em constante luta, um animal que se alimenta de carniça, e assim por diante. Em cada uma dessas asserções subtendidas tem-se agora de ser feita a adequação ao objeto principal (o homem) ou em sentido normal ou anormal. Se a metáfora for de todo adequada, isso pode ser feito - ao menos até certo ponto.

Um ouvinte apropriado será conduzido pelo sistema-lobo de suposições a construir um sistema correspondente de suposições sobre o objeto principal. Mas essas suposições *não* serão aquelas incluídas nos lugares comuns *normalmente* pressupostos pelos usos literais de “homem”. As novas suposições têm de ser determinadas pelo padrão de pressuposições associado ao uso literal da palavra “lobo”. Quaisquer traços humanos sobre os quais podemos sem esforço excessivo falar na “língua-lobo” serão retribuídos prontamente, e quaisquer traços que não possam ficarão em segundo plano. A metáfora-lobo suprime alguns detalhes e enfatiza outros – em suma, *organiza* a nossa concepção de homem.

Considere que eu olhe para o céu noturno através de um pedaço de vidro embaçado no qual certas linhas foram deixadas claras. Assim, verei somente as estrelas que podem repousar naquelas linhas que foram previamente deixadas claras sobre o vidro, e as estrelas que eu vejo serão vistas como organizadas pelo modo como o vidro está. Podemos pensar na metáfora como esse vidro e no sistema de “lugares-comuns associados” da palavra foco como uma rede de linhas sobre o vidro. Podemos dizer que o objeto principal é “visto através” da expressão metafórica – ou, se preferirmos, que o objeto principal é “projetado sobre” o campo do objeto subsidiário. (Na última analogia, o sistema de suposições da expressão foco tem de ser tomado como aquilo que determina a “lei da projeção”).

Ou tomemos outro exemplo. Considere que sou incumbido de descrever uma batalha usando, tanto quanto possível, palavras retiradas do vocabulário do xadrez. Esses últimos termos determinam um sistema de suposições cujo procedimento é controlar a minha descrição da batalha. A escolha forçada do vocabulário do xadrez conduzirá a alguns aspectos da batalha a serem enfatizados, outros a serem negligenciados, e tudo isso a ser organizado de um modo que causaria muito mais força que em outros modos de descrição. O vocabulário do xadrez filtra e transforma: ele não apenas seleciona,

apresenta aspectos da batalha que poderiam não ser vistos por outro meio, (Estrelas que não podem ser vistas, exceto através de telescópios.)

Também não devemos negligenciar as mudanças de atitude que regularmente resultam do uso da linguagem metafórica. Um lobo é (convencionalmente) um objeto detestável e aterrador; assim, chamar um homem de lobo é sugerir que ele também é detestável e aterrador (e, desse modo, sustentar e reforçar atitudes dialógicas). Novamente, o vocabulário do xadrez tem seu uso primário em um cenário altamente artificial no qual toda a expressão de sentimento é formalmente excluída: descrever uma batalha como se ela fosse um jogo de xadrez é, de acordo com isso, excluir, pela escolha da linguagem, todos os aspectos mais emocionalmente perturbadores da guerra. (Subprodutos similares não são raros nos usos filosóficos da metáfora.)

Uma objeção razoavelmente óbvia à “perspectiva interacionista” recém esboçada é que ela tem de sustentar que alguns dos “lugares-comuns associados” sofrem mudança de significado metafórico no processo de transferência do objeto secundário para o principal. E essas mudanças, caso ocorram, dificilmente podem ser explicadas pela abordagem fornecida. A metáfora primária, poder-se-ia dizer, foi analisada dentro de um conjunto de metáforas subordinadas, de modo que a explicação dada ou é circular ou conduz a um regresso ao infinito.

Poder-se-ia responder a tal negando-se que *todas* as mudanças de significado nos “lugares-comuns associados” tenham de ser consideradas como mudanças metafóricas. Muitas delas são mais bem descritas como *extensões* do significado, pois não envolvem conexões aprendidas entre dois sistemas de conceitos. Não me dei ao trabalho de explicar como tais extensões ou mudanças ocorrem no geral, e não penso que qualquer explicação simples caiba a todos os casos. (É muito fácil resmungar “analogia”, mas um exame mais

atento logo exhibe todos os tipos de “bases” para mudanças de significado com contexto – e, às vezes, nenhuma).

Em segundo lugar, eu não negaria que uma metáfora pudesse envolver certa quantidade de metáforas subordinadas entre suas pressuposições. Penso, contudo, que geralmente tenciona-se que tais metáforas subordinadas sejam entendidas menos “enfaticamente”, *i.e.*, com menos ênfase sobre *suas* pressuposições. (As pressuposições de uma metáfora são como as nuances de um acorde; dar “importância” demais a elas é como tentar fazer o som secundário soar tão alto quanto as notas principais – e tão inútil quanto). Em qualquer caso, as metáforas primárias e subordinadas irão normalmente pertencer ao mesmo campo de discurso, de modo que elas mutuamente reforçam um e o mesmo sistema de pressuposições. Reciprocamente, quando metáforas substancialmente novas aparecem, à medida que a metáfora primária vai sendo esclarecida, há sérios riscos de confusão de pensamento (*c.f.* a costumeira proibição contra “metáforas misturadas”).

A explicação precedente da metáfora, contudo, necessita de correção para que seja razoavelmente adequada. A referência aos “lugares-comuns associados” funcionará nos casos mais comuns em que o autor joga apenas com o estoque de conhecimento comum (e informações falsas comuns) presumivelmente compartilhados pelo leitor e por ele próprio. Mas num poema, ou numa prosa contínua, o escritor pode estabelecer um novo padrão de pressuposições para os usos literais de expressões chave, anterior ao uso delas como meio para suas metáforas. (Um autor pode fazer muito para suprimir pressuposições indesejáveis da palavra “contrato”, através de uma explícita discussão de seu significado pretendido, antes de prosseguir no desenvolvimento de uma teoria contratual da soberania. Ou um naturalista que realmente conheça de lobos pode nos contar muito acerca deles a ponto de *sua* descrição de homem como um lobo divergir notoriamente do estoque de usos daquela figura). As

metáforas podem ser auxiliadas por sistemas de pressuposições especialmente construídos, tanto quanto por lugares-comuns aceitos; elas podem ser feitas por métrica e não precisam ser batidas.

Foi uma simplificação, novamente, falar como se o sistema de pressuposição da expressão metafórica permanecesse inalterado pelo enunciado metafórico. A natureza da aplicação pretendida ajuda a determinar o caráter do sistema a ser aplicado (como se as estrelas pudessem determinar parcialmente o caráter do vidro através do qual nós as olhamos). Se chamar um homem de lobo é colocá-lo numa posição especial, não podemos nos esquecer de que a metáfora faz o lobo parecer mais humano do que de outro modo seria.

Espero que complicações como essas possam ser acomodadas dentro do esboço da “perspectiva interacionista” que tentei apresentar.

## **VI**

Uma vez que fiz bastante uso de exemplos e ilustrações, seria bom apresentar explicitamente (e por meio de sumário) alguns dos aspectos principais pelos quais a perspectiva “interacionista” recomendada difere de uma perspectiva “substitutivista” ou “comparativista”.

Na forma que a venho expondo, a “perspectiva interacionista” está comprometida com as sete afirmações seguintes:

- (1) Um enunciado metafórico tem *dois* objetos distintos – um objeto “principal” e outro “auxiliar”.<sup>64</sup>
- (2) Esses objetos são muitas vezes melhor considerados como “sistemas de coisas”, do que propriamente “coisas”.

---

<sup>64</sup> Essa observação foi feita diversas vezes. *E.g.*: “Como em expressões metafóricas, que é uma grande excelência em estilo, quando é usada com propriedade, para lhe dar duas ideias numa só” (Samuel Johnson, citado por Richard, *op. cit.*, p.93). A escolha de etiquetas para o “objeto” é problemática. Veja a “Nota sobre terminologia” anexada a este artigo.

- (3) A metáfora funciona pela aplicação, no objeto principal, de um sistema de “pressuposições associados” características do objeto auxiliar.
- (4) Essas pressuposições consistem comumente de “lugares-comuns” sobre o objeto auxiliar, mas podem, em casos apropriados, consistir de pressuposições desviantes estabelecidas *ad doc* pelo escritor.
- (5) A metáfora seleciona, enfatiza, suprime e organiza características do objeto principal através de enunciados pressupostos sobre ele que normalmente se aplicam ao objeto auxiliar.
- (6) Isso envolve mudanças no significado das palavras que pertencem à mesma família ou sistema que as expressões metafóricas; e algumas dessas mudanças, ainda que nem todas, podem ser transferências metafóricas. (As metáforas subordinadas têm, contudo, de ser lidas menos “enfaticamente”.)
- (7) Não há, no geral, “base” simples para as mudanças necessárias de significado – nenhuma razão abrangente do porquê de algumas metáforas funcionarem e outras não.

Veremos, após análise, que o ponto (1) é incompatível com as formas mais simples de uma “perspectiva substitutivista”, e que o ponto (7) é formalmente incompatível com a “perspectiva comparativista”; ao passo que os pontos remanescentes fornecem razões para considerar a “perspectiva comparativista” como inadequada.

É fácil, porém, exagerar nos conflitos entre essas três perspectivas. Se insistíssemos que somente exemplos que satisfizessem todas as sete afirmações listadas acima deveriam contar como metáforas “genuínas”, restringiríamos o uso correto da

palavra “metáfora” a um número de casos muito pequeno. Isso seria defender uma definição persuasiva de “metáfora” que tenderia a fazer todas as metáforas interessantemente complexas.<sup>65</sup> E tal desvio dos usos correntes da palavra “metáfora” deixar-nos-ia sem uma etiqueta conveniente para os casos mais triviais. Ora, apenas em tais casos triviais é que as perspectivas “substitutivista” e “comparativista” algumas vezes parecem mais próximas de estarem corretas do que “perspectivas interacionista”. Poderíamos resolver a questão *classificando* as metáforas como instâncias de substituição, comparação ou interação. Somente o último tipo é importante na filosofia.

Pois as metáforas da substituição e da comparação podem ser trocadas por traduções literais (com possível exceção no caso de catacrese) – sacrificando um pouco do charme, da vivacidade ou da sutileza do original, embora sem perda do conteúdo *cognitivo*. Mas as “metáforas da interação” não são esgotáveis. Seu modo de operação requer que o leitor use um sistema de pressuposições (um sistema de “lugares-comuns” – ou um sistema especial estabelecido para uma finalidade em mãos) como um meio para selecionar, enfatizar e organizar relações em um campo diferente. Esse uso de um “objeto auxiliar” para favorecer a compreensão do “objeto principal” é uma operação *intelectual* distintiva (ainda que bastante familiar através de nossas experiências de aprendizagem do que quer que seja), demandando atenção simultânea de ambos os objetos, mas não redutível a qualquer *comparação* entre os dois.

Suponha que tentemos enunciar o conteúdo cognitivo de uma metáfora de interação em “linguagem comum”. Até certo ponto, podemos ter sucesso em enunciar um número de relações relevantes

---

<sup>65</sup> Tendo a concordar com a posição de Empson de que “O termo [‘metáfora’] melhor corresponderia àquilo que os próprios falantes percebem como sendo um uso rico, ou sugestivo, ou persuasivo de uma palavra, ao invés da inclusão de usos como a *perna* de uma mesa” (*The Structure of Complex Words*, p. 333). Mas há também o perigo oposto de tornar as metáforas demasiado importantes por definição, e, por conseguinte, restringir excessivamente a nossa visão do objeto.

entre os dois objetos (embora da perspectiva da extensão do significado acompanhando a mudança no sistema de pressuposições do objeto auxiliar não se deva esperar muito da paráfrase literal). Mas o conjunto de enunciados literais assim obtidos não terá o mesmo poder de informar e iluminar como o original. Por um motivo, as pressuposições, previamente deixadas para que um leitor apropriado – sensível às suas relativas prioridades e graus de importância – as infira por si próprio, são agora apresentadas explicitamente como tendo igual importância. A paráfrase literal inevitavelmente diz muito – e com a ênfase errada. Um dos pontos que me convém mais salientar é o de que a perda em tais casos é uma perda em conteúdo *cognitivo*; o ponto fraco da paráfrase literal não é que ela possa ser desagradavelmente prolixa ou tediosamente explícita – ou deficiente em qualidades de estilo; ela falha em ser uma tradução porque falha em fornecer a *compreensão* que a metáfora deu.

Mas a “explicação”, ou elaboração das bases da metáfora, ainda que não seja considerada como uma substituta cognitiva adequada para o original, pode ser extremamente valiosa. Uma metáfora poderosa não será mais prejudicada por tal investigação do que uma obra musical pela análise de sua estrutura harmônica e melódica. Sem dúvida as metáforas são perigosas – e talvez especialmente na filosofia. Mas uma proibição contra seus usos seria uma restrição intencional e prejudicial à nossa capacidade de investigação.<sup>66</sup>

---

<sup>66</sup> (*Uma nota sobre a terminologia*): Para as metáforas que se ajustam às perspectivas substitutivista ou comparativista, os fatores que precisam ser distinguidos são: - (i) alguma palavra ou expressão  $E$ , (ii) que ocorre em algum “enquadramento” verbal  $F$ , de modo que (iii)  $F(E)$  é o enunciado metafórico em questão; (iv) o significado  $m'(E)$  que  $E$  tem em  $F(E)$ ; (v) que é o mesmo que o significado literal,  $m(X)$ , de algum sinônimo literal  $X$ . Um vocabulário técnico suficiente seria: “expressão metafórica” (para  $E$ ), “enunciado metafórico” (para  $F(E)$ ), “significado metafórico” (para  $m'$ ) e “significado literal” (para  $m$ ).

Onde a perspectiva da interação é apropriada, a situação é mais complicada. Podemos também precisar referir (vi) ao objeto principal de  $F(E)$ , digamos  $P$  (grosso modo, sobre aquilo que o enunciado “realmente” é), (vii) ao objeto auxiliar,  $S$  (sobre o que  $F(E)$  seria se interpretado literalmente); (viii) ao sistema de pressuposições relevante,  $I$ , conectado a  $S$ , e (ix) ao sistema de atribuições resultante,  $A$ , afirmado de  $P$ . Temos de aceitar toda essa complexidade caso concordemos que o significado de  $E$  em sua composição  $F$  depende da transformação de  $I$  em  $A$  pelo uso da linguagem, normalmente aplicada a  $S$ , aplicada ao invés disso a  $P$ .

Richards sugeriu usar as palavras “teor” e “veículo” para os dois “pensamentos” que, do seu ponto de vista, são “ativos conjuntamente” (para “as duas ideias que a metáfora, em sua forma mais

## Conhecimento *a priori*\*

Albert Casullo

Tradução de Luiz Helvécio Marques Segundo<sup>67</sup> e Alexandre Meyer Luz\*\*  
Revisor: Eduardo Benkendorf

### 1- Introdução<sup>68</sup>

A proeminência do *a priori* na epistemologia tradicional se deve em grande medida à influência de Immanuel Kant. Na Introdução da *Crítica da Razão Pura*,<sup>i</sup> ele introduz um enquadramento conceitual que envolve três distinções: (1) a distinção *epistêmica* entre conhecimento *a priori* e *a posteriori*; (2) a distinção *metafísica* entre proposições necessárias e contingentes; e (3) a distinção *semântica* entre proposições analíticas e sintéticas. Dentro desse enquadramento, Kant faz quatro perguntas:

- 1- O que é conhecimento *a priori*?
- 2- Há conhecimento *a priori*?

---

simples, nos dá”, *Op. cit.*, p. 96, itálico meu) e insiste que reservemos “a palavra ‘metáfora’ para toda a dupla unidade” (Ib.). Porém, essa imagem de duas *ideias* influenciando uma a outra é uma ficção inconveniente. E é sugestivo que Richards rapidamente deslize ao falar de “teor” e “veículo” como “coisas” (e.g. na p. 118). O “veículo” de Richards vacila em referência entre a expressão metafórica (*E*), o objeto auxiliar (*S*) e o sistema de pressuposições conectado (*I*). É menos claro o que seu “teor” significa: às vezes representa o objeto principal (*P*), às vezes as pressuposições conectadas ao objeto (que não simbolizei anteriormente), às vezes, apesar das intenções do próprio Richard, o significado *resultante* (ou como podemos dizer o “significado pleno”) de *E* em seu contexto *F(E)*.

Provavelmente não há esperança de se obter uma terminologia consensual dado que os autores que lidam com o tema ainda estão em grande desacordo.

\* “A priori knowledge”, in *The Oxford Handbook of Epistemology*, org. Paul K. Moser. Oxford University Press, pp. 95-143.

<sup>67</sup> Bolsista CAPES, doutorando pela UFSC.

\*\* Alexandre Meywe Luz é Professor da UFSC. E o revisor Eduardo Benkendorf é mestrando da UFSC.

<sup>68</sup> As traduções das notas deste texto encontram-se no final deste volume.

- 3- Qual a relação entre o *a priori* e o necessário?
- 4- Há conhecimento sintético *a priori*?

Essas perguntas ainda continuam centrais no debate contemporâneo.

Kant sustenta que o conhecimento *a priori* é “absolutamente independente de toda experiência”.<sup>ii</sup> Essa caracterização não é completamente clara, uma vez que ele concede que tal conhecimento possa depender da experiência em *alguns* aspectos. Por exemplo, de acordo com Kant, sabemos *a priori* que toda alteração tem sua causa, a despeito do fato de o conceito de alteração ser derivado da experiência. Contudo, ele não é explícito sobre em que aspecto tal conhecimento tem de ser independente da experiência.

Uma vez que Kant não oferece uma análise completamente articulada do conceito de conhecimento *a priori*, ele não está em posição de argumentar diretamente a favor de sua existência mostrando que algum conhecimento satisfaz as condições em sua análise. Ao invés, ele responde a segunda pergunta *indiretamente* procurando *critérios* do *a priori*. Os critérios fornecem condições suficientes para o conhecimento *a priori* que não estão incluídos na análise do conceito. Kant oferece dois critérios, a necessidade e a estrita universalidade, que ele diz serem inseparáveis um do outro. Os principais argumentos de Kant a favor do *a priori* recorrem ao primeiro. Por exemplo, ele argumenta que uma vez que as proposições matemáticas são necessárias e uma vez que conhecemos algumas delas, segue-se que temos conhecimento *a priori*.

A tese de Kant de que a necessidade é um critério do *a priori* compromete com a seguinte tese sobre a relação entre o *a priori* e o necessário:

(K<sub>1</sub>) Todo conhecimento de proposições necessárias é *a priori*.

Ele também parece sustentar

(K<sub>2</sub>) Todas as proposições conhecidas *a priori* são necessárias.

Embora Kant seja geralmente apresentado como sustentando que as categorias do *a priori* e do necessário são coextensivas, a conjunção de K<sub>1</sub> e K<sub>2</sub> não apoia essa atribuição, uma vez que não implica que todas as proposições necessárias sejam conhecidas ou conhecíveis *a priori*. K<sub>1</sub> liga a terceira pergunta à segunda, uma vez que fornece a premissa central do único argumento de Kant a favor da existência do

conhecimento *a priori*. Nem  $K_1$  e nem  $K_2$  se ligam diretamente à primeira pergunta, uma vez que Kant não afirma que a necessidade é um constituinte do conceito de conhecimento *a priori*.

Kant sustenta que todas as proposições da forma “Todo A é B” são ou analíticas ou sintéticas: analíticas se o predicado estiver contido no sujeito, e sintéticas se não. Utilizando essa distinção, ele argumenta que

(A<sub>1</sub>) Todo conhecimento de proposições analíticas é *a priori*, e

(A<sub>2</sub>) Algumas proposições conhecidas *a priori* são sintéticas.

Em apoio a (A<sub>2</sub>) Kant novamente recorre à matemática, argumentando que os termos predicado de “ $7+5=12$ ” e “A distância mais curta entre dois pontos é a linha reta” não estão ocultamente contidos em seus respectivos sujeitos. Nem (A<sub>1</sub>) ou (A<sub>2</sub>) têm uma ligação direta com as duas primeiras perguntas, uma vez que Kant não afirma que a analiticidade é um constituinte do conceito de conhecimento *a priori* e não as invoca como premissa em seus argumentos a favor da existência do conhecimento *a priori*. Kant considera (A<sub>2</sub>) como significativa porque ela prepara o terreno para o seu principal projeto epistêmico, que é explicar como tal conhecimento é possível. O projeto, no entanto, pressupõe que o conhecimento *a priori* de proposições analíticas e o conhecimento *a priori* de proposições sintéticas são fundamentalmente diferentes, uma pressuposição que Kant não defende explicitamente.

A discussão contemporânea do *a priori* gira em torno das quatro perguntas de Kant. Philip Kitcher oferece uma articulação da caracterização de Kant do conhecimento *a priori*.<sup>iii</sup> Ele sustenta que uma crença é justificada independentemente da experiência se e somente se for irrevogável por evidências da experiência. Baseado no trabalho de W. V. Quine<sup>iv</sup> e Hilary Putnam,<sup>v</sup> que defendem que nenhuma crença está imune à revisão frente a experiência recalcitrante, Kitcher conclui que o conhecimento matemático não é *a priori*. Em resposta, diversos teóricos rejeitam a tese de Kitcher de que o conceito de conhecimento *a priori* envolve uma condição de irrevogabilidade e oferecem propostas alternativas.<sup>vi</sup>

Paul Benacerraf ataca a estratégia de Kant de argumentar que há conhecimento *a priori*.<sup>vii</sup> Benacerraf sustenta que alguém sabe uma afirmação somente se estiver causalmente conectado às entidades referidas por suas condições de verdade e que as condições de verdade das afirmações matemáticas fazem referência

a entidades abstratas. A tradição, no entanto, sustenta que as condições de verdade de todas as verdades necessárias fazem referência a entidades abstratas.<sup>viii</sup> Assim, o argumento de Benacerraf, se cogente, estabelece que o conhecimento de verdades necessárias não é possível. O argumento desencadeou uma série de investigações sobre a questão geral do conhecimento de entidades abstratas e sobre a questão mais específica de papel propriamente dito das condições causais numa teoria do conhecimento plausível.<sup>ix</sup>

Os resultados metafísicos e semânticos de Saul Kripke renovaram o interesse na abordagem de Kant da relação entre o *a priori* e o necessário.<sup>x</sup> Kripke argumenta convincentemente que o conceito de conhecimento *a priori* é epistêmico, ao passo que o conceito de verdade necessária é metafísico e, por conseguinte, não se pode presumir sem um argumento que sejam coextensivos. Ademais, ele sustenta que há proposições necessárias conhecidas *a posteriori* e proposições contingentes conhecidas *a priori*. Os argumentos de Kripke geraram uma literatura ampla voltada para seus exemplos particulares tanto quanto para a questão mais geral da relação entre o *a priori* e o necessário.<sup>xi</sup>

A tese de Kant de que há conhecimento sintético *a priori* dominou a discussão do *a priori* nos últimos cinquenta anos. A controvérsia é abastecida por duas reações relacionadas a A<sub>2</sub>. A primeira se deve aos proponentes do empirismo lógico que argumentam que apenas as proposições analíticas são conhecíveis *a priori*.<sup>xii</sup> A segunda se deve a W. V. Quine que rejeita essa doutrina central do empirismo lógico rejeitando a cogência da distinção analítico-sintético.<sup>xiii</sup> Embora a conclusão de Quine seja semântica, é largamente considerada como tendo implicações mais amplas para a existência do conhecimento *a priori*. Os teóricos estão reavaliando tanto a cogência dos argumentos de Quine contra a distinção, quanto, o que é mais importante, a ligação, se houver alguma, da rejeição dessa distinção com a questão de se há conhecimento *a priori*.<sup>xiv</sup>

O escopo das questões levantadas pelas quatro perguntas de Kant é enorme, cobrindo grande parte das áreas centrais da investigação filosófica contemporânea. O foco deste ensaio é mais limitado. Meu propósito é tratar da questão de saber se existe conhecimento *a priori*. Uma vez que não se pode determinar se tal conhecimento existe sem saber o que tal conhecimento é, começo por fornecer uma análise do conceito de conhecimento *a priori*. Utilizo essa análise para mostrar que os argumentos tradicionais, tanto a favor quanto contra o *a priori*, não são convincentes. Concluo

oferecendo uma estratégia alternativa para se defender a existência do conhecimento *a priori*. Embora as questões sobre a relação entre o *a priori* e os conceitos não epistêmicos de necessidade e analiticidade não sejam meus alvos primários, trato-os na medida em que são relevantes para a análise do conceito de conhecimento *a priori* ou para determinar se tal conhecimento existe.

## **2- O conceito de conhecimento *a priori***

Há duas vias para se analisar o conceito de conhecimento *a priori*. A primeira, que é *reductiva*, analisa-o em termos do conceito de justificação *a priori*. De acordo com essa via, S sabe *a priori* que *p* apenas no caso de (a) a crença de S de que *p* ser justificada *a priori*, e (b) as outras condições para o conhecimento forem satisfeitas. O alvo principal dessa análise é o conceito de *justificação a priori*. A segunda, que é *não reductiva*, fornece uma análise do conceito que não inclui condições que envolvem o conceito de *a priori*. O alvo principal dessa análise é o conceito de *conhecimento a priori*.

As condições do conhecimento *a priori* propostas pelos epistemólogos contemporâneos têm sua inspiração em Kant. Enquadram-se em duas amplas categorias: *epistêmica* e *não epistêmica*. Há três tipos de condições epistêmicas. A primeira impõe condições sobre a *fonte* de justificação; a segunda impõe condições acerca da revogabilidade da justificação; e a terceira recorre à *força* da justificação. As condições da fonte e da irrevogabilidade são inspiradas pela caracterização de Kant do conhecimento *a priori* como independente de toda a experiência. As condições de força derivam da associação frequente que Kant fazia da *certeza* com o *a priori*.<sup>xv</sup> As duas condições não epistêmicas têm desempenhado um papel proeminente na análise do *a priori*. Alguns teóricos incluem a *necessidade*, que Kant usou como um critério do conhecimento *a priori*, na análise do conceito. Outros, reagindo contra Kant, negam que o conhecimento sintético *a priori* seja possível, e incluem a *analiticidade* na análise do conceito.

As análises do conceito de conhecimento *a priori* se enquadram em três categorias. As análises epistêmicas puras incluem apenas condições epistêmicas. As análises epistêmicas impuras incluem, além disso, algumas condições não epistêmicas. As análises não epistêmicas consistem em apenas condições não epistêmicas. Voltar-nos-emos primeiro às análises não epistêmicas.<sup>xvi</sup>

### **2.1 Análises não epistêmicas**

As análises não epistêmicas sustentam que ou a necessidade ou a analiticidade fornece as condições necessárias e suficientes para o conhecimento *a priori*. Há uma razão geral para considerá-las com suspeita. O *analysandum* em questão é epistêmico. É um tipo de justificação. Uma análise informativa, no entanto, deveria realçar aquilo que é distintivo acerca de tal justificação. Uma análise em termos de necessidade ou analiticidade realça aquilo que é distintivo sobre as proposições assim justificadas ao invés da própria justificação. Portanto, ela não será informativa.

As análises não epistêmicas tipicamente envolvem a expressão “verdade *a priori*” ou “proposição *a priori*”. Isso introduz uma complicação, uma vez que essas expressões não tem um significado fixo. Muitos autores as introduzem como abreviação de “verdade (proposição) que pode ser conhecida *a priori*”.<sup>xvii</sup> De acordo com esse uso, “*a priori*” permanece um predicado epistêmico, um predicado cuja aplicação primária é ao conhecimento ou à justificação ao invés da verdade. Alguns, no entanto, usam a expressão aplicando-a primariamente a verdades. Assim, por exemplo, Anthony Quinton sustenta que

“*A priori*” significa ou, amplamente, “não empírico” ou, estritamente, seguindo Kant, “necessário”.<sup>xviii</sup>

O uso de Quinton do termo “não empírico” sugere que, de acordo com essa perspectiva, a aplicação primária de “*a priori*” é epistêmica, uma vez que “empírico” é tipicamente um predicado epistêmico cuja aplicação primária é a itens do conhecimento ou da justificação. Mas o uso de Quinton de “não empírico” é também enganadora. Ele sustenta explicitamente que o termo *não* é um predicado *epistêmico*:

A ideia de empírico é um desenvolvimento do contingente. Visa explicar como uma afirmação pode dever sua verdade a algo mais, que condições esse algo mais tem de satisfazer se for para conferir verdade a uma afirmação.<sup>xix</sup>

Para Quinton, “empírico” tem sua aplicação primária às condições de verdade ou à fonte de verdade. Embora ele caracterize seu propósito como uma defesa da tese de que todas as afirmações *a priori* sejam analíticas, ele continua a sustentar que “o conteúdo essencial da tese é o de que todas as verdades *necessárias* são analíticas”.<sup>xx</sup> Para Quinton, os sentidos estritos de “*a priori*”, “necessário” e “analítico” são idênticos em significado.

O desfecho é que o termo “*a priori*” é ambíguo. É um predicado cuja aplicação primária ou é para tipos de justificação ou bases para a verdade. Portanto, uma análise epistêmica do *a priori* pode ter ambos como alvo. Se seu alvo for o último, então a análise não está aberta ao meu argumento inicial, pois ela não se direciona ao conceito epistêmico. É direcionada ao conceito metafísico pertencente às condições de verdade. O nosso interesse, no entanto, é com a análise do conceito epistêmico.

Há análises não epistêmicas do conceito epistêmico? R. G. Swinburne defende as seguintes teses:

(S<sub>1</sub>) Uma proposição é *a priori* se, e somente se, é necessária e pode ser conhecida como sendo necessária.

(S<sub>2</sub>) Uma proposição é *a priori* se, e somente se, é analítica e pode ser conhecida como sendo analítica.<sup>xxi</sup>

Ao contrário de Quinton, Swinburne sustenta que o termo “*a priori*” tem sua aplicação primária ao conhecimento. Uma proposição *a priori* é aquela que pode ser conhecida *a priori*. Portanto, parece que ele está a propor uma análise não epistêmica para um conceito epistêmico.

Um exame mais cuidadoso revela que Swinburne não está a propor ou S<sub>1</sub> ou S<sub>2</sub> como uma análise do conceito de conhecimento *a priori*. Ao invés, ele defende a análise de Kant do conhecimento *a priori* como absolutamente independente de toda a experiência, sustentando que Kant quis dizer com isso “conhecimento que nos é dado através da experiência, mas que não tem contribuição da experiência”.<sup>xxii</sup> O interesse de Swinburne, no entanto, é com a questão de como *reconhecer* tal conhecimento. Ele propõe que S<sub>1</sub> captura a resposta de Kant a essa questão.

O resultado aqui é que nem toda bicondicional da forma:

(AP) Uma proposição é *a priori* se e somente se...,

em que “*a priori*” é um predicado epistêmico, é uma análise do conceito epistêmico designado por esse termo. As bicondicionais dessa forma podem ser propostas em resposta a diferentes questões. A pergunta de Swinburne

(Q1) Como identificamos os itens que satisfazem algumas análises do conhecimento *a priori*?

é diferente da pergunta

(Q2) Qual é a análise do conhecimento *a priori*?

Uma resposta a Q<sub>1</sub> pressupõe, ao invés de fornecer, uma resposta a Q<sub>2</sub>.

As aparentes análises não epistêmicas do *a priori* têm de ser examinadas ao longo de duas dimensões. Qual é o alvo da análise? Que pergunta está sendo feita sobre o alvo? O meu alvo é o conceito de justificação *a priori* como oposto ao conceito de verdade *a priori*. A minha tese é a de que uma análise não epistêmica da primeira não pode ser bem-sucedida. Uma vez que o conceito é fundamentalmente epistêmico, qualquer análise satisfatória tem de identificar a característica saliente de tal justificação. Essa tese não implica que não haja características não epistêmicas comuns a todas e apenas aquelas proposições justificáveis *a priori*. Implica apenas que não é em virtude de possuir tais características que tais proposições são justificáveis *a priori*.

A ideia de que uma análise adequada do conceito de justificação *a priori* tem de incluir uma condição epistêmica deixa em aberto a possibilidade de que também ela inclua alguma condição não epistêmica. Voltar-nos-emos para a questão de saber se alguma condição não epistêmica é necessária para a justificação *a priori*. O meu foco é sobre as condições envolvendo o conceito de necessidade, uma vez que são as mais comuns.

## **2.2. Análises epistêmicas impuras**

As análises do conceito de justificação *a priori* que incluem o conceito de necessidade enquadram-se em duas categorias. Algumas incluem a necessidade como uma componente de uma condição epistêmica. Outros a incluem como uma condição independente. Laurence Bonjour oferece a seguinte versão da concepção racionalista tradicional de *a priori*:

uma proposição é justificada *a priori* quando e apenas quando o agente é capaz, ou diretamente ou via alguma série de passos individualmente evidentes, de “ver” intuitivamente ou apreender que a sua verdade é uma característica invariante de todos os mundos possíveis, que não há mundo possível em que ela seja falsa.<sup>xxiii</sup>

A concepção consiste numa única condição com duas componentes: a fonte de justificação *a priori*, a apreensão intuitiva, e o conteúdo de tais apreensões, as verdades necessárias.

É complicado avaliar as implicações da análise, uma vez que envolve um uso metafórico do termo “ver”. Tomada literalmente, a locução “S vê que *p*” (por exemplo, que há um coelho no jardim) implica “S acredita que *p*”.<sup>xxiv</sup> Presumindo que o uso metafórico de “ver” preserva as características lógicas da locução literal, “S intuitivamente ‘vê’ que *p* é verdadeira em todos os mundos possíveis” implica “S acredita que *p* é verdadeira em todos os mundos possíveis”. Por conseguinte, de acordo com a concepção racionalista tradicional, “A crença de S de que *p* é justificada *a priori*” implica “S acredita que necessariamente *p*”.

A concepção enfrenta três objeções. A primeira é devido à *deficiência conceitual*. Muitos, incluindo alguns matemáticos, não estão familiarizados com a distinção metafísica entre proposições necessárias e contingentes. Considere um matemático, S, que acredita num teorema T com base numa demonstração geralmente aceita. A crença de S de que T é justificada. Suponha que S não possui o conceito de necessidade e, como consequência, não acredita que necessariamente T. É implausível sustentar que a crença de S de que T não está justificada *a priori* simplesmente porque S não possui o conceito que sequer é um constituinte do conteúdo da crença de S.

A primeira objeção pode ser evitada enfraquecendo-se a concepção por exigir que S acredite que necessariamente *p* desde que S possua o conceito de necessidade. Restam duas objeções. A primeira se deve ao *ceticismo modal*. Dentre os filósofos familiarizados com o conceito de verdade necessária, alguns negam (suponhamos que erroneamente) sua cogência. Como consequência, evitam crenças modais, tais como que necessariamente  $2+2=4$ . Mas é implausível sustentar que nenhuma de suas crenças matemáticas sejam justificadas *a priori* apenas com base no fato de que eles tenham uma metafísica errônea. Segundo, a concepção está aberta ao regresso. Suponha que S acredite que necessariamente *p*. Tem a crença de S de que necessariamente *p* de ser justificada ou não? Se não, então é difícil ver por que é uma condição necessária para a justificação *a priori* da crença de S de que *p*. Se sim, então presumivelmente a sua justificação tem de ser *a priori*. Mas para que sua justificação seja *a priori* S tem de ver que necessariamente *p* é verdadeira em todos os mundos possíveis, o que, por sua vez, requer acreditar que necessariamente necessariamente *p*. Mas agora há a

ameaça de regresso, uma vez que podemos novamente fazer a pergunta: Tem a crença de S de que necessariamente necessariamente  $p$  de ser justificado ou não?

Chisholm fornece uma análise do *a priori* em que a necessidade é oferecida como uma condição necessária independente. Considere as seguintes definições:

$D_1$ :  $h$  é um axioma  $=_{df}$   $h$  é necessariamente tal que (i) é verdadeira e (ii) para todo S, se S aceita  $h$ , então  $h$  é certo para S.<sup>xxv</sup>

$D_2$ :  $h$  é axiomático para S  $=_{df}$  (i)  $h$  é um axioma e (ii) S aceita  $h$ .<sup>xxvi</sup>

$D_3$ :  $h$  é conhecido *a priori* por S  $=_{df}$  Há um  $e$  tal que (i)  $e$  é axiomático para S, (ii) a proposição,  $e$  implica  $h$ , é axiomática para S, e (iii) S aceita  $h$ .<sup>xxvii</sup>

O conhecimento *a priori* se restringe aos axiomas e às suas consequências axiomáticas. Para ser um axioma uma proposição tem de satisfazer duas condições *independentes*: tem de ser necessariamente verdadeira e certa para qualquer um que a aceite. Essas condições são independentes, pois nenhuma acarreta a outra. Uma vez que os axiomas são verdades necessárias e as consequências axiomáticas dos axiomas se seguem necessariamente dos axiomas, todo conhecimento *a priori* é conhecimento de verdades necessárias.

Que apoio Chisholm oferece para a sua análise? Ele inicia a sua discussão do *a priori* com as seguintes observações:

Há proposições que são necessariamente verdadeiras e tais que, uma vez entendidas, vê-se que são verdadeiras. Tais proposições têm sido tradicionalmente chamadas de *a priori*. Leibniz observa, “Encontrarás centenas de lugares em que os filósofos escolásticos disseram que essas proposições são evidentes, por seus termos, tão logo que entendidas”.<sup>xxviii</sup>

Essa passagem envolve duas afirmações: (1) algumas proposições têm tanto a propriedade metafísica de serem necessariamente verdadeiras quanto a propriedade epistêmica de ser tal que, se alguém as entende, então vê que são verdadeiras; e (2) tais proposições têm sido tradicionalmente chamadas de *a priori*. A citação de Leibniz, que invoca a autoridade dos escolásticos,

menciona apenas a segunda. Não há menção à propriedade metafísica. Por conseguinte, se o argumento de Chisholm estiver baseado no precedente histórico, sua análise deveria estar em termos de propriedades epistêmicas apenas.

A inclusão que Chisholm faz da condição metafísica na análise não é apenas desmotivada, mas também tem consequências indesejáveis. Primeiro, ou a análise é incompleta ou exclui a possibilidade de crenças *falsas* justificadas *a priori*.  $D_3$  fornece uma análise do conhecimento *a priori*. Se as condições de Chisholm para o conhecimento *a priori* são também condições para a *justificação a priori*, então a justificação *a priori* garante a verdade. Se não são, então a sua abordagem do *a priori* está incompleta. Segundo, a análise exclui por estipulação a possibilidade do conhecimento *a priori* de verdades contingentes. Porém, Kripke e Kitcher sustentam haver tal conhecimento.<sup>xxix</sup> Terceiro, a análise deixa de fora a possibilidade do conhecimento *a posteriori* de axiomas. Suponha que S aceite o axioma A com base no testemunho. Ou A é certo para S ou não é. Se é, então A é axiomático para S e S sabe *a priori* que A. Se não é, então A não é um axioma, pois não satisfaz a condição ii em  $D_1$ .

A análise de Chisholm do conhecimento *a priori* axiomático, ou não inferencial, também inclui uma condição epistêmica: a certeza.<sup>xxx</sup> Essa condição leva à consequência implausível de que é impossível que (1) S saiba axiomáticamente que  $1+1=2$ ; (2) que S saiba axiomáticamente que  $7+5=12$ ; (3) que a primeira crença é *mais* justificada do que a última. Chisholm, porém, oferece uma justificação para excluir a possibilidade de se diferenciar graus de justificação não inferencial *a priori*. Além do mais, isso também acarreta que se S sabe axiomáticamente que  $p$ , e que S sabe *a posteriori* que  $p$ , então a primeira crença é *mais* justificada do que a última. Não é óbvio, no entanto, que a crença de alguém que  $7+5=12$  seja *mais* justificada do que a sua crença de que existe.

### **2.3. Análises epistêmicas puras**

As análises epistêmicas puras mais comuns da justificação *a priori* são em termos de *fonte* de justificação. A divisão maior é entre análises negativas e positivas. As primeiras especificam as fontes *incompatíveis* com a justificação *a priori*, as últimas especificam as fontes que *fornece* tal justificação. A análise negativa mais familiar é

(N<sub>1</sub>) A crença de S de que  $p$  é justificada *a priori* se e somente se a justificação da crença de S que  $p$  não depende da experiência.

Os críticos das análises negativas sustentam que elas não são suficientemente informativas.<sup>xxxii</sup> Na melhor das hipóteses, especificam o que a justificação *a priori* não é ao invés do que é. O problema pode ser evitado optando-se por uma análise positiva com a forma

(P<sub>1</sub>) A crença de S de que  $p$  é justificada *a priori* se, e somente se, a crença de S de que  $p$  é justificada por  $\varphi$ ,

onde “ $\varphi$ ” designa alguma fonte específica de justificação. Por exemplo, de acordo com Panayot Butchvarov, designa descobrir a falsidade de uma crença impensável em quaisquer circunstâncias.<sup>xxxii</sup> Mas, de acordo com Laurence Bonjour, designa um aparente *insight* racional nas características necessárias da realidade.<sup>xxxiii</sup>

Uma análise do conceito de justificação *a priori* que enumere as fontes de tal justificação é bastante teoricamente dependente. Não se pode rejeitar a fonte de justificação *a priori* oferecida por tal análise sem se rejeitar a existência da justificação *a priori*. Por exemplo, dada a análise de Butchvarov, não se pode rejeitar (como Bonjour o faz) a tese de que descobrir a falsidade de uma crença impensável em quaisquer circunstâncias é a fonte da justificação *a priori*, sem rejeitar a existência do *a priori*. De acordo com essa análise, a justificação *a priori* é e a justificação baseada em tais descobertas. Seria possível para os proponentes do *a priori*, no entanto, discordarem sobre a fonte da justificação *a priori* sem com isso discordar da existência de tal justificação. Além do mais, ainda que alguma versão particular da análise positiva seja extensionalmente adequada, a análise não é informativa. Diz-nos que  $\varphi$  é uma fonte de justificação *a priori*, mas não nos dá indicação do porquê  $\varphi$  é uma fonte *a priori*. Não realça as características em virtude das quais  $\varphi$  se qualifica como uma fonte *a priori*.

Há uma análise positiva *geral* do *a priori* que evita o problema da dependência teórica:

(P<sub>2</sub>) A crença de S de que  $p$  é justificada *a priori* se, e somente se, a crença de S de que  $p$  é justificada em virtude de alguma fonte não empírica.

P<sub>2</sub> permite que os proponentes do *a priori* concordem que haja justificação *a priori* a despeito do desacordo acerca de sua fonte. Ademais, identifica a característica das fontes de justificação em virtude da qual elas se qualificam como *a priori*.

(C<sub>1</sub>) A justificação de S para a crença de que *p* não depende da experiência

não especifica o *aspecto* no qual a justificação de S tem de ser independente da experiência. Há, no entanto, duas possibilidades: a fonte de *justificação* para a crença de S de que *p* e a fonte dos *revogados* potenciais para a justificação de S. Alguns defendem que C<sub>1</sub> é equivalente a

(C<sub>2</sub>) A crença de S que *p* é justificada não experiencialmente.

Outros defendem que é equivalente à conjunção de C<sub>2</sub> e

(C<sub>3</sub>) A crença justificada de S de que *p* não pode ser revogada pela experiência.

Evidentemente, se a crença de S de que *p* é experiencialmente justificada, então a justificação de S depende da experiência. O que se pode dizer a favor de C<sub>3</sub>? Philip Kitcher argumenta que

se experiências alternativas pudessem colocar em causa o conhecimento de alguém, então haveria características da experiência atual de alguém que são relevantes ao conhecimento, designadamente, aquelas características cuja *falta* transformaria a experiência atual em experiência subversiva.<sup>xxxiv</sup>

De acordo com Kitcher, se evidências empíricas podem revogar a justificação de S para a crença de que *p*, então a justificação de S depende da *falta* dessas evidências empíricas.

A tese de Kitcher de que a justificação *a priori* é incompatível com revogados empíricos potenciais deveria ser distinguida da condição intimamente relacionada, embora mais forte, defendida por Hilary Putnam:

Há verdades *a priori*? Em outras palavras, há afirmações verdadeiras que (1) é racional aceitar [...], e (2) que subsequentemente nunca seria racional rejeitar não importa como o mundo venha a ser (epistemicamente)? De modo mais

simples, há afirmações cuja verdade não estaríamos justificados em negar em qualquer mundo *epistemicamente* possível?<sup>xxxv</sup>

De acordo com Putnam, a crença de S de que *p* é justificada *a priori* somente se

(C<sub>4</sub>) A crença de S de que *p* não puder ser revogada por quaisquer evidências.

C<sub>4</sub>, no entanto, não é uma condição plausível para a justificação *a priori*, uma vez que acarreta que se a crença de S de que *p* é revogável apenas por evidências *não empíricas*, então *não* é justificada *a priori*. Contudo, se a crença de S de que *p* é justificada apenas por evidências não empíricas, então não depende de modo algum da experiência. Portanto, C<sub>4</sub> separa o conceito de justificação *a priori* da ideia central de que tal justificação é independente da experiência.

Uma vez que Kitcher liga C<sub>3</sub> a C<sub>1</sub>, não se pode rejeitá-lo tão facilmente quanto C<sub>4</sub>. Ao invés, temos de distinguir duas versões diferentes da análise negativa:

(N<sub>2</sub>) A crença de S de que *p* é justificada *a priori* se, e somente se, a crença de S que *p* for justificada não empiricamente; e

(N<sub>3</sub>) A crença de S de que *p* é justificada *a priori* se, e somente se, a crença de S de que *p* for não empiricamente justificada e não puder ser revogada pela experiência.<sup>xxxvi</sup>

Uma vez que C<sub>2</sub> é equivalente a

(C<sub>5</sub>) A crença de S de que *p* é justificada por *alguma* fonte não empírica,

N<sub>2</sub> e P<sub>2</sub> são equivalentes. Portanto, ficamos com duas análises da justificação *a priori*. Meu propósito final é argumentar que N<sub>2</sub> é a análise superior.

#### **2.4. Um argumento a favor de N<sub>2</sub>**

N<sub>3</sub>, mas não N<sub>2</sub>, é incompatível com o critério de adequação amplamente aceito. Saul Kripke apresenta o ponto com se segue:

Algo pode pertencer ao reino de tais afirmações que *podem* ser conhecidas *a priori* e ainda assim pode ser conhecido por uma pessoa particular com base na experiência.<sup>xxxvii</sup>

Kitcher, repetindo esse ponto, sustenta que

Um apriorista perspicaz deveria admitir que as pessoas podem ter conhecimento empírico de proposições que podem ser conhecidas *a priori*.<sup>xxxviii</sup>

De acordo com o critério de adequação, uma análise do conceito de justificação *a priori* deveria conceder a seguinte possibilidade:

(CA) S sabe empiricamente que  $p$  e S pode saber *a priori* que  $p$ .

$N_3$ , no entanto, exclui essa possibilidade.

Antes de apresentar o argumento, um ponto precisa ser estabelecido.  $N_3$  *não* envolve uma condição forte. Não requer do conhecimento *a priori* um grau de justificação maior do que o minimamente exigido para o conhecimento em geral. Outro modo de se colocar o mesmo ponto é que  $N_3$  não requer do conhecimento *a priori* um grau de justificação maior do que o exigido para o conhecimento *a posteriori*. Estabeleçamos esse ponto explicitamente como a *Tese da Igualdade de Força*:

(IF) O grau de justificação minimamente suficiente para o conhecimento *a priori* se iguala ao grau de justificação minimamente suficiente para o conhecimento em geral.

Para manter o ponto explícito no decurso do argumento, chamemos uma crença justificada ao grau minimamente suficiente para o conhecimento de crença justificada<sub>c</sub>.

Voltemos ao argumento. Começemos por presumir que

(A) S sabe empiricamente alguma proposição matemática de que  $p$  e S pode saber *a priori* que  $p$ .

Da conjunta esquerda de A, segue-se que

(1) A crença de S de que  $p$  é empiricamente justificada<sub>c</sub>.

Diversas fontes empíricas supostamente justificam proposições matemáticas: (a) contar coleção de objetos, (b) ler manuais, (c)

consultar matemáticos, e (d) computar os resultados. Concedamos que cada uma delas pode justificar a crença matemática de que  $p$ . Cada uma dessas fontes é falível num aspecto importante. A justificação de cada uma garante que uma crença de que  $p$  é revogável por um revogador *refutante* empiricamente justificado: isto é, por uma crença empiricamente justificada de que não- $p$ . Suponha que a crença de S de que  $p$  é justificada pela contagem de uma coleção de objetos e pelo alcance de um resultado particular. É possível que S reconte a coleção e chegue a um resultado diferente. Se S assim o fizesse, a justificação original de S seria revogada por um revogador refutante empiricamente justificado. Suponha que a crença de S de que  $p$  é justificada por um manual (matemático, resultado computado) que diz que  $p$ . É possível que S encontre um manual diferente (matemático, resultado computado) que diz que não- $p$ . Em cada caso, se S o fizesse, a justificação original de S seria revogada por um revogador refutante empiricamente justificado. Por conseguinte, dado o caráter fiável da justificação empírica, segue-se que

- (2) A justificação<sub>c</sub> empírica para a crença de que  $p$  é revogável por uma crença empiricamente justificada de que não- $p$ ,

onde “justificação<sub>c</sub>” abrevia “justificação ao grau minimamente suficiente para o conhecimento”.

Uma difícil questão surge a esta altura. Quais as condições sob as quais a crença justificada de S de que  $p$  é revogada pela crença justificada de S de que não- $p$ ? Para os nossos presentes propósitos é suficiente notar que as condições sob as quais a crença justificada de S de que não- $p$  revoga a justificação de S para a crença de que  $p$  é uma função do grau relativo de justificação de que cada uma dispõe. Não precisamos decidir entre abordagens rivais do grau mínimo de justificação que a crença de S de que não- $p$  tem de dispor a fim de revogar a crença justificada<sub>c</sub> de S de que  $p$ . Introduzamos “g” para representar esse grau de justificação, seja qual for, e chamemos uma crença justificada ao grau g, uma crença justificada<sub>g</sub>. Podemos agora introduzir o princípio neutro:

- (G\*) A crença justificada de S de que não- $p$  revoga (pode revogar) a crença justificada<sub>c</sub> de que  $p$  se, e somente se, a crença de S de que não- $p$  é pelo menos justificada<sub>g</sub> (justificável<sub>g</sub>),

Onde “justificado<sub>g</sub>” e “justificável<sub>g</sub>” abreviam, respectivamente, “justificado ao grau g” e “justificável ao grau g”.

Retornando agora ao argumento, a conjunção de G\* e 2 acarreta

(3) A crença de S de que não-*p* é pelo menos empiricamente justificável<sub>g</sub>.

Ademais, a conjunção de N<sub>3</sub> e a conjunta da direita de A acarreta

(4) Não é o caso que a justificação<sub>c</sub> não empírica de S para a crença de que *p* é revogável pela crença empiricamente justificada de que não-*p* de S.<sup>xxxix</sup>

A conjunção de 4 e G\* acarreta

(5) Não é o caso de que a crença de S de que não-*p* seja pelo menos empiricamente justificável<sub>g</sub>.

A conjunção de 3 e 5 é uma contradição. Por conseguinte, N<sub>3</sub> não satisfaz o critério de adequação proposto. N<sub>2</sub>, por outro lado, satisfaz o critério, uma vez que não exclui a possibilidade de revogadores de qualquer tipo. Concluo que N<sub>2</sub> fornece uma análise superior.

Meu argumento contra N<sub>3</sub> realça uma diferença importante entre revogadores refutantes e enfraquecedores.<sup>xi</sup> Não é em geral verdadeiro que se a crença justificada de S de que *q* revoga a justificação conferida à crença de S de que *p* pela fonte A, também revoga a justificação conferida à crença de S de que *p* pela fonte B. Por exemplo, embora a crença justificada de S de que ele sofre de diplopia revogue a justificação conferida à sua crença de que 2+2=4 pelo processo de contagem de objetos, não afeta a justificação conferida a essa crença por intuição ou testemunho. Mais geralmente, os revogadores enfraquecedores para a crença justificada de S de que *p* são *sensíveis à fonte*. Os revogadores refutantes, no entanto, são *neutros à fonte*. Se a crença justificada<sub>g</sub> de S de que não-*p* revogar a justificação<sub>c</sub> conferida à crença de S de que *p* pela fonte A, então também revoga a justificação<sub>c</sub> conferida à crença de S de que *p* por *qualquer outra* fonte. Por exemplo, suponha que a crença de S de que a lista de compra em cima da mesa de centro é justificada<sub>c</sub> pela memória, mas uma experiência perceptiva ulterior, que justifica<sub>g</sub> a sua crença de que a lista não está sobre a mesa de centro, revoga a

sua justificação original. Tivesse a crença de S de que a lista de compras está em cima da mesa de centro sido originalmente justificada<sub>c</sub> por testemunho, a crença perceptivamente justificada<sub>d</sub> de que não está sobre a mesa de centro teria ainda revogado a sua justificação original.

### **3- Argumentos que apoiam a existência do conhecimento *a priori***

Há três vias para se argumentar a favor do *a priori*. A primeira é oferecer uma análise do conceito de conhecimento *a priori* e argumentar que algum conhecimento satisfaz as condições da análise. A segunda é identificar critérios do *a priori* e mostrar que algum conhecimento satisfaz esses critérios. A terceira é argumentar que as teorias empiristas radicais do conhecimento são deficientes em algum ponto e que o único remédio para essa deficiência é adotar o *a priori*.<sup>xii</sup>

#### **3.1. Argumentos conceituais**

Hilary Putnam adota a primeira estratégia. Ele defende uma concepção de justificação *a priori* que envolve uma condição de irrevogabilidade. Argumentamos na seção 2.3 que nem  $C_3$  e nem  $C_4$  são *necessários* para a justificação *a priori*. Todavia, se essa condição proposta for *suficiente* para tal justificação, ela pode ser utilizada na defesa da existência do *a priori*. Por conseguinte, duas perguntas têm de ser respondidas. A concepção fornece um conjunto de condições suficientes para uma justificação *a priori*? Quaisquer crenças satisfazem as condições propostas? Meu interesse principal é na primeira pergunta.

Putnam sustenta que uma afirmação *a priori* é aquela que “nunca seria *racional* abrir mão”.<sup>xlii</sup> Ele prossegue argumentando que o Princípio Mínimo de Contradição (PMC), Nem toda afirmação é verdadeira e falsa, é racionalmente irrevisível. O seu argumento é dirigido contra os seus próprios argumentos anteriores de que nenhuma afirmação são racionalmente irrevisíveis.<sup>xliii</sup> De acordo com a sua perspectiva anterior, os proponentes tradicionais do *a priori* confundiam a propriedade de ser *a priori* com a propriedade relacionada, mas diferente, de ser *contextualmente a priori*. A fonte

da confusão é uma falha em reconhecer dois tipos de bases para a revisão racional. As bases *diretas* para se revisar racionalmente alguma crença de que  $p$  consistem em alguma observação cujo conteúdo justifica a crença de que não- $p$ . As bases *teóricas* consistem num conjunto de observações que é melhor explicado por uma teoria que não contém a afirmação de que  $p$  do que por qualquer teoria que contém a afirmação de que  $p$ . Uma afirmação é contextualmente *a priori* no caso em que é racionalmente irrevisível em bases diretas, mas racionalmente revisível em bases teóricas. Uma afirmação é *a priori* apenas no caso em que é racionalmente irrevisível em quaisquer bases. Os proponentes tradicionais do *a priori* identificaram as afirmações que não são racionalmente revisíveis em bases diretas, e acreditavam que não são racionalmente revisíveis em quaisquer bases. Putnam, no entanto, argumenta que as afirmações *a priori* pretendidas são racionalmente revisíveis em bases teóricas.

O ponto crucial desse seu argumento é que não há bases teóricas possíveis para se revisar racionalmente PMC. Como podemos excluir a possibilidade de que alguma teoria física futura, talvez uma que não pudéssemos conceber agora, implicasse na rejeição do PMC, e não obstante fosse aceita porque explica um domínio diverso de fenômenos, produzindo previsões surpreendentes que são subsequentemente verificadas e aumenta o nosso entendimento do mundo? Podemos assim fazer, de acordo com Putnam, porque sabemos agora que tal teoria terá de consistir de toda afirmação e sua negação. Mas uma teoria que nada exclui não é afinal de contas uma teoria. Portanto, não há circunstâncias sob as quais seria racional aceitá-la.

A proposta de Putnam não é clara num aspecto crucial. Ele não é explícito na questão de saber se uma crença justificada *a priori* em princípios lógicos, tal como PMC, requer evidências que a suportem e, caso requeira, qual a natureza dessa evidência. Há pelo menos três interpretações possíveis da condição que ele propõe para a justificação *a priori*:

- (A)  $p$  é racionalmente irrevisível e S acredita que  $p$ ;
- (B)  $p$  é racionalmente irrevisível e S está justificado em acreditar que  $p$ ;
- (C)  $p$  é racionalmente irrevisível e S está justificado em acreditar que  $p$  é racionalmente irrevisível.

(A) não é suficiente para a justificação *a priori*; é compatível com S não ter justificação para a crença de que  $p$ . De acordo com (A), alguém que acreditasse que PMC por qualquer razão que seja, por mais esquisito que seja, estaria com isso justificado *a priori* em acreditar que  $p$  (presumindo que PMC é de fato racionalmente irrevisível). Mas, como argumentamos anteriormente, a justificação *a priori* para a crença de que  $p$  requer justificação não experiencial para essa crença.

(B) é também insuficiente para a justificação *a priori*, uma vez que é compatível com S ter justificação *experiencial* para a crença de que  $p$ . Por exemplo, suponha que Hilary olha para sua mão, nota a quantidade de dedos e, com base nisso, vem a acreditar que a afirmação “Minha mão tem cinco dedos” é verdadeira e que essa afirmação não é falsa. Hilary está justificado, com bases *a posteriori*, em acreditar que alguma afirmação não é verdadeira e falsa ao mesmo tempo.

Putnam, no entanto, rejeita isso pelas seguintes razões:

Poderia acontecer que não há cinco dedos em minha mão. Por exemplo, minha mão pode ter sido amputada e o que estou vendo pode ser uma prótese plástica [...] Mas ainda que acontecesse de eu ter uma mão, ou que minha mão tivesse apenas quatro dedos, ou sete dedos, ou seja lá quantos forem, descobrir que eu estava errado acerca do relato de observação não abalaria a fé que tenho em minha crença de que esse relato de observação não é verdadeiro e falso ao mesmo tempo.<sup>xliv</sup>

Esse argumento não é adequado. Suponha, por exemplo, que Hilary acredita baseado em olhar para sua mão que a frase “A minha mão tem cinco dedos” é verdadeira, mas, quando olha novamente, descobre que sua mão tem apenas quatro dedos. A observação posterior de que sua mão tem apenas quatro dedos o justifica em acreditar que a frase “A minha mão tem cinco dedos” é falsa e que a afirmação não é verdadeira. Assim, a sua fé na crença de que o relato de observação original não é verdadeiro e falso ao mesmo tempo permaneceria inabalada, uma vez que a observação posterior também justifica essa crença. A intenção de Putnam pode ser a de que o seu reconhecimento de que nenhuma situação epistemicamente possível abalaria a sua fé de que PMC é verdadeiro justifica a sua crença de que PMC é verdadeiro. Essa interpretação do argumento conduz a (C).

(C) não é suficiente para a crença de S de que *p* ser justificada *a priori*, uma vez que (C) é compatível com ter S ter justificção *experencial* para acreditar que *p* é racionalmente irrevisível. Por exemplo, um estudante pode acreditar que PMC seja racionalmente irrevisível apenas pelo testemunho de um tutor de filosofia. Mas se a justificção do estudante para acreditar que PMC é verdadeiro com base na crença justificada de que PMC é racionalmente irrevisível, então, se essa última crença é justificada *a posteriori*, a primeira é também justificada *a posteriori*. Além do mais, ainda que S acredite que PMC seja racionalmente irrevisível com base na determinação das consequências da negação de PMC e por achar algumas dessas consequências inaceitáveis, ainda não se segue que a crença de S de que PMC é racionalmente irrevisível seja justificada *a priori*. Há dois problemas relacionados. Primeiro, ao determinar as consequências de se negar PMC tem-se de empregar *outros* princípios da lógica. Mas, a fim de estar justificado *a priori* em acreditar que PMC é racionalmente irrevisível, tem-se de estar justificado *a priori* em acreditar nos princípios lógicos que se utiliza ao derivar as consequências de se negar PMC. Putnam, no entanto, não pode apelar a (C) para estabelecer que os princípios lógicos usados para derivar as consequências da negação de PMC são em si *a priori*. Tal apelo provoca um regresso, uma vez que tem de considerar as consequências de se negar tais princípios, o que exigirá outros princípios da lógica. Segundo, a fim de se estar justificado *a priori* em acreditar que PMC é racionalmente irrevisível, tem-se de estar justificado *a priori* em acreditar numa teoria que nada exclui não é uma teoria genuína. Putnam, no entanto, não diz se os princípios *metodológicos*, enquanto opostos aos lógicos, são justificados *a priori*.

### **3.2. Argumentos Criteriais**

Os argumentos criteriais têm uma estrutura comum. Identificam alguma característica das proposições que supostamente sabemos e alegam que não podemos conhecer *a posteriori* proposições que tenham essa característica, do que se segue que o conhecimento dessas proposições tem de ser *a priori*. Os argumentos criteriais diferem dos argumentos conceituais, uma vez que não alegam que a característica alegada como suficiente para o conhecimento *a priori* esteja incluída na análise do conceito de conhecimento *a priori*.

Kant fornece o mais conhecido e influente argumento criterial. Ele sustenta que a necessidade é *critério* do *a priori*: “se temos uma

proposição que ao ser pensada é pensada como *necessária*, ela é um juízo *a priori*; [...]”.<sup>xlv</sup> Essa tese se baseia na observação de que “A experiência nos ensina que uma coisa é de tal e tal modo, mas não que não pode ser de outro modo”.<sup>xlvi</sup> Kant continua a argumentar que “as proposições matemáticas estritamente falando são sempre juízos *a priori*, e não empíricos; pois levam à necessidade, que não pode ser derivada da experiência.”<sup>xlvii</sup> Portanto, conclui ele, o conhecimento das proposições matemáticas é *a priori*.

O argumento de Kant, o *Argumento da Necessidade*, pode ser apresentado com se segue:

- (1) As proposições matemáticas são necessárias.
- (2) Não se pode conhecer uma proposição necessária com base na experiência.
- (3) Portanto, não se pode conhecer as proposições da matemática com base na experiência.

A primeira premissa é controversa. Alguns questionam a cogência do conceito de verdade necessária. Outros sustentam que as frases modais não expressam verdades ou falsidades. Para os nossos propósitos concedamos que (1) expressa uma verdade a fim de tratar das questões epistêmicas que suscita.

A expressão “conhecer uma proposição necessária” em (2) é ambígua. Introduzamos as seguintes distinções:

- (A) S conhece o *estatuto modal geral* de  $p$  apenas no caso de S saber que  $p$  é uma proposição necessária ou S saber que  $p$  é uma proposição contingente.
- (B) S conhece o *valor de verdade* de  $p$  apenas no caso de S saber que  $p$  é verdadeira ou S saber que  $p$  é falsa.
- (C) S conhece o *estatuto modal específico* de  $p$  apenas no caso de S saber que  $p$  é necessariamente verdadeira ou S saber que  $p$  é necessariamente falsa ou S saber que  $p$  é contingentemente verdadeira ou S saber que  $p$  é contingentemente falsa.

(A) e (B) são logicamente independentes. Pode-se saber que  $p$  é uma proposição matemática e que todas as proposições matemáticas são necessárias, e não saber se  $p$  é verdadeira ou falsa. A conjectura de Goldbach fornece um exemplo. Alternativamente, pode-se saber que alguma proposição matemática é verdadeira, e não saber se é uma

verdade necessária ou contingente. (C), no entanto, não é independente de (A) e (B). Não se pode saber o estatuto modal específico de uma proposição a menos que se saiba seu estatuto modal geral e seu valor de verdade.

Utilizando essas distinções podemos agora ver que o Argumento da Necessidade se parte em dois argumentos distintos. O primeiro, o *Argumento Kantiano*, corre com se segue:

- (1) As proposições matemáticas são necessárias.
- (2\*) Não se pode conhecer o *estatuto modal geral* de uma proposição necessária com base na experiência.
- (3\*) Portanto, não se pode conhecer o valor de verdade das proposições matemáticas com base na experiência.

Kant argumenta do seguinte modo. Ele admite que a experiência pode fornecer evidências de que uma coisa é de tal e tal modo, ou de maneira mais perspicaz, que é o caso. O que ele nega é que a experiência possa fornecer evidências de que algo *tenha de ser* o caso, ou, de maneira mais perspicaz, de que é necessário. (2\*) articula essa interpretação. Kant conclui com base nisso que o conhecimento de que  $7+5=12$  (*não* o conhecimento de que “ $7+5=12$ ” é necessário) é *a priori*.

O Argumento Kantiano envolve a seguinte suposição:

- (4) Se o estatuto modal geral de  $p$  é conhecível apenas *a priori*, então o valor de verdade é conhecível apenas *a priori*.

(4), no entanto, é falso. Considere uma proposição contingente como a de que este copo é branco. Se só se pode saber *a priori* que essa proposição é necessária, então só se pode saber *a priori* que tal proposição é contingente. As evidências relevantes para determinar a última são as mesmas para determinar a primeira. Por exemplo, se determino que “ $2+2=4$ ” é necessária tentando e não conseguindo conceber sua falsidade, determino que “Este copo é branco” é contingente tentando e sendo bem-sucedido em conceber sua falsidade. Mas se o meu conhecimento de que “Este copo é branco” é contingente for *a priori*, não se segue que o meu conhecimento de que este copo é branco seja *a priori*. Pelo contrário, é *a posteriori*. Portanto, (4) tem de ser rejeitada.

Os proponentes do argumento poderiam nesse ponto recuar a uma versão mais fraca de (4):

(4\*) Se  $p$  é uma proposição necessária e se o estatuto modal geral de  $p$  é conhecível apenas *a priori*, então o valor de verdade de  $p$  é conhecível apenas *a priori*.

Há, no entanto, contraexemplos plausíveis a (4\*). Se Kripke estiver correto sobre a semântica dos nomes próprios, então afirmações de identidade verdadeiras envolvendo nomes próprios diferentes são verdades necessárias.<sup>xlviii</sup> O conhecimento de que tais proposições são necessárias se baseia em experimentos mentais: a incapacidade de conceber que algum objeto seja diferente de si mesmo. Mas o conhecimento de que são verdadeiras é baseado na experiência, no caso de Vésper e Fósforo observações astronômicas. Outro exemplo familiar surge quando acreditamos, e aparentemente conhecemos, proposições matemáticas com base no testemunho de um professor ou da autoridade de um manual.<sup>xlix</sup>

A segunda versão do Argumento da Necessidade, o *Argumento Modal*, prossegue assim:

- (1) As proposições matemáticas são necessárias.
- (2\*) Não se pode conhecer o *estatuto modal geral* de uma proposição necessária com base na experiência.
- (3\*\*) Portanto, não se pode conhecer o *estatuto modal geral* das proposições matemáticas com base na experiência.

O Argumento Modal é menos ambicioso que o Argumento Kantiano e, como consequência, não está aberto às objeções levantadas contra esse último. Por outro lado, é demasiado fraco para estabelecer que o conhecimento matemático difere do conhecimento científico. Se sólido, estabelece que o conhecimento do estatuto modal geral das proposições matemáticas e científicas é *a priori* e compatível com a perspectiva de que o valor de verdade de ambas é *a posteriori*.

Não obstante, uma vez que é incompatível com a tese mais geral de que *todo* conhecimento é *a posteriori*, o Argumento Modal merece um exame cuidadoso. O que pode ser dito em apoio de (2\*)? A manobra comum é invocar a tese kantiana de que a experiência pode apenas nos ensinar o que é o caso, ou a sua contraparte leibniziana de que a experiência pode fornecer conhecimento apenas do mundo *efetivo*, mas não de outros mundos possíveis.<sup>1</sup> Se essa tese estiver certa, então (2\*) é plausível. Mas um bocado do nosso conhecimento prático comum e a maior parte do nosso conhecimento científico fornecem contraexemplos claros a essa tese. O meu

conhecimento de que a minha caneta cairá caso eu a solte não me fornece informação sobre o que é o caso, pois a antecedente é contrária aos fatos. Fornece-me informação sobre alguns outros mundos possíveis além do mundo efetivo. As leis científicas não são meras descrições do mundo efetivo. Elas suportam condicionais contrafactuais e, por conseguinte, fornecem informação para além daquilo que é verdadeiro no mundo efetivo. Na falta de apoio adicional para a premissa (2\*), o Argumento Modal deveria também ser rejeitado.

### **3.3. Argumentos Criteriais: Irrefutabilidade**

Ao defender a existência do conhecimento *a priori*, Kant deu atenção à suposta necessidade das proposições matemáticas. Os proponentes do empirismo lógico, que reagiam contra a tese de John Stuart Mill de que conhecemos proposições matemáticas, como  $3+2=5$ , com base em generalização indutiva de casos observados, deram atenção a uma característica diferente das proposições matemáticas: a sua suposta irrefutabilidade pela experiência. Carl Hempel a expõe assim:

considere agora uma “hipótese” simples da aritmética:  $3+2=5$ . Se essa é de fato uma generalização empírica de experiências passadas, então tem de ser possível apresentar que tipo de evidência nos obrigaria a conceder que a hipótese não era afinal geralmente verdadeira. Se alguma evidência infirmadora para a dada proposição puder ser cogitada, o seguinte caso poderia ser perfeitamente típico: colocamos alguns micróbios numa lâmina, pondo primeiro três e depois dois. Posteriormente contamos todos os micróbios para testar se nesse caso 3 e 2 somaram 5. Suponha agora que contamos 6 micróbios no todo. Consideraríamos isso como uma infirmação empírica da proposição dada, ou pelo menos como uma prova de que ela não se aplica a micróbios? Claramente que não; ao invés, presumiríamos que cometemos um erro na contagem ou que um dos micróbios se dividiu em dois entre a primeira e a segunda contagem.<sup>li</sup>

Uma vez que Hempel sustenta que não consideramos qualquer evidência experiencial como infirmando uma proposição matemática, ele conclui que tais proposições não são confirmadas pela experiência.

O argumento de Hempel, o *Argumento da Irrefutabilidade*, pode ser exposto com se segue:

- (1) Nenhuma evidência experiencial pode infirmar proposições matemáticas.
- (2) Se evidências experienciais não podem infirmar proposições matemáticas, então não podem confirmar tais proposições.
- (3) Portanto, evidências experienciais não podem confirmar proposições matemáticas.

Esse argumento é válido e a segunda premissa incontroversa. A premissa (1), no entanto, não é obviamente verdadeira. Além do mais, a defesa que Hempel faz de (1) não é muito forte. Ele considera apenas o caso mais fraco possível de evidência infirmadora experiencial potencial.

A fim de apresentar esse ponto mais claramente, notemos primeiro duas características familiares da prática indutiva: (a) as nossas avaliações do grau ao qual um caso particular confirma ou infirma uma generalização é uma função das evidências disponíveis totais; e (b) os casos infirmadores aparentes de uma generalização podem sempre ser explicados de modo que não afetem a hipótese original. A defesa de Hempel de (1) é fraca em vários aspectos. Primeiro, não leva em conta o número de instâncias confirmadoras aparentes da proposição em questão. Segundo, envolve apenas uma única instância infirmadora da proposição. Terceiro, as hipóteses que são invocadas para explicar a aparente instância infirmadora não estão sujeitas a teste empírico independente. Em tal situação, dado um pano de fundo de evidências que apoiam a generalização, é razoável desconsiderar as instâncias infirmadoras como aparentes e explicá-las por alguma outra base empírica mais plausível.

O argumento contra a premissa (1) pode ser consideravelmente reforçado revisando-se o cenário de Hempel com se segue: (a) o número de instâncias infirmadoras da proposição é aumentado de modo que seja alto em relação ao número de instâncias confirmadoras; e (b) as hipóteses invocadas para explicar as aparentes instâncias infirmadoras sejam submetidas à investigação independente e se mostrem infundadas. Suponhamos que experimentamos um grande número de aparentes instâncias infirmadoras da proposição de que  $3+2=5$  e, ademais, que as investigações empíricas das hipóteses invocadas para explicar essas instâncias infirmadoras produziram muito pouco, se é que

produziram, apoio para as hipóteses. Dadas essas revisões, o proponente do Argumento da Irrefutabilidade pode continuar a sustentar a premissa (1) apenas pelo preço ou de divorciar a matemática das aplicações empíricas ou sustentando crenças empíricas que estão em desacordo com as evidências disponíveis.

Esse ponto pode ser apresentado mais claramente considerando-se o seguinte conjunto de proposições:

- (a) A proposição matemática  $3+2=5$  é aplicável a micróbios;
- (b) O procedimento empírico de contar micróbios fornece evidências infirmadoras *apenas aparentes* para a proposição de que  $3+2=5$ ;
- (c) Os resultados da investigação empírica independente não apoiam as hipóteses auxiliares introduzidas para explicar as evidências infirmadoras como apenas aparentes.

Embora (c) não implique não-(b), fornece fortes razões para se rejeitar (b). Claramente, o proponente do Argumento da Irrefutabilidade não pode simplesmente vindicar (b), pois simplesmente vindicar (b) sem apoio independente é cometer uma petição de princípio contra o empirista radical. Mas (c) estabelece que as razões independentes oferecidas no apoio de (b) são infundadas. Portanto, (b) tem de ser rejeitada. O proponente do Argumento da Irrefutabilidade, no entanto, não pode aceitar (a) e não-(b) ao mesmo tempo. Se as evidências infirmadoras fornecidos pelo procedimento de contagem de micróbios *não* é meramente aparente, então é genuíno. Portanto, restam apenas duas alternativas: ou (i) rejeitar (a) e sustentar que a matemática não é aplicável a micróbios, ou (ii) continuar a sustentar (b) a despeito de (c). Nenhuma das alternativas é palatável, uma vez que (i) divorcia de fato a matemática de suas aplicações empíricas, ao passo que (ii) nos deixa em posição de sustentar uma crença que vai contra as evidências disponíveis. A alternativa mais plausível é aceitar (a) e rejeitar (b). Mas rejeitar (b) é rejeitar a premissa (1) do Argumento da Irrefutabilidade. Portanto, o argumento não atinge o alvo.

### **3.4. Argumentos da Deficiência**

Lawrence Bonjour oferece três argumentos que pretendem expor deficiências no empirismo radical. O primeiro alega que o empirismo radical conduz ao ceticismo. Pressupõe que algumas crenças são diretamente justificadas apenas pela experiência. Tais

crenças são “particulares ao invés de gerais em seu conteúdo e se confinam a situações observáveis a lugares e momentos específicos e totalmente delimitados.”<sup>lii</sup> Ou algumas crenças cujo conteúdo vá além da experiência direta são justificadas ou o ceticismo é verdadeiro. A justificação das crenças cujo conteúdo vai além da experiência requer inferência a partir das crenças justificadas. Uma vez que os princípios de inferência são *gerais*, não podem ser diretamente justificados pela experiência.

Os outros dois argumentos são direcionados ao empirismo radical de W. V. Quine. O primeiro sustenta que para que uma pessoa esteja justificada em acreditar que *p*, essa pessoa tem de possuir uma razão para pensar que *p* é provavelmente verdadeira. De acordo com Quine, um sistema de crenças que satisfaz padrões como simplicidade, escopo, fecundidade, adequação e conservadorismo está justificado. Mas, pergunta Bonjour,

Que razão pode ser oferecida para se pensar que um sistema de crenças que é mais simples, mais conservador, explicativamente mais adequado, etc., seja por isso mais provavelmente verdadeiro, que seguir tais padrões seja pelo menos conduza de algum modo à verdade?<sup>liii</sup>

Há duas opções. Ou tal razão é *a priori* ou é empírica. A primeira é incompatível com o empirismo radical. A última é uma petição de princípio, uma vez que em última instância tem de apelar para alguns dos padrões que está tentando justificar.

A objeção final alega que os padrões de Quine para a revisão de crença não impõe quaisquer restrições à justificação epistêmica:

Afinal, tal padrão, uma vez que em bases quonianas não pode estar justificado ou visto como epistemicamente relevante independentemente de considerações de ajustes à experiência, é em si apenas mais um fio (ou nó?) na teia, e por isso igualmente aberto à revisão.<sup>liv</sup>

Assim, quando esses padrões parecerem ditar que alguma crença devesse ser revisada, tal revisão pode ser evitada revisando-se os próprios padrões. Quine não pode responder que tal revisão não está justificada, uma vez que tal resposta se baseia ou nos próprios padrões, o que é circular, ou em algum outro padrão, que é em si revisível.

Os argumentos de Bonjour fornecem uma base para se preferir o seu racionalismo moderado em detrimento de seus rivais empiristas radicais apenas se o primeiro evitar as supostas deficiências que assolam os últimos. Uma vez que o racionalismo moderado sofre das mesmas deficiências, como mostrarei agora, seus argumentos não fornecem razões para preferi-lo em detrimento do empirismo radical. A primeira objeção alega que uma vez que o conteúdo da experiência é *particular*, a experiência não pode justificar diretamente princípios *gerais*. O racionalismo moderado está aberto à mesma objeção a menos que possa mostrar que o conteúdo do *insight* racional não se limita aos objetos particulares. Bonjour sustenta que embora experienciemos apenas objetos particulares, apreendemos as propriedades dos objetos.<sup>lv</sup> O termo “apreende” sugere uma analogia com a percepção, que requer contato causal com o objeto percebido. As propriedades, porém, não podem estar em relações causais. Bonjour sustenta que a metáfora perceptual é enganadora. Assim, a fim de subscrever a tese de que apreendemos as características gerais dos objetos, ele tem de fornecer um tratamento não metafórico dessa suposta capacidade cognitiva.

Bonjour propõe explicar a apreensão das propriedades em termos de uma teoria mais geral de como um pensamento pode ser acerca de, ou ter como conteúdo, alguma propriedade particular. Um pensamento tem como conteúdo alguma propriedade particular em virtude de caráter *intrínseco* ao invés de em virtude de alguma *relação*, quase-perceptiva ou qualquer outra coisa, com essa propriedade. Para um pensamento ser acerca de uma propriedade particular, a triangularidade digamos, essa propriedade tem de ser um constituinte de seu caráter intrínseco:

A ideia central de tal perspectiva seria a de que é um fato necessário, quase-lógico, que um pensamento que instancia um universal complexo envolvendo a triangularidade universal de maneira apropriada [...] seja sobre *coisas* triangulares.<sup>lvi</sup>

A explicação de Bonjour não atinge seu objetivo. Seu objeto é explicar como um pensamento pode ter como conteúdo alguma *propriedade* particular, como por exemplo, a triangularidade. Ele fornece, ao invés disso, apenas um mero esboço de como um pensamento pode ter como conteúdo *objetos* triangulares *particulares*. Uma vez que não oferece uma explicação de como um pensamento pode ter como conteúdo alguma *propriedade*, ele não fornece uma explicação da apreensão das propriedades. Por

consequente, o racionalismo moderado está aberto à primeira objeção.

A segunda objeção repousa na ideia de que estar epistemicamente justificado em acreditar que  $p$  exige ter uma razão para se pensar que  $p$  é provavelmente verdadeira. A expressão “ter uma razão para pensar que  $p$  é provavelmente verdadeira” é ambígua. Distingamos dois sentidos:

- (B) S tem uma razão *básica* R para acreditar que  $p$  se, e só se, S tem R torna provável que  $p$  seja verdadeira;
- (M) S tem uma *meta* razão R para acreditar que  $p$  se, e só se, S tem R e S tem razão para acreditar que R torna provável que  $p$  seja verdadeira.

Seja  $\phi$  o conjunto de condições que Quine defende ser suficiente para a justificação. Suponha que pertencer a um sistema de crenças que satisfazem  $\phi$  torna provável que  $p$  é verdadeira. Se S aprende cognitivamente o fato de que  $p$  pertence a tal sistema, então S tem uma razão *básica* para acreditar que  $p$ . O ataque de Bonjour é que o empirismo radical não pode oferecer um argumento para mostrar que tais razões são verocondutoras. Assim, o problema diz respeito a ter uma *meta* razão para acreditar que  $p$ .

O racionalismo moderado se sai melhor nesse ponto? Suponha que ter um aparente *insight* racional de que  $p$  torna provável que  $p$  seja verdadeira. Assim, se S tem um aparente *insight* racional de que  $p$ , então S tem uma razão *básica* para acreditar que  $p$ . Bonjour enfrenta agora a seguinte questão:

Que razão pode ser oferecida para se pensar que uma crença baseada num aparente *insight* racional seja por isso mais provavelmente verdadeira?

A sua resposta é que a exigência de uma *meta* razão é circular porque, de acordo com sua abordagem, o aparente *insight* racional é por si próprio uma excelente razão para se aceitar uma crença:

Equivale simples e obviamente a uma recusa de tomar seriamente o *insight* racional como uma base para a justificação: uma recusa pela qual a presente objeção não pode oferecer outra razão, e que, por isso, é uma petição de princípio.<sup>lvii</sup>

Os empiristas radicais, no entanto, podem oferecer uma resposta similar à segunda objeção de Bonjour. Eles podem sustentar que a sua exigência por uma meta razão é circular, uma vez que se recusa a considerar seriamente que pertencer a um sistema de crenças que satisfaçam  $\phi$  seja por si próprio uma excelente razão para se aceitar uma crença. Assim, o empirismo radical não se sai *pior* do que o racionalismo moderado com respeito a exigência de meta razões.

A terceira objeção de Bonjour repousa em dois princípios:

- (P1) Crenças justificadas pela experiência são revisíveis; e
- (P2) Os padrões de revisão de crenças justificadas pela experiência são eles próprios justificados pela experiência.

Desses princípios segue-se que

- (P3) Os padrões de revisão de crenças justificadas pela experiência são eles próprios revisíveis.

Mas o racionalismo moderado endossa análogos desses princípios:

- (P1\*) Crenças justificadas por aparente *insight* racional são revisíveis; e
- (P2\*) Os padrões para revisão de crenças justificadas por aparente *insight* racional são eles próprios justificados por aparente *insight* racional.

Por conseguinte, o racionalismo moderado está comprometido com

- (P3\*) Os padrões para revisão de crenças justificadas por aparente *insight* racional são eles próprios revisíveis.

O restante do argumento de Bonjour se aplica com igual força ao racionalismo moderado e ao empirismo radical. Qualquer tentativa de bloquear a revisão dos padrões de revisão de crença ou recorrer aos próprios padrões, o que é circular, ou invoca algum outro padrão, que é em si revisível. Portanto, mais uma vez, o racionalismo moderado não se sai *melhor* do que o empirismo radical.

#### **4. Argumentos opostos à existência do conhecimento *a priori***

Os argumentos contra a existência do conhecimento *a priori* se enquadram em três amplas categorias. Os da primeira oferecem uma análise do conceito de conhecimento *a priori* e alegam que nenhum caso de conhecimento satisfaz as condições na análise. Os da segunda oferecem abordagens empiristas radicais do conhecimento de proposições supostamente conhecíveis apenas *a priori*. Os argumentos da terceira categoria sustentam que o conhecimento *a priori* é incompatível com condições plausíveis para uma teoria do conhecimento adequada.

#### **4.1. Argumentos Conceituais**

Hilary Putnam e Philip Kitcher fornecem exemplos claros da primeira abordagem. Ambos sustentam que o conceito de justificação *a priori* inclui uma condição de irrevogabilidade. De acordo com Putnam, uma afirmação *a priori* é uma afirmação “que nunca seria racional abandonar [...]”<sup>lviii</sup> Kitcher insiste que para um processo justificar crenças *a priori*, tem de ser capaz de “garantir aquelas crenças contra o pano de fundo de uma experiência apropriadamente recalitrante.”<sup>lix</sup> Eles continuam a argumentar que as crenças tradicionalmente consideradas com justificadas *a priori* não cumprem a condição de irrevogabilidade exigida. Argumentamos que o conceito de justificação *a priori* não inclui uma condição de irrevogabilidade. Portanto, o fato de uma crença não satisfazer uma condição de irrevogabilidade não acarreta *imediatamente* que não seja *a priori*. Permanece, porém, a possibilidade de uma conexão *mediada*.

Chamemos a tese geral de que uma justificação *a priori* acarreta a irrevisibilidade racional *Tese da Irrevisibilidade* (TI), e distingamos entre uma versão forte e uma fraca:

(TIFt) Necessariamente, se a crença de S de que *p* é justificada *a priori*, então a crença de S de que *p* é racionalmente irrevisível frente *qualquer* evidência; e

(TIFr) Necessariamente, se a crença de S de que *p* é justificada *a priori*, então a crença de S de que *p* é racionalmente irrevisível frente a qualquer evidência *experencial*.

Meu propósito é argumentar que tanto (TIFt) quanto (TIFr) deveriam ser rejeitados.

Começamos considerando um exemplo que tira as consequências mais explícitas de (TIFt). Suponha que Mary é uma estudante universitária que teve algum treino em lógica. Como

resultado, ele é capaz de discriminar confiavelmente entre inferências elementares válidas e inválidas com base no pensamento reflexivo. Mary hoje se pergunta se “ $p \supset q$ ” acarreta “ $\neg p \supset \neg q$ ”. Ela reflete sobre as afirmações em questão e com base nessa reflexão conclui que a primeira de fato não acarreta a segunda. Após chegar a essa conclusão, um contraexemplo ocorre a ela. A ocorrência do contraexemplo resulta em sua rejeição da primeira conclusão e na crença de que “ $p \supset q$ ” acarreta “ $\neg p \supset \neg q$ ”. As características salientes do exemplo são as seguintes: (a) a crença inicial de Mary se baseia num processo não-experiencial que é confiável embora não infalível; (b) um processo do *mesmo tipo* leva Mary a concluir que a crença inicial está errada e a alcançar a conclusão correta; e (c) as conclusões de Mary como expostas em (b) são crenças justificadas. Já agora algumas afirmações mais controversas: (d) a crença original de Mary de que “ $p \supset q$ ” acarreta “ $\neg p \supset \neg q$ ” é também *justificada*; e (e) a crença original de Mary está justificada *a priori* a despeito de sua revisão subsequente.

O que pode ser dito em favor de (d) e (e)? (d) parece ser similar em todos os aspectos relevantes ao seguinte caso. Mary vê uma tira de papel sobre a mesa e com base nisso forma a crença de que essa tira é quadrada. Um segundo exame visual mais próximo releva que dois dos lados são ligeiramente maiores do que os outros dois. Com base nisso Mary rejeita sua primeira crença sobre a forma do papel e passa a acreditar que é retangular. Uma vez que as circunstâncias sob as quais Mary percebeu a página foram normais e Mary é um discriminador confiável de formas, a sua crença inicial está justificada. O fato de as nossas capacidades discriminatórias às vezes nos trair não acarreta que as crenças baseadas na percepção da forma não são justificadas. Ademais, se tais crenças são tipicamente justificadas, não selecionamos casos particulares como injustificadas *meramente* em virtude do fato de serem falsas. Alguma outra diferença relevante tem de ser citada, como por exemplo, a de que o sujeito estava prejudicado ou o ambiente influenciado. Assim, a falha habitual da habilidade de Mary de discriminar formas, confiável em outras circunstâncias, não acarreta que a sua crença de que o papel é quadrado seja injustificada a despeito do fato de ser falsa. Similarmente, a falha habitual da habilidade de Mary de discriminar inferências válidas, confiável em outras circunstâncias, não acarreta que sua crença de que “ $p \supset q$ ” acarreta “ $\neg p \supset \neg q$ ” seja injustificada a despeito do fato de ser falsa.

A única questão remanescente é se a crença original é justificada *a priori* ou *a posteriori*. Um proponente de (TIFt) tem de

sustentar que a crença é justificada *a posteriori meramente* em virtude do fato dela ter sido revisada. Esse ponto pode ser posto mais claramente pela introdução da noção de um *processo autocorretivo*:

(PAC) Um processo  $\phi$  é autocorretivo para S apenas no caso em que, para qualquer crença falsa de que  $p$  produzida em S por  $\phi$ ,  $\phi$  pode também justificar a crença de que não- $p$  para S.<sup>lx</sup>

(TIFt) acarreta

(1) Se um processo  $\phi$  é autocorretivo e justifica a S uma crença falsa de que  $p$ , então  $\phi$  não justifica *a priori* a crença de S de que  $p$ .

(1) é implausível. É insensível à questão central de se o processo justificativo em questão é experiencial ou não-experiencial. Portanto, sustentar (1) é divorciar a noção de justificação *a priori* da noção de independência das evidências experienciais. É mais plausível rejeitar (1) com base no fato de que a crença original de Mary tanto quanto a crença que a leva a revisar a crença original se baseiam em evidências não-experienciais. Uma vez que rejeitamos (1), (TIFt) tem também de ser rejeitado.

(TIFr) evita o problema inicial com (TIFt). Distingue entre revisões baseadas em evidências experienciais como opostas a evidências não-experienciais, e sustenta que apenas as primeiras são incompatíveis com a justificação *a priori*. Não obstante, (TIFr) está também aberta à objeção.

Começamos considerando um exemplo. Suponha que Pat é um lógico que regular e consistentemente chega a resultados interessantes. Pat, no entanto, se incomoda com o fato de que embora seja um produtor *confiável* de resultados interessantes, não é um produtor *infalível* de tais demonstrações. Acontece que ela tem uma colega, May, que fez um trabalho pioneiro nas bases neurofisiológicas dos processos cognitivos. Com recursos radicais para automelhoria, Pat pede May para conduzir um estudo de seus esforços em construir demonstrações a fim de ver se ela consegue descobrir alguma causa neurofisiológica, esperançosamente reversíveis, de suas infrequentes demonstrações erradas. A investigação revela que (a) um padrão de interferência particular está presente no cérebro de Pat quando e apenas quando ele constrói demonstrações erradas; e (b) quando Pat constrói uma demonstração sob a influência desse padrão, e o padrão é subsequentemente

erradicado por intervenção neurofisiológica, ele consegue ver a falha na demonstração original e corrigi-la. Finalmente, há uma teoria neurofisiológica disponível que apoia a hipótese de que tal padrão deveria causar lapsos cognitivos. Suponha agora que Pat acredita que  $p$  acarreta  $q$  com base na construção de uma demonstração que ele cuidadosamente examinou e achou aceitável. A despeito de seu exame cuidadoso, a demonstração é falha. Ele mais tarde descobre num encontro com May que (a) ela estava monitorando, com um sensor remoto, a sua atividade cerebral no momento em que a demonstração foi construída; (b) que o sensor indicou que o padrão de interferência estava presente; e (c) que os testes padrão indicaram que todo o equipamento estava funcionando adequadamente. Pat ainda não consegue descobrir a falha em sua prova, muito embora conclua, com base nas descobertas empíricas de May, que a sua demonstração é falha e evita a crença de que  $p$  acarreta  $q$ .

As características salientes do exemplo são: (a) a crença inicial de Pat de que  $p$  acarreta  $q$  se baseia num processo de pensamento reflexivo que é confiável embora não infalível; (b) a crença inicial de Pat de que  $p$  acarreta  $q$  é justificada pelo processo não-experencial de pensamento reflexivo; e (c) a justificação que o processo de pensamento reflexivo confere à sua crença é posteriormente revogada pela evidência empírica que indica a presença do padrão de inferência. (a) é incontroverso. (b) é mais controverso, uma vez que envolve a ideia de que crenças falsas podem ser justificadas *a priori*. Essa ideia foi defendida anteriormente na discussão do exemplo de Mary. Aceitamos (c) com o propósito de avaliar (TIFr). Finalmente, considere (d) a crença inicial de Pat de que  $p$  acarreta  $q$  é justificada *a priori* a despeito da última revisão frente a evidências experienciais. (d) parece ser uma consequência direta de (b). Uma vez que a crença de Pat é justificada por um processo não-experencial, é justificada *a priori*. Um proponente de (TIFr) pode resistir a essa conclusão apenas por insistir que uma vez que as evidências experienciais revogam a justificação conferida à crença pelo processo não-experencial, a crença é justificada, pelo menos em parte, pelas evidências experienciais.

A defesa proposta de (TIFr) invoca a seguinte simetria entre evidências justificativas e evidências revogadoras:

(TS) Se evidências do tipo A podem revogar a justificação conferida à crença de S de que  $p$  pela evidência do tipo B, então a crença de S de que  $p$  é justificada pela evidência do tipo A.

(TS), no entanto, não é muito plausível. Considere, por exemplo, o conhecimento introspectivo de sensações corporais como dores e coceiras. Alguns sustentam que o conhecimento introspectivo é indubitável. Não há razões possíveis para se duvidar da verdade de uma crença introspectiva sobre as sensações corporais. Essa ideia foi desafiada pelo chamado argumento EEG.<sup>ixi</sup> A ideia básica é que embora a introspecção forneça presentemente as nossas únicas evidências a favor da presença das sensações corporais, é possível que a neurofisiologia evolua a ponto de que leituras eletroencefalográficas também forneçam tais evidências. Ademais, em circunstâncias apropriadas, as leituras EEG podem cancelar as nossas evidências introspectivas em favor de uma crença acerca da presença de uma sensação corporal. O nosso propósito aqui não é avaliar esse argumento. Suponha que aceitamos

(N) E  
vidências neurofisiológicas podem revogar a justificação conferida a uma crença sobre sensações corporais por introspecção.

Claramente não se segue que a minha presente crença justificada de que tenho uma leve dor de dentes se baseie, ainda que em parte, em evidências neurofisiológicas. Consequentemente, (TS) tem de ser rejeitada. E uma vez que rejeitamos (TS), (TIFr) tem também de ser rejeitada.

## **4.2. Abordagens Empíricas**

Uma estratégia comum de se argumentar contra a existência do conhecimento *a priori* é considerar os exemplos mais proeminentes de proposições supostamente conhecíveis apenas *a priori* e argumentar que tais proposições são conhecidas empiricamente. Foquemo-nos no conhecimento matemático, já que recebeu maior parte da atenção. As abordagens empiristas do conhecimento matemático podem ser divididas em duas amplas categorias: indutiva e holista. A ideia central das teorias indutivas é que proposições matemáticas *epistemicamente básicas* são justificadas diretamente pela observação e pela generalização indutiva. As proposições matemáticas não-básicas são indiretamente justificadas através de suas relações lógicas e explicativas com as proposições matemáticas básicas. O empirismo holista nega que algumas proposições matemáticas sejam justificadas diretamente pela observação e pela generalização indutiva. Todas as proposições matemáticas são parte

de uma teoria explicativa mais ampla que inclui princípios científicos e metodológicos. Apenas teorias inteiras, ao invés de proposições individuais, são confirmadas ou infirmadas pela experiência.

John Stuart Mill é o mais proeminente defensor do indutivismo. No caso da matemática, seu interesse primário é com os princípios primeiros, os axiomas e as definições, da aritmética e da geometria. A sua perspectiva, sucintamente posta, é que esses princípios são justificados indutivamente com base na observação. Tal perspectiva enfrenta obstáculos formidáveis. Por exemplo, as definições não parecem exigir justificação empírica. Além do mais, as propriedades conotadas por alguns termos matemáticos não parecem ser exemplificadas pelos objetos da experiência. Mill sustenta, no entanto, que as definições dos termos matemáticos asserem a existência de objetos que exemplificam as propriedades conotadas pelos termos nas definições e que as definições matemáticas são apenas aproximadamente verdadeiras dos objetos da experiência.<sup>lxii</sup>

Poucos consideram plausível a abordagem de Mill. O meu propósito aqui não é defendê-la. Ao invés, proponho que aceitemos sua cogência a fim de determinarmos se pode ser considerada como um argumento contra o *a priori*. Se Mill estiver correto, segue-se que todas as proposições matemáticas básicas são justificadas com base na observação e na generalização indutiva. Além do mais, todas as outras proposições matemáticas justificadas com base nessas proposições são também justificadas com base na experiência. Não obstante, o sucesso da abordagem não estabelece que não haja conhecimento *a priori* de proposições matemáticas. Tirar tal conclusão é negligenciar a possibilidade da *sobredeterminação epistêmica*: a possibilidade de que proposições matemáticas sejam (ou possam ser) justificada tanto experiencialmente quanto não-experiencialmente.

Mill está ciente desse hiato em seu argumento e apelou para uma versão do princípio da simplicidade:

Onde então está a necessidade de se pressupor que o nosso reconhecimento dessas verdades tenha uma origem diferente do restante do nosso conhecimento quando a sua existência é perfeitamente explicada através da suposição de que a sua origem é a mesma? Quando as causas que produzem crenças em todas as outras instâncias existem nessa instância, e num grau de força um tanto superior ao que existe em outros casos, que a intensidade da própria crença é superior?<sup>lxiii</sup>

Mill sustenta que não há necessidade de supor que haja conhecimento *a priori* para dar conta do nosso conhecimento da matemática. Mas o apelo à simplicidade é mal-usado. O propósito de uma teoria epistemológica não é oferecer a explicação *mais simples* do nosso conhecimento de algum conjunto determinado de proposições. O propósito é oferecer uma explicação *acurada* do nosso conhecimento: uma explicação que forneça uma imagem completa de nossas faculdades cognitivas que dizem respeito ao domínio de verdades em questão. É uma questão aberta se, dadas nossas faculdades cognitivas, temos mais do que uma única fonte de justificação para as crenças num dado domínio. A suposição de que, para qualquer domínio do conhecimento humano, há apenas uma única fonte de justificação não tem fundamento. O princípio de simplicidade exclui a sobredeterminação da justificação. Assim, na falta de um argumento contra sobredeterminação epistêmica, o empirismo de Mill, ainda que cogente, não pode ser oferecido com um argumento contra o *a priori*.

O empirismo holista enfrenta uma dificuldade parecida. A apresentação clássica da posição é fornecida por W. V. Quine.<sup>lxiv</sup> Há (pelo menos) dois modos de se interpretar o seu argumento. A interpretação tradicional é a de que ele está a fornecer um argumento unificado contra a cogência da distinção analítico-sintético. A segunda, que se deve a Hilary Putnam, é que Quine está a fornecer dois argumentos distintos em “Os Dois Dogmas”: um argumento (malsucedido) nas quatro primeiras seções visando a distinção analítico-sintético, e um argumento (bem-sucedido) nas duas seções finais visando a existência do conhecimento *a priori*.<sup>lxv</sup>

Na interpretação de Putnam, quando Quine argumenta que nenhuma afirmação está imune à revisão frente a experiência recalcitrante, ele visa o conhecimento *a priori*. O alvo do ataque é a perspectiva de que há algumas afirmações confirmadas seja como for. A alegação de Putnam é que o conceito de uma afirmação confirmada seja como for não é um conceito de analiticidade mas um conceito de aprioricidade. Quine se confundiu ao pensar que era um conceito de analiticidade por causa das suposições positivistas sobre o significado. Assim, de acordo com Putnam, se o argumento de Quine for sólido, estabelece que não há conhecimento *a priori*.

Meu propósito aqui não é exegético. Aceitemos a interpretação de Putnam da estrutura do argumento de Quine. Aceitemos, além do mais, que Quine foi bem-sucedido em estabelecer que nenhuma afirmação está imune a revisão. Meu interesse é se a conclusão de

Quine pode ser usada num argumento contra a existência do conhecimento *a priori* como alega Putnam. É claro que por si só essa conclusão não é suficiente para tal. A premissa adicional de que a justificação *a priori* acarreta a irrevisibilidade racional é também necessária. Mas, a despeito da alegação de Putnam do contrário, a premissa adicional é falsa.

Retornemos à interpretação mais tradicional do argumento de Quine e, novamente, concedamos que Quine foi bem-sucedido em mostrar que a distinção analítico-sintético não é cogente. Esse resultado fornece aos empiristas os recursos necessários para se argumentar que não há conhecimento *a priori*? Novamente, essa premissa apenas não é suficiente. Argumentamos na seção 2.1 que a análise do conceito de justificação *a priori* não inclui o conceito de analiticidade. Assim, não há conexão imediata ou óbvia entre os dois conceitos. Resta a possibilidade de que haja alguma conexão mediada. Mas se houver tal conexão, algum argumento tem de ser oferecido para tal.

### **4.3. Argumentos de Incompatibilidade**

Paul Benacerraf fornece o exemplo clássico de um argumento que se enquadra na terceira categoria.<sup>lxvi</sup> Ele sustenta que a nossa melhor teoria da verdade fornece condições de verdade para as afirmações matemáticas que referem a entidades abstratas e a nossa melhor abordagem do conhecimento requer uma relação causal entre sujeitos cognoscentes e as entidades referidas pelas condições de verdade das afirmações que conhecem. Dado que entidades abstratas não podem manter relações causais, há uma tensão entre nossa melhor explicação da verdade matemática e nossa melhor explicação do conhecimento matemático. Uma vez que é amplamente defendido que a maior parte, se não todo, conhecimento *a priori* é de verdades necessárias e que as condições de verdade para tais afirmações referem a entidades abstratas, o argumento de Benacerraf levanta uma questão mais acerca da possibilidade do conhecimento *a priori*.

Alguns rejeitam o argumento pela razão de que sua premissa epistêmica, que endossa uma condição causal para o conhecimento, assenta-se amplamente rejeitada teoria causal do conhecimento. Os proponentes do argumento sustentam que a condição causal endossada pela premissa epistêmica do argumento tem seu apoio das condições de uma epistemologia naturalizada ao invés das de uma teoria causal do conhecimento.<sup>lxvii</sup>

É complicado avaliar a afirmação de que o naturalismo é incompatível com o conhecimento de entidades abstratas, uma vez que há muitas versões rivais dessa perspectiva. Com o risco de simplificação, identifiquemos duas variedades gerais. A primeira, o naturalismo *científico*, deve-se a W. V. Quine, que rejeita o projeto epistemológico tradicional de fornecer uma justificação filosófica *a priori* e oferece, em seu lugar, uma visão da epistemologia como um ramo da ciência.<sup>lxviii</sup> A segunda, o naturalismo *filosófico*, defende a introdução de condições naturalistas para os projetos filosóficos tradicionais. No caso de análise conceitual, por exemplo, exige-se que o *analisis* de um conceito inclua apenas conceitos naturalisticamente respeitáveis. Deparamo-nos agora com dois problemas. O naturalismo filosófico ou o científico excluem a possibilidade do conhecimento de entidades abstratas?

#### **4.4. Naturalismo Filosófico**

Se o naturalismo filosófico exclui o conhecimento de entidades abstratas, exclui em virtude das exigências de um descendente mais promissor da teoria causal. O mais promissor é confiabilismo processual. Alvin Goldman sustenta que

- (G) A  
crença de S de que  $p$  em  $t$  é justificada se e somente se  
(a) A crença de S de que  $p$  em  $t$  é permitida por um sistema  
correto de regras-J, e  
(b) Essa permissão não é solapada pelo estado cognitivo de S  
em  $t$ .<sup>lxix</sup>

De acordo com (G), qualquer crença produzida por um processo psicológico confiável satisfaz (a). Tal crença é justificada contanto que S não possua evidências revogadoras como a de que de a crença é falsa ou que é produzida por um processo infiel. (G) também parece ser compatível com a possibilidade de crenças justificadas sobre entidades abstratas, uma vez que nem (a) e nem (b) envolvem quaisquer condições causais. (G), no entanto, está aberta à objeção.

Suponha que Maud pertença a uma organização cujos líderes acreditam, por bases fracas, que a clarividência é uma fonte confiável de conhecimento.<sup>lxx</sup> Ademais, suponha que um extenso trabalho empírico de investigação desse fenômeno fora feito, tendo sido os resultados negativos, e que essa informação está *presente* na comunidade epistêmica de Maud. Os outros estão cientes dessa

informação. Os jornais, as revistas, os livros e a televisão reportam amplamente essa informação. Além do mais, Maud teve *pronto acesso* a essa informação. Os outros com que elas interagem têm essa informação e a compartilhariam caso fossem perguntados. Os jornais e as revistas que ela às vezes lê relatam a informação. Os livros e os periódicos da biblioteca que ela frequentemente visita documentam a informação. Os programas televisivos transmitem a informação nos canais que ela assistiu recentemente. Os líderes da organização estão cientes da evidência negativa, do fato que é amplamente divulgado, e do fato de que muitos de seus seguidores têm pronto acesso à informação. Como consequência, insistem continuamente que seus seguidores ignorem a informação de fontes externas sobre o assunto. Maud adere a seus pedidos e prossegue na formação de muito poucas crenças acerca da clarividência além daquelas promulgadas na organização. Suponha agora que ela é de fato uma clarividente, que o processo é confiável, e que Maud forma a crença verdadeira de que  $p$  através desse processo. A crença de Maud não é justificada, uma vez que a evidência que ela tem a favor da confiabilidade da clarividência é fraca e ela escolhe ignorar evidências abundantes a favor do contrário. Contudo, sua crença satisfaz (a) e (b) em (G). (b) é satisfeita porque Maud não está justificada em acreditar que a clarividência não é um processo confiável de formação de crenças. O seu sistema de crenças é demasiado empobrecido para justificar essa crença a despeito do fato de que ela tem pronto acesso às evidências que a apoiam.

(G) é vulnerável ao caso de Maud porque presume que apenas as evidências que *alguém possui* sejam relevantes para a justificação de suas crenças. Não leva em conta a dimensão *social* da justificação. Contudo, como o caso de Maud indica, não se pode ignorar as evidências prontamente disponíveis, e tais evidências, ainda que ignoradas, podem ser relevantes para a justificação das crenças. Assim, qualquer tratamento plausível das evidências solapadoras tem de levar em conta as evidências que alguém *efetivamente não* possui, mas que estão *presentes* na comunidade epistêmica e que se tem *pronto acesso*. A condição (b) de Goldman tem de ser substituída por

(b\*) Essa permissão não é solapada pelo estado cognitivo de S em  $t$  ou pela evidência presente na comunidade epistêmica de S a qual S tenha pronto acesso em  $t$ .<sup>lxxi</sup>

(b\*) produz o resultado correto de que a justificação de Maud é solapada pela evidência prontamente disponível em sua comunidade epistêmica acerca da infiabilidade da vidência. A principal consequência de (b\*) é que a informação na comunidade epistêmica de S tanto sobre a possibilidade quanto da confiabilidade de um processo de formação de crenças é relevante para se esse processo justifica as crenças que produzem em S. Portanto, a questão que temos de tratar é se o caráter causalmente inerte das entidades abstratas fornece uma base para se questionar a possibilidade ou confiabilidade do processo que supostamente produz crenças sobre tais entidades.

Os proponentes do *a priori* sustentam que têm acesso cognitivo a entidades abstratas através de um processo não-experiencial, chamado *intuição*, e que esse processo justifica as crenças sobre essas entidades. Junto do processo estão os estados cognitivos com uma fenomenologia única aos quais seus proponentes reconhecem. A experiência de tais estados fenomenologicamente distintos lhes fornece alguma razão para acreditar que têm acesso cognitivo a entidades abstratas. Mas há também evidência contrária de dois tipos. Primeiro, há controvérsia acerca da existência e confiabilidade da intuição. Alguns sustentam que não têm o estado cognitivo em questão, ao passo que outros reconhecem ter tais estados embora neguem que forneçam acesso cognitivo às entidades abstratas. Além disso, há outros que questionam a confiabilidade das crenças baseadas na intuição, e há movimentos nos campos da matemática e da filosofia para afastar tais apelos. Frente à evidência que outros não têm tais experiências, os proponentes do *a priori* têm de acreditar ou que têm um equipamento cognitivo único, ou que os outros têm o mesmo equipamento, mas que não funciona, ou que são menos confiáveis enquanto relatores dos fatos de suas próprias vidas cognitivas. Há pouca evidência que apoie quaisquer das alternativas. Frente aos supostos casos de erro intuitivo, os proponentes conseguem fornecer apenas evidência anedótica no apoio da afirmação que os casos citados são anômalos e que o processo é geralmente confiável.

Segundo, esses problemas são reforçados pelo fato de que pouco se sabe acerca dos mecanismos neurofisiológicos pelos quais a intuição produz crenças. Uma vez que tomamos por garantido que todos os processos cognitivos têm uma base neurofisiológica, a falta de evidência neurofisiológica aumenta a desconfiança da existência do processo. Além do mais, o caráter causalmente inerte das entidades abstratas garante que elas não desempenhem qualquer

papel na geração de crenças sobre elas. Assim, se a intuição é um processo confiável, a sua confiabilidade não pode ser explicada do mesmo modo que a confiabilidade de nossos melhores processos cognitivos já entendidos. Mas dado que os processos causais subjacentes são desconhecidos, não estamos em posição de oferecer uma explicação alternativa. A crença de que a intuição é um processo confiável introduz um hiato explicativo que reforça as dúvidas acerca da confiabilidade do processo.

O problema diante de nós é saber se o caráter causalmente inerte das entidades abstratas levanta um obstáculo à satisfação de (b\*) em (G) através de processos, como a intuição, que produzem crenças cujas condições de verdade referem entidades abstratas. As crenças produzidas pela intuição satisfazem (b\*) apenas se não houver evidência prontamente acessível presente na comunidade epistêmica que coloque em xeque a possibilidade ou a confiabilidade da intuição. Argumentei que há tal evidência. Chegar a uma solução, no entanto, requer uma investigação mais detalhada de duas questões: o escopo e a qualidade da evidência; e o quão fortemente um revogador potencial tem de ser apoiado a fim de suspender a justificação conferida a uma crença produzida de maneira confiável. Essa investigação mais detalhada vai além do escopo da presente discussão. A minha conclusão principal é que, dentro do enquadramento do confiabilismo processual, o caráter causalmente inerte das entidades abstratas constitui uma *ameaça* à justificação *a priori*. Embora o confiabilismo processual não exclua a *possibilidade* de que processos tais como a intuição justifiquem crenças cujas condições de verdade se referem a entidades abstratas, a falta de uma explicação de como esses processos podem produzir tais crenças de maneira confiável gera revogadores potenciais para tal justificação.

#### **4.5. Naturalismo Científico**

O problema final do qual trataremos é se o naturalismo científico exclui o conhecimento de entidades abstratas. Penelope Maddy fornece a tentativa mais articulada de mostrar que a inatividade causal de tais entidades apresenta um problema genuíno para o conhecimento matemático de uma perspectiva naturalizada quineana. O problema não é conceitual, mas sim explicativo. Quando os matemáticos, como por exemplo, R. M. Solovay, formam opiniões sobre questões matemáticas, geralmente estão corretos. Assim

Ainda que o confiabilismo não venha a ser a análise correta do conhecimento e da justificação, de fato, ainda que conhecimento e justificação venham a ser noções dispensáveis, permanecerá o problema de explicar o fato inegável da confiabilidade dos especialistas. Em particular, mesmo de uma perspectiva completamente naturalizada, o platonista ainda nos deve uma explicação de como e por que as crenças de Solovay sobre conjuntos são indicadores confiáveis da verdade sobre conjuntos.<sup>lxxii</sup>

O caráter causalmente inerte das entidades abstratas, alega Maddy, é um entrave para a explicação da confiabilidade das crenças matemáticas de Solovay.

De uma perspectiva completamente naturalizada, a ciência é uma disciplina autônoma que não está sujeita a exigências filosóficas de fora. Assim, se o entrave à explicação da confiabilidade de Solovay surge dentro de uma perspectiva completamente naturalizada, tem de surgir dentro da ciência. Maddy oferece a seguinte razão para se pensar que o caráter causalmente inerte das entidades matemáticas levanta um obstáculo no fornecimento de uma resposta cientificamente aceitável:

Obviamente, estamos nos insurgindo aqui contra outra versão, menos específica, da mesma convicção vaga que torna a teoria causal do conhecimento tão persuasiva: para ser seguro o processo pelo qual venho a acreditar em afirmações sobre  $x$ 's tem em última instância de se conectar de algum modo apropriado aos  $x$ 's efetivos.<sup>lxxiii</sup>

O suposto empecilho na explicação da confiabilidade de Solovay é uma condição causal da formação de crenças confiável:

(M) O processo pelo qual S vem a ter crenças sobre  $x$ 's é confiável (seguro) apenas se esse processo for apropriadamente conectado aos  $x$ 's.

Se o suposto empecilho surge dentro da ciência, tem de haver evidência de algum ramo relevante da ciência que apoie (M). Maddy oferece três considerações para o apoio:

- (a) a analogia da matemática/ciência;
- (b) a crença de que todas as explicações são em última instância causais;

(c) uma forma forte de fisicismo.

Todas as três considerações parecem ser filosóficas em caráter. Além do mais, Maddy não tenta afastar as aparências oferecendo alguma evidência de que o comprometimento com (a), (b) ou (c) surge dentro da ciência. Ela, porém, sustenta que há apoio para (a) dentro da matemática.

De acordo com a analogia da matemática/ciência, a matemática se assemelha à ciência natural em dois aspectos importantes:

- (a1) Algumas crenças matemáticas são básicas e não inferenciais;
- (a2) As crenças matemáticas são produzidas por um mecanismo “como a percepção” que muito provavelmente é causal.<sup>lxxiv</sup>

O apoio que Maddy oferece para a analogia da matemática é que

os matemáticos não estão dispostos a pensar que a justificação para as suas afirmações está nas atividades nos laboratórios de física. Ao invés, os matemáticos têm uma gama completa de práticas justificacionais que vão de demonstrações e evidências intuitivas até argumentos de plausibilidade e defesas em termos de consequências.<sup>lxxv</sup>

A descrição de Maddy da prática matemática, se tomada ao pé da letra, apoia (a1); apoia a concepção da matemática como uma disciplina autônoma com seus próprios procedimentos de justificação, alguns dos quais são não inferenciais. A principal consequência epistêmica dessa descrição é que ela coloca em causa a abordagem holista de Quine ao conhecimento matemático. Nada na descrição apoia (a2). O único apoio que Maddy oferece a (a2) é a opinião, se bem que significativa, de um matemático: Kurt Gödel. Mas do fato de um matemático endossar (a2) não se segue que seja apoiada pela *prática da matemática*. Os matemáticos podem ter opiniões sobre questões que não surgem na matemática, e nem todas as questões sobre a matemática surgem da matemática. Algumas surgem da epistemologia tradicional. O que é preciso mostrar é que Gödel está a responder uma questão que surge da prática da matemática e que sua resposta é geralmente aceita na prática da matemática.

Em conclusão, os naturalistas científicos têm de fornecer evidência da ciência que indique que o conhecimento das entidades abstratas é problemático. O nosso exame da posição de Maddy revela

que a evidência que ela oferece é filosófica ao invés de científica. Assim, ela não forneceu uma razão para acreditarmos que o naturalismo científico não acomoda tal conhecimento.

## **5. Rumo à Solução**

Os resultados das seções 3 e 4 são inconclusivos. Nem os proponentes e nem os oponentes do *a priori* oferecem argumentos convincentes a favor de suas posições. Além do mais, suas estratégias são tipicamente negativas: cada um argumenta, principalmente em bases *a priori*, que a posição oposta é deficiente em algum aspecto. O resultado é um impasse. Avançar o debate para além desse impasse requer o fornecimento de evidência a favor de uma das posições que seja convincente a ambas as partes. A estratégia mais promissora para desenvolver a defesa do *a priori* é reunir apoio empírico a favor da tese de que há fontes não-experienciais de justificação.<sup>lxxvi</sup>

Essa estratégia é atraente por duas razões. A primeira é dialética. Um argumento a favor do *a priori* que se baseie em evidências e princípios metodológicos endossados por empiristas radicais tem de ser reconhecido por eles. Essa vantagem dialética persiste ainda que haja algum argumento rival *a priori* e não circular pronto para ser usado. A segunda é estratégica. Ao se limitar a argumentos *a priori*, os proponentes do *a priori* se colocam numa posição desnecessariamente desvantajosa. Eles reconhecem que temos crenças justificadas tanto *a priori* quanto *a posteriori*, embora não empreguem as últimas ao defender sua posição. Na falta de uma objeção voltada ao emprego do apoio *a posteriori*, é simplesmente um erro negligenciá-lo.

Que evidência empírica é relevante para se estabelecer que há fontes não-experienciais de justificação? Antes que os proponentes do *a priori* possam reunir apoio empírico para essa tese, ela tem de ser mais cuidadosamente articulada. Chamemos a isto *Projeto de Articulação* (PA).

Fornecer (a) uma descrição geralmente aceita, pelo menos a nível fenomenológico, dos estados cognitivos que justificam não inferencialmente crenças *a priori*; (b) o tipo de crenças que

justificam; e (c) as condições sob as quais elas justificam as crenças em questão.<sup>lxxvii</sup>

Podemos agora examinar as três componentes de (AP).

Grande parte da controvérsia sobre o *a priori* se concentra nos estados cognitivos que supostamente justificam *a priori*. Os empiristas radicais afirmam que tais estados são enigmáticos ou mesmo misteriosos. Os seus proponentes respondem que são familiares e oferecem descrições fenomenológicas. Contudo, se inspecionamos essas descrições, uma enorme variação é encontrada.

Alvin Plantinga apela a uma analogia com a percepção para caracterizar a fonte do conhecimento *a priori*: “um modo de se acreditar *a priori* que *p* é ver que é verdadeiro”.<sup>lxxviii</sup> Ademais, ele alega que

[Esse “ver”] consiste, primeiro (sugiro), em você estar completamente convencido de que a proposição em questão é verdadeira. Segundo, porém, em você estar completamente convencido que essa proposição não é apenas verdadeira, mas que *não poderia ter sido falsa*.<sup>lxxix</sup>

De acordo com Plantinga, a analogia perceptual pode ser articulada em termos de algum estado cognitivo mais familiar. A reflexão fenomenológica revela que o “ver” que subjaz à justificação *a priori* não é de todo misterioso. Consiste em se estar convencido de que *p* é necessariamente verdadeira.

Plantinga não está sozinho ao recorrer à analogia perceptual para caracterizar a fonte de justificação *a priori*. Laurence Bonjour também apela a tal analogia ao articular a sua abordagem da justificação *a priori*. Ele oferece a seguinte descrição do *insight* racional, a suposta fonte de tal justificação:

quando cuidadosa e refletidamente considero a proposição (ou inferência) em questão, simplesmente sou capaz de ver ou compreende ou apreender que a proposição é *necessária*, que tem de ser verdadeira em qualquer mundo ou situação possível (ou alternativamente, que a conclusão da inferência tem de ser verdadeira se as premissas o forem).<sup>lxxx</sup>

Embora endosse a analogia perceptual, Bonjour discorda de Plantinga num ponto fundamental. Ele insiste que os *insights a priori* são aparentemente *irredutíveis*: “são aparentemente incapazes de serem

reduzidos a ou compostos de alguma constelação de passos discursivos ou elementos cognitivos mais simples de algum outro tipo".<sup>lxxxix</sup> A metáfora perceptual não pode ser articulada em termos de algum estado cognitivo mais familiar. Plantinga, alega Bonjour, simplesmente deturpa os fatos fenomenológicos.<sup>lxxxii</sup>

A despeito das diferenças, Bonjour e Plantinga parecem concordar num ponto: o estado cognitivo que justifica *a priori* a crença de que *p* inclui a crença de que *p*. George Bealer, porém, discorda até desse ponto. De acordo com Bealer, a justificação *a priori* está arraigada na *intuição a priori*:

[Por intuição] não queremos referir a um poder mágico ou a uma voz interna ou a algo do tipo. Pois ter uma intuição de que *A* é simplesmente *parecer* a você que *A*. "Parece" aqui é entendido não como um termo de precaução ou de "proteção", mas em seu uso enquanto termo para um tipo genuíno de episódio consciente. [...] Esse tipo de aparência é, com certeza, *intelectual*, e não sensível ou introspectivo (ou imaginativo). O assunto aqui é a intuição *a priori* (ou racional).<sup>lxxxiii</sup>

Uma aparência intelectual de que *p* tem de ser distinguida de uma crença de que *p*. Por exemplo, pode parecer a alguém que a versão ingênua do axioma da compreensão da teoria de conjuntos é verdadeira embora ele não acredite que seja verdadeiro. Conversamente, há teoremas matemáticos que alguém acredita com base não construção de uma demonstração, mas que não parece nem verdadeiro nem falso.

Ernest Sosa concorda com Bealer que uma intuição de que *p* não precisa envolver a crença de que *p* ou, no que diz respeito ao assunto, qualquer crença afinal. Não obstante, ele sugere que tais aparências poderiam ser analisáveis em termos daquilo que alguém *acreditaria* em certas circunstâncias:

As aparências, portanto, sejam sensíveis ou intelectuais, poderiam ser vistas como inclinações para crença com base na experiência (sensível) direta ou no entendimento (intelectual) e a despeito de qualquer raciocínio paralelo, na memória, ou na introspecção onde os objetos da aparência *intelectual* também se apresentam como necessários.<sup>lxxxiv</sup>

Sosa e Bealer diferem em dois aspectos significantes. Primeiro, oferecem descrições fenomenológicas diferentes das aparências. Sosa sustenta que uma aparência intelectual de que *p* é

uma inclinação para acreditar que  $p$  com base na compreensão de que  $p$ . Bealer insiste que “intuição é atitude proposicional *sui generis*, irreduzível e natural que ocorre episodicamente”.<sup>lxxxv</sup> Assim, Bealer concorda com Bonjour que o estado cognitivo que justifica *a priori* é irreduzível embora discorde dele acerca do caráter desse estado. Sosa, por outro lado, concorda com Plantinga que o estado é redutível a um estado cognitivo mais familiar embora discorde dele acerca do caráter do estado redutor. Segundo, embora ambos concordem que há aparências sensíveis e intelectuais, discordam acerca de como essas aparências diferem.<sup>lxxxvi</sup> Bealer sustenta que as aparências sensíveis e as aparências intelectuais são estados conscientes fenomenologicamente distintos. Ambos envolvem uma inclinação para acreditar que  $p$ , mas diferem na base dessa inclinação: as aparências sensíveis se baseiam na experiência direta, ao passo que as aparências intelectuais se baseiam no entendimento.

Os proponentes do *a priori* enfrentam um dilema. Ou têm acesso introspectivo direto aos estados cognitivos que fornecem justificção não inferencial *a priori*, ou não. Se sim, os proponentes favoráveis à posição deveriam ser capazes de concordar sobre a descrição correta desses estados. Se não, então alguma base alternativa tem de ser oferecida para apoiar a ideia de que há tais estados. A falta de consenso entre os proponentes fornece apoio à alegação dos empiristas radicais de que mais precisa ser dito aqui.

Voltando à segunda componente de (PA), há também uma ampla variação entre os proponentes acerca do escopo das crenças justificadas *a priori*. Essas diferenças tipicamente não estão manifestas em contextos epistemológicos, uma vez que o foco é sobre o estoque de exemplos tais como as proposições da lógica elementar ou da matemática, verdades analíticas simples, e alguns casos familiares de supostas verdades sintéticas *a priori*. Poucos proponentes, no entanto, sustentam que o conhecimento *a priori* se limita a esses casos. Consequentemente, não podem responder efetivamente a questão da transmissão de verdade dos estados cognitivos que supostamente justificam *a priori* focando-se exclusivamente em casos incontroversos. De fato, eles têm de fornecer uma especificação mais completa do domínio de crenças supostamente justificadas por tais estados. Na falta de uma articulação mais completa do escopo do *a priori*, a questão crucial da transmissão de verdade permanecerá uma questão de especulação, apoiada ou rejeitada por evidência anedótica.

Há uma questão acerca do escopo da justificação *a priori* que requer atenção especial. Os exemplos de conhecimento *a priori* tipicamente citados pelos proponentes são verdades necessárias. Mas, como apresentado na seção 3, temos de ser cuidadosos ao distinguir entre o conhecimento do *valor de verdade* de uma proposição necessária como oposto ao conhecimento de seu *estatuto modal geral*. Uma questão crucial surge aqui. Qual o alvo da justificação *a priori*: o estatuto modal geral de uma proposição, seu valor de verdade, ou ambos? Se a justificação *a priori* se estende ao valor de verdade das proposições, duas outras questões surgem. São as crenças acerca do valor de verdade das proposições necessárias e de seu estatuto modal geral justificadas por um único estado cognitivo ou por estados cognitivos diferentes? Pode-se ter uma crença justificada *a priori* de que uma proposição contingente é verdadeira?

A terceira componente de (PA) diz respeito às condições nas quais crenças são justificadas *a priori*. Há dois conjuntos distintos de questões aqui. O primeiro é uma especificação das condições nas quais as crenças são *prima facie* justificadas pelo estado cognitivo proposto como a fonte da justificação *a priori*. Bonjour, por exemplo, sustenta que há certas condições de fundo que têm de ser satisfeitas para que um *insight* racional tenha sua força justificativa: a proposição tem de ser considerada com razoável precaução, a pessoa tem de ter uma compreensão aproximada do conceito de necessidade, e a razão tem de não estar obscurecida pelo dogmatismo ou pelo preconceito.<sup>lxxxvii</sup> Duas questões emergem. A lista está completa? Estão as condições suficientemente articuladas de modo que se possa determinar se foram satisfeitas? Uma condição é que o agente cognoscente tenha uma *compreensão adequada* do conceito de necessidade. Tal compreensão requer familiaridade com os princípios da lógica modal? Um cético modal carece do conhecimento *a priori*.

O segundo é uma especificação das condições nas quais a justificação *a priori prima facie* é revogável. Os revogadores se enquadram em duas amplas categorias: revogadores refutantes e revogadores enfraquecedores. Há duas questões principais no caso dos revogadores refutantes. Primeiro, em que condições, se as houver, conflitos de *insight* racional coloca em causa a justificação baseada em tal *insight*. Segundo, pode haver revogadores refutantes empiricamente justificados para crenças justificadas *a priori*? Questões paralelas surgem no caso dos revogadores enfraquecedores. Um registro de crenças conflitantes ou erros

baseados em *insight* racional coloca em causa a justificação baseada em tal *insight*? Podem crenças justificadas *a priori* serem revogadas por crenças empiricamente justificadas acerca dos processos cognitivos que subjazem ao *insight* racional?

Uma vez que as principais peças do Projeto de Articulação foram postas, o projeto de oferecer evidência empírica a favor do *a priori* pode ser implementado. Chamemos a isto *Projeto Empírico* (PE):

Fornecer (a) evidência de que os estados cognitivos identificados em nível fenomenológico estejam associados aos processos de um tipo único ou tipos relevantemente similares; (b) evidência de que os processos envolvidos desempenhem um papel na produção ou sustentação das crenças que supostamente justificam; (c) evidência de que os processos envolvidos sejam veros condutores; e (d) uma explicação de como os processos envolvidos produzem as crenças que supostamente justificam.

Examinaremos brevemente as quatro áreas de investigação iluminadas por (PE).<sup>lxxxviii</sup>

A tese central dos defensores do *a priori* é que as fontes de justificação são de dois tipos significativamente diferentes: experiencial e não-experiencial. Inicialmente, essa diferença é assinalada a nível fenomenológico. Os proponentes identificam certos estados fenomenologicamente distintos como a fonte da justificação *a priori*. O fato de que os estados são fenomenologicamente distintos, no entanto, não assegura que sejam produzidos exclusivamente por processos de um único tipo, e nem que, caso o sejam, esses processos difiram significativamente dos processos experienciais. Contudo, o caráter dos processos que produzem o estado é relevante para saber se o estado justifica *a priori*.

Por exemplo, suponha que as aparências intelectuais tenham um caráter fenomenológico distinto e prontamente identificável. Além disso, suponha que um tutor ensine a uma criança “ver” que  $4 \times 4 = 16$  utilizando técnicas como as empregadas no *Menon* e também ensine a criança que bolas rolam em planos inclinados tendo a criança feito experimentos com bolas e planos. Finalmente, suponha que a criança se esqueça posteriormente das lições do tutor, mas que, como resultado delas, ambas as proposições, quando consideradas, apareçam com verdadeiras à criança. É implausível sustentar que ambas as crenças sejam justificadas *a priori* para a

criança. A primeira é justificada *a priori*, uma vez que é baseada num processo “racional” ou não-experiencial, mas a segunda é justificada *a posteriori*, uma vez que é baseada num processo perceptivo ou experiencial. Assim, o fato de que alguns estados cognitivos tenham uma fenomenologia distinta, uma fenomenologia diferente daquelas associadas a processos experienciais familiares como a percepção, a memória, ou a introspecção, não assegura que esses estados sejam produzidos por um tipo único de processo e nem que o processo que os produz seja não experiencial.

A segunda área de investigação avalia a tese de que as crenças supostamente justificadas *a priori* sejam produzidas e/ou sustentadas por processos que envolvam o estado cognitivo que supostamente fornece tal justificação. Se uma teoria epistêmica tem de fornecer uma explicação de com as nossas crenças são *de fato* justificadas, então os processos aos quais a teoria apela tem de desempenhar de fato algum papel na aquisição ou sustentação das crenças em questão. A investigação empírica pode oferecer apoio a essa tese. Embora um proponente do *a priori* pudesse se contentar com a tese mais fraca de que os processos em questão *podem* justificar crença *a priori*, considerações empíricas são ainda relevantes de três formas. Primeiro, se a tese mais fraca envolver mais do que uma asserção de mera possibilidade lógica, é necessário evidência para mostrar que os processos cognitivos em questão podem, num sentido mais robusto, desempenhar um papel na produção ou sustentação das crenças em questão. Segundo, o estatuto epistêmico das nossas crenças *efetivas* acerca do assunto em questão tem de ser tratado. Os processos que efetivamente produzem as crenças em questão também as justificam? São as nossas crenças efetivas epistemicamente sobredeterminadas ou injustificadas? Terceiro, uma explicação do porquê os processos não-experienciais não são empregados pelos agentes cognoscentes faz-se necessária. É porque os processos só podem ser empregados por especialistas? É porque os processos são cognitivamente dispensáveis? As respostas a essas questões são necessárias para fornecer uma imagem acurada do papel de tais processos em nossa economia cognitiva.

A terceira área de investigação diz respeito à questão da transmissão de verdade. Essa questão desempenha um papel duplo. Se a transmissão de verdade é uma condição necessária para a justificação epistêmica, como muitos proponentes do *a priori* alegam,<sup>lxxxix</sup> ou, se é uma condição necessária para a justificação *a priori*, como outros alegam,<sup>xc</sup> então para se oferecer evidências em

apoio à tese de que um processo cognitivo particular é uma fonte de justificação *a priori*, tem-se de oferecer evidência em apoio à tese de que as crenças baseadas nesse processo são provavelmente verdadeiras. Mesmo aqueles que negam que a transmissão de verdade seja uma condição necessária para a justificação epistêmica concedem que a evidência a favor de uma fonte particular de crenças que transmite erro revoga a justificação que tal fonte confere às crenças que esse processo produz. Se vamos oferecer evidência em apoio à tese de que um processo particular é uma fonte de justificação *a priori*, temos de oferecer evidência em apoio à tese de que não está disponível qualquer evidência revogadora que coloca em causa a capacidade dessa fonte de justificar algumas crenças. A tese de que um processo é verocondutor ou, pelo menos, que é errocondutor é uma tese geral contingente que pode ser apoiada pela investigação empírica.

A investigação empírica pode desempenhar um segundo papel importante na avaliação das credenciais de um processo cognitivo. A fim de avaliar a verotransmissividade de um processo de formação de crença, é preciso ter alguma proximidade com todo o domínio de crenças que se pode produzir/ou sustentar pelo processo em questão. O *a priori* é tipicamente introduzido e defendido usando-se um domínio estreito de exemplos. Os empiristas radicais geralmente atacam o *a priori* argumentando que alguns desses exemplos, como os princípios da geometria euclidiana, por exemplo, mostram-se falsos. Apontar essas disputas apenas não apoia ou refuta convincentemente o *a priori*, uma vez que o âmbito dos casos em consideração é bastante limitado. As investigações históricas e psicológicas, no entanto, podem fornecer uma imagem mais completa da extensão das crenças produzidas por tal processo.

A quarta área de investigação empírica, que se foca em considerações explicativas, oferece um panorama para o desenvolvimento de argumentos a favor do *a priori* de diferentes frentes. Primeiro, se tal investigação revelar que os processos cognitivos associados aos estados que supostamente justificam o *a priori* são de um único tipo ou de tipos similares em aspectos semelhantes, então a identificação das características distintivas desses processos poderia fornecer a base para a articulação da distinção experiencial/não-experiencial. O resultado seria uma compreensão mais profunda do conceito de justificação *a priori*. Segundo, tais investigações podem fornecer uma compreensão melhor de como os processos em questão produzem crenças verdadeiras sobre seu conteúdo. Essa compreensão, por sua vez, é a

chave para se oferecer uma explicação perceptiva não-causal de como os estados em questão fornecem acesso cognitivo ao conteúdo das crenças que produzem e por que são vericondutores. Terceiro, assim que alcançamos uma melhor compreensão desses processos, nossas teorias epistemológicas e psicológicas tornam-se mais integradas. O fato da nossa teoria epistemológica concordar bem com as teorias psicológicas para as quais temo apoio independente aumenta o apoio geral da primeira teoria.

## **6. Conclusão**

Argumentei a favor de três conclusões principais. A primeira é uma concepção mínima da justificação *a priori*: a justificação *a priori* é justificação não-experiencial. Segundo, os argumentos tradicionais baseados amplamente em considerações *a priori*, tanto a favor quanto contra a existência do conhecimento *a priori* são inconclusivos. Finalmente, a estratégia mais promissora no avanço de argumentos a favor do *a priori* é oferecer evidência empírica a favor da tese de que há fontes não-empíricas de justificação.

# O ANALÍTICO E O SINTÉTICO, O NECESSÁRIO E O POSSÍVEL, O A PRIORI E O A POSTERIORI\*

Scott Soames

Tradução de Luiz Helvécio Marques Segundo<sup>69</sup>

## PRÉVIA DO CAPÍTULO

### 1. *O contexto*

Quine e a herança dos positivistas

### 2. *O argumento da circularidade contra a distinção analítico/sintético*

Tentativa de Quine de estabelecer que a distinção é ilegítima mostrando que a analiticidade só pode ser definida em termos de conceitos que a pressupõe.

### 3. *Avaliando o argumento da circularidade*

---

\* “The Analytic and the Synthetic, the Necessary and the Possible, the A Priori and the A Posteriori”. *Philosophical Analysis in Twentieth Century, vol. 1: The Dawn of Analysis*. Princeton University Press, 2003. Capítulo 16.

<sup>69</sup> Bolsista CAPES, doutorando pela UFSC.

Por que o argumento é bem-sucedido somente se adotamos as concepções dos positivistas de necessidade, aprioricidade e analiticidade.

4. *Uma definição alternativa de sinonímia (e, por conseguinte, de analiticidade)*

Como a sinonímia pode ser definida em termos de substitutividade em construções envolvendo atitudes proposicionais, produzindo uma concepção defensável, porém filosoficamente menos significativa de analiticidade; a resposta radical de Quine.

5. *A resposta a Quine de Grice e Strawson*

Crítica 1: as interpretações forte e fraca

O argumento de que a uniformidade na aplicação do analítico e do sintético demonstra que há uma distinção genuína entre os dois; *insights* e defeitos do argumento.

Crítica 2: ceticismo sobre o significado

Por que a posição de Quine leva a um ceticismo insustentável sobre o significado.

### **O contexto**

Willard van Orman Quine lecionou em Harvard, primeiro como instrutor, depois como professor, de 1936 até a sua aposentadoria aos 70 anos em 1978, continuando ainda a escrever e dar conferências sobre filosofia por mais de vinte anos. Ele começou a sua vida acadêmica estudando lógica, e sua primeira grande publicação filosófica foi o conhecido artigo "Truth by Convention" publicado em 1936. No início da década de 40 ele era uma figura importante no cenário filosófico, especialmente na América. Com a publicação, em 1951, de seu celebrado artigo "Os dois dogmas do empirismo" ele se tornou o filósofo dominante na América, permanecendo até janeiro de 1970, quando Saul Kripke, que estudara

com Quine como graduando em Havard, dera as três conferências em Princeton que se tornaram o *Naming and Necessity*. Mesmo após a emergência de Kripke a influência de Quine na filosofia analítica permaneceu forte por mais de vinte e cinco anos.

No capítulo 12, discutimos o argumento de Quine em “Truth by Convention”, e no volume 2 examinaremos as doutrinas céticas sobre o significado e a referência desenvolvidas em suas principais obras, *Palavra e Objeto*, publicada em 1960, e *Ontological Relativity and Others Essays*, publicada em 1969.<sup>70</sup> Neste capítulo e no próximo nos concentraremos principalmente em seu “Os dois dogmas do Empirismo”, que, dentre todos os seus artigos, foi o mais conhecido e amplamente discutido.<sup>71</sup> Nele Quine oferece uma crítica ao positivismo lógico. Ele isola aquilo que toma como sendo as ideias centrais por trás do positivismo; indica quais daquelas ideias considera corretas e quais considera incorretas; e delinea brevemente as ideias centrais de uma nova perspectiva filosófica que mantém grande parte do espírito e do legado do positivismo, evitando, ele espera, seus problemas fundamentais.

Embora as ideias centrais apresentadas em “Os dois dogmas” tenham sido parte do pensamento de Quine por quase toda a década de 40, o artigo não foi publicado até 1951. Nessa época o positivismo já tinha acabado. Não obstante, muitos filósofos continuaram a acreditar que havia algo correto no positivismo que deveria ser mantido, algo que no fim poderia ter de ser rejeitado. O problema era como separar o bom do ruim, e muitos filósofos viram o artigo de Quine fazendo exatamente isso. Ao olhar agora para a história desse período, não se pode evitar em pensar que uma razão importante de “Os dois

---

<sup>70</sup> Quine, *Word and Object* (Cambridge, MA: MIT Press, 1960/ Trad. Brasileira: *Palavra e Objeto*, Petrópolis: Vozes, 2010); *Ontological Relativity and Other Essays* (Nova York e Londres: Columbia University Press, 1969/ Trad. Brasileira: *Relatividade Ontológica e Outros Ensaio*, in *Os Pensadores*, São Paulo: Abril Cultural, 1975).

<sup>71</sup> Quine, “Two Dogmas of Empiricism”, *Philosophical Review* 60 (1951); reimpresso em Quine, *From a Logical Point of View* (Cambridge, MA: Harvard University Press, 1953, 1961, 1980/ Trad. Brasileira: *De um Ponto de Vista Lógico*, São Paulo: Editora Unesp, 2011). A menos que eu indique o contrário, as citações serão da edição de 1980.

dogmas” ter se tornado tão importante e influente foi ter oferecido uma perspectiva a um grupo de filósofos precisamente num período em que eles precisavam. Ao examinar o artigo tentaremos construir essa perspectiva passo a passo.

A melhor maneira de abordar o artigo é dividi-lo em três partes. A primeira parte inclui as seções 1-4, em que Quine discute e rejeita a distinção entre frases ou afirmações analíticas e sintéticas. A segunda parte consiste na seção 5; onde ele discute as suposições centrais por trás do critério verificacionista (ou empirista) do significado, como também as por trás de uma doutrina que ele chama de *reducionismo* (essencialmente aquilo que chamamos de *teoria das construções lógicas*). Quine sugere quais dessas suposições deveriam ser mantidas, quais deveriam ser rejeitadas, e que novas suposições deveriam ser adicionadas. A parte final do artigo, a seção 6, contém um breve esboço de sua teoria do significado positiva, e de suas perspectivas sobre a natureza da ciência.

Neste capítulo trataremos da parte 1, em que Quine discute a distinção filosófica tradicional entre frases ou afirmações analíticas e sintéticas. Vimos que tanto os positivistas lógicos quanto o Wittgenstein do *Tractatus* dividiam a classe das frases, ou afirmações, dotadas de significado em duas classes – frases analíticas (ou tautologias), que supostamente são verdadeiras em virtude do significado apenas, e frases sintéticas, cuja verdade ou falsidade era concebida como dependendo não apenas do que significam, mas também do modo com o mundo é. Como enfatizamos no capítulo 12, era crucial tanto para Wittgenstein quanto para os positivistas que a distinção entre verdades analíticas e sintéticas deveria coincidir exatamente com a distinção necessário/contingente, e com a distinção *a priori/a posteriori*. De acordo com eles, todas as verdades necessárias e *a priori* são analíticas, e é apenas porque são analíticas que são necessárias e *a priori*.

Para Wittgenstein, a fonte dessa perspectiva repousa em sua tese de que para que uma frase diga algo, para que forneça alguma informação, a sua verdade tem de excluir certos estados possíveis em que o mundo poderia estar. Uma vez que as verdades necessárias nada excluem, nada dizem, e uma vez que nada dizem sobre o modo como o mundo é, o modo como o mundo é não dá qualquer contribuição para a sua verdade. Assim, sua verdade tem de ser devido a seus significados apenas. Para os positivistas todo conhecimento acerca do mundo é dependente da observação e da experiência sensível. Segue-se que uma vez que as verdades *a priori* podem ser conhecidas independente da observação e da experiência sensível, não têm de ser sobre o mundo; e se não nos dizem algo sobre o mundo, sua verdade tem de ser devido aos seus significados apenas. Dada a suposição de fundo de que todas e apenas as verdades *a priori* são necessárias, os positivistas viam sua identificação do *a priori* com o analítico coincidindo com a identificação de Wittgenstein do necessário com o analítico.

Além disso, os positivistas insistiam num tipo de prioridade explicativa; a **razão** para a necessidade ou aprioricidade de uma frase qualquer tem de ser encontrada em sua analiticidade. Eles pensavam que simplesmente não há como explicar o que a necessidade é, como conhecemos alguma verdade como sendo necessária, ou como podemos saber algo *a priori* sem recorrer ao nosso conhecimento de que certas afirmações são verdadeiras em virtude do significado. Assim, de seu ponto de vista, seria melhor que as verdades necessárias e as *a priori* fossem analíticas, uma vez que, se não fossem analíticas, então não se poderia dar qualquer explicação delas. Ironicamente, o peso teórico que os positivistas depositaram na noção de analiticidade deixou a doutrina deles sobre a analiticidade, a necessidade e a aprioricidade vulnerável a uma crítica potencialmente devastadora. Se se pudesse mostrar que a analiticidade não pode desempenhar o papel explicativo que eles a atribuíam, então seu comprometimento como a necessidade, com a

aprioricidade, e talvez até mesmo a própria analiticidade poderia ser ameaçada. Essa foi a estratégia por trás do ataque de Quine.

No capítulo 12 examinamos aquilo que pode ser visto, pelo menos em retrospecto, como a primeira parte desse ataque. Lá esmiuçamos o raciocínio aparente por trás da tese positivista de que o conhecimento de que certas afirmações são verdadeiras em virtude do significado (ou verdadeiras por convenção) pode ser usado para explicar todo conhecimento *a priori*. Concluimos, com base num argumento tirado do “Truth by Convention” de Quine, que esse raciocínio é fundamentalmente falho, pois qualquer explicação dessas pressupõe certo conhecimento *a priori* anterior que não pode ser explicado linguisticamente. Se isso estiver correto, então os positivistas nunca foram bem-sucedidos em mostrar que a analiticidade poderia fazer o trabalho que eles consideravam tão importante. Isso em si não prova que haja algo de ilegítimo com a analiticidade, a aprioricidade, ou a necessidade. No entanto, no contexto das suposições centrais dos positivistas que conectam essas noções, seria o bastante torná-las claramente difíceis de tratar.

Seja por qual for a razão, eles demoraram a aprender a lição do “Truth by Convention”. Assim, quinze anos depois, em 1951, Quine apresentou um novo ataque, mais direto e mais amplo. Ele concordava com a premissa fundamental dos positivistas de que não há como explicar a necessidade e a aprioricidade sem se apelar à analiticidade. No entanto, ele desafiou a ideia de que qualquer distinção genuína poderia ser traçada entre o analítico e o sintético sem pressupor as próprias distinções de que necessitavam para explicar. Ele concluiu, portanto, que não há maneira de explicar e legitimar a necessidade e a aprioricidade – ou mesmo a analiticidade. Isso quer dizer que não há distinção genuína a ser traçada entre o analítico e o sintético, entre o necessário e o contingente, ou entre o *a priori* e o *a posteriori*. De fato, a ideia de que tais distinções existem é um dos “dois dogmas” que são alvos do seu artigo. O ônus das

seções 1-4 do artigo é demonstrar que esse dogma deveria ser rejeitado.

### **O argumento da circularidade contra a distinção analítico/sintético**

Quine começa dizendo que não tentará explicar a analiticidade em termos de necessidade, uma vez que a distinção entre verdade necessária e contingente precisa tanto de explicação quanto a distinção entre o analítico e o sintético (se não mais). Poder-se-ia tentar fazer a distinção dizendo-se que uma frase analítica é uma frase que é **verdadeira em virtude do significado apenas**, ao passo que uma verdade sintética é **verdadeira em virtude dos fatos**. Mas, Quine se pergunta, o que queremos com essa conversa sobre significado? Certamente, pensa ele, não precisamos supor que há **coisas** que são os significados das frases e de outras expressões – entidades obscuras que de algum modo mediam as palavras, por um lado, e os objetos que as palavras representam, ou se aplicam, por outro. Pelo contrário, Quine sustenta que falar de significado não é falar de coisas; ao falar sobre significado queremos saber que frases e outras expressões são dotadas de significado, e quais são sinônimas uma das outras. Isso o leva a pensar que o melhor modo de dar sentido à ideia de uma frase sendo verdadeira em virtude do significado é interpretá-la como a ideia de que uma frase possa ser transformada numa verdade lógica trocando-se sinônimos por sinônimos. Uma vez que uma verdade lógica é aquela que se mostra verdadeira não importa como as palavras não lógicas sejam entendidas, segue-se que qualquer frase que satisfaça essa condição será garantida como verdadeira.

Essa interpretação também dá sentido à outra caracterização tradicional de analiticidade. Tradicionalmente, diversos filósofos fizeram a distinção entre o analítico e o sintético dizendo que uma

afirmação analítica é aquela cuja negação é contraditória. Na interpretação de Quine, uma vez que uma frase analítica é aquela que pode ser transformada numa verdade lógica trocando-se sinônimos por sinônimos, é também uma frase cuja negação pode ser transformada numa falsidade lógica - *i.e.* uma frase logicamente equivalente a uma contradição simples  $A \ \& \ \neg A$  - trocando-se sinônimos por sinônimos. Assim, a interpretação de Quine de analiticidade oferece um modo plausível de entender a doutrina de que a negação de uma frase analítica é contraditória.

### **DEFINIÇÃO PROPOSTA DE ANALITICIDADE**

S é analítica sse (i) S for uma verdade lógica, ou (ii) S puder ser transformada numa verdade lógica trocando-se sinônimos por sinônimos.

Para que essa definição seja bem-sucedida temos de ser capazes de dar sentido a duas noções cruciais: *verdade lógica* e *sinonímia*. Quine toma a noção de uma verdade lógica como não problemática, desde que nos seja dado antecipadamente o inventário das constantes lógicas - *e.g.*, e, ou, não, pelo menos um, a condicional e a bicondicional materiais. Dado tal inventário, podemos definir as *verdades lógicas* como se segue.

### **VERDADE LÓGICA**

S é uma verdade lógica sse é uma instância substitucional de um esquema em que todas as instâncias substitucionais são verdadeiras.

Um esquema, como *S ou não S*, é uma fórmula construída usando-se letras esquemáticas mais as constantes lógicas. Uma instância substitucional de um esquema é uma frase que resulta da substituição das letras esquemáticas por expressões da linguagem - *e.g.*, substituir as letras esquemáticas predicativas por predicados da



linguagem *extensional* é uma linguagem em que as expressões que referem, ou se aplicam, aos mesmos objetos podem sempre ser substituídas umas pelas outras sem alterar os valores de verdade das frases em que elas ocorrem. Exemplos de tais linguagens são a linguagem da lógica formal de Russell, as linguagens da matemática, e, em algumas abordagens, grande parte da linguagem na qual a ciência física é feita. Em contraste, uma linguagem *intensional* é uma linguagem em que a substituição das expressões que referem, ou se aplicam, aos mesmos objetos às vezes altera os valores de verdade das frases em que elas ocorrem. As linguagens naturais, como o Português, são intensionais. Consideraremos primeiro como a definição proposta de sinonímia se sai quando aplicada a uma linguagem extensional, e depois como se sai quando aplicada a linguagens intensionais.

Ao aplicar a definição a uma linguagem extensional, é útil considerar os seguintes pares de expressões.

*o planeta visto no céu matutino*                      *o planeta visto no céu noturno*

*Ben Franklin*    *o primeiro diretor geral dos correios dos Estados Unidos*

*é uma criatura com rins*    *é uma criatura com coração*

A descoberta astronômica mostrou que um e o mesmo planeta é visto no céu matutino (em determinado lugar e tempo) e no céu noturno (em determinado lugar e tempo). Assim, as descrições definidas singulares *o planeta visto no céu matutino* e *o planeta visto no céu noturno* são correferenciais. Numa linguagem extensional uma pode sempre ser substituída pela outra sem alteração do valor de verdade de qualquer frase. O mesmo vale para o nome *Ben Franklin* e a descrição definida singular *o primeiro diretor geral dos correios dos Estados Unidos*. Os predicados *é uma criatura com rins* e *é uma criatura com coração* ilustram outra versão do mesmo ponto. Parece

ser um fato contingente da biologia toda criatura com coração ser uma criatura com rins, e vice-versa. (De qualquer modo, Quine o considera como sendo.) Portanto, os dois predicados se aplicam exatamente aos mesmos objetos. Numa linguagem extensional isso significa que um predicado pode sempre ser substituído pelo outro em qualquer frase sem afetar o valor de verdade.

Segue-se que se a definição proposta de sinonímia for aplicada a uma linguagem extensional, então todos os três pares serão classificados como pares de sinonímias, e as frases (4) e (5) serão consideradas como analíticas.

- 4a. Para qualquer objeto que seja, esse objeto é o planeta visto no céu matutino sse é o planeta visto no céu noturno.
- b. Qualquer descendente de Bem Franklin é um descendente do primeiro diretor geral dos correios dos Estados Unidos.
5. Toda criatura com coração é uma criatura com rins.

Mas esses resultados estão errados. Os defensores da analiticidade argumentariam que uma vez que a verdade de (4) e (5) são contingentes, e não necessárias, e uma vez que não podem ser conhecidas como verdadeiras simplesmente com base no conhecimento dos significados das palavras nelas contidas, não são analíticas.

Suponha, no entanto, que a definição de sinonímia aplicada a uma linguagem intensional (como o Português), e em particular a uma linguagem que contém a construção linguística (6).

6. É uma verdade necessária que...

Note que a substituição dos termos que referem ao mesmo objeto às vezes altera o valor de verdade nesse tipo de construção.

- 7a. É uma verdade necessária que se um planeta é visto no céu matutino, então o planeta visto no céu matutino é visto de manhã. (verdadeiro)

- b. É uma verdade necessária que se um planeta é visto no céu matutino, então o planeta visto no céu noturno é visto de manhã. (falso)

Isso significa que os termos *o planeta visto no céu matutino* e *o planeta visto no céu noturno* são corretamente caracterizados como não-sinônimos pela definição proposta quando aplicada a uma linguagem que contem a construção (6). O mesmo ocorre com os outros pares de expressões que vimos.

- 8a. É uma verdade necessária que Ben Franklin era Ben Franklin. (verdadeiro)  
b. É uma verdade necessária que Ben Franklin foi o primeiro diretor geral dos correios dos Estados Unidos. (falso)
- 9a. É uma verdade necessária que uma criatura com coração é uma criatura com coração. (verdadeiro)  
b. É uma verdade necessária que uma criatura com coração é uma criatura com rins. (falso)

Em geral, a definição proposta de sinonímia terá a consequência de que o compartilhamento da referência (extensão) não é suficiente para a sinonímia quando uma linguagem contiver construções como (6).<sup>74</sup> Em contraste, sinonímias genuínas como *solteiro* e *homem não casado* são substituíveis em (6) sem alteração do valor de verdade.

- 10a. É uma verdade necessária que se alguém é um solteiro, então essa pessoa é um homem não casado. (verdadeiro)  
b. É uma verdade necessária que se alguém é um homem não casado, então essa pessoa é um homem não casado. (verdadeiro)

Assim, a definição as caracteriza corretamente como sinônimos.

Quine conclui disso que a definição proposta é uma explicação adequada da sinonímia, que por sua vez nos permite definir a analiticidade, desde que a linguagem da qual estejamos falando

---

<sup>74</sup> A extensão de um termo singular é a coisa que ela denota, a extensão de um predicado é a classe das coisas às quais ele se aplica.

inclua a noção de necessidade. Mas e quanto a necessidade? O que ela significa? De acordo com Quine,

11. É uma verdade necessária que S.

Significa apenas,

12. A afirmação de que S é analítica.

Mas agora andamos em círculo. Para explicar a analiticidade temos primeiro de dar sentido à sinonímia. Para dar sentido à noção de sinonímia propriamente dita temos de recorrer a uma noção de necessidade anteriormente entendida. Mas para explicar a necessidade, Quine pensa que temos de pressupor a analiticidade, o que significa que não chegamos a lugar algum. Dado qualquer um dos termos na família - *analiticidade*, *sinonímia*, *necessidade* - poderíamos definir os outros. Mas uma vez que não podemos explicar quaisquer desses termos, exceto por usar os outros, e uma vez que Quine pensa que todos eles precisam de explicação, ele conclui que todas essas noções têm de ser rejeitadas.

### **Avaliando o argumento da circularidade**

Começamos com a observação de que o argumento de Quine é eficaz, na melhor das hipóteses, apenas contra as posições que aceitam duas das teses fundamentais dos positivistas.

- T1. Todas as verdades necessárias (e *a priori*) são analíticas.  
(Para todas as frases S, se S expressa uma verdade necessária (*a priori*), então S é analítica.)
- T2. A analiticidade é necessária para explicar e legitimar a necessidade (e a aprioricidade).

O argumento é destinado a mostrar que nenhuma dessas posições pode estar correta, uma vez que o único modo de dar sentido à

analiticidade envolve a pressuposição das próprias noções de necessidade e aprioricidade que pretendia explicar. Ao avaliar esse argumento, e compreender seu impacto, nada é mais importante do que ter em mente esse pano de fundo histórico. Muitos poucos filósofos hoje aceitariam T1 ou T2, que parecem agora decididamente antiquadas. Hoje em dia a perspectiva predominante – substancialmente derivada do *Naming and Necessity* de Kripke (que será discutido no volume 2) – é que a necessidade e a aprioricidade são, respectivamente, noções metafísica e epistemológica que podem ser sustentadas por si próprias; além do mais, embora algumas verdades sejam necessárias e *a priori*, há muitos exemplos de cada uma delas que não são exemplos umas das outras. Quanto à analiticidade, as opiniões variam; muitos hoje presumem que quando frases que contém expressões indexicais (e.g., *eu, agora, aqui, efetivamente*) são excluídas, as verdades analíticas são um subconjunto das verdades que são necessárias e *a priori*. (As coisas se tornam mais complicadas quando os indexais são introduzidos.<sup>75</sup>) Por essas razões, a tentativa dos positivistas de explicar a necessidade e a aprioricidade em termos de analiticidade parece agora muitíssimo errada. O argumento da circularidade de Quine dificilmente se sai melhor. Uma vez que pressupõe a suposição errada dos positivistas de que a necessidade (aprioricidade) e a analiticidade fazem sentido apenas se T1 e T2 estiverem corretas, compartilha o erro deles. No máximo, ele é bem-sucedido em colocar em causa uma concepção particular de analiticidade e um conjunto particular de teses que os positivistas e outros sustentavam.

Para que não pareça uma pequena conquista, é preciso lembrar que T1 e T2 eram bastante populares na época em que Quine escreveu. Por uma coisa, a influência dos positivistas, e dos primeiros positivistas, permaneceu forte, especialmente na América. E por outra, T1 e T2 foram também aceitas por muitos não-positivistas. Não

---

<sup>75</sup> Para uma discussão iluminante dos indexicais, veja David Kaplan, “Demonstratives”, em J. Almog, J. Perry e H. Wettstein, eds., *Themes From Kaplan* (Nova York e Oxford: Oxford University Press, 1989).

apenas suas raízes podem ser traçadas até o *Tractatus*, mas Wittgenstein as manteve em sua filosofia tardia, incluindo as *Investigações Filosóficas*. Além do mais, a identificação da necessidade e da aprioricidade, e a crença na fonte linguística de ambas, persistiu na filosofia da linguagem comum, concentrada em Oxford do final da década de 40 até o início de 60, que foi muito influenciada pelo último Wittgenstein. Todo esse trabalho estava concentrado no alvo do argumento da circularidade de Quine. Assim, o fato de o argumento ter sido uma objeção poderosa à concepção de analiticidade então dominante, e ter sido considerado como tal, não foi uma conquista histórica pequena.

Certamente, houve respostas contemporâneas a Quine que levantaram objeções cruciais importantes. Já que a examinaremos mais adiante neste capítulo, não precisamos, a esta altura, prejudicar o quão bem-sucedido foi o seu argumento da circularidade contra os originalmente atacados. Antes de entrar nessa questão, examinarei a ideia de tentar definir a analiticidade sem presumir T1 ou T2. Que força, poder-se-ia perguntar, as considerações de Quine teriam sobre tais tentativas?

### **Uma definição alternativa de sinonímia (e, por conseguinte, de analiticidade)**

Ao procurar por uma definição alternativa de analiticidade, podemos começar com a noção de sinonímia (numa linguagem particular). Como vimos, ao apresentar o argumento da circularidade, Quine diz que se nos for dada a noção de necessidade, podemos definir a sinonímia em termos de substitutividade, preservando o valor de verdade na construção linguística (6) - *é uma verdade necessária que...*. Em particular, Quine pensa que se pudéssemos dar sentido aos exemplos como (13), e se algumas dessas afirmações fossem verdadeiras, então os predicados A e B contariam como sinônimos, como também seriam os termos singulares *n* e *m*.

- 13a. É uma verdade necessária que todos e apenas os A's são B's.  
b. É uma verdade necessária que  $n = m$ .

No entanto, essa não é a nossa noção normal de sinonímia, como é indicado em (14a-b).

- 14a. É uma verdade necessária que todos e apenas os *triângulos equiláteros são triângulos equiangulares*.  
b. É uma verdade necessária que  $2^{10} = 1024$ .

Embora cada uma dessas frases seja verdadeira, as expressões em itálico **não** são normalmente tomadas como significando a mesma coisa - *i.e.*, como sendo sinônimas.

Uma razão importante pela qual não tomamos essas expressões como sinônimas é que tomamos as afirmações em (14) como sendo descobertas significantes. Uma pessoa poderia saber que um triângulo fosse equilátero sem saber que fosse equiangular. Similarmente, uma pessoa poderia saber que um livro tem 1024 páginas sem saber que o número de páginas desse livro era  $2^{10}$ . O que esses exemplos mostram é que há certas expressões que podem sempre ser substituídas uma pela outra sem alterar o valor de verdade na construção (6), mas que não podem sempre ser substituídas sem alterar o valor de verdade nas construções em (15).

15. x sabe/acredita/pensa/diz que... .

Muitos filósofos sustentam que a nossa noção comum de sinonímia requer que as sinonímias sejam permutáveis não apenas em construções modais como (6), mas também em construções epistêmicas como (15). Assim, eles sustentariam que a definição proposta de sinonímia em termos de substitutividade dará resultados corretos apenas se a substituição nas construções em (15) forem incluídas. Ao serem incluídas,  $2^{10}$  e 1024 são corretamente caracterizados como não sendo sinônimos. Esses filósofos adicionariam que *solteiro* e *homem não casado* passam no teste da

substitutividade e, por isso, são corretamente caracterizados como sinônimos. A ideia é que alguém que acredite que Jones é um homem não casado, acredita por isso que ele é um solteiro, e vice-versa. Acreditar que alguém é um solteiro não requer qualquer inferência; é nada mais do que acreditar que ele é um homem não casado.

De acordo com essa perspectiva, a noção de sinonímia que Quine define em termos de necessidade é uma noção diferente e mais fraca do que a nossa noção comum de sinonímia, que pode ser definida em termos de substitutividade nas construções em (15). Quando a definição de sinonímia é entendida desse modo, ela faz um bom trabalho de capturar a nossa noção de mesmo significado sem pressupor qualquer coisa sobre a necessidade. Assim, se uma frase analítica continua a ser definida como uma frase que pode ser transformada numa verdade lógica trocando sinônimos por sinônimos, então temos uma definição de analiticidade que não pressupõe a necessidade, e escapa, portanto, do argumento da circularidade de Quine.

Com certeza, a concepção de analiticidade que resulta da nova definição é muito mais estrita do que a dos positivistas. Do meu ponto de vista, o melhor modo de entendê-la é vê-la como deixando de lado as teses T1 e T2 que eram bases comuns tanto para Quine quanto para os filósofos que ele estava a criticar. Por exemplo, de acordo com essa definição, poucas verdades necessárias *a priori* da aritmética acabariam sendo classificadas como analíticas. O mesmo vale para teses filosóficas, mesmo quando são verdadeiras. Isso é importante para certa concepção da filosofia. Muitos positivistas lógicos, assim como outros filósofos analíticos da época de Quine, pensavam que uma vez que as teses filosóficas não são empíricas, têm de ser analíticas, caso sejam verdadeiras. O trabalho do filósofo era concebido como consistindo em desenterrar verdades analíticas escondidas, embora significantes, usando o método da análise linguística ou conceitual. Se a analiticidade se mostrar muito mais

estrita do que originalmente concebida, então essa concepção de filosofia é indefensável. Assim, se algo como a posição modificada sobre a analiticidade que acaba de ser delineada for a posição a qual os defensores da analiticidade são forçados a adotar frente ao argumento de Quine, então tem-se de considerar seu argumento como tendo sido substancialmente bem-sucedido – muito embora ele possa ter errado ao pensar que não há uma distinção de todo a ser traçada entre o analítico e o sintético.

No entanto, o próprio Quine não ficou contente com essa vitória limitada. Em “Os dois dogmas do empirismo” ele não discute a estratégia de definir a sinonímia em termos de substitutividade em outras construções além de *é uma verdade necessária que*. No entanto, ele discute essa possibilidade dezenove anos depois em seu livro *Filosofia da Lógica*, publicado em 1970, onde ele discute os exemplos (16) e (17).<sup>76</sup>

16. Necessariamente cordados são cordados.
17. Tom pensa que cordados são cordados.

Ao discutir esses exemplos, Quine usa *cordado* como uma abreviação para *criatura com coração* e *renado* como abreviação para *criatura com rins*. Supostamente é uma verdade da biologia que todas e apenas as criaturas com coração são criaturas com rins, de modo que as duas expressões, *cordado* e *renado*, supostamente se aplicam às mesmas coisas, sem, com certeza, serem sinônimas. É engraçado que Quine indique que *cordado* é a abreviação para *criatura com coração*. O que ele quer dizer, com certeza – embora não o diga – é que como ele usa esses termos, eles são **sinônimos**. Mas se isso é o que ele quis dizer, então tem de haver, afinal, tal coisa como a sinonímia. Assim, o seu próprio exemplo parece pressupor a posição a qual ele usa como exemplo para argumentar contra.<sup>77</sup>

---

<sup>76</sup> Willard Van Orman Quine, *The Philosophy of Logic* (Englewood Cliffs, NJ: Prentice Hall, 1970/ Trad. Brasileira: *Filosofia da Lógica*, Rio de Janeiro: Zahar Editores, 1972), pp. 8-10.

<sup>77</sup> Quine inadvertidamente permitiu que a mesma pressuposição se infiltrasse na seção 2 de “Dois Dogmas”. A seção trata da noção de definição, e o principal ponto de Quine é que uma vez que as definições mais familiares – *e.g.*, as definições de dicionário e as *explicações* do filósofo – ou reportam,

Seja como for, após apontar como a substituição do termo *renado* por uma das ocorrências do termo *cordado* em (16) alteraria o valor de verdade, Quine diz o seguinte:

De fato outros exemplos poderiam ser citados. O exemplo [17] serve tão bem quanto o [16], uma vez que Tom poderia **não** pensar que todos os cordados são renados, embora ainda reconhecesse que todos os cordados têm coração.<sup>78</sup>

O que Quine está dizendo aqui é que no caso de (17), como no caso de (16), a substituição de *renados* por uma das ocorrências de *cordados* pode mudar o valor de verdade – o que significa que poderíamos definir sinonímia em termos de substitutividade em (17), como oposto a (16), e ainda obter o resultado desejado de que *cordado* e *renado* não são sinônimos. Mas depois Quine continua a dizer o seguinte:

E [17] tem a vantagem de repousar numa linguagem mais inocente que [16], com seu imaginado sentido de necessidade. Porém, inocência é uma coisa, clareza é outra. A expressão

---

ou dependem da crença em instâncias pré-existentes de sinonímia, uma tentativa de definir uma verdade analítica como sendo uma que, em virtude das definições de suas palavras, é equivalente a uma verdade lógica não evitaria de repousar na noção problemática de sinonímia. Contudo, num estágio dessa discussão, Quine nota uma exceção à ideia de que as definições repousam numa crença nas instâncias pré-existentes de sinonímia. Nas pp. 25-26 ele diz:

Permanece, contudo, um tipo extremo de definição que não remete de modo algum a sinonímias anteriores: a saber, a introdução explicitamente convencional de novas notações para propósitos meramente abreviativos. Assim, o *definiendum* se torna sinônimo do *definiens* porque foi criado especialmente para o propósito de ser sinônimo do *definiens*. Temos aqui um caso realmente transparente de sinonímia criado por definição; quem dera todas as espécies de sinonímia fossem assim inteligíveis.

Embora o ponto principal aqui esteja de fato livre de objeção a alguém que acredite na sinonímia, Quine aqui parece ter se esquecido que a questão que é central a seu argumento geral não é como as sinonímias são criadas, mas se a noção de sinonímia – *i.e.*, o mesmo significado – faz sentido. A sua posição é que não. Mas se não, então conceder que sinônimos explicitamente estipulados são genuinamente sinônimos é dizer algo inconsistente com a sua conclusão geral. É dizer, penso, que ainda que a dedicação de Quine ao seu propósito argumentativo negativo mais amplo não foi suficiente para evitar que a verdade negada reaparecesse.

<sup>78</sup> *The Philosophy of Logic*, p. 9. Note a substituição implícita aqui. Na frase final da passagem Tom é descrito como (i) não acreditando que todos os cordados são renados embora (ii) acredite que todos os cordados tenham coração – *i.e.*, que todos os cordados são criaturas com coração. Quine considera que isso mostra como a substituição em (17), *Tom acredita que cordados são cordados*, pode mudar o valor de verdade. Isso tudo faz sentido somente se Quine supuser que acreditar que todos os cordados são criaturas com coração é acreditar que todos os cordados são cordados, o que, por sua vez, depende da sua suposição que *criatura com coração* é sinônimo de *cordado*. Novamente, a discussão de Quine pressupõe a sinonímia ao mesmo tempo em que a despreza.

*pensa* em [17], apesar de comum, é herdeira de todas as obscuridades da noção de sinonímia [...] e mais.<sup>79</sup>

Diante disso, essa passagem parece ser uma *reductio ad absurdum* da posição de Quine. Ele diz que *pensa*, e presumivelmente outros verbos como *acredita* e *sabe*, têm a obscuridade da sinonímia e da necessidade **e mais**. Mas ele rejeita as noções de sinonímia e necessidade porque são obscuras. Se as noções de pensar, acreditar e saber são ainda mais obscuras, então, pelo mesmo raciocínio, deveriam também ser rejeitadas. Mas, de acordo com Quine, disso se seguiria que quando dizemos que alguém pensa, acredita ou sabe que tal e tal, estaríamos a dizer algo incorreto, ininteligível e certamente não verdadeiro. Se essa é a sua posição, então certamente é absurda.

Há razões para pensar que essa era a posição de Quine. No período entre a publicação de “Os dois dogmas” e a publicação de *Filosofia da Lógica*, Quine produziu seu livro mais notável e influente, *Palavra e Objeto*, publicado em 1960.<sup>80</sup> Nessa obra ele desenvolveu um argumento independente, baseado numa doutrina chamada de *Indeterminação da Tradução*, que leva à conclusão de que não há tal coisa como significado, referência, ou crença no sentido que comumente entendemos essas noções. A implicação da conclusão de Quine é que na medida em que estamos interessados em descrever acuradamente a realidade, as nossas noções de significado, referência e crença têm de ser substituídas por substitutos comportamentais drasticamente mais fracos e limpos. No volume 2 discutirei essas conclusões em detalhe, e explicarei por que são mal motivadas e em última instância autoderrotantes. Por agora, simplesmente noto que ao discutir a possibilidade de se definir a analiticidade e a sinonímia em termos de substituição em contextos de crença, Quine parecia, em 1970, estar completamente sob a influência dessas conclusões radicais e, por essa razão, disposto ao

<sup>79</sup> *The Philosophy of Logic*, p. 9.

<sup>80</sup> Quine, *Word and Object*.

que de outro modo pareceria ir longe demais para rejeitar qualquer tentativa de definir sinonímia e, por conseguinte, analiticidade em termos de crença.

É uma pena que Quine estivesse disposto a extrapolar, pois há algo que poderia ser feito a seu favor e que não é preciso ir tão longe. Considere (18) e (19).

- 18. Jones é um homem não casado.
- 19. Jones é solteiro.

Certamente, é concebível que se poderia mostrar essas duas frases a alguém, perguntando-lhe se ele acredita no que elas dizem, e receber dele como resposta que ele acredita em (18) mas não em (19). Mas então, poder-se-ia perguntar, como pode a substituição na construção *x acredita que...* ser um teste adequado para a sinonímia? Muitos filósofos responderiam essa pergunta distinguindo entre *dar assentimento sincero a uma frase*, por um lado, e *acreditar naquilo que a frase expressa*, por outro. Por exemplo, se eu perguntasse a Manuel, um falante monolíngue do Espanhol, se ele acredita no que é dito pela frase *A terra é redonda*, ele não saberia se responderia *sim* ou *não*, uma vez que não saberia o que essa frase em português significa. Contudo, provavelmente seria correto descrever Manuel como acreditando que a terra é redonda, especialmente se ele aceitasse a frase correspondente no Espanhol. Consequentemente, o fato de alguém não dar assentimento à frase S nem sempre mostra que esse alguém não acredita naquilo que a frase expressa. Alguém pode acreditar no que S expressa, e mesmo assim não dar assentimento a S porque não sabe o que S significa.

Muitos defensores da sinonímia aplicariam um raciocínio similar a (18) e (19). Eles diriam que qualquer um que dê assentimento a (18) mas não a (19) mostraria com isso que ou não entende a expressão *homem não casado* ou que não entende *solteiro*, ou que não entende ambas. Mas se ele não entende o significado dessas expressões, então seu assentimento ou dissentimento não será um indicador

fiável do que ele de fato acredita. Em particular, se ele dá assentimento a (18), entende o que ela significa, então ele de fato acredita que Jones é um homem não casado. E se ele acredita nisso, por conseguinte acredita que Jones é um solteiro, entenda ele a palavra *solteiro* e dê assentimento a (19) ou não.

É assim que muitos defensores da sinonímia argumentariam. Duas coisas sobre essa posição precisam ser notadas. Primeiro, essa defesa que usa a substitutividade nas construções que envolvem crença para definir a sinonímia depende de se presumir que haja uma distinção genuína a ser feita entre **não aceitar uma frase devido a não se entender o que ele significa** e **não aceitar uma frase devido a não se acreditar no que ela diz**. Embora eu pense que seja razoável presumir que haja tal distinção, Quine a rejeitaria se pressionado, acredito. (Tanto pior para Quine.) Segundo, essa defesa que usa a substituição em construções que envolvem crença para definir sinonímia funcionará, e nos dará os resultados que queremos – como o resultado de que *solteiro* e *homem não casado* são sinônimos – apenas se estivermos certos de que alguém que acredita que Jones é um homem não casado acredita que Jones é um solteiro (mesmo uma pessoa que aceite a frase (18) e rejeite a frase (19)). Mas se nos perguntarmos por que estamos tão certos de que isso é assim, é difícil resistir à seguinte resposta: uma vez que (18) e (19) significam a mesma coisa, alguém que acredita naquilo que uma delas expressa tem de acreditar naquilo que a outra expressa – nesse caso, alguém que acredita que Jones é um homem não casado tem de acreditar que Jones é um solteiro, seja lá o que ele diga. Mas se isso é de fato a nossa razão para pensar que *solteiro* é sempre substituível por *homem não casado* sem alteração do valor de verdade nas construções que envolvem crença, então a nossa definição de sinonímia em termos de substitutividade em tais contextos, pressuporá uma apreensão e uma aplicação anteriores da própria noção que supostamente estamos a definir – o mesmo significado.

Assim, pareceria que há um tipo de círculo aqui afinal, ainda que não seja o que Quine estava interessado.

O que deveríamos concluir disso? A conclusão a ser tirada é que as nossas noções de *crença, conhecimento, asserção, o que algo diz, significado, o que uma frase significa ou diz, e o que uma expressão significa* são interdependentes. As verdades sobre cada uma dessas noções estão ligadas com as verdades sobre as outras. Perguntas, falta de clareza, ou indeterminações envolvendo quaisquer dessas noções traduzem-se em perguntas, falta de clareza e indeterminações sobre as outras. Todas as noções são genuínas e inteligíveis, mas nem a família de atitudes proposicionais – *crença, conhecimento, asserção, o que alguém diz* – e nem a família semântica – *significado, o que uma frase significa ou diz, o que uma expressão significa* – é conceitualmente anterior a outra. Para cada noção há uma distinção genuína entre casos que definitivamente se enquadram a ela e casos que definitivamente não se enquadram. Além disso, para cada noção há casos em que é duvidoso, ou até mesmo indeterminado, se a noção se aplica. Assim, se definirmos analiticidade em termos de sinonímia, e se relacionamos a sinonímia a noções como *crença*, então haverá sempre algumas frases que definitivamente contam como analíticas, algumas que definitivamente não são analíticas, e algumas em que é duvidoso ou mesmo indeterminado se são analíticas. Com certeza, ao dizer isso, fui além tanto do argumento explícito de Quine e do que ele aceitaria. Não obstante, é a isso que o famoso argumento da circularidade de Quine na primeira das quatro seções de “Os dois dogmas” naturalmente leva, uma vez divorciado das suposições de fundo, T1 e T2, que Quine compartilhava com os positivistas.

### **Resposta a Quine de Paul Grice e Peter Strawson**

A resposta crítica a Quine que acabei de delinear não foi dada no período após a publicação de “Os dois dogmas”, em grande parte, suspeito, porque exigiria abandonar as teses T1 e T2, que naquela

altura eram aceitas tanto por Quine quanto por seus oponentes. Além do mais, adotar essa crítica teria significado aceitar uma concepção bastante reduzida de analiticidade, tanto quanto abandonar a concepção de filosofia como análise puramente linguística visando a descoberta de verdades analíticas significantes e iluminantes. A maioria dos oponentes de Quine naquela altura estava comprometida demais com essa concepção de filosofia para abandoná-la. Por essa razão, eles não compreenderiam as críticas mais eficazes que pudessem ser feitas aos seus argumentos.

Não obstante, algumas respostas críticas contemporâneas interessantes e impressionantes foram dadas. Duas em particular distinguem-se por levantar questões das quais há algo importante a se aprender. Ambas foram dadas num artigo de Paul Grice e Peter Strawson chamado “In Defense of a Dogma”, publicado em 1956.<sup>81</sup> A primeira crítica começa com uma tentativa de clarificar qual é de fato a posição de Quine. No primeiro parágrafo de “Os dois dogmas” Quine anuncia que a *“crença numa separação fundamental entre verdades que são analíticas, ou fundadas nos significados independentemente das questões de fato, e verdades que são sintéticas, ou fundadas nos fatos”*, é um dogma sem bases que deveria ser abandonado.<sup>82</sup> No final da seção 4 ele conclui seu argumento contra a distinção dizendo: *“Mas apesar da sua razoabilidade a priori, uma fronteira entre afirmações analíticas e sintéticas simplesmente não foi traçada. Que haja tal distinção a ser traçada de todo em todo é um dogma não empírico dos empiristas, uma artigo de fé metafísico.”*<sup>83</sup> O que é dizer que a ideia de que há uma distinção entre afirmações analíticas e sintéticas não possui bases, e que essa distinção deveria ser abandonada? Qual é precisamente a conclusão de Quine? Grice e

---

<sup>81</sup> H. P. Grice e P. F. Strawson, “In Defense of a Dogma”, *Philosophical Review* 65 (1965), reimpresso em James F. Harris, Jr. e Richard H. Severens, eds., *Analyticity* (Chicago: Quadrangle Books, 1970). Outra importante crítica que incentivou Quine a desenvolver as teses apresentadas em *Word and Object* se encontra em Rudolph Carnap, “Meaning and Synonymy in Natural Languages”, Appendix D, segunda edição de *Meaning and Necessity* (Chicago: University of Chicago Press, 1956). Isso é considerado no volume 2.

<sup>82</sup> “Two Dogmas of Empiricism”, p. 20.

<sup>83</sup> *Ibid.*, p. 37.

Strawson mostram que se pode dar a ela tanto uma interpretação forte quanto uma interpretação fraca.

### ***Crítica 1: As interpretações Forte e Fraca***

De acordo com a interpretação forte, o que Quine afirma é que **não** há distinção entre afirmações analíticas e sintéticas - *i.e.*, nenhuma diferença de todo em todo entre a classe de afirmações a qual os filósofos afixam a etiqueta *analítico* e a classe de afirmações a qual eles afixaram a etiqueta *sintético* (talvez porque não haja realmente afirmações analíticas ou sintéticas). Justamente por isso não há distinção entre expressões sinônimas e não-sinônimas - *i.e.*, não há diferença de todo entre pares de expressões que se diz terem o mesmo significado e pares de expressões que se diz terem significados diferentes. Similarmente, não há distinção entre verdades necessárias e contingentes. De acordo com a interpretação fraca, Quine não está a negar que essas distinções existam. Ao invés, seu ponto é que embora haja diferenças genuínas indicadas por essas distinções, a natureza dessas diferenças e, por conseguinte, das razões para se fazer as distinções, foram mal compreendidas pelos filósofos que falavam delas. De acordo com essa interpretação, **há** um tipo de diferença entre as afirmações que foram caracterizadas como *analíticas* e as que foram caracterizadas como *sintéticas*, mas que os filósofos descreveram mal.

Tendo distinguido essas duas interpretações, Grice e Strawson continuam a argumentar que a perspectiva expressa pela interpretação forte é falsa. Há certamente, eles dizem, uma diferença entre frases ou afirmações analíticas e sintéticas, como também uma diferença entre expressões sinônimas e não-sinônimas. A existência dessas diferenças é exibida pelo fato de que, em cada caso, há uma prática estabelecida caracterizada pelo amplo acordo sobre que exemplos pertencem a uma categoria e quais pertencem a outra. Tome a distinção analítico/sintético, por exemplo. Há um amplo acordo de que *Solteiro são não casados, Triângulos têm três lados,*

*Chove ou não chove*, e *Se Sam deu um relógio a Maria, então Sam deu a Maria um relógio*, pertencem a uma classe, ao passo que *O livro está em cima da mesa*, *Chove em Seattle*, *Há pessoas na sala*, e *Tenho uma mão* pertencem à outra. Além do mais, e isto é crucial, as frases atribuídas às duas classes não formam uma lista fechada. Ao invés, diferentes pessoas familiares com a distinção classificam novas frases aproximadamente da mesma maneira – muito embora possam nunca ter encontrado esses exemplos particulares antes, e certamente não terem dito se são analíticas ou não ao terem aprendido a distinção. Isso mostra que ao aprender a distinção analítico/sintético, as pessoas não memorizam simplesmente uma pequena lista de frases a que os filósofos afixaram etiquetas arbitrárias. Ao invés, adquirem uma habilidade genuína – aproximadamente a mesma para cada pessoa – que as permite diferenciar dois tipos diferentes de frases dada uma lista aberta de novos exemplos.

De acordo com Grice e Strawson, esse fato – de que diferentes pessoas fazem notadamente discriminações similares – precisa de explicação. Certamente, argumentam, a explicação natural é que algumas características das próprias frases têm de trazer à tona juízos similares da parte de diferentes agentes. Algumas características comuns às frases que os agentes classificam como sintéticas têm de ser responsáveis pelo fato de que diferentes agentes as classificam do mesmo modo, e as outras características, comuns às frases que os agentes classificam como analíticas, têm de ser responsáveis pelo fato de que diferentes agentes concordem em agrupá-las conjuntamente. Assim, é errado afirmar que **não** há distinção entre essas duas classes de frases ou afirmações – *i.e.*, entre as afirmações que os filósofos chamam *analíticas* e as que chamam *sintéticas*. Podem muito bem haver sérias questões sobre qual é exatamente a distinção, e como deve ser descrita. No entanto, sem dúvida alguma há uma distinção a ser feita, sustentam Grice e

Strawson. Assim, a perspectiva expressa pela interpretação forte da conclusão de Quine é falsa.

Façamos uma pausa por um momento para avaliar esse argumento antes de prosseguir no exame do que Grice e Strawson têm a dizer sobre a interpretação fraca da conclusão de Quine. O argumento repousa no suposto fato de que as frases classificadas pelos filósofos como analíticas são uma coleção mais ou menos homogênea que seria reconhecida virtualmente por qualquer um que a que se tivesse uma introdução rudimentar à distinção. Sendo filósofos, Grice e Strawson não fizeram qualquer pesquisa empírica para validade essa tese. Ao invés, presumiram simplesmente que sua observação estava além de qualquer dúvida. Embora possa muito bem haver um grão de verdade na suposição deles, há uma razão para se pensar que essas questões são mais complicadas do que eles pensavam.

Quanto ao grão verdade, se começarmos com duas listas - uma consistindo de **verdades necessárias simples, é óbvio que são também conhecíveis a priori**, e a outra de **verdades contingentes simples e óbvias que só são conhecíveis a posteriori** - e então apresentarmos novos exemplos de frases aleatoriamente tiradas de ambas as categorias às pessoas, suspeito que seja bastante provável que encontremos um grau considerável de acordo na classificação dos novos exemplos. É também verdadeiro que esse resultado apoiaria a tese de que há uma distinção a ser traçada entre as afirmações agrupadas numa classe pelos falantes e as afirmações agrupadas na outra. Se, enquanto filósofo, você simplesmente tomar por garantido que o necessário, o *a priori*, e o analítico são um e o mesmo, e para serem contrastados como o contingente, o *a posteriori*, e o sintético, que são também um e o mesmo, então você poderia naturalmente tomar essas observações como fornecendo apoio à sua tese de que tem de haver uma distinção a ser traçada entre o analítico e o sintético.

No entanto, há duas sérias restrições que limitam fortemente a força desse argumento. Primeiro, como repetidamente tenho mostrado, essas identificações – do necessário, do *a priori* e do analítico, por um lado, e do contingente, do *a posteriori* e do sintético, por outro – não são nem inevitáveis, e por fim, nem mesmo naturais. Ao invés, eram artefatos paroquiais de um período particular na filosofia analítica. Da perspectiva de hoje em dia, podemos ver que nem todas as verdades necessárias são *a priori*, nem todas as verdades *a priori* são necessárias, e nem todos os membros de ambas as classes são assim transparentes. Em muitos casos é preciso análise e argumentos cuidadosos para se chegar à classificação correta. Por essas razões, simplesmente **não** se espera que os falantes comuns, dadas apenas introduções rudimentares da distinção necessário/contingente, e da distinção *a priori/a posteriori*, e oferecidos novos exemplos aleatoriamente selecionados das quatro categorias resultantes, classificasse-as com altos graus de exatidão e uniformidade.

Segundo, há limites àquilo que pode ser concluído mesmo do teste experimental mais cuidadoso das uniformidades nos juízos dos falantes do tipo que Grice e Strawson imaginam. Suponha que P é alguma palavra ou expressão de uso comum, e que um teste do estilo de Grice e Strawson de uniformidade nos juízos do falante revelou que os falantes aplicam P de modo fiável a novos casos aleatoriamente selecionados de maneira bastante uniforme. Isso mostraria (i) que há uma distinção genuína entre objetos aos quais os falantes aplicariam P e objetos aos quais não aplicariam, e, por conseguinte, (ii) que há propriedades que os objetos possuem na primeira classe que os distinguem dos objetos na segunda classe. No entanto, a uniformidade dos juízos do falante não mostraria sempre (iii) que essas propriedades constituem o significado de P, (iv) que P se aplica de fato aos objetos que os falantes chamam de P, ou (v) que a frase **Há P's é verdadeira**. Para ver isso, imagine que P é o predicado *é uma bruxa*, que os falantes numa certa comunidade aplicam esse predicado de modo fiável a certos tipos de mulheres e

não a outras, e que é parte da definição de *bruxa* que ser uma é ser uma mulher cujo pacto com o Demônio dá-lhe poderes sobrenaturais. Embora de acordo com essa interpretação não há bruxas de fato, há uma distinção genuína entre indivíduos aos quais os falantes aplicariam a palavra (talvez por causa de alguma vocalização ou comportamento que pareça suspeito) e indivíduos aos quais não aplicariam. Pelo mesmo raciocínio, o quiniano poderia dizer que o mero fato de haver uniformidade nos modos pelos quais os falantes aplicam termos como *analítico* e *necessário* a novos casos (supondo que isso poderia ser estabelecido) não mostra que haja quaisquer verdades analíticas ou necessárias, ou que haja qualquer distinção genuína entre o analítico e o sintético, ou o necessário e o contingente – ainda que haja uma distinção entre frases que os falantes chamariam *analíticas* e frases que chamariam *sintéticas*, ou frases que eles chamariam *necessárias* vs. frases que chamariam *contingentes*.<sup>84</sup>

Essa resposta quiniana está correta nesse aspecto; no entanto, não é a última palavra. Um fator crucial no exemplo da *bruxa* é que a estipulação de que para ser uma bruxa uma mulher tem de satisfazer uma condição que ninguém de fato satisfaz. É apenas porque os falantes têm crenças falsas, que são aceitas como definicionais, que a distinção genuína entre mulheres que são chamadas de *bruxas* pelos falantes e mulheres que não são se traduz numa distinção genuína entre bruxas e não-bruxas. Para aplicar o mesmo raciocínio ao *analítico*, o quiniano teria de mostrar que os falantes têm crenças similarmente falsas, que são propriamente consideradas como definicionais de *analítico*; e o argumento da circularidade não faz isso – a menos que seja construído no caso de os falantes aceitarem as teses filosóficas T1 e T2, o que é bastante improvável a menos que sejam filósofos profissionais.<sup>85</sup> O desfecho de tudo isso é que nem

---

<sup>84</sup> Essa é, essencialmente, a resposta de Gil Harman ao argumento de Grice e Strawson. Agradeço a Jeff Speaks pela profícua discussão sobre isso.

<sup>85</sup> Além disso, o quiniano teria de explicar como, de acordo com a sua posição extrema, algo pode ser definicional de algo mais.

Quine e nem seus críticos, Grice e Strawson, são completamente bem sucedidos. O argumento da circularidade não estabelece que não haja distinção analítico/sintético, e nem quaisquer crenças falsas sobre a analiticidade (à parte de T1 e T2). O argumento de Grice e Strawson sobre a uniformidade da aplicação do falante estabelece no máximo a suposição de que há alguma distinção a se fazer, mas não exclui a possibilidade de que possa não haver verdades analíticas. Com T1 e T2 fora de jogo simplesmente nada havia para se extrair desse argumento por si próprio.

Isso nos deixa com uma avaliação mista do argumento de Grice e Strawson contra perspectiva expressa pela interpretação forte da conclusão de Quine. Por um lado, o argumento deles da classificação uniforme de uma lista aberta de exemplos é falho por não notar as complicações que lhes eram invisíveis devido a compartilharem as pressuposições problemáticas de Quine, e por não perceber completamente as complicações do passo da classificação uniforme envolvendo um par de termos à existência de uma distinção genuína indicada por esses termos. Por outro lado, pode muito bem ser possível construir versões restritas do argumento deles que conferem um grau de apoio à tese de que há uma distinção a ser feita entre o necessário *a priori* e o contingente *a posteriori*, e uma distinção diferente entre expressões sinônimas e não-sinônimas.<sup>86</sup> Se pudermos fazer isso para apoiar a ideia de que há uma distinção genuína entre o analítico e o sintético, tudo vai depender de como a analiticidade é definida, e do que ela é usada para explicar. Grice e Strawson não nos ajudam com isso.

Com esse veredito sobre o argumento deles contra a interpretação forte da conclusão de Quine sob nossa inspeção, podemos passar para o que eles dizem sobre a interpretação fraca de sua conclusão. De acordo com essa interpretação, **há** uma distinção entre o analítico

---

<sup>86</sup> Mesmo no caso de sinônimos e não-sinônimos há razão para pensar que a distinção não é inteiramente transparente aos falantes comuns. Veja o capítulo 3 do meu *Beyond Rigidity* (Nova York: Oxford University Press, 2002).

e o sintético, o necessário e o contingente, e o sinônimo e o não-sinônimo, mas foram amplamente mal interpretadas e descritas. De acordo com a interpretação fraca, esse era o objetivo de Quine. No entanto, se era esse seu objetivo, então Grice e Strawson perguntam, *O que exatamente está errado como os modos pelos quais os outros filósofos entenderam e descreveram essas distinções?* É errado, de acordo com a interpretação fraca, sustentar que as verdades analíticas se reduzem a verdades lógicas pela troca de sinonímias por sinonímias? Sustentar que são é, com certeza, apelar à noção de sinonímia. Mas, de acordo com a interpretação fraca, está tudo bem, pois, de acordo com essa interpretação, há uma distinção genuína entre expressões que são sinônimas e expressões que não são. Concedido isso, poderíamos querer ter uma compreensão mais completa e exata da sinonímia, da necessidade, da analiticidade, mas isso é outra questão.

Grice e Strawson concluem que, de acordo com a interpretação fraca de Quine, o máximo que ele mostrou com o argumento da circularidade é que essas noções formam uma família de noções interdefiníveis. Se ele estiver correto sobre isso, então não é errado dizer que as verdades analíticas se reduzem a verdades lógicas trocando-se sinonímias por sinonímias; não é errado dizer que a sinonímia é definível em termos da necessidade; e assim por diante. Ao invés, todas essas afirmações são verdadeiras. Assim, se admitirmos que todas essas distinções existem, e que é apenas uma questão de corrigir as afirmações falsas que os filósofos têm feito sobre as distinções, então temos também de concluir que o argumento da interdefinibilidade de Quine nas seções 1-4 não identifica tais afirmações. Essa é a lição que Grice e Strawson tiraram da interpretação fraca.

Novamente, embora tenham um objetivo, é preciso qualificá-lo. Certamente que mostrar que um conjunto de noções é interdefinível não é, em geral, mostrar que essas noções sejam questionáveis. Nem

é mostrar que foram descritos ou entendidos de maneira errada, a menos que uma dessas noções tenha sido tomada como conceitualmente anterior às outras, e consideradas como constituindo a base para endente-las. No entanto, é assim que a analiticidade foi tratada, não só pelos positivistas, mas pelos filósofos da linguagem comum, a qual Grice e Strawson estavam ligados. Uma vez que foi esse o alvo de Quine, seu argumento da interdefinibilidade era o objetivo.

Reiterando, o máximo que se pode dizer do argumento da intedefinibilidade é que ele mostra que a conjunção de T1 e T2 é falsa - não é o caso que todas as verdades necessárias (e *a priori*) são analíticas e que a analiticidade pode ser usada para explicar e legitimar a necessidade (e a aprioricidade). Naturalmente, se não se aceitar T1 e T2 desde o início - como muitos filósofos agora não fazem - não se concluirá que o argumento de Quine estabelece que não há distinção analítico/sintético, e nem que estabelece que as outras distinções não existam. Pode-se conceder, no entanto, que ele mostra que certa concepção da relação entre o analítico, a necessidade e a aprioricidade é incoerente. Embora esse resultado seja muito menos ambicioso do que o que Quine estabeleceu a si mesmo, não é pouca coisa. A razão pela qual Grice e Strawson erraram é que eles compartilhavam (erradamente) as suposições problemáticas dele, T1 e T2, embora continuassem (corretamente) convencidos de que há de fato algumas verdades necessárias, algumas verdades *a priori*, e algumas expressões sinônimas.

### ***Crítica 2: Ceticismo sobre o Significado***

A segunda crítica que Grice e Strawson fazem ao argumento da circularidade se foca no significado e na sinonímia ao invés da analiticidade e da necessidade. O ponto deles é que o ceticismo de Quine sobre a sinonímia conduz diretamente a um ceticismo absurdo sobre o significado em geral. A esta altura, parece-me, a crítica deles além de precisa, foi historicamente profética. Eles argumentaram que

é absurdo rejeitar a noção de sinonímia como ininteligível porque isso exigiria rejeitar completamente a noção de significado – uma posição que é obviamente insustentável. O argumento que estabelece essa concepção é simples. Se as expressões podem ter significados de todo em todo, então certamente tem de haver, em princípio, respostas verdadeiras à pergunta *O que essa ou aquela expressão significa?* Mas se há respostas verdadeiras a essas perguntas, então podemos identificar as expressões sinônimas com aquelas cujas respostas a essas perguntas são as mesmas. Eis o que Grice e Strawson dizem num contexto de algum modo mais amplo.

Dizer que duas expressões, x e y, são cognitivamente sinônimas parece corresponder, em alguma medida, aproximadamente àquilo que comumente exprimiríamos ao dizer que x e y têm o mesmo significado ou que x significa o mesmo que y. Se Quine for consistente com sua adesão à tese extrema [a interpretação forte], então parece que tem de sustentar não apenas que a distinção que supomos ser distinta pelo uso dos termos “analítico” e “sintético” não existe, mas também que a distinção que supomos ser distinta pelo uso das expressões “significa o mesmo que”, “não significa o mesmo que” também não existe. Pelo menos ele tem de sustentar isso na medida em que a noção de *significa o mesmo que*, em sua aplicação a expressões predicativas, supostamente difere e vão além da noção de *ser verdadeiro para os mesmos objetos*. [...] Contudo, a negação de que a distinção (tomada como diferente da distinção entre o coextensional e o não-coextensional) realmente exista é extremamente paradoxal. [...] Mas o paradoxo é mais violento que isso. Pois frequentemente falamos da presença ou falta das relações de sinonímia entre tipos de expressões – por exemplo, conjunções, partículas de muitos tipos, frases completas – em que não parece haver qualquer substituto óbvio para a noção comum de sinonímia no modo em que se diz que a coextensionalidade é um substituto para a

sinonímia de predicados. **Toda essa conversa é destituída de significado? Toda essa conversa de tradução correta ou incorreta de frases de uma linguagem em frases de outra é destituída de significado?** É difícil acreditar que seja. Mas se nos esforçamos com êxito para acreditar nisso, temos ainda abdições mais difíceis diante de nós. **Se falar de sinonímia frásica é destituído de significado, então parece que falar de frases tendo um significado de todo tem também de ser destituído de significado. Pois se faz sentido falar de uma frase tendo um significado, ou significando algo, então presumivelmente faria sentido perguntar “O que isso significa” e se faz sentido perguntar “O que isso significa” para uma frase, então a sinonímia frásica poderia ser aproximadamente definida como se segue: Duas frases são sinônimas se, e somente se, qualquer resposta verdadeira à pergunta “O que isso significa?” feita a uma delas, for uma resposta verdadeira à mesma pergunta feita a outra.**<sup>87</sup>

O argumento aqui é poderoso: podemos abandonar a sinonímia apenas se estivermos dispostos a abandonar inteiramente o significado e a tradução. Além do mais, Quine parece ter sentido seu poder, uma vez que quatro anos após essa crítica ter aparecido, ele publicou *Palavra e Objeto*, onde ele defende – de maneira um tanto ambígua – o abandono do significado e da tradução inteiramente. Temos de esperar até o volume 2 para explicar essa escolha desastrosa.

---

<sup>87</sup> Grice e Strawson, “In Defense of a Dogma”, pp. 60-62, em *Analyticity*, minha ênfase.

# A descoberta do mundo: O Ser Intemporal\*

Reinhardt Grossman  
Tradução de L. H. Marques Segundo<sup>88</sup>  
Revisão de Renato Mendes Rocha

## *A natureza da classificação*

A ontologia pergunta e tenta responder duas questões relacionadas. Quais as *categorias* do mundo? E quais são as *leis* que regem essas categorias? Na química, por comparação, procuramos pelos elementos químicos e pelas leis da química; na física, pelas partículas elementares e suas leis. As categorias são para a ontologia o que esses blocos de construção básicos são para as ciências naturais. Mas a ontologia não é uma ciência entre outras ciências. Para ver como a ontologia difere da ciência, temos de, em primeiro lugar, entender a noção de uma categoria. Nossa primeira questão é, portanto: o que é uma categoria?

Os filósofos gregos, como por exemplo, Empédocles (no século quinto a.C), desenvolveram a teoria de que tudo é composto de quatro elementos, terra, água, fogo e ar. Esses quatro tipos básicos de coisas se combinam em várias proporções para compor, por exemplo, a cadeira sobre a qual estou sentado, um cabelo em minha cabeça, e o sol que brilha fora de minha janela. Como esses elementos se distinguem um dos outros? Bem, o fogo, é claro, é

---

\* Capítulo 1 de *The Existence of the World: An introduction to ontology*, Routledge, 1992.

<sup>88</sup> Marques Segundo é bolsista CAPES, doutorando pela UFSC. Renato Mendes Rocha é professor na UFC

quente, e a água é úmida. O que quer que seja quente, o que quer que tenha essa *propriedade* é fogo; e o que quer que seja úmido é uma porção de água. O caso do ar e da terra não é intuitivamente tão claro. Mas uma vez que esses filósofos também acreditavam que os quatro elementos formavam opostos, a resposta aceita foi a de que o ar, como o oposto da água, é seco, e que a terra, como o oposto do fogo, é fria. O que quer que agora possamos pensar sobre essa amostra rudimentar de especulação química, duas coisas sobressaem-se. Primeiramente, estamos de fato lidando com uma teoria da química, simplista como nos parece hoje. Segundo, os quatro elementos são distinguidos pelas *propriedades características* que alegadamente têm: o fogo por ser quente; a água por ser úmida; o ar por ser seco; e a terra por ser fria.

A química percorreu um longo caminho desde o tempo de Empédocles. A última vez que olhei num livro de química havia 106 elementos. Esses elementos não foram distinguidos uns dos outros por propriedades como a quentura ou a umidade, mas por propriedades completamente diferentes. O hidrogênio, por exemplo, é um gás em temperaturas comuns, mas se liquefaz a baixas temperaturas; e tem certa densidade para uma dada temperatura e pressão. O *princípio* da classificação dos elementos, no entanto, é o mesmo como no tempo de Empédocles: *As coisas são distinguidas umas das outras por meio das propriedades que têm*. Chamemos isso de “o princípio de classificação”. Esse princípio não se restringe à química. Qualquer que seja a classificação de coisas individuais, sejam elas elementos químicos, partículas elementares, plantas, animais, pessoas, ou o que quer que seja, repousa em uma distinção entre essas coisas individuais por um lado, e de suas propriedades por outro. As baleias, por exemplo, são classificadas não como peixes, mas como mamíferos, por causa da propriedade de gestar seu filhote.

Alguém deve ter percebido que a base de todas as classificações de coisas individuais, nomeadamente, a distinção entre

essas coisas e suas propriedades, é ela própria uma classificação. Mas é uma classificação não de *coisas individuais* – porções individuais de água ou terra, ou bocados de ouro ou ferro, ou baleias individuais ou carpas – mas de *entidades* em geral. É uma classificação de qualquer tipo de existente. Divide tudo o que há em dois amplos grupos de existentes, nomeadamente, em coisas individuais por um lado e suas propriedades por outro. Toda classificação “comum” repousa nessas propriedades. A fim de distinguir essas classificações de todas as outras, devemos falar de “categorização”. As *entidades*, como diremos, são *categorizadas*. Os tipos de coisa que a categorização distingue são então chamados “categorias”. Sabemos que há pelo menos duas categorias, isto é, dois tipos de entidades (existentes), nomeadamente, as *coisas individuais* e as *propriedades* das coisas individuais.

Certas coisas individuais – coisas muito pequenas! – são classificadas como elétrons, pósitrons, nêutrons, etc. Outras são classificadas como ferro, hidrogênio, oxigênio, etc. Outras ainda são classificadas como mamíferos, répteis, pássaros, etc. E por aí vai. Essas, como já observado, são classificações de coisas individuais por meio de suas propriedades. Mas podemos também classificar, não coisas individuais, mas coisas em geral – o que chamei de “existentes” ou “entidades” – em dois grupos de coisas individuais e propriedades de coisas individuais. A fim de distinguir essas classificações fundamentais daquelas que são construídas em cima delas, falei de uma categorização. Mas essa categorização levanta imediatamente uma importante questão: uma vez que todas as classificações repousam sobre o princípio da classificação, nossa categorização tem de repousar sobre esse princípio, e temos de perguntar: que propriedade (ou propriedades) distingue as categorias de coisa individual e propriedade de coisa individual? Como as coisas individuais em geral diferem das propriedades das coisas individuais? A essa questão, outro filósofo grego ofereceu a resposta mais surpreendente.

*Os dois reinos de Platão*

Platão (cerca de 427-347 a.C) fez da distinção entre as coisas individuais e suas propriedades uma pedra angular de sua filosofia. Distinguiu entre dois reinos: o reino das coisas individuais *mutáveis* e o reino das propriedades *imutáveis*. Ele não usou esses termos, mas sua perspectiva acaba por dar na mesma coisa, nomeadamente, que indivíduos são diferentes de suas propriedades pelo fato de que o primeiro, mas não o último, pode mudar. Considere uma maçã comprada ontem e guardada em minha geladeira. Há um mês ou mais ela era muito menor do que é agora. Desde então ela tem mudado seu tamanho. Mudou também sua cor: há algum tempo, quando ainda não estava madura, era verde; agora é vermelha. Temos aqui uma coisa individual mutável. Mas considere agora certa tonalidade de cor, certa tonalidade de vermelhidão que a maçã acabou por ter agora. Essa tonalidade de vermelhidão muda? Não vejo como muda ou mesmo como poderia mudar. É claro, a cor *da maçã* pode mudar: primeiro a maçã era verde, agora é vermelha. Mas essa é uma mudança na cor da maçã, não uma mudança da cor vermelha em si. Como seria para a tonalidade de cor mudar? Bem, essa tonalidade particular de vermelho é uma cor. Tem a propriedade de ser uma cor. Para mudar, poderia trocar sua propriedade de ser uma cor por outra propriedade, assim como a maçã mudou de ser verde para ser vermelha. Por exemplo, a tonalidade de cor poderia cessar de ter essa propriedade e poderia adquirir a propriedade de ser uma forma. Tão logo que pomos deste modo, vemos que a cor talvez não pudesse mudar. Essa tonalidade de vermelho talvez não possa mudar de ser uma cor para ser uma forma.

Deixe-me citar duas passagens do *Phaedo* de Platão em que ele faz a distinção de coisas individuais e suas propriedades, e em que ele afirma que as propriedades, em distinção às coisas individuais, não mudam. O contexto da primeira passagem é este: Sócrates, que fala através de Platão, argumenta que a alma é imortal e que

sabemos coisas antes de nascermos. Para construir seu exemplo, Sócrates compara coisas iguais com a propriedade da igualdade, e avança à afirmação de que a alma conhece a propriedade antes do nascimento:

Veja agora se isto é verdadeiro, continuou. Não acreditamos na existência da igualdade – não a igualdade das peças de madeira ou de pedra, mas algo além delas – da igualdade em abstrato? Deveríamos dizer que há tal coisa ou não?

De fato que sim, disse Simmias, devemos dizer mais enfaticamente.

E sabemos o que essa igualdade abstrata é?

Certamente, respondeu.

D'onde obtivemos conhecimento disso? Não foi vendo as peças iguais de madeira, e de pedras, e as demais, das quais falávamos ainda agora? Não formamos delas a ideia de igualdade abstrata, que é diferente delas? Ou pensas que não é diferente? Considere a questão desta maneira. Peças de madeira e pedra não nos parecem iguais e às vezes desiguais, apesar do fato delas permanecerem as mesmas todo o tempo?

Certamente que sim.

Mas parece-te que os iguais absolutos sejam desiguais, ou que a igualdade abstrata seja desigual?

Não, nunca Sócrates.

Então as coisas iguais, disse, não são o mesmo que a igualdade abstrata?

Não, certamente que não Sócrates? (Phaedo, 78)

O exemplo de Platão aqui é a “propriedade” da igualdade, e claramente ele distingue entre coisas iguais e “igualdade abstrata”, isto é, da *propriedade* da igualdade. Na próxima citação Platão afirma que as propriedades abstratas são imutáveis, ao passo que as coisas individuais mudam:

A igualdade absoluta, a beleza absoluta, e todas as outras existências absolutas admitem qualquer mudança? Ou a existência absoluta em cada caso, sendo essencialmente uniforme, permanece a mesma e imutável e nunca, em qualquer caso, admite qualquer tipo ou espécie de mudança?

Tem de permanecer a mesma e imutável, Sócrates, disse Cebes.

E as muitas coisas belas como homens, cavalos, roupas, e os demais, e de tudo o que compartilha os nomes de ideias, se iguais, ou belos, ou qualquer outra coisa mais? Permanecem o mesmo ou acontece exatamente o oposto com elas? Em suma, nunca permanecem o mesmo de modo algum em si mesmas ou em suas relações?

Essas coisas, disse Cebes, nunca permanecem as mesmas. (Phaedo, 78)

A fim de obter uma diferença mais fundamental entre os indivíduos mutáveis e as propriedades imutáveis temos de considerar a natureza da mudança. A maçã de nosso primeiro exemplo mudou sua cor de verde para vermelho. Isso significa que num momento foi verde, enquanto noutro momento (posterior) foi vermelho. Para uma coisa individual mudar, tem de ter diferentes propriedades em *diferentes momentos de sua existência*. A mudança, portanto, pressupõe que uma coisa perdura no tempo, que existe no tempo, que tem uma duração. Apenas as coisas que existem no tempo podem talvez mudar. Segue-se disso que todas as coisas individuais

têm de existir no tempo. Têm de ser, como direi para encurtar, *temporais*. Se uma coisa for temporal, tem de sofrer mudança? Platão parece ter pensado que sim. Ele parece ter sustentado não apenas que coisas individuais têm de ser temporais, mas também que as propriedades têm de ser atemporais. Se assim for, segue-se, então, que as propriedades são atemporais; não estão no tempo; não existem no tempo; não têm duração. Presumirei que Platão está correto: *Todas as coisas individuais são temporais, enquanto que todas as propriedades são atemporais.*

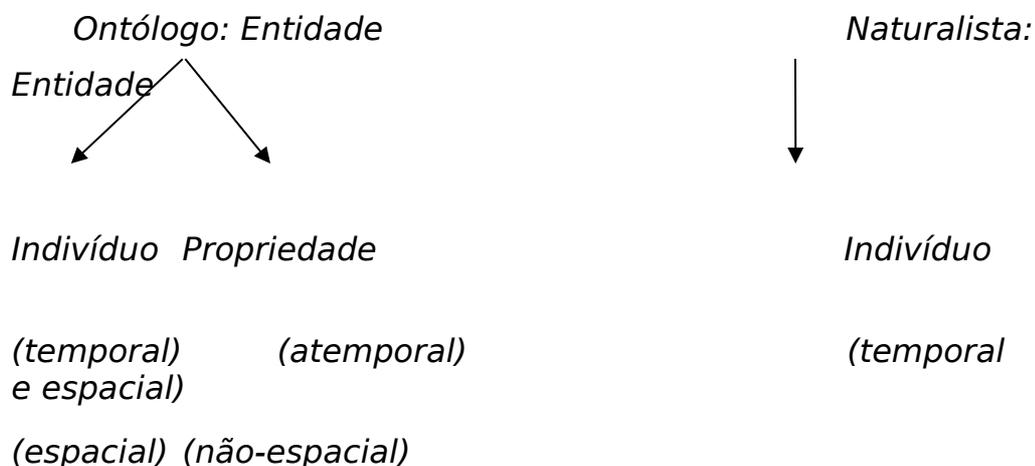
De acordo com Platão, como vimos, há dois reinos: o reino das coisas temporais, das coisas que existem no tempo, e o reino das coisas atemporais, das coisas que não existem no tempo. Ao primeiro reino pertencem as coisas individuais à nossa volta; ao segundo, suas propriedades.

A questão que naturalmente surge é a de saber se é também o caso que todas as coisas individuais que estão no espaço são espaciais, ao passo que todas as propriedades que não existem no tempo não são espaciais. Em outras palavras, a distinção entre coisas temporais e atemporais coincide com a distinção entre coisas espaciais e não-espaciais? A maçã, nosso exemplo de indivíduo, obviamente existe no espaço: está localizada em diferentes momentos em diferentes lugares; primeiro na árvore, três milhas ao sul de minha casa, e agora em minha geladeira, duzentas milhas a sudeste de Chicago. Tem também certas propriedades espaciais. Por exemplo, era pequena em tamanho quando estava verde e ainda crescia na árvore, enquanto que agora está muito maior. Também tem certa forma agora; é aproximadamente esférica. Em suma, num dado momento a maçã estava localizada no espaço e tinha uma forma e um tamanho. Todas as coisas individuais, no que diz respeito a isso, são como a maçã? Todas estão localizadas no espaço e todas têm forma e tamanho? Há muitos filósofos, e eu sou um deles, que acreditam que há coisas individuais que não são espaciais. Um

exemplo seria um pensamento; por exemplo, o pensamento de que me esqueci de levar o meu almoço para o escritório. Agora, esse pensamento me ocorre num dado momento; está localizado no tempo. Ocorreu cinco para o meio-dia, logo quando estava pronto para almoçar. Está localizado no tempo, mas não está no espaço. É claro que não tem forma e nem tamanho. O pensamento de que me esqueci de levar o meu almoço não é nem redondo nem quadrado. Nem tem certo comprimento ou certo diâmetro. Está localizado no espaço? Penso que não. Esse pensamento não está nem “na minha cabeça” e nem ocorre a muitas milhas ao sul de Chicago. O que está em minha cabeça propriamente dito não é meu pensamento, mas meu cérebro e tudo aquilo que se passa em meu cérebro, toda a química e todos os outros tipos de processos. Pode-se apenas se concluir que o pensamento em si está localizado “em minha cabeça” se se aceitar que o meu pensamento é *idêntico* a tudo o que acontece em meu cérebro. Mas quanto a isso há uma longa e complicada disputa filosófica. Não podemos discutir essa questão agora. Desejo apenas apontar que de acordo com alguns filósofos as coisas temporais são de dois tipos: algumas coisas, como a maçã, são espaciais, ao passo que outras, como meu pensamento, não são espaciais. Outros filósofos sustentaram que todas as coisas temporais são também espaciais e que, portanto, aquilo que é temporal coincide com o que é espacial.

Essa coincidência, com certeza, existe apenas se também assumirmos que nenhuma coisa atemporal é espacial. Penso que essa é de fato uma suposição verdadeira: *Todas as coisas atemporais são não-espaciais*. Um exemplo novamente nos ajudará a tornar essa posição clara. Considere a tonalidade da cor da maçã em minha geladeira, uma certa tonalidade de vermelho. Essa tonalidade tem um tamanho ou uma forma? Obviamente que não. A tonalidade de cor não é redonda ou quadrada, nem tem certo comprimento ou certa circunferência. A maçã, é claro, tem uma forma e um tamanho, como notamos anteriormente, mas a maçã não é a cor que a maçã tem. Por

mais que seja óbvio, como penso, que a cor não tenha forma ou tamanho, não é tão óbvio assim que não esteja localizada no *espaço*. Não está “onde a maçã está”, logo ali, na geladeira? A questão de saber se as propriedades estão localizadas no espaço nos ocupará por muitas páginas. É uma das mais importantes discussões em ontologia. Por agora, devo apenas dizer que alguns filósofos, e especialmente Platão, sustentaram que todas as propriedades são não-espaciais, ao passo que outros sustentaram que são espaciais. De acordo com o primeiro, a cor da maçã não está localizada em qualquer lugar no espaço, ao passo que de acordo como os últimos, está localizada “onde a maçã está”. De acordo com a primeira perspectiva, todas as propriedades são atemporais e não-espaciais. De acordo com a segunda, as propriedades são espaciais. Uma vez que são espaciais, têm também de ser temporais. Deixe-me fazer um diagrama dessas duas perspectivas:



Façamos um balanço. Vimos que toda classificação pressupõe uma distinção entre duas categorias, de coisa individual e de propriedade de coisa individual. Platão pensa que essas duas categorias diferem no que as coisas individuais mudam, ao passo que as propriedades não mudam. Mas isso significa que as coisas individuais são temporais, ao passo que as propriedades não são temporais. Depois consideramos o papel do espaço. Os filósofos,

como de costume, diferem sobre esse ponto. Alguns acreditam que a distinção entre as coisas temporais e as coisas atemporais coincide com a distinção entre as coisas espaciais e não-espaciais. Outros acreditam que, ainda que seja verdadeiro que todas as coisas atemporais sejam também não-espaciais, algumas coisas temporais, coisas como pensamentos, são não-espaciais.

É tempo de introduzir a nossa distinção terminológica mais importante. Platão, vimos, falou de “igualdade abstrata”. Vou falar de *coisas abstratas* (entidades, existentes) em geral. Uma coisa abstrata é uma coisa que não é nem temporal nem espacial. Uma *coisa concreta*, por outro lado, é uma coisa que é temporal e/ou espacial. O “e/ou” é necessário, pois há a possibilidade de haver coisas temporais, como pensamentos, que não são espaciais. Em termos dessa distinção a questão mais importante da ontologia é: *há coisas abstratas?*

#### *O mundo versus o universo*

Passemos agora do mundo dos dois reinos de Platão ao universo físico. O universo (ou cosmo) é a totalidade de matéria e energia em existência. Consiste de todas as partículas elementares que há. Essas partículas formam todos os átomos existentes. Os átomos, por sua vez, combinam-se em moléculas, e essas moléculas compõe as coisas a nossa volta: a maçã de nosso exemplo, nossos corpos, as plantas e os animais sobre a terra, as montanhas e os rios. Mas o universo contém não apenas a terra e tudo que está nela, mas também a lua, o sol e os planetas dos sistemas solares. Há muitos sóis e sistemas planetários. Eles formam as galáxias de estrelas. Nosso sistema solar, por exemplo, é um membro da galáxia Via Láctea. E as galáxias ainda formam grandes aglomerados de galáxias. Dentre as estrelas, há as gigantes vermelhas, as supernovas, as anãs brancas, e as estrelas de nêutrons. Em suma, o universo é um todo gigantesco espaço-temporal, consistindo de partículas elementares e de todas as suas configurações. Estima-se que tem cerca de dez

bilhões de anos luz de diâmetro e que existe há cerca de oito a dez bilhões de anos. É claro, penso, que o universo pertence ao reino das coisas concretas de Platão; pois é uma entidade espaço-temporal. É também claro que tudo o que pertence ao universo, que é parte do universo, é uma coisa concreta; pois é uma parte espaço-temporal do universo, e, portanto, é em si espaço-temporal. *O universo é uma coisa concreta, e também cada parte dele.*

Mas as propriedades, como presumimos na última seção, são coisas abstratas; não são espaço-temporal. Segue-se disso que não pertencem ao universo. Não são parte do universo. A tonalidade de vermelho da qual falamos, por exemplo, por incrível que pareça, não é uma parte (espaço-temporal) do universo. E o que vale para essa propriedade particular vale para todas as outras: nenhuma dessas coisas são partes do universo. Concordemos em dizer que tudo que há, que todo existente, pertença ou não ao universo, pertence ao *mundo*. Platão, ao descobrir que as propriedades são abstratas, descobriu que há coisas que não pertencem ao universo. *Descobriu que há um mundo e não apenas um universo.*

Já no tempo de Platão, havia filósofos que tinham afirmado que não há nada mais além do universo. Argumentaram que não há tal coisa como o mundo. Sustentaram que não há coisas abstratas. Chamarei tais filósofos de “naturalistas”. Esses, com certeza, são os gigantes da citação de Platão. Por outro lado, havia também os filósofos que tinham defendido a existência de entidades abstratas. Afirmavam que o universo é apenas uma parte do mundo. O mundo, na opinião deles, é muito mais rico que o universo. *A estrutura do mundo, afirmam, é o assunto apropriado para o estudo dos filósofos, ao passo que o universo é o assunto apropriado para o estudo dos cientistas.* Chamarei esses filósofos, não surpreendentemente, de “ontólogos”. Esses, com certeza, são os deuses de Platão. Por mais de dois mil anos travou-se uma batalha intelectual sem igual na história da mente humana entre naturalistas e ontólogos. Nenhum dos lados,

contudo, obteve uma vitória decisiva. Nem é de ser esperar que tal vitória ocorra no futuro. As questões são bastante complexas para uma solução definitiva. Mais importante, a batalha entre naturalistas e ontólogos é em grande parte uma batalha entre dois temperamentos. Há, por um lado, o temperamento científico que favorece uma concepção de filosofia como, na pior das hipóteses, irmã da poesia, e na melhor, subserviente da ciência. Por outro lado, há o espírito ontológico, de acordo com o qual a ontologia tem uma perspectiva completamente diferente das ciências e que oferece vislumbres de verdades que os cientistas nunca sonharam. De acordo com o naturalista, tudo o que não é científico cheira a misticismo. A ontologia não pode ser nada mais do que poesia disfarçada. Abandonar a ciência é viajar através de uma noite sem estrelas, através de uma escuridão sem luz. De acordo com o ontólogo, abandonar a ontologia é desertar um campo no qual crescem as mais fascinantes flores da verdade. Descrevo essa batalha nesses termos emocionais, pois é uma batalha emocional. Em todas essas batalhas a escolha nunca é meramente entre duas posições racionais.

A descoberta do mundo levanta a questão de saber se há ou não quaisquer outras coisas abstratas. Lembre-se do exemplo de Platão da igualdade abstrata. Tratei a igualdade como se fosse uma propriedade, mas não é. É uma relação entre coisas. Há muitas dessas relações. Há relações espaciais: à esquerda de, entre, dentro de, etc. Há também relações temporais: mais tarde que, mais cedo que, simultâneo com, etc. Há também relações entre pessoas: pai de, tio de, esposo de, etc. E assim por diante. Mais importante, há a relação que conecta as coisas individuais com suas propriedades. Chamá-la-ei de “exemplificação”. Essa relação é representada pela cópula na frase: “A maçã é vermelha”. Em português temos os tempos verbais: “Essa maçã *foi* verde há um mês”. Mas sempre que atribuímos uma propriedade a uma coisa individual, estabelecemos que a propriedade se relacionou pela relação de exemplificação a um indivíduo. Platão, poder-se-ia notar, viu a importância da

exemplificação. Sem ela seus dois reinos não formariam um único mundo unificado, antes se separariam. Sem a exemplificação o mundo de Platão se dividiria num universo e num reino de propriedades. Mas é claro, as coisas individuais do universo *têm* essas propriedades. Qual a natureza da exemplificação? Platão não poderia se decidir sobre essa questão. (Veja seu Parmênides). Nem qualquer outro poderia até mais recentemente. Mas não temos de nos desviar. O que nos interessa no momento é saber se as relações são abstratas ou não. Penso que são. Considere a relação de *estar entre* que existe entre três pontos a, b e c sob uma linha traçada. Esses pontos, agora, são coisas individuais; estão localizados no espaço (e no tempo). Mas a relação não está em lugar algum. Certamente não está onde quaisquer uns dos três pontos estejam. Nem está entre, digamos, a e b. Não está localizada no espaço. Nem está localizada no tempo: é atemporal. Concluimos, portanto, que é abstrata.

Platão, devo adicionar, tentou mostrar que não há relações. Tentou argumentar que as afirmações relacionais são realmente as propriedades das coisas. Por exemplo, o fato de Tom ser mais alto que Henry é realmente uma conjunção de dois fatos, nomeadamente, do fato de Tom ter certa altura e do fato de Henry ter certa altura. Veremos mais tarde que isso não é assim. O que quero apontar aqui é que um naturalista tem uma boa razão para seguir os passos de Platão no que diz respeito às relações. Ele tentará “reduzir” as relações às propriedades, e então tentará mostrar que as propriedades não são abstratas.

Já falei do *fato* de Tom ter certa altura. São os fatos concretos ou são abstratos? Bem, há muitos tipos de fatos. Considere o fato de que a maçã de nosso exemplo é vermelha num certo tempo *t*. Onde está o fato? Um naturalista pode querer dizer que está onde a maçã está, mas isso seria insincero. Equivaleria a nada mais do que o acordo arbitrário em localizar um fato onde quer que a coisa individual do fato esteja posta. Ademais, há fatos de outras coisas

além das coisas individuais. É um fato, por exemplo, que o azul escuro é mais escuro que o amarelo limão. Nesse caso, o naturalista tem de descobrir uma localização para essas tonalidades de cor antes de poder invocar a regra já mencionada. Há, contudo, fatos sobre relações, por exemplo, o fato de a relação de ser esposo de alguém ser simétrica: se existe entre a e b, então também existe entre b e a. (A relação de ser pai de alguém, por comparação, não é simétrica; é assimétrica). Onde está a relação de ser esposo? De estar a vinte milhas de Chicago? De estar sobre a lua. O naturalista terá de afirmar que está em algum lugar próximo onde os seres humanos estão. Mas e o fato de dois mais dois ser quatro. Certamente que esse fato não depende dos seres humanos. A relação de soma existe entre quatro, dois e dois, “na lua” tanto quanto “na terra”.

Se as relações e os fatos são abstratos, como argumentei brevemente, então não pertencem ao universo, mas pertencem ao mundo, e o mundo é ainda “maior” do que podemos ter pensado no início. Contém não apenas propriedades, mas relações e fatos também. O universo, é claro, contém muitos diferentes *tipos de indivíduos*, de partículas elementares a galáxias, e a ciência está interessada em todas elas. O mundo, por outro lado, é o repositório dos *tipos de existentes*, e como já vimos há pelo menos três desses tipos: as propriedades, as relações e os fatos. E essas nem são todas, como veremos mais tarde. Há, em adição, os números, os conjuntos e as estruturas (totalidades).

#### *Um sobre muitos*

Há outro modo de se olhar para a descoberta de Platão do mundo. Nele, também, as propriedades são contrastadas com as coisas individuais. O que as torna especiais, no entanto, não é por serem atemporais, mas por pertencerem a muitas coisas individuais. Vários indivíduos podem *compartilhar* a mesma propriedade. Considere duas bolas de bilhar brancas. Chame-as de A e B. Essas são *duas* coisas individuais, mas sua cor, sua tonalidade de branco, é a

mesma: elas compartilham *uma* tonalidade de cor. E uma vez que as bolas têm a mesma forma, há também apenas uma forma: ambas as bolas são esféricas. A propriedade da brancura é uma sobre muitos. É um *universal*. As bolas de bilhar individuais, por outro lado, são chamadas *particulares*.

Às vezes, a distinção entre universais e particulares é introduzida em termos de tipos e espécimes. Veja as seguintes palavras: vermelho, vermelho. Aqui você vê dois espécimes do mesmo tipo, nomeadamente, da palavra “vermelho”. A palavra realmente é, como se pode ver, de certa forma complicada, e qualquer inscrição particular como essa forma é um espécime do tipo “vermelho”. Nesse caso, a forma é a propriedade na qual todos os espécimes dos tipos “vermelho” compartilham. De maneira resumida, os tipos são os universais e os espécimes são os particulares do mesmo tipo.

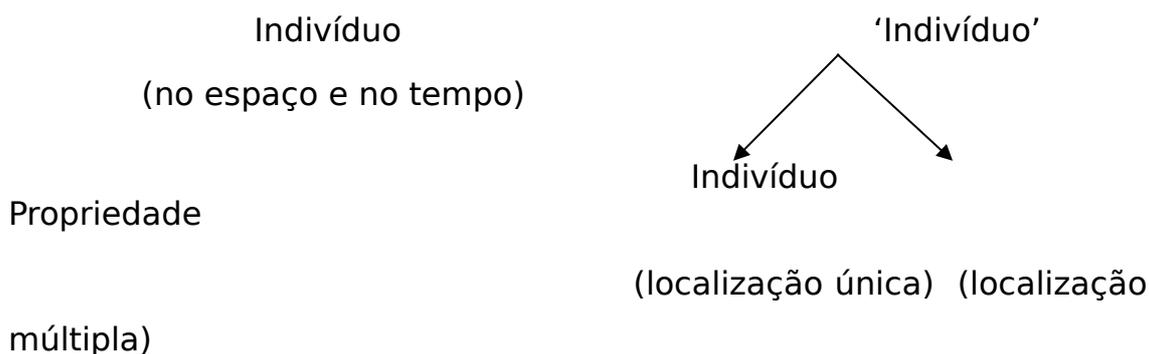
A brancura das duas bolas de bilhar é literalmente a mesma? Há apenas uma entidade que é exemplificada por ambas as bolas? Ou cada bola tem sua própria brancura? Esse é o chamado problema dos universais. Em nosso exemplo da palavra “vermelho”, a questão crucial é esta: a forma de uma inscrição de “vermelho” é a mesma forma que a da outra? Os filósofos que acreditam que a cor da bola de bilhar A é a mesma que a da bola de bilhar B são chamados de *realistas*. Os que negam isso são chamados de *nominalistas*.

Podemos e temos de distinguir entre estas duas questões completamente diferentes embora intimamente relacionadas: (i) as propriedades são abstratas? e (ii) as propriedades são universais? Como expliquei, a primeira questão é: estão as propriedades localizadas no espaço e no tempo? A segunda é completamente diferente: pode uma e a mesma propriedade ser uma propriedade de várias coisas? Parece, agora, que ser um naturalista não ajuda, mas antes um nominalista. Pois se a cor branca está localizada no espaço, por exemplo, então não pode ser a mesma para as duas bolas de

bilhar. Cada bola tem de ter sua própria brancura. A bola A é branco<sub>1</sub> e a bola B é branco<sub>2</sub>; e o branco<sub>1</sub> está localizado num lugar, onde A está, ao passo que o branco<sub>2</sub> está localizado noutra lugar, onde B está. Mas essa impressão está errada: um naturalista pode ser um realista, pois pode sustentar que há coisas que, apesar de estarem localizadas no espaço, podem existir simultaneamente em muitos lugares diferentes. De acordo com essa alternativa, A tem a mesma brancura que B. Não existem duas brancuras. Mas essa uma e única brancura está, não obstante, localizada no espaço: existe simultaneamente onde A está e também onde B está (e também onde outros indivíduos dessa tonalidade de cor estão).

De acordo com essa perspectiva as propriedades estão localizadas no espaço e no tempo. São concretas. Não existe algo além do universo. Porém o padrão monolítico do naturalismo é quebrado: mesmo apesar de tudo que há ser um indivíduo em nosso sentido definido, há, contudo, dois tipos de indivíduos, nomeadamente, aqueles que podem apenas existir num lugar e num tempo e aqueles que podem ter múltiplas localizações ao mesmo tempo. Os primeiros são as “coisas individuais” de nosso entendimento comum, as últimas são as propriedades. O padrão é quebrado, eu disse, porque admitir que há coisas que possam existir em muitos lugares diferentes ao mesmo tempo é admitir uma categoria de coisas completamente distintas dos indivíduos comuns de nossa experiência comum. Seja como for, aprendemos que há pelo menos dois tipos de naturalismo, nomeadamente, o naturalismo puro e o naturalismo impuro. O primeiro é uma combinação do naturalismo com o nominalismo, ao passo que o último consiste do naturalismo conjunto com o realismo. Eis dois diagramas para essas duas perspectivas:





O que separa o naturalismo puro do impuro é uma suposição fundamental. Uma vez que essa suposição é muito importante, vou lhe dar um nome e chamá-lo “o axioma da localização”:

Nenhuma entidade que seja pode existir em diferentes lugares ao mesmo tempo ou em intervalos de tempo interrompidos.

O naturalismo impuro é imposto ao naturalismo por causa do debate nominalismo-realismo. Um naturalista que percebe que o nominalismo é insustentável não tem escolha, além de oferecer o axioma da localização e, portanto, adotar o naturalismo impuro. Essa é a conexão essencial entre o debate nominalismo-realismo, por um lado, e a disputa naturalismo-ontologismo, por outro: uma refutação do nominalismo ou conduz a um colapso do naturalismo ou a um abandono do axioma da localização. Isso mostra o quão importante é o debate nominalismo-realismo para o naturalista e para o ontólogo. Um naturalista pode bem acreditar que renunciar ao axioma da localização é renunciar à sua concepção de universo. Ele, portanto, defenderá o nominalismo com toda a sua força. O ontólogo, por outro lado, ao perceber a conexão íntima entre os dois debates, atacará a posição do naturalista, não diretamente, mas em seu ponto mais vulnerável, atacando o nominalismo.

# O DIREITO A UM ELEITORADO COMPETENTE\*

Jason Brennan  
Tradução de Aluizio Couto  
Revisão de L. H. Marques Segundo<sup>89</sup>

*A prática do sufrágio universal irrestrito é injusta. Os cidadãos têm o direito a que qualquer poder político exercido sobre eles seja exercido por pessoas competentes de um modo competente. O sufrágio universal viola esse direito. Para satisfazê-lo, o sufrágio universal tem, na maioria dos casos, de ser substituído por uma epistocracia moderada, na qual o sufrágio é restrito a cidadãos dotados de suficiente competência política. Em si, a epistocracia parece colidir com a exigência de aceitabilidade qualificada, segundo a qual o poder político deve ser distribuído de maneiras contra as quais não existam quaisquer objeções qualificadas. No entanto, ela é intrinsecamente menos injusta do que a democracia com sufrágio universal e provavelmente produz melhores resultados. Assim, a epistocracia é mais justa do que a democracia, mesmo não sendo perfeitamente justa.*

---

\* “The Right to a Competent Electorate”, *The Philosophical Quarterly*, Vol. 61 n° 245.

<sup>89</sup> Marques Segundo é bolsista CAPES, doutorando pela UFSC. Aluizio Couto é Mestrando pela UFMG, com bolsa CNPq.

## I. INTRODUÇÃO

Muitos de meus concidadãos são incompetentes, ignorantes, irracionais e moralmente irrazoáveis sobre política. A despeito disso, eles têm poder político sobre mim. Podem compor repartições dotadas de bastante poder e usar a autoridade coercitiva do estado contra mim. Podem forçar-me a fazer coisas que não quero fazer ou não tenho boas razões para fazer.

Como uma pessoa inocente, eu não deveria ter de tolerar isso. Seria errado colocar-me à força sob a faca de um cirurgião incompetente ou forçar-me a navegar no navio de um capitão incompetente. Da mesma forma, é errado que eu seja submetido à força às decisões de eleitores incompetentes. As pessoas que exercem poder sobre mim, incluindo os outros eleitores, deveriam fazê-lo de modo competente e moralmente razoável. Do contrário, por uma questão de justiça, elas deveriam ser impedidas de ter poder político, o que inclui o poder do voto.<sup>90</sup> As democracias modernas concedem a cada cidadão adulto o direito legal ao voto (algumas democracias excluem alguns cidadãos, como os criminosos e os loucos). Muitas vezes pessoas foram impedidas de votar por razões moralmente arbitrárias como a cor da pele ou o sexo. Essa prática era injusta. Mesmo assim, por mais que várias pessoas tenham sido impedidas por más razões, pode haver boas razões para negar o poder a muitos. Como paralelo, seria injusto impedir cidadãos ateus de dirigir porque eles são ateus. Entretanto, mesmo que essa lei fosse injusta, disso não se seguiria que todas as restrições ao direito legal de dirigir seriam também injustas. O mesmo sucede com os direitos políticos. As democracias costumavam negar poder às pessoas por más razões, mas talvez elas deveriam excluir cidadãos por boas razões.

---

<sup>90</sup>Em meu artigo “Polluting the Polls: When Citizens Should Not Vote”, *Australasian Journal of Philosophy*, 87 (2009), pp. 535-49, argumento que cidadãos incompetentes têm a obrigação moral de não votar, mas não que cidadãos incompetentes não devem ter o direito ao voto. Este artigo vai além e argumenta que os cidadãos devem ser impedidos de votar caso sejam incompetentes.

Neste artigo, argumento em favor do sufrágio restrito. É injusto garantir a certos cidadãos o direito legal de exercer poder político sobre os outros. Uma forma de epistocracia com sufrágio restrito é moralmente superior à democracia com sufrágio universal. De modo geral, um regime é epistocrático na medida em que o conhecimento e a competência são requisitos legais para a posse do poder político. Platão defendeu o governo de reis filósofos, uma forma extrema de epistocracia. Todas as democracias modernas negam às crianças o voto e a ocupação de cargos com base na ideia de que crianças são incompetentes. Nesse sentido, todas as democracias são fracamente epistocráticas. Neste artigo, argumento em favor de uma posição epistocrática moderada. Nas democracias contemporâneas, os cidadãos devem ter competência moral e epistêmica suficientes para ter o direito ao voto.

Estou, dessa forma, argumentando em favor de algo que poderia ser chamado de um *sistema eleitoral de elite*. Sistemas eleitorais de elite têm mecanismos políticos similares àqueles encontrados nas democracias contemporâneas, mas restringem o poder eleitoral aos cidadãos que podem demonstrar competência. Sistemas eleitorais de elite são moderadamente epistocráticos. Não argumento, no entanto, que a epistocracia moderada ou qualquer tipo de epistocracia é a forma *mais* justa ou *idealmente* justa de governo. Meu objetivo é limitado: todo o resto permanecendo igual, nas democracias contemporâneas, o sufrágio restrito seria um *progresso moral* em relação ao sufrágio universal incondicional. Dito isso, o sufrágio restrito pode ainda ser injusto – melhor do que o sufrágio universal, mas não bom o suficiente para se qualificar como justo. O sufrágio restrito pode ser injusto, mas é *menos injusto* do que o sufrágio universal incondicional.

Este artigo está disposto da seguinte maneira. Na seção II argumento que o sufrágio universal é injusto porque viola o princípio da competência. O princípio da competência exige que quando uma decisão é importante e involuntariamente imposta sobre os outros por

meio da força, ela deve ser tomada por pessoas razoáveis e competentes de um modo razoável e competente. Na seção III examino um argumento cuja proposta é mostrar que o sufrágio restrito é injusto porque viola a exigência de aceitabilidade qualificada. A exigência de aceitabilidade qualificada exige que qualquer base para a distribuição de poder político deve ser justificável perante todos os pontos de vista qualificados. Se os argumentos das seções III e IV forem ambos bem-sucedidos, a democracia com sufrágio universal e um regime igual em todos os outros aspectos, mas com sufrágio restrito são ambos injustos, cada um por diferentes razões. A democracia viola o princípio da competência, mas a epistocracia viola a exigência da aceitabilidade qualificada. No entanto, na seção V argumento que o sufrágio restrito é intrinsecamente menos injusto do que o sufrágio universal. Na seção VI argumento que o sufrágio restrito provavelmente produziria consequências melhores do que a democracia, embora não saibamos disso com certeza. Sendo o caso, o sufrágio restrito seria no máximo moralmente superior ao que é, na prática, o sufrágio universal, mesmo que nenhum sistema seja *justo*.

## II. O ARGUMENTO BÁSICO EM FAVOR DO SUFRÁGIO RESTRITO

Nesta seção argumento que o sufrágio universal é injusto porque ele viola o direito do cidadão de não ser sujeito a decisões importantes feitas por pessoas incompetentes e moralmente irrazoáveis. Nas seções finais, argumento que o sufrágio restrito é moralmente superior ao sufrágio universal.

Os filósofos, entre outros, por vezes asseveram que qualquer um que esteja sujeito ao poder político deve ter voz em como esse poder é usado.<sup>91</sup> Garantir aos cidadãos poder político pode, em alguma medida, reconciliá-los com seu próprio governo, uma vez que isso reduz o grau em que o governo é uma imposição sobre eles. No

<sup>91</sup>Ver J. Waldron, "Participation: the Right of Rights", *Proceedings of The Aristotelian Society*, 98 (1998), pp. 307-37, na p. 310.

entanto, o direito de votar não dá aos cidadãos poder apenas sobre si próprios, mas também sobre outras pessoas. Temos o direito básico de nos governar e decidir o que será de nós, mas não temos qualquer direito básico de mandar e decidir pelos outros. Não temos qualquer direito básico de conduzir as vidas de outras pessoas e impor regras sobre elas.<sup>92</sup>

A democracia, entre outras coisas, é um método particular de tomada de decisões. A democracia *política*, entre outras coisas, é um método para decidir quando, como e de que formas o governo irá ameaçar as pessoas com violência a fim de induzir conformidade às regras. Possuir o direito ao voto é possuir algum grau de poder político, mesmo que pequeno. Tal poder é exercido sobre os *outros*, e não apenas sobre si mesmo.

Estar imbuído de tal poder, por pouco que seja, é algo que exige justificação. Uma democracia com uma população eleitoral de tamanho  $n$  confere, *de jure*, a cada cidadão,  $1/n$  da autoridade política total (digo “*de jure*” porque o poder político efetivo ou *de facto* dos cidadãos varia dependendo das circunstâncias. Obama e eu temos os mesmos direitos de voto, mas ele tem muito mais poder político efetivo do que eu). Assim, por exemplo, em uma democracia de três cidadãos, cada um deles detém, *de jure*, um terço do poder total. Para uma democracia de grandes proporções, um poder político de  $1/n$  é uma quantia bem pequena. Mesmo assim, sair de uma situação em que se tem 0 poder político para uma em que se tem  $1/n$  é algo moralmente significativo. Posso apontar para cada um de meus concidadãos e, sensatamente, perguntar: “Por que *aquele cara* tem  $1/n$  de poder sobre mim? Quem fez daquele cara  $1/n$  meu chefe?”.

### *O princípio da competência e a analogia do júri*

---

<sup>92</sup>Ver R. Arneson, “The Supposed Right to a Democratic Say”, in T. Christiano e J. Christman (eds.), *Contemporary Debates in Political Philosophy* (Oxford: Wiley-Blackwell, 2009), pp. 197-212, e “Democracy is Not Inherently Just”, in K. Dowding *et al.* (eds.), *Justice and Democracy* (Cambridge University Press, 2004), pp. 40-58.

A justificação básica para o sufrágio restrito pode ser ilustrada por analogia com a atuação de um júri. Em casos criminais, os júris detêm um sério poder sobre o réu. A decisão do júri pode alterar significativamente as perspectivas de vida do réu e infligir a privação de sua propriedade, liberdade e vida.

Seguem três júris hipotéticos.

1. *O júri ignorante*: o júri não presta atenção ao que se passa no julgamento. Quando solicitados a deliberar, os membros são ignorantes acerca dos detalhes do caso, mas pensam que, de qualquer forma, o réu é culpado. Após o julgamento, eles admitem que decidiram o caso dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).
2. *O júri irracional*: o júri presta alguma atenção aos detalhes do caso. No entanto, os membros pensam que o réu é culpado não com base na evidência, mas com base em pensamento mágico e em várias teorias da conspiração nas quais acreditam. Após o julgamento, eles admitem que decidiram o caso dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).
3. *O júri moralmente irrazoável*: o júri acha que o réu é culpado porque ele é muçulmano; o júri é composto de membros cristãos que pensam que os muçulmanos pervertem a Palavra de Deus. Após o julgamento, eles admitem que decidiram o caso dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).

Esses júris não têm autoridade ou legitimidade (um júri tem autoridade sobre o réu quando o réu tem a obrigação moral de acatar a decisão do júri *porque* é a decisão do júri. Um júri tem legitimidade sobre o réu quando é moralmente permissível ao governo fazer valer a decisão do júri contra o réu). Réus que sabem que foram sujeitos a

um desses júris não têm qualquer obrigação moral de considerar a decisão como dotada de autoridade. Seria injusto por parte do governo conscientemente garantir a aplicação dessas decisões –isso privaria o cidadão de sua propriedade, liberdade/ou vida por razões inaceitáveis, com base em uma decisão tomada de modo inapropriado.

Nos casos de julgamentos, é plausível que os réus tenham direito a um júri competente, como expresso a seguir.

### *O princípio da competência*

É injusto privar cidadãos de sua vida, liberdade ou propriedade, ou alterar significativamente suas perspectivas de vida pela força ou ameaças de força quando isso é resultado de uma decisão feita por um corpo deliberativo incompetente ou moralmente irrazoável, ou como resultado de decisões feitas de um modo incompetente ou moralmente irrazoável.

O princípio da competência desqualifica a composição do júri aqueles indivíduos dotados de uma quantidade suficiente de mau caráter epistêmico e moral (mesmo que tais pessoas possam agir competentemente em outros casos particulares). Ele também desqualifica decisões individuais de um júri (mesmo se o júri, de maneira geral, tem bom caráter epistêmico e moral) quando essas decisões são tomadas de modo inapropriado. O princípio implica que:

- A. Pessoas dotadas de mau caráter epistêmico e moral não devem participar de júris;
- B. Quando se sabe que decisões particulares foram tomadas de modo incompetente (mesmo que as decisões tenham sido tomadas por pessoas que, de modo geral, são competentes e razoáveis), essas decisões não devem ser efetivadas, e os réus não têm qualquer dever de submissão a elas.

Em suma, o princípio da competência exige que cada decisão de um dado tipo seja tomada competentemente por pessoas competentes.

Não poderíamos justificar o cumprimento da decisão de um júri incompetente mostrando que a maioria dos júris são competentes. O réu pode objetar que esse, neste caso, não foi. Não podemos privar alguém de sua liberdade, propriedade ou vida por causa de uma decisão incompetente meramente porque *outros* júris são competentes.

Também não poderíamos justificar o cumprimento de uma decisão tomada de modo incompetente mesmo se o júri em particular que decidiu o caso seja geralmente competente. Suponha que o mesmo painel de jurados participe de 100 casos. Eles decidem 99 casos de maneira racional, bem informada e moralmente razoável, mas em um caso decidem de modo irracional, ignorante, mal informado e/ou moralmente irrazoável que o réu é culpado. Não poderíamos dizer a esse réu que “certamente o júri foi incompetente no seu caso, mas ele foi competente em outros casos. Assim, devemos garantir o cumprimento sua decisão e você deve se submeter a ela.” O réu poderia objetar dizendo “que bom que o júri fez um trabalho em todos esses *outros* julgamentos, mas é a da *minha* vida e da *minha* liberdade que estamos falando. Este júri decidiu o *meu* caso de modo incompetente e irrazoável.” A objeção do réu parece-me decisiva.

O princípio da competência não implica que júris têm autoridade e legitimidade apenas quando tomam decisões *corretas*. Em vez disso, ele afirma que lhes falta autoridade e legitimidade quando chegam a respostas de modos inaceitáveis, independentemente de sua resposta ser correta ou incorreta. O princípio da competência não desqualifica as decisões do júri com base em seu conteúdo substantivo. Em vez disso, desqualifica pessoas da composição do júri por causa de seu mau caráter moral e epistêmico, e desqualifica decisões individuais de júris por causa do mau (ou nenhum) raciocínio usado para chegar até elas.

O princípio da competência exige que corpos deliberativos sejam moralmente razoáveis e que tomem decisões de um modo moralmente razoável. O termo “moralmente razoável” não quer dizer que os membros do corpo deliberativo tenham de ter crenças morais *corretas* e ajam segundo essas crenças. Em vez disso, minha intenção é invocar uma ideia comum às filosofias políticas que se inserem no interior da tradição do liberalismo político (ou liberalismo de “razão pública”). Liberais políticos sustentam, entre outras coisas, que se é permitido o exercício da razão a indivíduos razoáveis, eles tendem a discordar sobre muitas questões básicas sobre justiça, moralidade e vida boa.

Liberais políticos sustentam que há discordâncias razoáveis sobre esses aspectos, embora nem todas as discordâncias sejam razoáveis. Diferentes liberais traçam a fronteira entre razoável e irrazoável de modos diferentes. O princípio da competência exige que os corpos deliberativos devem ser razoáveis e tomar decisões de um modo moralmente razoável, seja lá qual for a perspectiva correta da razoabilidade moral. Não vou aqui defender qualquer teoria particular acerca de tal distinção. Para os propósitos deste artigo, uso o termo “moralmente razoável” como uma variável a ser preenchida pela verdade, seja lá o que ela for.

O princípio da competência pode em si parecer plausível. Mas argumentos podem ser apresentados em seu favor. Uma justificação para o princípio da competência é o fato de ser injusto expor pessoas a riscos indevidos. No caso acima, os membros do júri agem de forma negligente para com o réu. Do ponto de vista do réu, a decisão do júri é importantíssima e o resultado é imposto involuntariamente. Nesses tipos de casos, um júri tem a obrigação de ser adequadamente cuidadoso ao decidir.

Há casos paralelos. Tenho bronquite severa. Meu médico consulta um curandeiro em busca de conselhos para o tratamento. O curandeiro queima alguma gordura animal, atira nela uma sopa de letrinhas e lê os padrões formados pelas letras. Por acaso, as letras

revelam uma medicação e meu médico a prescreve. Independentemente do fato de a medicação, no fim das contas, ser a correta (e.g., prednisona) ou a errada (e.g., moxonidina), o médico fez algo errado ao usar um método de decisão altamente não confiável para chegar à prescrição. O uso desse método me coloca sob sério risco de dano. Se o médico tivesse o poder de me *forçar* a tomar a medicação (tal como os júris têm o poder de impor suas decisões sobre os réus), isso seria intolerável.

Até certo grau, os Estados Unidos tentam respeitar o princípio da competência no que diz respeito às decisões de júris. Os potenciais jurados são escolhidos aleatoriamente entre todos os cidadãos adultos no interior de uma área geográfica. No entanto, jurados por vezes são desqualificados por exibirem viés ou alguns tipos de incompetência. Pela lei, uma vez que condenações criminais podem privar os réus de sua propriedade, liberdade e (às vezes) vida, os réus têm direito a um julgamento justo e a um júri imparcial. Após o julgamento, caso se revele que o júri tomou sua decisão de modo corrupto ou incompetente, tal como nos três casos acima, pode haver base legal para reverter a decisão do júri.

#### *Aplicando o princípio da competência ao governo*

O princípio da competência tem um amplo escopo de aplicação. Ele não se aplica apenas a decisões de júris.

Governos democráticos, tal como júris, também podem privar os cidadãos (e outros) de sua liberdade, propriedade e vida. De fato, eles frequentemente o fazem. Governos democráticos podem impor políticas que alteram significativamente, para o bem ou para o mal, as perspectivas de vida do cidadão. Se os réus têm direito a júris competentes, podem haver razões similares para sustentar que cidadãos têm direito a governantes competentes.

Se um policial, juiz ou político toma uma decisão caprichosa, irracional ou maliciosa, o cidadão não pode simplesmente ir embora.

Em geral, ele pode escolher se submeter à decisão ou ser penalizado (por meio de coerção) por desobediência (às vezes, se tiver sorte, ele pode obter uma reparação após o fato).

Os governos fazem mais do que escolher cor de bandeiras e melodias para hinos nacionais. Eles elaboram políticas e escolhem cursos de ação que podem ter consequências muitíssimo importantes e até mesmo desastrosas para os cidadãos. Por exemplo, se um banco central ou um tesouro adotar más políticas monetárias e se o governo impor barreiras comerciais altas, uma recessão pode se tornar uma profunda depressão. Se líderes militares inflarem ou deturpam informações a eles fornecidas pelas agências de inteligência, o resultado pode ser uma guerra custosa, destrutiva e inumana.

Decisões governamentais tendem a ter dois aspectos fundamentais:

1. Os resultados das decisões são impostos involuntariamente por meio de violência ou ameaças de violência. Cidadãos e outros indivíduos dentro do domínio do governo são forçados a obedecer, mesmo se eles tiverem excelentes razões para não obedecer e mesmo se eles sabem que as decisões foram tomadas de modo incompetente;
2. Decisões governamentais tendem a ser de grande importância. Elas podem alterar significativamente as perspectivas de vida dos cidadãos e privá-los de sua vida, liberdade e propriedade.

À luz de (1) e (2), os cidadãos podem exigir, como um direito, competência dos oficiais do governo e dos responsáveis por tomar decisões. Infelizmente, esse direito é não raro e talvez quase sempre impossível de se fazer aplicar, mas, para se ter um direito, não é exigido que a sua aplicação seja garantida. Afinal, quando Hitler e Stálin assassinaram milhões de pessoas, nós diríamos que Hitler e Stálin violaram o direito delas à vida por mais que a essas pessoas

esse direito não se fizesse valer. Dizer que os cidadãos podem exigir competência dos oficiais do governo e daqueles responsáveis por tomar decisões como um direito é dizer que, *prima facie*, tomadas de decisão incompetentes ou decisões tomadas por pessoas incompetentes são coisas injustas.

O princípio da competência se aplica não apenas aos júris, mas também a outras pessoas dotadas de poder político, como a polícia, os burocratas e ministros, os juízes e os políticos investidos de cargos públicos. Na prática, até certo ponto, os governantes tentam se conformar ao princípio da competência. Muitas posições de poder exigem dos candidatos certas qualificações para que o poder seja concedido. Não fazemos de qualquer pessoa um policial, tal como não é qualquer um que pode dirigir o FED. Os juízes devem ser formados em direito, e mesmo os políticos estão sujeitos a exigências. Essas exigências são impostas para eliminar incompetências grosseiras.

### *Aplicando o princípio da competência ao eleitorado*

Nas democracias, os detentores últimos do poder são os eleitores. Geralmente, eles determinam como serão preenchidos os cargos políticos. Os eleitores escolhem os governantes e esses então usam o poder coercitivo do estado contra cidadãos inocentes, incluindo cidadãos que justificadamente se opõem às ações do estado. Se os eleitores escolherem mal, as consequências podem ser muito más.

Não devemos subestimar o dano que o mau voto pode causar. Ele pode ser desastroso. Mesmo que nos EUA ou na Grã-Bretanha candidatos desastrosos raramente tenham alguma chance de vencer, não devemos nos esquecer que em várias partes do mundo muitos candidatos desastrosos foram eleitos. Os eleitores que colocaram os Nacional Socialistas no poder na Alemanha em 1933 não podem ser

apontados como responsáveis por tudo que o governo fez. Mas muito do que o governo fez era previsível para qualquer pessoa razoável e bem informada, o que tornam culpáveis os seus apoiadores.

Apesar do mau voto, podemos ainda ter boas políticas; e podemos ter más políticas apesar do bom voto. Mesmo assim, o voto faz diferença. Em geral, quanto mais baixa a qualidade epistêmica e moral do eleitorado, piores as políticas do governo tendem a ser. Eleitorados de baixa qualidade tendem a fazer escolhas piores nos pleitos: eles são piores ao escolher bons líderes e tendem a escolher as piores políticas nos referendos. Eleitorados de baixa qualidade também tendem a reduzir a qualidade dos candidatos que aparecem na cédula de votação. Um eleitorado de má qualidade reduz a qualidade dos resultados antes mesmo da eleição ocorrer.

O princípio da competência se aplica igualmente bem aos eleitorados e aos júris. Seguem três eleitorados hipotéticos:

1. *O eleitorado ignorante*: o eleitorado não presta atenção aos detalhes da eleição ou às questões em jogo. Durante a eleição, eles escolhem aleatoriamente um candidato em particular. Eles admitem que decidiram a eleição dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).
2. *O eleitorado irracional*: a maioria presta alguma atenção aos detalhes da eleição ou às questões em jogo. No entanto, eles não votam com base nas evidências, mas em pensamento mágico e em várias teorias infames nas quais calham de acreditar. Eles admitem que decidiram a eleição dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).
3. *O eleitorado moralmente irrazoável*: simplesmente por racismo a maioria escolhe o candidato branco e não o negro. Eles admitem que decidiram a eleição dessa maneira (ou temos outras fontes robustas que evidenciam que foi assim mesmo que eles tomaram sua decisão).

Suponha, em cada um desses casos, que a maioria não representa a todos na sociedade. Pode haver, por exemplo, uma minoria de eleitores bem informada, racional e moralmente razoável, ou pode haver pessoas inocentes que não votam, como crianças e estrangeiros residentes. Se for o caso, então a maioria dos eleitores fez algo profundamente injusto: impôs um governante sobre pessoas inocentes sem ter bases adequadas para a decisão (além disso, se os eleitores tendem a ser ignorantes, irracionais ou moralmente irrazoáveis, há não apenas a tendência de as escolhas serem ruins, mas também a de que, logo de início, os candidatos disponíveis sejam de baixa qualidade).

Os governados têm o direito de não ser expostos a riscos indevidos na seleção das políticas ou dos governantes que farão as políticas. Quando as eleições são decididas com base em procedimentos epistêmicos não confiáveis ou com base em atitudes morais irrazoáveis, os governados são expostos indevidamente ao risco de sofrerem danos. Uma vez que os governados são *forçados* a obedecer às decisões do eleitorado, a tomada de decisões negligente é intolerável. O eleitorado tem a obrigação de não expor os governados a riscos indevidos.

Isso conclui o argumento básico em favor do sufrágio restrito. Quando decisões importantes são impostas sobre pessoas inocentes, o princípio da competência exige que cada decisão individual seja tomada de modo competente e razoável por pessoas competentes e razoáveis. Ele se aplica não apenas a decisões de júri, mas a qualquer decisão importante tomada por aqueles que detêm o poder político.

Alguém poderia objetar que há um hiato entre afirmar (a) que o eleitorado (como um todo) deve ser competente e tomar decisões competentes, e (b) que os eleitores individuais que formam o eleitorado devem ser competentes e tomar decisões competentes; dessa forma, o princípio da competência exige apenas que o eleitorado como um coletivo tome suas decisões de modo

competente, mas isso não implica que os eleitores individuais devam ser competentes.

Uma forma dessa objeção pode sustentar que o eleitorado como um coletivo tende a fazer excelentes escolhas mesmo que muitos eleitores ou a maioria deles sejam incompetentes. Por exemplo, o “milagre da agregação” supostamente mostra que quando um vasto eleitorado é composto quase inteiramente de eleitores incompetentes mas tem, ao mesmo tempo, uma pequena minoria de eleitores informados, ele toma as mesmas decisões que seriam tomadas por um eleitorado composto apenas de eleitores informados. No entanto, a maioria dos teóricos da democracia agora sustentam que esses modelos matemáticos, embora interessantes, não se aplicam à democracia real.<sup>93</sup> Não pretendo discuti-los mais aqui.

Pode-se também objetar que o eleitorado como um todo tem poder, ao passo que os eleitores individualmente não o têm, e por isso o princípio da competência não se aplica a eleitores individuais. Não pretendo me alongar na resposta a esta objeção, mas parece valer a pena explorar o espaço conceitual em causa. Indicarei apenas algumas respostas rápidas à objeção.

Em primeiro lugar, ela parece confundir ou misturar a *importância da consequência* com *poder político*. O direito ao voto parece ser um tipo de poder político. Não se espera que o meu voto ou o de David Duke mude o resultado de qualquer eleição. Nossos votos também não têm utilidade ou desutilidade esperada significativa. No entanto, em virtude de termos o direito ao voto, cada um de nós ainda goza de um tipo de status. Estamos autorizados, em conjunção com os outros, a tomar decisões políticas fundamentais. Qualquer um que diga o contrário, que ter o direito ao voto não é um exercício de poder

---

<sup>93</sup>Para críticas à aplicação do milagre da agregação à democracia real, ver S. Althaus, *Collective Preferences in Democratic Politics: Opinion Surveys and the Will of the People* (Cambridge University Press, 2003); B. Caplan, *The Myth of the Rational Voter* (Princeton University Press, 2007); Brennan, *The Ethics of Voting* (Princeton University Press, 2011), pp. 169-75. Para uma discussão sobre por que o teorema do júri de Condorcet não se aplica a democracias reais, ver D. Estlund, *Democratic Authority* (Princeton University Press, 2008), cap. 11.

político, parece se comprometer com a tese de que restringir o sufrágio não subtrairia o poder de ninguém.

Em segundo lugar, teríamos dificuldade de aceitar essa objeção no caso dos júris. Suponha que um júri de 12 indivíduos pensa que um réu negro é culpado. Após a condenação, nós descobrimos que 5 jurados são racistas fanáticos, mas que os 7 restantes são pessoas moralmente razoáveis. Isso enfraqueceria a autoridade do júri.

A inexistência de consequências de votos individuais é uma razão pela qual o eleitorado como um todo viola o princípio da competência tão frequentemente. Independentemente de suas razões serem egoístas ou altruístas, os eleitores têm poucos incentivos para se informar bem sobre política, ou até mesmo para formar suas crenças políticas de modo racional. Os eleitores são racionalmente ignorantes, ou talvez até mesmo racionalmente irracionais. Os custos de adquirir informação relevante e processar essa informação de modo epistemicamente racional superam o benefício esperado de votar bem (seja a motivação altruísta ou egoísta). A inexistência de consequências de votos individuais oferece aos eleitores um incentivo à irresponsabilidade. À luz disso, garantir a aplicação do princípio da competência exigiria de nós barrar os eleitores incompetentes.<sup>94</sup>

Até aqui, argumentei que o sufrágio universal, tal como praticado nas democracias contemporâneas, é injusto porque viola o princípio da competência. Sistemas democráticos de governo, na prática, usam ameaças de violência para sujeitar pessoas inocentes ao poder político exercido por governantes incompetentes e moralmente irrazoáveis. Nessa medida, eles são injustos. Não fiz ainda quaisquer propostas positivas de políticas com base nesse ponto.

---

<sup>94</sup>Nota de passagem que o trabalho empírico sobre o comportamento dos eleitores sugere que eles tendem a ser altruístas mas mal informados quando seus votos não contam muito. Tendem, no entanto, a se tornar mais egoístas e bem informados nos casos raros em que seus votos contam muito. Ver, e.g., T Feddersen *et al.*, "Moral Bias in Large Elections: Theory and Experimental Evidence", *American Political Science Review*, 103 (2009), pp. 175-92.

### III. OBJEÇÕES AO SUFRÁGIO RESTRITO

O princípio da competência não é o único princípio pelo qual julgar a distribuição de poder político. Na melhor das hipóteses, ele fornece uma condição necessária embora não suficiente para a alocação de poder político. Pode haver outros princípios deontológicos que restrinjam ou determinem a alocação de poder. Alguns modos de distribuir o poder tendem a produzir resultados políticos melhores (incluindo resultados mais justos) que outros. Presumivelmente, as consequências de diferentes alocações são também algo importante.

Até aqui não argumentei que, permanecendo as outras coisas iguais, o sufrágio restrito é moralmente superior ao sufrágio universal. Isto é, ainda não mostrei que alguma forma de epistocracia seria moralmente superior a uma opção em todos os outros aspectos igual, só que mais democrática. O sufrágio universal é moralmente objetável porque viola o princípio da competência. Mesmo assim, o sufrágio restrito poderia *também* ser moralmente objetável por outras razões. Talvez o sufrágio restrito viole algum outro princípio de justiça. Pode ser que o sufrágio universal incondicional e o sufrágio restrito sejam objetáveis e injustos *para o tanto*. Se for esse o caso, a moralidade pode exigir uma terceira alternativa de alocação de poder ou, caso não haja uma terceira alternativa viável ou defensável, a escolha do melhor dos dois sistemas.

Não vou neste artigo tentar determinar se há sistemas políticos alternativos (ou sistemas anárquicos) superiores à democracia com sufrágio universal e ao sistema eleitoral de elite com sufrágio restrito (por exemplo, uma outra resposta ao problema da incompetência democrática é limitar bastante o escopo do controle democrático ao reservar à constituição certas questões políticas, o que as colocaria, portanto, além do alcance das maiorias democráticas). Em vez disso, nesta seção pretendo examinar um argumento cuja proposta é mostrar que o sufrágio restrito é injusto. Nas duas seções seguintes,

argumentarei que mesmo que o sufrágio restrito seja injusto, é *menos* injusto do que o sufrágio universal. Se tivéssemos de escolher o menor de dois males, deveríamos escolher o sufrágio restrito.

Não disponho de espaço para responder a cada objeção interessante ao sufrágio restrito. A teoria democrática é um amplo subcampo da filosofia política e da teoria política. A maioria dos teóricos da democracia apoiam, ao invés de rejeitar, o sufrágio universal, e a maioria dos teóricos possui seu próprio argumento particular em favor da democracia com sufrágio universal. Não posso tratar desse campo inteiro aqui. O melhor que posso fazer é oferecer uma forte crítica contra o sufrágio universal e então responder a algumas das melhores críticas ao sufrágio restrito.

Assim, nesta seção irei considerar uma objeção feita por David Estlund em seu recente livro *Democratic Authority*. Há duas boas razões para levar em conta suas objeções e não as de outro. A primeira é que o próprio Estlund articulou um conjunto de objeções que enfraquecem uma grande variedade de teorias cujo propósito é defender a democracia. A segunda é que ele aceita várias das premissas básicas aceitas por mim, de modo que vale a pena investigar as razões pelas quais ele não chega às mesmas conclusões.

Acerca do primeiro ponto: muitos (talvez a maioria) dos teóricos da democracia oferecem defesas *procedimentalistas* da democracia. Eles apoiam o sufrágio universal porque pensam que é o único método de decisão justo. Estlund tem objeções poderosas a essas justificações procedimentalistas: “a ideia de justiça procedimental (...) é um valor muito frágil e ocasional para explicar, sem qualquer apelo a padrões de bons resultados independentes do procedimento, a importância moral da democracia.”<sup>95</sup> Ele argumenta que tomar decisões a partir do lançamento de uma moeda ou pela loteria é, em princípio, tão justo quanto tomar decisões por meio de uma democracia com sufrágio universal. Eu acrescentaria que, na prática, o lançamento de

---

<sup>95</sup>Estlund, “Replies to Saunders, Lister, and Quong”, *Representation*, 46 (2010), pp. 53-67, na p. 53.

moedas é certamente mais justo do que a democracia com sufrágio universal, uma vez que é menos sujeita ao rentismo, à manipulação, à corrupção e à demagogia do que os procedimentos democráticos de votação. Estlund conclui que não podemos justificar a democracia invocando a noção de justiça. O objetivo de manter os procedimentos justos não seleciona unicamente a democracia em preferência a outros métodos alternativos de tomada de decisão. Colocarei de lado as defesas procedimentalistas do sufrágio universal pelas razões que Estlund especifica.

Acerca do segundo ponto: Estlund aceita várias das premissas que aceito. Ele acredita que democracias podem tomar decisões melhores ou piores, que as democracias às vezes decidem de maneira imoral ou estúpida, que alguns cidadãos são muito mais competentes do que outros, que alguns cidadãos são moralmente irrazoáveis, que as pessoas não têm qualquer direito natural ao mando ou à posse de poder político, que parte da justificação de qualquer sistema político é a qualidade das políticas que implementa e que a qualidade das decisões políticas poderia ser aprimorada (talvez bastante) pelo sufrágio restrito. Poder-se-ia esperar que, assim, Estlund fosse favorável ao sufrágio restrito, mas ele é favorável ao sufrágio universal. Ele aceita várias das minhas premissas, mas não pensa que essas premissas justificam o sufrágio restrito. A esse respeito, ele é um grande oponente da posição que defendo aqui.

Uma razão pela qual vou para um lado enquanto Estlund vai para o outro é o fato de meu projeto ser diferente do dele. Estlund quer explicar como a democracia sob as condições certas poderia ser plenamente justa. Ele admite que essas condições podem ser um tanto ideais e que podem exigir dos eleitores muito mais do que eles provavelmente estão dispostos a dar.<sup>96</sup> Logo, mesmo que a teoria de Estlund sobre a legitimidade democrática esteja correta, ela não poderia justificar nenhuma democracia real, uma vez que nenhuma

<sup>96</sup>O livro *Democratic Authority*, cap. 14, “Utophobia: Concession and Aspiration in Democratic Theory”, explica a razão pela qual Estlund pensa que seu projeto normativo não é ameaçado se os Sistema que ele defende se revelar inviável. Ele admite (p. 275) que não está oferecendo uma teoria do segundo melhor, mas aqui estou oferecendo uma teoria do segundo melhor.

democracia real satisfaz as condições sob as quais a democracia é plenamente justificada. Sendo este o caso, teríamos de elaborar uma teoria não ideal, uma teoria do segundo melhor, para saber quais instituições defender quando as melhores instituições são inviáveis. Estlund não diz muito sobre isso (ver, e.g., p. 265). Neste artigo estou trabalhando em uma teoria não ideal. Entre dois sistemas políticos viáveis, mas imperfeitos e injustos, quero determinar qual é o preferível.

Estlund (p. 30) afirma que defesas da epistocracia tipicamente repousam em três princípios: um princípio da verdade, um do conhecimento e um da autoridade.

- *Princípio da verdade:* há respostas corretas para (ao menos algumas) questões políticas;
- *Princípio do conhecimento:* alguns cidadãos sabem mais dessas verdades do que outros;
- *Princípio da autoridade:* quando alguns cidadãos têm maior conhecimento justifica-se conceder a estes autoridade política sobre aqueles com menor conhecimento.

Ele pensa que deveríamos aceitar os princípios da verdade e do conhecimento. Alguns teóricos da democracia rejeitam esses princípios, mas suas razões para fazê-lo são profundamente implausíveis. Em vez disso, diz Estlund, deveríamos rejeitar o princípio da autoridade. O princípio da autoridade comete o que ele chama de “falácia do *expert*/chefe”. Cometer uma falácia do *expert*/chefe é pensar que ser um *expert* é razão suficiente para que uma pessoa detenha poder político sobre os outros. Mas possuir conhecimento superior não é suficiente para justificar *qualquer* poder, quanto mais um poder maior sobre os outros. Sempre podemos dizer aos *experts* “você pode até saber mais que eu, mas quem fez de você meu chefe?”. Um nutricionista, por exemplo, não pode me obrigar a obedecer a dieta mesmo sabendo que ela seria boa para mim. Você

não pode me forçar a ouvir o último álbum da Celine Dion mesmo se tiver uma prova indisputável de que eu o amaria. E assim por diante.

Mas o argumento que estou desenvolvendo em favor da epistocracia não se assenta no princípio da autoridade, mas no princípio da *anti*-autoridade.

- 3\*. *O princípio da anti-autoridade*: quando alguns cidadãos são moralmente irrazoáveis, ignorantes ou incompetentes sobre política justifica-se *não conceder* a eles poder político sobre os outros.

O princípio da competência é uma versão do princípio da anti-autoridade. Enquanto que o princípio da autoridade especifica qualificações para se ter poder, o princípio da anti-autoridade especifica desqualificações. Ao jogar o princípio da autoridade nas costas dos epistocratas, Estlund faz com que a defesa da epistocracia pareça mais difícil do que realmente é. Os epistocratas não precisam argumentar que os *experts* devem ser os chefes: eles apenas precisam argumentar que aquelas com pouca expertise não devem ser os chefes. O princípio da competência não diz que os *experts* deveriam ser os chefes; diz que pessoas incompetentes e irrazoáveis não deveriam ser impostas aos outros como chefes. O princípio da competência deixa em aberto se deveríamos ou não ter chefes e quais seriam as bases para tornar chefes algumas pessoas. Uma vez que meu argumento não se assenta no princípio da autoridade, não cometo a falácia do expert/chefe.

Estlund, no entanto, tem outro argumento contra a epistocracia. Ele concorda comigo (p. 262) que “remover os assuntos certos do controle democrático e entregá-los ao controle dos *experts* certos levaria a melhores decisões políticas e a mais justiça e prosperidade.” Mas ele acrescenta: “O truque é saber e justificar publicamente em quais *experts* confiar e quais assuntos entregar.” Ele não pensa que nós podemos resolver esse problema.

Estlund pensa que o sufrágio restrito viola um princípio de justiça ao qual chama de *exigência de aceitabilidade qualificada*. A exigência de aceitabilidade qualificada estabelece que qualquer base para a distribuição de poder político tem de ser aceitável para todos os pontos de vista qualificados.<sup>97</sup> O princípio, ou algo parecido com ele, é amplamente aceito por filósofos políticos liberais (eu o aceito também). De fato, alguma versão dele é um aspecto central ou definidor da filosofia política liberal.

Como um sistema eleitoral de elite (que veda o poder aos incompetentes) pode violar a exigência da aceitabilidade qualificada? Em uma tentativa de satisfazer o princípio da competência, suponha que as instituições das democracias modernas sejam mais ou menos as mesmas, com a diferença de que o sufrágio é agora restrito. Um modo pelo qual poderíamos fazer isso é a imposição de um exame de qualificação eleitoral, semelhante ao exame de direção, mas cuja função fosse testar em nível geral o básico das ciências sociais relevantes e um conhecimento básico acerca dos candidatos. O propósito do exame seria o de impedir o voto dos cidadãos severamente incompetentes. Isso seria feito barrando os cidadãos severamente ignorantes ou mal informados sobre a eleição, ou aqueles que não possuíssem expertise em ciências sociais para avaliar as políticas propostas por um candidato.

Um exame eleitoral escrito não é o único modo de tentar garantir a aplicação do princípio da competência. Consigo pensar em outros modos, mas a maioria deles é proibitivamente caro ou irrealista. Assim, discuto aqui os exames eleitorais porque eles são uma maneira óbvia de fazer com que o princípio da competência seja aplicado.

Estlund afirma que desqualificar ao poder alguns cidadãos com base em qualquer exame desse tipo violaria a exigência de aceitabilidade qualificada. Todas as pessoas razoáveis poderiam

<sup>97</sup>“Ninguém tem autoridade ou poder coercitivo legítimo sobre o outro sem uma justificação que poderia ser aceita por todos os pontos de vista qualificados”: Estlund, p. 33. Ele não especifica quais pontos de vista são qualificados e quais não o são, mas ele quer dizer que há um âmbito não especificado de perspectivas em relação as quais a autoridade política deve ser justificada.

aceitar que há uma distinção entre pessoas competentes e incompetentes e entre decisões tomadas de modo competente e incompetente. No entanto, não podemos esperar que todas as pessoas razoáveis concordem sobre o lugar onde a linha entre competência e incompetência deva ser traçada (isso não significa que inexistente uma verdade a respeito de como traçar a linha, mas apenas que essa verdade não será reconhecida como tal por todas as pessoas cujas opiniões são relevantes). Quase qualquer maneira de traçar a linha estará, na prática, sujeita a uma objeção vinda de alguém com um ponto de vista qualificado. Não há um critério facilmente identificável para distinguir as pessoas competentes das incompetentes que pudesse ser aceito por todos os pontos de vista qualificados. O sufrágio restrito dividiria as pessoas em duas classes - os aptos a exercer o poder e os não aptos - sobre uma base que nem todas as pessoas razoáveis poderiam aceitar. É, portanto, injusto usar a competência como uma base para atribuição de poder político.

Vou agora esboçar algumas razões, algumas de Estlund e outras minhas, para pensar que distribuir o poder com base no exame de competência eleitoral violaria a exigência de aceitabilidade qualificada. Por exemplo, algumas pessoas razoáveis poderiam objetar que o exame é imperfeito: alguém pode ser um eleitor competente mesmo que não passe no exame. O exame pode gerar muitos falsos negativos. Suponha que Sally nada conheça os candidatos, mas sabe que sua irmã inteligente e razoável apoia certos candidatos. Se Sally sabe que sua irmã tem boas credenciais, então o fato de que sua irmã prefere um candidato pode ser uma evidência suficientemente forte para Sally acreditar que o candidato é bom. Isso pode ser suficiente para fazer dela uma eleitora competente. No entanto, nós não podemos testar com facilidade esse tipo de competência.

Um exame poderia também atribuir pesos a diferentes fragmentos do conhecimento de um modo que não seria aceito por todos. Suponha, por exemplo, que há dois candidatos para a próxima

eleição. Suponha que A enfraqueceria as liberdades civis mas ajudaria a encorajar o crescimento econômico; B protegeria as liberdades civis mas causaria uma recessão. Suponha agora que o exame desqualifica alguns eleitores porque eles não têm noções básicas de economia. O conhecimento econômico é relevante para quase qualquer eleição, de modo que há boas razões para incluí-lo no exame. Uma pessoa razoável, entretanto, poderia acreditar que o candidato A é tão pior que o candidato B no quesito liberdades civis que, para esta eleição em particular, o conhecimento de economia não é necessário. Se uma pessoa razoável pode acreditar nisso, suspeito que Estlund afirmaria que nós não podemos usar nesse caso a ignorância em economia para desqualificar os cidadãos ao poder político.

Os exames poderiam também estar sujeitos ao que Estlund (p. 215) chama de *a objeção demográfica*:

A porção da população [que tem condições de ser aprovada no exame] pode desproporcionalmente ter características epistemicamente danosas a ponto de elas contrabalançarem os benefícios epistêmicos aceitos de se [ter o conhecimento exigido pelo exame].

(Sua objeção original era ao esquema de voto plural de Mill (que conferiria votos extras aos bem-educados), mas eu a modifiquei para usá-la aqui). Vou referir-me ao segmento da população que pode ser aprovado no exame como “eleitorado restrito”. Pode ocorrer de os membros do eleitorado restrito pertencerem de modo desproporcional a um gênero, uma raça ou classe. Estlund (p. 20) pensa, e concordo, que isso em si mesmo não seja objetável (nós concordamos que ninguém tem um direito básico ao exercício do poder e que não é inerentemente justo que membros de todos os grupos demográficos tenham os mesmos direitos a tal exercício). No entanto, uma pessoa razoável poderia temer que o eleitorado restrito, apesar de ser mensuravelmente mais competente do que a população em geral, possa possuir certos vieses latentes ou desconhecidos, preconceitos ou limitações cognitivas, de modo que ao se atribuir poder somente

ao eleitorado restrito, ele fará escolhas ainda piores do que as da população em geral sob o sufrágio universal. A objeção demográfica não afirma que um eleitorado restrito faria de fato escolhas piores do que um eleitorado irrestrito. Pelo contrário, ele provavelmente faria escolhas melhores. Mas Estlund argumenta que uma pessoa razoável *poderia* temer que o eleitorado restrito fizesse escolhas piores sistematicamente. Não podemos mostrar decisivamente que esse temor é infundado. Isso é suficiente, pensa Estlund, para mostrar que o sufrágio restrito não pode satisfazer a exigência de aceitabilidade qualificada.

Isso sumariza as objeções de Estlund ao sufrágio restrito. Se eu quisesse argumentar que a epistocracia pode ser perfeitamente justa, teria de refutar todas essas objeções e mostrar que a epistocracia pode em princípio satisfazer a exigência de aceitabilidade qualificada. Mas tudo o que quero neste artigo é mostrar que o sufrágio universal é moralmente *pior* do que um certo tipo de sufrágio restrito. Assim, para fins de argumentação irei aceitar que o sufrágio restrito é injusto e moralmente objetável exatamente pelas razões especificadas por Estlund (de fato penso que essas objeções podem ser superadas, mas demonstrar isso é um projeto mais extenso).

Suponha que eu esteja correto a respeito do princípio da competência e do que ele implica, e suponha que Estlund está correto a respeito da exigência de aceitabilidade qualificada e do que ela implica. Sendo assim, podemos então concluir que

- A. A democracia (na prática) é injusta porque viola o princípio da competência;
- B. Um sistema eleitoral de elite, no qual o direito ao voto é restrito àqueles que possam demonstrar competência, seria (na prática) injusto porque violaria a exigência de aceitabilidade qualificada.

Um democrata poderia objetar que nós poderíamos ao menos imaginar uma democracia na qual é sabido que todos os cidadãos são competentes, razoáveis e que tomam todas as decisões de modo competente e razoável. Na prática, não podemos esperar por isso, mas não se pode ao menos contar uma ficção científica coerente na qual a democracia não viola o princípio da competência? Talvez, mas também podemos imaginar uma ficção científica coerente na qual o exame de competência é feito de tal modo que nenhuma pessoa razoável pudesse oferecer objeções. Isto é, também podemos imaginar uma epistocracia que não viola a exigência de aceitabilidade qualificada, mesmo que não possamos esperar que isso ocorra na prática. Dessa forma, a democracia e a epistocracia estão no mesmo barco. Em suas formas ideais, elas não estariam sujeitas às objeções que apresentei aqui, mas em suas formas realistas elas estariam sujeitas a elas.

#### IV. ESCOLHENDO A MENOR INJUSTIÇA

Suponha, pelas razões especificadas acima, que a democracia com sufrágio universal e um sistema eleitoral de elite com sufrágio restrito sejam ambos injustos. Suponha que por alguma razão tivéssemos de escolher um ou o outro; não podemos escolher uma terceira alternativa mais justa. Qual deveríamos escolher?

O tipo de epistocracia que estou concebendo tem duas vantagens sobre a democracia:

1. A epistocracia viola a exigência de aceitabilidade qualificada, ao passo que a democracia viola o princípio da competência; no entanto, o modo como a democracia viola o princípio da competência é intrinsecamente pior do que o modo como a epistocracia viola a exigência de aceitabilidade qualificada;

2. Mantendo-se iguais todas as outras instituições, o sufrágio restrito irá produzir políticas melhores, mais justas e mais capazes de atingir a prosperidade e vários objetivos humanitários.

Nesta seção defenderei o ponto (1); na seguinte defenderei o ponto (2).

Violar o princípio da competência significa colocar a vida, liberdade e propriedade dos cidadãos, à força, nas mãos de pessoas irrazoáveis e incompetentes. Estamos acostumados com essa prática, mas ela é objetável pelas razões que especifiquei acima. Não devemos tolerar isso em julgamentos. Espero que os leitores consigam ver o paralelo entre democracias e júris e reconheçam que tomadas de decisão incompetentes em política são tão injustas (às vezes mais, às vezes menos) quanto em julgamentos. Infelizmente, estou lutando uma batalha difícil. A maioria dos leitores foi ensinada desde a infância que as democracias e o sufrágio universal são especialmente justos. Embora os argumentos que os professores do jardim de infância e os pais ofereceram não tenham sido particularmente bons – se fossem, os teóricos da democracia teriam pouco trabalho a fazer – a maioria dos leitores foi ainda treinada a ter respostas emocionais positivas à democracia (de modo similar, os súditos britânicos em 1550 teriam sido treinados a ter reações emocionais positivas pela monarquia, e qualquer um que tenha argumentado em favor das virtudes da democracia também lutou uma batalha difícil). Não posso esperar que os leitores sintam pelas democracias incompetentes a mesma repulsa moral que sentem por júris igualmente incompetentes, mesmo nos casos em que a importância e as consequências são as mesmas. Ainda assim, espero que eles possam enxergar a analogia tão longe quanto ela vai, e ela vai bem longe.

Suponha que nenhum sistema de exame de competência satisfaça a exigência de aceitabilidade qualificada. Não importa qual sistema concebamos para separar os cidadãos competentes e razoáveis dos

incompetentes e irrazoáveis, haverá ainda algumas objeções qualificadas ao sistema. Ao usarmos um exame para excluir cidadãos, quão ruim é essa injustiça? Bem, isso depende. Depende do quão bom nosso é nosso exame e de quão boa é a evidência que podemos fornecer em favor da qualidade do exame. Quanto melhor pudermos mostrar que o exame rastreia a diferença real entre competência e incompetência, menos objetável é o exame. Afinal, se tivéssemos evidência indisputável de que o exame rastreia realmente a distinção, nenhuma pessoa razoável poderia objetá-lo e dessa forma ele satisfaria a exigência de aceitabilidade qualificada.

Se queremos determinar o quão intrinsecamente injusto um sistema de exame de competência seria, podemos compará-lo com as leis existentes cuja função é impedir que certos cidadãos votem. De fato, a maior parte das democracias já exclui alguns cidadãos porque sustentam que eles são incompetentes para votar. Os critérios específicos que as democracias de fato usam para separar os competentes dos incompetentes não são justificáveis perante todos os pontos de vista qualificados.

A maioria das democracias estabelece limites legais de idade. Elas exigem que os cidadãos atinjam uma determinada idade antes de poder votar. A justificativa principal é a de que as crianças não são competentes para exercer o poder. No entanto, as leis de idade não conseguem satisfazer a exigência de aceitabilidade qualificada pelas mesmas razões que os exames de competência não a satisfazem.

Toda pessoa razoável pode aceitar que as crianças tendem a ser incompetentes para votar e que a maioria dos adultos poderia votar de modo competente se eles se esforçassem. Toda pessoa razoável pode aceitar que a transição (gradual) da incompetência à potencial competência política tende a ocorrer no fim da adolescência. Mas as leis de idade não rastreiam essa transição perfeitamente. Em vez disso, elas traçam uma luz vermelha brilhante que separa todos os cidadãos em eleitores e não eleitores, detentores e não detentores do poder, independentemente das habilidades individuais dos cidadãos.

Em alguns países, a idade para votar é dezesseis; em outros, dezoito. As leis de idade separam os cidadãos em classes de quem detém e não detém o poder com base em distinções específicas que não são justificáveis perante todos os pontos de vista qualificados. A menos que estabeleçamos um limite extremamente baixo (tal como 2 anos, de modo a não excluir crianças prodígio), alguma pessoa razoável pode oferecer objeções a essas leis. Esse modo de distribuir o poder político não satisfaz a exigência de aceitabilidade qualificada.<sup>98</sup>

Uma pessoa razoável também poderia oferecer a objeção demográfica de Estlund contra as leis de idade. Um temor é o de que adultos com mais de 18 anos, apesar de serem mensuravelmente mais competentes do que crianças com menos de 18, podem ter certos vieses latentes ou desconhecidos, preconceitos ou limitações cognitivas, e que um eleitorado limitado a cidadãos com mais de 18 anos faria escolhas piores do que a população em geral caso todos, incluindo as crianças, pudessem votar (por exemplo, adultos mais velhos têm um incentivo a legar a seus filhos um gasto público injustificadamente grande). Não afirmo que um eleitorado composto apenas de adultos realmente tomaria decisões piores do que um eleitorado irrestrito que incluísse crianças. Afirmo apenas que uma pessoa razoável *poderia* temer que um eleitorado mais velho fizesse escolhas sistematicamente piores. Estlund avança esse tipo de argumento contra outras bases propostas para a epistocracia, mas ele se aplica também aqui.

As leis de idade são, nessa medida, injustas. Não estou brincando: é realmente uma injustiça que certos adolescentes de 17 anos sejam declarados incompetentes para votar, especialmente quando David Duke, antigo grão-mestre da Ku Klux Klan, é declarado competente. Mesmo assim, comparada a outras injustiças, isso não é uma injustiça *horrível*. As leis de idade fazem bem o trabalho de rastrear uma distinção moral importante. No fim, os cidadãos abaixo da idade podem adquirir o direito ao voto simplesmente ficando mais velhos.

---

<sup>98</sup>Ver, e.g., S. Olsson, “Children’s Suffrage: a Critique of the Importance of Voter Knowledge for the Well-being of Democracy”, *International Journal of Children’s Rights*, 16 (2008), pp. 55-76.

As leis de idade são problemáticas porque traçam uma linha vermelha brilhante e artificial entre competentes e incompetentes de um modo que pessoas razoáveis poderiam oferecer objeções. Mas isso é o que também faria um bom exame de competência.

Um sistema bem administrado de exames de competência seria aproximadamente tão injusto quanto as leis de idade. Ele tentaria rastrear uma distinção moralmente importante, reconhecida como tal por pessoas razoáveis, mas o faria imperfeitamente, razão pela qual não poderia ser aceito por todas as pessoas razoáveis. Estlund (p. 37) teme que “Sob o sufrágio desigual, algumas pessoas estão formal e permanentemente sujeitas ao poder de outras.” Pelo contrário, um bom sistema de exames de competência não negaria permanentemente a posse do poder a ninguém (com exceção talvez das pessoas com problemas mentais severos). Em geral, caso houvesse disposição para se esforçar, qualquer um poderia se qualificar como um eleitor. Em geral, um sistema de exames de competência nos permitiria até mesmo eliminar as leis de idade. Se prodígios de três anos podem demonstrar competência, deixe-os votar! Veríamos até mesmo companhias como a *Princeton Review* (uma companhia que oferece aulas para melhorar o desempenho nos exames de admissão universitários nos Estados Unidos) oferecendo aos cidadãos cursos para aprovação no exame de desempenho.

A esta altura, penso que forneci uma base para decidir o que é mais intrinsecamente injusto entre o sufrágio restrito e o sufrágio universal. Meu argumento é este:

1. Tanto (a) leis de idade e (b) uma política que garante o cumprimento das decisões tomadas por júris incompetentes e irrazoáveis são intrinsecamente injustas, mas (b) é mais injusta do que (a);
2. Um sufrágio restrito àqueles que podem demonstrar competência seria tão injusto quanto (a);

3. O sufrágio universal seria aproximadamente tão injusto quanto (b);
4. Portanto, o sufrágio universal é intrinsecamente mais injusto do que o sufrágio restrito.

O sufrágio restrito é aproximadamente tão injusto quanto as leis de idade. Ele cria uma relação de poder entre diferentes classes de cidadãos baseada em uma distinção que todas as pessoas podem aceitar em abstrato, mas sobre a qual, na prática, haverá discordância razoável. Em contraste, o sufrágio universal é aproximadamente tão injusto quanto a imposição de uma política geral de cumprimento de decisões de júris até mesmo quando temos, em casos particulares, indícios conclusivos para pensar que os jurados foram incompetentes ou tomaram sua decisão de modo incompetente. O sufrágio universal, dessa forma, parece ser intrinsecamente mais injusto do que o sufrágio restrito.

Poder-se-ia objetar que exames de competência, mesmo se feitos corretamente, teriam alguns maus efeitos previsíveis. Por exemplo, poderia ser o caso de que membros de certos grupos em desvantagem e minorias historicamente oprimidas fossem desproporcionalmente desqualificados por qualquer exame do tipo. No entanto, sendo o caso, essa desqualificação *refletiria* e *resultaria* de uma injustiça subjacente, o que não significa que ela em si mesma seria injusta. Suponha, para efeito de comparação, que nós exijamos que os cidadãos devam ter graduação em medicina para exercer o ofício de cirurgião; mas membros de certos grupos historicamente oprimidos, graças às condições desvantajosas de seu desenvolvimento, não têm condições de frequentar uma escola de medicina. Isso não tornaria injusta uma lei que exigisse a graduação em medicina. Injusto é que membros de certos grupos estejam em desvantagem devido a opressões passadas.

Assumi acriticamente aqui que a democracia satisfaz a exigência de aceitabilidade qualificada. No entanto, pode ser o caso de que a

própria objeção que apresento neste artigo mostre que a democracia não satisfaz tal exigência. Minha objeção seria uma “objeção qualificada” à democracia (Estlund não nos fornece uma teoria das objeções qualificadas, mas a minha objeção parece ser uma boa candidata).<sup>99</sup> Estlund diz que democracias têm autoridade e legitimidade caso se saiam melhor do que decisões tomadas ao acaso. Minha objeção vai no sentido de que as decisões políticas devem ser tomadas competentemente por pessoas competentes: não é suficiente que elas sejam melhores que o acaso. Embora isso soe como dupla contagem, talvez a democracia, de fato, tenha contra si dois aspectos. O primeiro é que a democracia com sufrágio universal viola o princípio da competência ao sustentar que as decisões têm autoridade e legitimidade mesmo quando tomadas de modo incompetente ou por pessoas incompetentes. O segundo é que ao violar o princípio da competência, a democracia também viola a exigência de aceitabilidade qualificada, uma vez que uma pessoa qualificada pode exigir que a democracia satisfaça o princípio da competência.

## V. AS CONSEQUÊNCIAS DOS SUFRÁGIOS UNIVERSAL E RESTRITO

Argumentei que o sufrágio universal é intrinsecamente pior do que o sufrágio restrito. Mas tanto o princípio da competência quanto a exigência de aceitabilidade qualificada parecem ser (algo como) condições necessárias para a alocação de poder político. Talvez seja melhor, até certo ponto, produzir resultados justos de uma maneira injusta do que produzir resultados injustos de uma maneira justa. Decidir na prática entre sufrágio universal e sufrágio restrito requer o exame de suas consequências esperadas. Qual tipo de regime tende na prática a produzir políticas melhores?

Os sistemas eleitorais de elite pretendem eliminar eleitores de baixa qualidade e assim aprimorar a qualidade das decisões. Pode

---

<sup>99</sup>Agradeço a um avaliador anônimo por ter sugerido esse ponto.

parecer que restringir o sufrágio terá como consequência direta a obtenção de resultados (em geral) melhores. Nem tanto.

Afinal, exames de competência no mundo real serão redigidos e inspecionados por pessoas reais, e não por anjos morais perfeitamente competentes. Pessoas reais frequentemente respondem mal a maus incentivos. Na prática, o exame de competência é um convite ao abuso e à captura institucional. Os exames de competência poderiam ser usados para excluir as pessoas que votariam contra o partido no poder. Grupos de interesse especiais poderiam lutar para controlar a agência que supervisiona os exames. Por mais que o exame fosse equânime e justo em princípio, é pouco provável que seria administrado, na prática, de uma maneira perfeitamente equânime e justa. Se o sistema de exames de competência for mal dirigido, o sufrágio restrito pode produzir resultados até piores do que o sufrágio universal. Mesmo quando o governo garante por meio de lei que apenas os competentes terão poder, não há garantia de que a garantia legal terá sucesso.

Por outro lado, temos bons estudos sobre as patologias que envolvem a tomada de decisão democrática sob o sufrágio universal. Bryan Caplan argumenta que os eleitores são sistematicamente enviesados e estão sistematicamente equivocados sobre que políticas econômicas irão promover seus fins.<sup>100</sup> Sua situação é pior do que a ignorância econômica. Eles sabem menos do que nada. Scott Althaus usa conjuntos de dados diferentes dos usados por Caplan, mas chega às mesmas conclusões.<sup>101</sup> Ele argumenta que cidadãos bem informados têm preferências políticas sistematicamente diferentes das exibidas por cidadãos desinformados. Se Caplan e Althaus estiverem corretos, os eleitores em uma democracia têm

---

<sup>100</sup>B. Caplan, “Systematically Biased Beliefs about Economics: Robust Evidence of Judgmental Anomalies from the Survey of Americans and Economists on the Economy”, *Economic Journal*, 112 (2002), pp. 1-26, e *The Myth of the Rational Voter* (Princeton University Press, 2008).

<sup>101</sup>S. Althaus, *Collective Preferences in Democratic Politics: Opinion Surveys and the Will of the People* (Cambridge University Press, 2003). Althaus estima estatisticamente as preferências políticas de um “público ilustrado”, isto é, uma sociedade demograficamente idêntica aos Estados Unidos mas na qual todos os cidadãos têm conhecimento político completamente objetivo. Althaus conclui que o público ilustrado tem crenças sistematicamente diferentes daqueles do público real e não ilustrado, cujas preferências não são aleatoriamente distribuídas.

desempenho pior do que o acaso na tarefa de selecionar líderes e políticas. Andrew Healy produziu uma série de artigos mostrando que os eleitores nos Estados Unidos e em outros lugares culpam candidatos eleitos pelo mau tempo.<sup>102</sup> Diana Mutz mostrou que os cidadãos mais politicamente ativos tendem a ser (em minhas palavras, não nas dela) ideólogos caricaturais.<sup>103</sup> Drew Westen documentou casos de tomadas de decisão emocionalmente enviesadas [*Motivated reasoning*] e irracionalidade entre cidadãos politicamente ativos.<sup>104</sup> E assim por diante.<sup>105</sup>

Não vou fornecer aqui uma demonstração completa de que a incompetência política é disseminada. Este artigo é de filosofia normativa e não um trabalho de ciências sociais. Meu objetivo não é mostrar quais são os fatos sobre a performance dos eleitores, mas perguntar o que deveríamos fazer à luz dos fatos, sejam eles quais forem. Ainda assim, há evidência significativa de que muitos eleitores são politicamente incompetentes. Isso não é uma surpresa. Como Thomas Christiano (um ardente democrata) diz, “É difícil ver como os cidadãos podem satisfazer mesmo que moderadamente os padrões para crenças acerca de como melhor atingir seus objetivos políticos.”<sup>106</sup> Isso exigiria o conhecimento de “uma imensa quantidade de ciência social” e “de fatos particulares”, conhecimento que falta à maioria dos cidadãos. A aquisição desse conhecimento por parte dos cidadãos exigiria deles o abandono da divisão do trabalho na sociedade de modo que todos pudessem se tornar cientistas políticos, sociólogos ou economistas.

---

<sup>102</sup>Ver, e.g., A. Healy e N. Mahotra, “Random Events, Economic Losses, and Retrospective Voting: Implications for Democratic Competence”, *Quarterly Journal of Political Science*, 5 (2010), pp. 193-208. Ver também J. Wolfers, “Are Voters Rational? Evidence From Gubernatorial Elections” [2006], disponível em [http://works.bepress.com/justin\\_wolfers/1](http://works.bepress.com/justin_wolfers/1).

<sup>103</sup>Ver D. Mutz, *Hearing the Other Side* (Cambridge University Press, 2006), pp. 30-5, 128.

<sup>104</sup>D. Westen et al., “The Neural Basis of Motivated Reasoning”, *Journal of Cognitive Neuroscience*, 18 (2006), pp. 1947-58.

<sup>105</sup>Ver também D. Austen-Smith e W. H. Riker, “Asymmetric Information and the Coherence of Legislation”, *American Political Science Review*, 81 (1987), pp. 897-918; J. Zaller, *The Nature and Origins of Mass Opinion* (Cambridge University Press, 1992); R. R. Lau e D. P. Redlawsk, “Voting Correctly”, *American Political Science Review*, 84 (1997), pp. 585-98.

<sup>106</sup>T. Christiano, “Democracy”, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (2006), <http://plato.stanford.edu/entries/democracy>.

O sucesso do argumento em favor do sufrágio restrito depende em parte de uma questão empírica que não estou em condições de responder. Podemos estudar o quão mal os eleitores se comportam e assim determinar progressos *potenciais* que poderiam ser produzidos por um sistema de exames de competência eleitoral. Mas podemos apenas especular a respeito de quais seriam os progressos *efetivos*, uma vez que não temos certeza de quão bem funcionará o sistema de exame de competência ou qualquer outra tentativa de vedar o poder político aos incompetentes e irrazoáveis.

Não sabemos com certeza se os sistemas de exame de competência produziram resultados melhores ou piores do que a democracia com sufrágio universal. No entanto, como argumentei, esses sistemas são intrinsecamente menos injustos do que as democracias com sufrágio universal. Há boas razões para pensar que eles irão produzir melhores resultados do que os produzidos pela democracia com sufrágio universal, muito embora haja razões para temer que não será o caso. Uma vez que não temos certeza das consequências, mas temos ao mesmo tempo razões para esperar que elas seriam positivas, podemos, de início, experimentar sistemas de exame de competência em uma escala pequena. Seria ótimo, por exemplo, que um estado americano testasse o sistema primeiro (seria melhor, penso, começar com um estado relativamente não corrupto, como New Hampshire, ao invés de um estado corrupto, como Rhode Island). Se o experimento for bem-sucedido, as regras poderiam ser então estendidas. Afinal, há poucos séculos, as pessoas tinham pouca experiência com a democracia. Alguns defendiam a democracia em parte porque acreditavam que ela tendia a produzir resultados melhores e mais justos do que a monarquia. Outros, por sua vez, temiam que a democracia fosse ainda mais corrupta ou então que colapsasse no caos. À luz da falta de experiência, um democrata poderia, com razoabilidade, ter argumentado em favor da ideia de testar a democracia em escalas relativamente pequenas e, em caso de sucesso, estendê-la.

Há alguns argumentos até mesmo contra a realização de experimentos em pequena escala. Um democrata poderia, por exemplo, invocar uma forma robusta de conservadorismo burkiano. A advertência de Edmund Burke foi a de que é arriscado tentar reconstruir a sociedade com base em fantasias filosóficas. A sociedade é complexa, complexa demais para que uma simples teoria lhe dê conta, de modo que os planos podem ter más consequências imprevisíveis. Em contraste com isso, Burke afirma, há uma presunção em favor das instituições sociais pré-existentes. Embora essas instituições possam parecer injustas e imperfeitas com a reflexão filosófica, elas têm ao menos um histórico de funcionamento. Instituições legais e políticas evoluíram ao longo de gerações – com efeito, elas se adaptaram. Tal como devemos ter reservas quanto à interferência em um ecossistema, o conservador burkiano pensa que devemos ter reservas a respeito da substituição dos sistemas políticos existentes. O conservadorismo burkiano estava preocupado com a reconstrução, de uma só vez, da sociedade a partir de suas fundações. Ele não era contra tentar pequenas melhorias aqui e ali. Ele provavelmente seria favorável à ideia de experimentar em uma escala modesta, progredindo em caso de sucesso a escalas maiores.

## V. SUMÁRIO E CONCLUSÃO

A democracia tem feito muito bem às pessoas. Os melhores governos que existem, aqui e agora, são os democráticos com sufrágio universal. Isso não significa que a democracia seja perfeitamente justa, ou que ninguém poderia fazer ainda melhor. A democracia tem algumas características profundamente objetáveis – em particular, ela frequentemente impõe sobre as pessoas (por meio da força) políticas que são resultado de decisões feitas de modo incompetente.

Em qualquer forma realista, a democracia viola o princípio da competência, ao passo que a epistocracia viola a exigência de aceitabilidade qualificada (ou, ao menos, para fins de argumentação,

aceito que ela o faça). Tanto a democracia quanto a epistocracia são, na prática, injustas. No entanto, se violarmos a exigência de aceitabilidade qualificada, podemos esperar, no total, consequências melhores (incluindo políticas mais justas) do que se violarmos o princípio da competência. O modo como a democracia viola o princípio da competência é inerentemente mais objetável do que o modo como a epistocracia viola a exigência de aceitabilidade qualificada. Sendo assim, se tivéssemos de escolher entre a democracia com sufrágio universal ou uma epistocracia semidemocrática com sufrágio restrito, deveríamos escolher a última.

**TRADUÇÃO DAS NOTAS DE “CONHECIMENTO A *PRIORI*”, DE  
ALBERT CASULLO**

- <sup>i</sup> Immanuel Kant, *Critique of Pure Reason*, trad. N. K. Smith (New York: St. Martin's Press, 1965).
- <sup>ii</sup> *Ibidem*, 43.
- <sup>iii</sup> Philip Kitcher, *The Nature of Mathematical Knowledge* (New York: Oxford University Press, 1983), cap. 1.
- <sup>iv</sup> W. V. Quine, "Two Dogmas of Empiricism" in *From a Logical Point of View*, 2<sup>a</sup> ed. (New York, Harper and How, 1963).
- <sup>v</sup> Hilary Putnam, "'Two Dogmas' Revisited", in *Realism and Reason: Philosophical Papers*, vol. 3 (Cambridge: Cambridge University Press, 1983).
- <sup>vi</sup> Veja, por exemplo, Albert Casullo, "Revisability, Reliabilism, and A Priori Knowledge", *Philosophy and Phenomenological Research* 49 (1988): 187-213; Aron Edidin "A Priori Knowledge for Fallibilists," *Philosophical Studies* 46 (1984): 189-197; Bob Hale, *Abstract Objects* (Oxford: Basil Blackwell, 1987), chap. 6; and Donna Summerfield, "Modest A Priori Knowledge," *Philosophy and Phenomenological Research* 51 (1991): 39-66. Os artigos de Casullo, Edidin, e Summerfield foram reimpressos em *A Priori Knowledge*, ed. Albert Casullo (Aldershot: Dartmouth, 1999). Para bibliografias mais compreensivas sobre o *a priori*, veja Albert Casullo, "A Priori Knowledge Appraised," in *A Priori Knowledge*, ed. Casullo, and *A Priori Knowledge*, ed. Paul K. Moser (Oxford: Oxford University Press, 1987).
- <sup>vii</sup> Paul Benacerraf, "Mathematical Truth", *Journal of Philosophy* 70 (1973): 661-679.
- <sup>viii</sup> R. M. Chisholm, *Theory of Knowledge*, 3<sup>a</sup> ed (Englewood Cliffs, N. J.: Prentice Hall, 1989), 26-28, fornece uma abordagem cogente da perspectiva tradicional.
- <sup>ix</sup> Veja, por exemplo, Albert Casullo, "Causality, Reliabilism, and Mathematical Knowledge," *Philosophy and Phenomenological Research* 52 (1992): 557-584; Hartry Field, *Realism, Mathematics and Modality* (Oxford: Blackwell, 1989); Bob Hale, "Is Platonism Epistemologically Bankrupt?," *Philosophical Review* 103 (1994): 299-324; Jerrold J. Katz, "What Mathematical Knowledge Could Be," *Mind* 104 (1995): 491-522; e Penelope Maddy, "Mathematical Epistemology: What Is the Question?," *Monist* 67 (1984): 46-55. Os quatro artigos foram reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>x</sup> Saul Kripke, "Identity and Necessity", in *Identity and Individuation*, ed. M. K. Munitz (New York: New York University Press, 1971), e *Naming and Necessity* (Cambridge, Mass.: Harvard University Press, 1980).
- <sup>xi</sup> Veja, por exemplo, C. Anthony Anderson, "Toward a Logic of A Priori Knowledge," *Philosophical Topics* 21 (1993): 1-20; Albert Casullo, "Kripke on the A Priori and the Necessary," *Analysis* 37 (1977): 152-159; Keith S. Donnellan, "The Contingent A Priori and Rigid Designators," in *Contemporary Perspectives on the Philosophy of Language*, ed. P. French et al. (Minneapolis: University of Minnesota Press, 1979); Gareth Evans, "Reference and Contingency," *Monist* 62 (1979): 161-189; Philip Kitcher, "Apriority and Necessity," *Australasian Journal of Philosophy* 58 (1980): 89-101; and R. G. Swinburne, "Analyticity, Necessity, and Apriority," *Mind* 84 (1975): 225-243. Os artigos de Casullo, Kitcher e Swinburne foram reimpressos em *A Priori Knowledge*, ed. Moser. O artigo de Anderson foi reimpresso *A Priori Knowledge*, ed. Casullo.
- <sup>xii</sup> Veja, por exemplo, Carl Hempel, "On the Nature of Mathematical Truth", in *Necessary Truth*, ed. R. C. Sleigh (Englewood Cliffs, N.J.: Prentice Hall, 1972); e A. J. Ayer, *Language, Truth and Logic* (New York: Dover, 1952).
- <sup>xiii</sup> Quine, "Two Dogmas".
- <sup>xiv</sup> Veja, por exemplo, Paul A. Boghossian, "Analyticity Reconsidered," *Nous* 30(1996): 360-391; Laurence Bonjour, "A Rationalist Manifesto," *Canadian Journal of Philosophy*, suppl. vol. 18 (1992): 53-88; M. Giaquinto, "Non-Analytic Conceptual Knowledge," *Mind* 105 (1996): 249-268; Gilbert Harman, "Analyticity Regained?," *Nous* 30 (1996): 392-400; e Putnam, "'Two Dogmas' Revisited." O primeiro dos quatro artigos foi reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>xv</sup> Kant, *Critique of Pure Reason*, 42, diz que "tais modos universais de conhecimento, que ao mesmo tempo possuem o caráter da necessidade interna, têm de ser, independente da experiência, claros e certos. São denominados, portanto, de conhecimento *a priori*."
- <sup>xvi</sup> Para uma discussão mais compreensiva da análise do conceito de conhecimento *a priori*, veja os caps. 1-3 de Albert Casullo, *A Priori Justification* (New York: Oxford University Press, 2003).
- <sup>xvii</sup> Veja, por exemplo, Kitcher, *The Nature of Mathematical Knowledge*, e Swinburne.
- <sup>xviii</sup> Anthony Quinton, "The A Priori and the Analytic", in *Necessary Truth*, ed. Sleigh, 90.
- <sup>xix</sup> *Ibidem*, 92.
- <sup>xx</sup> *Ibidem*, 93. A ênfase é de Quinton.
- <sup>xxi</sup> Swinburne, in *A Priori Knowledge*, ed. Moser, 186-187.
- <sup>xxii</sup> *Ibidem*, 186.
- <sup>xxiii</sup> Laurence Bonjour, *The Structure of Empirical Knowledge* (Cambridge: Harvard University Press, 1985), 192. Bonjour já não defende mais essa concepção.
- <sup>xxiv</sup> Veja, por exemplo, Chisholm, *Theory of Knowledge*, 41.
- <sup>xxv</sup> *Ibidem*, 28.
- <sup>xxvi</sup> *Ibidem*.
- <sup>xxvii</sup> *Ibidem*, 29.
- <sup>xxviii</sup> *Ibidem*, 26. A passagem citada é de G. W. Leibniz, *New Essays on Human Understanding*, trad. e ed. Peter Remnant e Jonathan Bennett (New York: Cambridge University Press, 1982) livro IV, cap. 7.
- <sup>xxix</sup> Kripke, *Naming and Necessity*; e Kitcher, "Apriority and Necessity".

- <sup>xxx</sup> Chisholm, *Theory of Knowledge*, 12, diz que “ $p$  é certa para  $S =_{df}$  Para todo  $q$ , acreditar que  $p$  está mais justificada para  $S$  do que recusar  $q$ , e acreditar que  $p$  está pelo menos tão justificada para  $S$  quanto acreditar que  $q$ .”
- <sup>xxxi</sup> Veja Panayot Butcharov, *The Concept of Knowledge* (Evanston: Northwestern University Press, 1970), parte 1, seção 9; e John Pollock, *Knowledge and Justification* (Princeton: Princeton University Press, 1974), cap. 10.
- <sup>xxxii</sup> Butcharov, *The Concept of Knowledge*, 93.
- <sup>xxxiii</sup> Laurence Bonjour, *In Defense of Pure Reason* (Cambridge: Cambridge University Press, 1998), 106-110.
- <sup>xxxiv</sup> Kitcher, *The Nature of Mathematical Knowledge*, 89.
- <sup>xxxv</sup> Hilary Putnam, “Analyticity and Apriority: Beyond Wittgenstein e Quine”, in *Realism and Reason: Philosophical Papers*, vol. 3 (Cambridge: Cambridge University Press, 1983), 127.
- <sup>xxxvi</sup> É comum distinguir entre as crenças que são *diretamente* justificadas e as que são *indiretamente* justificadas por fontes não-experienciais. As que são indiretamente justificadas são justificadas exclusivamente por outras crenças que são diretamente justificadas por fontes não-experienciais. Para facilidade de exposição não introduzo essa distinção em minhas formulações. O leitor deveria as considerar como implícitas nessa e nas formulações subsequentes.
- <sup>xxxvii</sup> Kripke, *Naming and Necessity*”, 35.
- <sup>xxxviii</sup> Kitcher, *The Nature of Mathematical Knowledge*, 22. A plausibilidade do critério de Kitcher deriva da observação de que o seguinte argumento é intuitivamente inválido:  $S$  sabe que  $p$ . É possível que  $S$  saiba *a priori* que  $p$ . Portanto,  $S$  sabe *a priori* que  $p$ .
- <sup>xxxix</sup> Sigo Kitcher aqui, *The Nature of Mathematical Knowledge*, ao supor que as modalidades colapsam.
- <sup>xl</sup> A crença justificada de  $S$  de que não- $p$  é um revogador refutante para a crença justificada de  $S$  que  $p$ . A crença justificada de  $S$  de que a justificação de  $S$  para a crença de que  $p$  é inadequada ou defectiva é um revogador *refutante* para a crença justificada de  $S$  de que  $p$ .
- <sup>xli</sup> Há versões de cada um desses três tipos de argumento formuladas em termos de *justificação* ao invés de *conhecimento*. Para facilidade de exposição, ao oferecer caracterizações gerais, ofereço apenas a versão formulada em termos de *conhecimento*. Uso o termo “empirismo radical” para designar a perspectiva que nega a existência do conhecimento *a priori*, e o termo “apriorismo” para designar a perspectiva que *afirma* a existência de tal conhecimento. Similarmente, uso o termo “empirista radical” para designar uma pessoa ou teoria que endossa o empirismo radical, e “apriorista” para designar uma pessoa ou teoria que endossa o apriorismo. Para uma discussão mais compreensiva sobre os argumentos a favor, veja Casullo, *A Priori Justification*, cap. 4.
- <sup>xlii</sup> Hilary Putnam, “There is at Least One A Priori Truth”, in *Realism and Reason: Philosophical Papers*, vol. 3 (Cambridge: Cambridge University Press, 1983), 98.
- <sup>xliii</sup> Veja, por exemplo, Hilary Putnam, “The Analytic and the Synthetic”, in *Mind, Language and Reality: Philosophical Papers*, vol. 2 (Cambridge: Cambridge University Press, 1975).
- <sup>xliv</sup> Putnam, “There is at Least One A Priori Truth”, 106.
- <sup>xlv</sup> Kant, 43. A tese de Kant é repetida por Bertrand Russell, *The Problems of Philosophy* (Oxford: Oxford University Press, 1971), cap. 7; e Roderick Chisholm, *Theory of Knowledge*, 2ª ed. (Englewood Cliffs, N. J.: Prentice Hall, 1977), cap. 3.
- <sup>xlvi</sup> *Ibidem*.
- <sup>xlvii</sup> *Ibidem*, 52.
- <sup>xlviii</sup> Saul Kripke, “Identity and Necessity” e “Naming and Necessity”.
- <sup>xlix</sup> Tyler Burge, “Content Preservation”, *Philosophical Review* 102 (1993): 457-488, disputa essa ideia. O artigo de Burge foi reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>l</sup> Veja, por exemplo, R. M. Chisholm, *Theory of Knowledge*, 2ª ed., 37; e C. McGinn “*A Priori* and *A Posteriori* Knowledge”, *Proceedings of Aristotelian Society* 76 (1975-6), 204. Philip Kitcher, “Apriority and Necessity”, 100-101, também sustenta que a plausibilidade da versão Modal do Argumento da Necessidade depende dessa tese. Ele continua a rejeitar o argumento por razões diferentes da minha.
- <sup>li</sup> Carl Hempel, “On the Nature of Mathematical Truth”, 36. A. J. Ayer, *Language, Truth, and Logic*, 75-76, oferece um argumento similar.
- <sup>lii</sup> Laurence Bonjour, *In Defense of Pure Reason*, 4. Para uma discussão mais compreensiva da posição de Bonjour, veja Casullo, “The Coherence of Empiricism”, *Pacific Philosophical Quarterly* 81 (2000): 31-48.
- <sup>liii</sup> *Ibidem*, 91.
- <sup>liv</sup> *Ibidem*, 92.
- <sup>lv</sup> Bonjour, *ibidem*, 162, articula a perspectiva com se segue:  
Uma pessoa apreende ou compreende, por exemplo, as propriedades da vermelhidão e da verdura, e supostamente “vê” com base nessa apreensão que elas não podem ser conjuntamente instanciadas. Tal imagem claramente parece pressupor que como resultada dessa apreensão ou compreensão, as propriedades da vermelhidão e da verdura estão diante da mente de um modo que permite que suas naturezas e incompatibilidade mútua sejam aparentes.
- <sup>lvi</sup> *Ibidem*, 184. A ênfase é minha.
- <sup>lvii</sup> *Ibidem*, 145.
- <sup>lviii</sup> Hilary Putnam, “There is at Least One A Priori Truth”, 98. Putnam fornece um resumo lúcido de seu argumentos contra o *a priori* em “‘Two Dogmas’ Revisited”. Para uma discussão mais compreensiva dos argumentos contra, veja Casullo, *A Priori Justification*, cap. 5.
- <sup>lix</sup> Kitcher, *The Nature of Mathematical Knowledge*, 88.

- <sup>lx</sup> A autocorreção ocorre em graus. Uma forma mais fraca pode ser definida como se segue: para *alguma* crença falsa de que produzida em S por  $\phi$ ,  $\phi$  pode também justificar a crença para S de que não-*p*. Outras versões, evidentemente, tanto mais fortes quanto mais fracas, são possíveis. Uso a versão forte nesse contexto, uma vez que ela produz um argumento mais direto.
- <sup>lxi</sup> Veja, por exemplo, D. M. Armstrong, “Is Introspective Knowledge Incorrigible?”, *Philosophical Review* 72 (1963): 417-432.
- <sup>lxii</sup> John Stuart Mill, *A System of Logic*, ed. J. M. Robson (Toronto: University of Toronto Press, 1973), livro II, caps. V e VI.
- <sup>lxiii</sup> *Ibidem*, 41.
- <sup>lxiv</sup> Quine, “Two Dogmas of Empiricism”.
- <sup>lxv</sup> Putnam, “Two Dogmas’ Revisited”.
- <sup>lxvi</sup> Paul Benacerraf, “Mathematical Truth”, *Journal of Philosophy* 70 (1973): 661-679.
- <sup>lxvii</sup> W. D. Hart, “Review of Mark Steiner, *Mathematical Knowledge*”, *Journal of Philosophy* 74 (1977): 125-126, argumenta que “é um crime contra o intelecto tentar mascarar o problema de naturalizar a epistemologia da matemática com alvoroço filosófico. Preocupações superficiais acerca da higiene intelectual das teorias causais do conhecimento são irrelevante e enganadoras a esse problema, pois o problema não é tanto sobre causalidade quanto é sobre a própria possibilidade do conhecimento natural dos objetos abstratos”.
- <sup>lxviii</sup> W. V. Quine, “Epistemology Naturalized”, in *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969).
- <sup>lxix</sup> Alvin Goldman, *Epistemology and Cognition*, (Cambridge, Mass.: Harvard University Press, 1986), 63.
- <sup>lxx</sup> Essa é uma variação do argumento apresentado por Bonjour, *The Structure of Empirical Knowledge*, 40, e discutido por Goldman, *Epistemology and Cognition*, 111-112.
- <sup>lxxi</sup> Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973), cap. 9, e *Change in View* (Cambridge: MIT Press, 1986), cap. 5, chama forçosamente a atenção para a importância da evidência que não está disponível. Alvin Goldman, “What is Justified Belief?” in *Justification and Knowledge*, ed. George S. Pappas (Dordrecht: Reidel, 1979), 20, reconhece que a relevância dos processos de formação de crenças disponível numa abordagem anterior da evidência enfraquecedora. Essa abordagem, no entanto, é demasiado restritiva para dar conta do caso de Maud, uma vez que Goldman exclui explicitamente a coleta de nova evidência do escopo dos processos disponíveis.
- <sup>lxxii</sup> Penelope Maddy, *Realism in Mathematics* (Oxford: Oxford University Press, 1990), 43.
- <sup>lxxiii</sup> *Ibidem*, 44.
- <sup>lxxiv</sup> *Ibidem*, 45-46.
- <sup>lxxv</sup> *Ibidem*, 31.
- <sup>lxxvi</sup> Essa proposta é elaborada e defendida de maneira mais completa contra potenciais objeções em Casullo, *A Priori Justification*, cap. 6.
- <sup>lxxvii</sup> Nosso foco aqui, e na discussão subsequente, é sobre as fontes da justificação *a priori não-inferencial*, ou *básica*, uma vez que a justificação *a priori inferencial*, ou não básica, resulta da aplicação de princípios inferenciais que são (não-inferencialmente) justificados *a priori* a outras crenças que são (não-inferencialmente) justificadas *a priori*. Em última análise, portanto, todas as crenças justificadas *a priori* são em última instância justificadas por essas fontes.
- <sup>lxxviii</sup> Plantinga, *Warrant and Proper Function*, 106.
- <sup>lxxix</sup> *Ibidem*, 105.
- <sup>lxxx</sup> Bonjour, *In Defense of Pure Reason*, 106.
- <sup>lxxxii</sup> *Ibidem*, 108.
- <sup>lxxxiii</sup> *Ibidem*, nn. 12 e 13.
- <sup>lxxxiii</sup> George Bealer, “A Priori Knowledge and the Scope of Philosophy”, *Philosophical Studies* 81 (1996): 123, reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>lxxxiv</sup> Ernest Sosa, “Rational Intuition: Bealer on Its Nature and Epistemic Status”, *Philosophical Studies* 81 (1996): 154, reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>lxxxv</sup> George Bealer, “A Priori Knowledge: Replies to William Lycan and Ernest Sosa,” *Philosophical Studies* 81 (1996): 169, reimpresso em *A Priori Knowledge*, ed. Casullo.
- <sup>lxxxvi</sup> A ilusão Müller-Lyer fornece um exemplo de uma aparência sensível.
- <sup>lxxxvii</sup> Bonjour, em *In Defense of Pure Reason*, 133-137. Ele oferece duas descrições diferentes do que ocorre quando o um sujeito cognoscente não satisfaz a condição de fundo para a justificação por um *insight* racional aparente: (1) o sujeito cognoscente não tem sequer um *insight* racional aparente; e (2) a força justificativa do *insight* racional aparente é suspendida.
- <sup>lxxxviii</sup> Alvin Goldman, “A Priori Warrant and Naturalistic Epistemology”, *Philosophical Perspectives* 13 (1999): 1-28, argumenta que os estudos psicológicos são relevantes para a existência do conhecimento *a priori*. Seu foco é em se tais estudos apoiam a perspectiva de que as habilidades matemáticas e lógicas são inatas.
- <sup>lxxxix</sup> George Bealer, Alvin Plantinga e Ernest Sosa endossam tal condição, embora haja diferenças em suas posições. Bealer, “A prior knowledge and the Scope of Philosophy”, 129, endossa uma concepção confiabilista das *fontes básicas de evidência*: “algo é uma fonte básica de evidência se tem certo tipo de ligação com a verdade”. Plantinga, *Warrant and Proper Function*, 17, endossa uma exigência confiabilista da *garantia*: “o módulo do projeto que rege a sua produção tem de ser tal que seja objetivamente provável que uma crença produzida pelas faculdades cognitivas que

funcionam apropriadamente de acordo com esse módulo (num ambiente adequado) será verdadeira ou verossímil”. Sosa, “Modal and Other *A Priori* Epistemology: How Can We Know What is Possible and What Impossible?”, *Southern Journal of Philosophy* 38, suplem. (2000): 4, endossa uma condição confiabilista da *justificação epistêmica*: “A justificação epistêmica de uma crença B no instante de tempo *t* pode assim exigir a produção de B em *t* através de uma virtude V residente nesse sujeito. O que se exige para que uma disposição V seja uma virtude é que em circunstâncias normais V produzisse uma preponderância suficiente de crenças verdadeiras nos sujeitos como S”. Embora Laurence Bonjour, *In Defense of Pure Reason*, 1, rejeite as abordagens confiabilistas da justificação epistêmica, ele introduz a verocondutividade em sua caracterização das *razões epistêmicas*: “O conhecimento requer, ao invés, que a crença em questão seja justificada ou racional de um modo que esteja internamente conectada com o objetivo definido da atividade cognitiva, isto é, que haja uma razão que aumente, a um grau apropriado, as chances de que a crença seja *verdadeira*. A justificação desse tipo distintivo, verocondutora, será aqui referida como *justificação epistêmica*”.

<sup>xc</sup> O exemplo mais familiar sustenta que, no caso da crença *a priori* básica (ou *não inferencialmente*) justificada de que *p*, compreender que *p* é suficiente para “ver” que *p* é verdadeira.