

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL

ESCOLA DE HUMANIDADES

PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

LUCAS ROISENBERG RODRIGUES

**AS EPISTEMOLOGIAS MODAIS E SEUS CRÍTICOS**

PORTO ALEGRE

2017

LUCAS ROISENBERG RODRIGUES

**AS EPISTEMOLOGIAS MODAIS E SEUS CRÍTICOS**

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

Orientador: Dr. Cláudio Gonçalves de Almeida

PORTO ALEGRE

2017

## Ficha Catalográfica

R696e Rodrigues, Lucas Roisenberg

As epistemologias modais e seus críticos / Lucas Roisenberg  
Rodrigues . – 2017.

121 f.

Tese (Doutorado) – Programa de Pós-Graduação em Filosofia,  
PUCRS.

Orientador: Prof. Dr. Cláudio Gonçalves de Almeida.

1. Análise do conhecimento. 2. Teoria do rastreamento. 3. Teoria da  
segurança. 4. Epistemologia modal. 5. Princípio de sensibilidade. I.  
Almeida, Cláudio Gonçalves de. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS  
com os dados fornecidos pelo(a) autor(a).

LUCAS ROISENBERG RODRIGUES

**AS EPISTEMOLOGIAS MODAIS E SEUS CRÍTICOS**

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

Aprovada em: \_\_\_\_ de \_\_\_\_\_ de \_\_\_\_\_.

**BANCA EXAMINADORA:**

---

Prof. Dr. Tiegue Viera Rodrigues - UFMT

---

Prof. Dr. Rogel Esteves de Oliveira - PUCRS

---

Prof. Dr. Rodrigo Martins Borges – PUCRS

---

Prof. Dr. Luís Fernando Munaretti da Rosa - Munich Center for Mathematical  
Philosophy

---

Prof. Dr. Cláudio Gonçalves de Almeida - PUCRS

PORTO ALEGRE

2017

Dedico este trabalho ao meu pai, Telmo Elias Rodrigues.

## **AGRADECIMENTOS**

Agradeço ao meu orientador, Cláudio Gonçalves de Almeida, à minha esposa, Anelise, cuja ajuda e apoio foram imprescindíveis, aos meus colegas com quem tive proveitosas discussões todos esses anos e a todos aqueles que me incentivaram e contribuíram para que este trabalho viesse a ser realizado. Por fim, agradeço aos membros da comissão examinadora, Tiegue Viera Rodrigues, Rogel Esteves de Oliveira, Rodrigo Martins Borges e Luís Fernando Munaretti da Rosa, pelas valiosas observações e correções.

## RESUMO

Nesta tese, trato de um conjunto de teorias, aqui denominadas de *epistemologias modais*, que buscam explicar e/ou analisar o conhecimento a partir de condições ou princípios contrafactuais, isto é, princípios que dizem respeito ao que aconteceria em situações ou circunstâncias não atuais. Inicialmente, explico a teoria do rastreamento, proposta por Robert Nozick, que utiliza dois condicionais contrafactuais, os princípios da sensibilidade e aderência, para analisar o conhecimento e tentar solucionar difíceis problemas epistemológicos, notadamente o problema do ceticismo. Após explicar a proposta de Nozick, exponho diversas objeções à sua teoria, que foram aqui divididas em quatro tipos: restritivas, permissivas, violações do princípio de fecho epistêmico e relativas ao modo como Nozick entende o conceito de *método*. Pretendo mostrar que o conjunto das objeções aqui apresentadas sugere fortemente que a teoria de Nozick não é correta. Após tratar da teoria do rastreamento, apresento a principal sucessora e epistemologia modal concorrente da teoria de Nozick: a teoria da segurança. Os dois principais defensores da condição de segurança são Ernest Sosa e Duncan Pritchard, e dedico uma seção a cada um dos autores. No que diz respeito a Duncan Pritchard, alego que sua teoria não consegue explicar adequadamente nossas atribuições de ignorância e conhecimento em certas situações lotéricas. Também procuro mostrar que certas objeções, tais como o contraexemplo da festa de *Halloween*, proposto originalmente por Juan Comesaña, não foram adequadamente respondidas pelos defensores da segurança, e que as principais respostas disponíveis a essas objeções não são promissoras.

**Palavras-chave:** Análise do conhecimento. Teoria do rastreamento. Teoria da segurança. Epistemologia modal. Princípio de sensibilidade.

## ABSTRACT

This thesis is about some theories, here called *modal epistemologies*, that attempt to explain and/or analyze knowledge by means of counterfactuals principles or conditions, that is, principles that talk about what would be the case in non-actual circumstances. Initially, I explain the tracking theory, proposed by Robert Nozick, that employ two counterfactuals conditionals, the sensitivity and adherence principles, to analyze knowledge and solve difficult epistemological problems, such as the skeptical problem. After explaining tracking theory, I present multiple objections to Nozick, which were here divided in four main kinds: restrictive, permissive, violations of epistemic closure and problems regarding Nozick's conception of methods. I claim that the objections here presented strongly suggests that Nozick's theory cannot be correct. I also explain the main competitor with tracking: safety theory. The two main proponents of safety are Ernest Sosa and Duncan Pritchard, and I study each author within a separate section. Concerning Pritchard, I claim that he cannot explain knowledge and ignorance attributions in lottery examples. I also try to show that some replies to safety's objectors, such as the *Halloween's* party counterexample, originally made by Juan Comesaña, are not correct, and that the main response's available against Comesaña's objection are not very promising.

**Keywords:** Knowledge analysis. Tracking theory. Safety theory. Modal epistemology. Sensitivity principle.

## LISTA DE SÍMBOLOS

$\square\rightarrow$  - Condicional contrafactual ou subjuntivo

$\beta_m p$  - sujeito acredita que p com base em um método m

$\beta$  - sujeito acredita que

$Kp$  - sujeito sabe que p

$\neg$  - negação

$\vee$  - disjunção

$\in$  - elemento pertence ao conjunto

$\&$  - conjunção

$\diamond$  - possibilidade

$\rightarrow$  - condicional material

$\square$  - é necessário que

$\leftrightarrow$  - bicondicional

$\mapsto$  - relação causal

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
<b>2</b>	<b>EXPLICANDO A TEORIA DO RASTREAMENTO.....</b>	<b>17</b>
2.1	A REFERÊNCIA AOS MÉTODOS.....	18
2.2	COMO NOZICK INDIVIDUA OS MÉTODOS.....	20
2.3	A SEMÂNTICA DOS CONTRAFCTUAIS .....	23
2.3.1	As propostas de Lewis e Stalnaker .....	24
2.4	O PROBLEMA DE GETTIER.....	28
2.5	O PROBLEMA DA LOTERIA.....	34
2.6	A RESPOSTA DE NOZICK AO CETICISMO .....	35
2.6.1	A resposta de Nozick .....	37
2.6.2	Princípios de fecho do conhecimento .....	38
2.6.2.1	<b>Reformulando o paradoxo do ceticismo .....</b>	<b>42</b>
2.7	A CONDIÇÃO DE ADERÊNCIA .....	43
2.7.1	Aderência e verdades logicamente necessárias.....	44
2.7.2	A resposta de Nozick ao problema das verdades necessárias.....	46
2.7.3	Condição de aderência e os contrafactuais verdade-verdade.....	47
2.8	A TEORIA DE NOZICK DO CONHECIMENTO INFERENCIAL.....	49
<b>3</b>	<b>OBJEÇÕES À TEORIA DO RASTREAMENTO .....</b>	<b>51</b>
3.1	A TEORIA É MUITO RESTRITIVA .....	51
3.1.1	Indução.....	51
3.1.2	A resposta de Nozick .....	52
3.1.3	A crítica de Vogel .....	54
3.1.4	Reversão de contrafactuais .....	55
3.1.5	Um último recurso: a resposta de Kelly Becker .....	57
3.1.6	O conhecimento de segunda-ordem.....	59
3.1.6.1	<b>As respostas de Becker e Salerno .....</b>	<b>61</b>
3.1.7	Disjunções.....	67
3.1.7.1	<b>As testemunhas corroborantes.....</b>	<b>67</b>
3.1.8	O reator .....	68
3.1.9	Os contraexemplos de Peter Klein.....	70
3.1.10	Margens de erro .....	72

<b>3.1.10.1</b>	<b>A resposta de becker ao problema da margem de erro.....</b>	<b>73</b>
3.2	A TEORIA É MUITO PERMISSIVA.....	74
3.2.1	Conjunções.....	74
3.2.2	Klein sobre rastreamento e verdades necessárias .....	77
3.2.3	Conhecimento de segunda ordem e verdades necessárias .....	77
3.3	VIOLAÇÕES IMPLAUSÍVEIS DO FECHO EPISTÊMICO.....	79
3.3.1	O celeiro vermelho.....	79
3.3.2	Contrafactuais disjuntivos e o caso do celeiro vermelho.....	81
3.3.3	A resposta de Adams e Clarke à objeção de Kripke.....	83
3.3.4	O cachorro <i>daschund</i> .....	85
3.4	PROBLEMAS COM A CARACTERIZAÇÃO DO MÉTODO .....	86
3.5	CONSIDERAÇÕES FINAIS SOBRE O PRINCÍPIO DE SENSIBILIDADE	87
<b>4</b>	<b>A CONDIÇÃO DE SEGURANÇA .....</b>	<b>89</b>
4.1	A CONDIÇÃO DE SEGURANÇA EM SOSA .....	90
4.1.1	Segurança e conhecimento indutivo .....	92
4.1.2	O fecho epistêmico .....	93
4.1.3	Segurança e a resposta ao ceticismo .....	94
4.2	A CONDIÇÃO DE SEGURANÇA EM PRITCHARD .....	95
4.2.1	O Conceito de sorte.....	96
4.2.2	Sorte Epistêmica .....	98
4.2.3	O princípio de segurança .....	99
<b>4.2.3.1</b>	<b>Refinamentos da condição de segurança .....</b>	<b>100</b>
4.2.4	A crítica de McEvoy .....	103
4.2.5	Por que a incapacidade de responder ao exemplo da loteria é problemática	105
4.2.6	Por que loteria <sub>1</sub> não é um caso de sorte epistêmica .....	107
4.2.7	Segurança e verdades necessárias .....	109
4.3	ALGUNS CONTRAEXEMPLOS À CONDIÇÃO DE SEGURANÇA ...	111
4.3.1	Diagnosticando os contraexemplos.....	112
4.3.2	A resposta de Bogardus .....	113
4.3.3	Relativizar aos métodos e às circunstâncias não resolve o problema....	116
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>119</b>

<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>121</b>
--	------------

## 1 INTRODUÇÃO

Nos últimos 50 anos ocorreu o renascimento do debate a respeito do conceito de *conhecimento*. Na origem desse fenômeno, encontram-se alguns devastadores contraexemplos propostos por Gettier (1963) à denominada *análise tradicional do conhecimento*, segundo a qual conhecimento é crença verdadeira justificada. A lição que muitos filósofos tiraram dessas objeções é que a justificação, como quer que venhamos a entendê-la mais especificamente, não basta para excluir a acidentalidade ou sorte epistêmica, qual seja, a sorte ou coincidência que é incompatível com a posse do conhecimento<sup>1</sup>. Tal constatação originou o denominado *problema de Gettier*, que entendo da seguinte forma: que condição ou cláusula deve complementar ou substituir a justificação, a fim de excluir a sorte epistêmica? O princípio ou tese que exclui a sorte epistêmica é a denominada *condição anti-sorte*.

Nesta tese, tratamos de uma proeminente família de teorias epistemológicas - abarcadas aqui sob o rótulo de *epistemologias modais*- que surgiram na continuação das reflexões sobre o problema de Gettier. Elas sugerem que a condição anti-sorte, responsável pela geração do problema de Gettier, é algum tipo de condição contrafactual, que diz respeito ao que seria o caso em certas circunstâncias ou situações (ou, conforme diremos ao longo deste trabalho, mundos possíveis) não atuais.

São duas as principais epistemologias modais: a teoria do rastreamento e a teoria da segurança. A primeira delas, desenvolvida por Robert Nozick, identifica dois princípios contrafactuais, a sensibilidade e a aderência, embora seja a sensibilidade o mais importante e plausível entre os dois<sup>2</sup>.

A teoria do rastreamento pretende fornecer uma análise do conceito de *conhecimento*, e resolver outros enigmas filosóficos, além do já referido problema de Gettier. Em seu engenhoso tratamento do ceticismo, Nozick (1981) propõe que a rejeição de uma importante tese, algumas vezes denominada *princípio de fecho epistêmico*, é a chave para uma resposta ao ceticismo filosófico. Conforme veremos, sua proposta é essencialmente conciliatória, e pre-

---

<sup>1</sup> O primeiro filósofo a entender as objeções de Gettier dessa maneira foi Unger (1968). Desde então, é praxe compreender o problema sob esta perspectiva. Uma reflexão extensa sobre as relações entre sorte e conhecimento pode ser encontrada em Pritchard (2005).

<sup>2</sup> A teoria das razões conclusivas, proposta por Dretske (1970), é uma predecessora da teoria de Nozick, e contém diversos elementos em comum com ela. Todavia, aqui não abordo diretamente teoria das razões conclusivas de Dretske.

tende harmonizar as pretensões de conhecimento do senso comum com certos *insights* cétricos.

Ao longo desta tese, procurarei mostrar a inviabilidade da teoria epistemológica de Nozick, e para tanto apresentarei diversas objeções. Por conveniência, separei-as em quatro tipos principais: restritivas, permissivas, violações implausíveis do fecho epistêmico e relativas ao conceito de método<sup>3</sup>.

Direi que uma análise ou explicação do conhecimento é muito restritiva se e somente se ela classifica certos casos de conhecimento como casos de ignorância. As objeções que pretendem mostrar que a teoria do rastreamento é muito restritiva são as seguintes:

- O contraexemplo das crenças indutivas (seções 3.1.1 a 3.1.5);
- A objeção do conhecimento de segunda ordem (seção 3.1.6);
- O problema das disjunções (seção 3.1.7);
- O contraexemplo do reator (seção 3.1.8)
- O contraexemplo do tanque de gasolina (seção 3.1.9);
- A objeção das margens de erro (seção 3.1.10).

Direi que uma análise ou explicação do conhecimento é muito permissiva se e somente se ela classifica certos casos de ignorância como casos de conhecimento. As objeções que pretendem mostrar que a teoria do rastreamento é muito permissiva são:

- O problema das conjunções (seção 3.2.1);
- O problema das verdades necessárias (seções 3.2.2);
- Crenças de segunda ordem com verdades necessárias de primeira ordem (seção 3.2.5);

As objeções restritivas alegam que a teoria do rastreamento não propõe condições *necessárias* adequadas para a presença de conhecimento. As objeções permissivas alegam que a teoria do rastreamento não propõe condições conjuntamente *suficientes* adequadas para a presença de conhecimento.

Entre as objeções acima listadas, existem algumas que merecem destaque. Do lado das objeções restritivas, o problema das crenças indutivas é o mais difícil e complicado para o defensor da teoria do rastreamento. Do lado das objeções permissivas, os problemas das ver-

---

<sup>3</sup> Divisão semelhante, mas não idêntica, é realizada por Becker (2007).

dades necessárias, e das crenças de segunda ordem com verdades necessárias de primeira ordem, são bastante contundentes.

Além dos contraexemplos permissivos e restritivos, a teoria do rastreamento também enfrenta outras duas objeções. Uma delas alega que a teoria do rastreamento produz violações implausíveis do princípio de fecho epistêmico, e que estas violações não possuem qualquer base intuitiva ou independente (diferentemente do que acontece com a resposta de Nozick ao ceticismo). O principal contraexemplo relativo a casos como este é o contraexemplo do celeiro vermelho. Tratarei desta questão na seção 3.31.

A outra objeção é relativa ao conceito de *método*, visto que Nozick julgou pertinente inserir o conceito em sua análise do conhecimento. A absorção do conceito de *método* por uma análise do conhecimento parece trazer consigo certas dificuldades. Na seção 3.4, explicarei por que o critério de individuação de métodos proposto por Nozick é insatisfatório.

A segunda epistemologia modal de que trataremos, desenvolvida para suprir as deficiências originais da teoria do rastreamento, é a denominada *teoria da segurança*, defendida principalmente por Ernest Sosa e Duncan Pritchard<sup>4</sup>. O princípio de segurança recebeu diferentes formulações, nem todas consistentes entre si, e que frequentemente variam conforme o autor e do texto em questão. Veremos as diferenças específicas nos momentos apropriados. No capítulo 4 tratarei das posições dos dois principais defensores da teoria da segurança: Ernest Sosa e Duncan Pritchard.

Também apresentarei algumas objeções que considero relevantes. Em primeiro lugar, defenderei que Duncan Pritchard não consegue solucionar adequadamente o problema da loteria, tal como alega fazer. Em particular, argumentarei nas seções 4.2.4, 4.2.5, e 4.2.6 que a tentativa de solucionar o problema da loteria no quadro da teoria da segurança gera diversos problemas para Pritchard, e que estes dificilmente serão solucionados.

Na seção 4.3, destaquei dois contraexemplos: contraexemplo da festa de *Halloween* e contraexemplo da água contaminada. Conforme argumentarei, tais contraexemplos não podem ser respondidos pela teoria da segurança sem artifícios duvidosos. O resultado final é que a teoria da segurança não possui uma resposta adequada a tais problemas.

---

<sup>4</sup> Sosa defendeu condição de segurança, com formulações ligeiramente diferentes, em Sosa (1999b), Sosa (1999a) e Sosa (2002). Pritchard defendeu que a condição de segurança em diversos lugares, especialmente em Pritchard (2005). Williamson (2000) também defendeu uma variação da condição da segurança, ainda que não no contexto de uma análise do conhecimento. Não abordarei nesta tese o tratamento dado por Williamson à condição de segurança.

Apesar da condição de segurança representar um avanço em relação à sensibilidade, a condição de segurança possui defeitos graves que a inviabilizam, e que não foram afastados convincentemente pelos seus defensores. Assim, nenhuma das duas teorias modais examinadas parece forte o suficiente para resistir à pesada artilharia de objeções a elas dirigidas, e o projeto de uma epistemologia modal deve ser seriamente reconsiderado. É isto que tentarei mostrar nas páginas a seguir.

## 2 EXPLICANDO A TEORIA DO RASTREAMENTO

A primeira epistemologia modal de que trataremos é a teoria do rastreamento, defendida por Nozick (1981). O autor analisou o conceito de *conhecimento* em termos de condicionais subjuntivos ou contrafactuais, que correlacionam a verdade ou falsidade da proposição alvo e as respectivas crenças do sujeito<sup>5</sup>. A principal tese (mas não a única, diga-se de passagem) da teoria nozickiana é o princípio de sensibilidade, o qual afirma, *grosso modo*, ser uma condição necessária para que um sujeito S saiba que é verdadeira uma proposição *p*, que se fosse falso que *p*, então S não acreditaria que *p*. Quando uma crença de um sujeito S de que *p* satisfaz o princípio de sensibilidade, direi que a crença de S de que *p* é *sensível* à verdade (ou simplesmente sensível). Assim, o princípio da sensibilidade diz que S sabe que *p* somente se S crê sensivelmente que *p*.

O princípio de sensibilidade tem considerável força intuitiva. Muitos autores, incluindo-se muitos críticos da teoria nozickiana, reconhecem a plausibilidade da condição de sensibilidade, e o fato de que ela parece ser a explicação correta para muitas atribuições de ignorância (i.e., ausência de conhecimento). Assim, por exemplo, Ernest Sosa escreve<sup>6</sup>:

O requisito de que uma crença deve ser sensível se ela vier a ser (corretamente caracterizada) como "conhecimento" é geralmente considerado *prima facie* plausível: em muitos casos, julga-se intuitivamente que a falha de uma crença em ser (corretamente caracterizada) como "conhecimento" pode ser explicada pelo fato de que a crença permaneceria em seu lugar, mesmo que fosse falsa (nas circunstâncias determinadas pelo contexto de atribuição) (SOSA, 1999a, p.146).

Nozick adicionou um segundo contrafactual a sua análise do conhecimento: a condição de aderência. Direi que uma crença de um sujeito S de que *p* é aderente se e somente se: se fosse verdadeiro que *p*, então S acreditaria que *p*. A aderência desempenha um papel importante na teoria nozickiana. Se a sensibilidade, conjuntamente com a verdade e a crença, fosse suficiente para o conhecimento, então Nozick não teria meios de explicar, entre outras coisas, como é possível distinguir crenças em verdades necessárias de conhecimento de verdades necessárias, uma vez que é impossível que a proposição alvo seja falsa.

---

<sup>5</sup> Por *proposição alvo* entendo a proposição que é objeto de alguma atitude proposicional de um sujeito em um determinado contexto. Assim, por exemplo, se um sujeito S acredita que *p*, então a proposição alvo é *p*.

<sup>6</sup> Todas as citações contidas nesta tese consistem em traduções realizadas por mim, a partir do original em inglês.

Em uma versão simplificada, a teoria do rastreamento afirma que um sujeito *S* sabe que uma proposição *p* é verdadeira se e somente se as quatro seguintes condições são satisfeitas:

1. **Verdade** - é verdade que *p*;
2. **Crença** - *S* acredita que *p*;
3. **Sensibilidade simples** - se fosse falso que *p*, então *S* não acreditaria que *p*;
4. **Aderência simples** - se fosse verdade que *p*, então *S* acreditaria que *p*.

As duas primeiras cláusulas são as tradicionais condições da verdade e da crença, e normalmente integram todas as análises do conceito de *conhecimento*. Não podemos saber que uma proposição é verdadeira, quando ela não é verdadeira; não podemos saber que uma proposição é verdadeira, quando sequer acreditamos que ela é verdadeira. As duas últimas, por sua vez, são o núcleo da teoria, e constituem o que há de propriamente original e interessante na proposta. Quando todas as quatro cláusulas da análise são satisfeitas, direi que a crença do sujeito cognoscente rastreia a verdade.

Nozick julgou ser necessário alterar a versão simplificada acima da teoria do rastreamento, tendo em vista as dificuldades encontradas por esta versão simples. A adição mais relevante foi a inclusão da referência aos métodos de formação de crenças nos respectivos condicionais contrafactuais da sensibilidade e aderência.

## 2.1 A REFERÊNCIA AOS MÉTODOS

Consideremos o seguinte exemplo sugerido por Nozick (1981). Uma senhora idosa está conversando com o seu neto, que se encontra à sua frente. Não há nada de particularmente anormal ou estranho no ambiente (ela está observando o seu neto em boas condições de iluminação, distância, etc.), tampouco ela sofre de qualquer problema cognitivo ou existe qualquer deficiência em seu aparato perceptual. Ela parece saber que ele está vivo. Contudo, se o neto não estivesse vivo, os familiares, sabendo dos problemas de saúde dela, afirmariam que o neto está vivo, o que levaria, por sua vez, a senhora idosa a crer falsamente que ele está vivo. Ainda que nós digamos que ela saiba que o neto está vivo, a crença da senhora idosa não

é sensível à verdade: é falso que se o neto não estivesse vivo, então ela não acreditaria que ele está vivo.

Para responder a contraexemplos como esse, Nozick inseriu uma referência ao método de formação de crenças. Se o neto não estivesse vivo, a senhora idosa formaria a crença de que o neto está vivo com base no testemunho dos familiares. Contudo, no mundo atual, em que ela observa diretamente o próprio neto, ela forma sua crença de acordo com as experiências perceptuais que presentemente tem.

O exemplo da vovó sugere que devemos fixar o método ou maneira como a crença foi formada, quando avaliamos se os contrafactuais relevantes são satisfeitos. A referência ao método coloca-nos diante da seguinte questão, que adiciona mais complexidade à teoria: como avaliar os casos em que existem diferentes métodos operando simultaneamente<sup>7</sup>? Consideremos o seguinte exemplo mencionado por Nozick, que teve sua origem em Armstrong (1973):

Um pai acredita que seu filho é inocente tanto por causa da fé em seu filho, como (agora) porque ele viu evidências apresentadas no tribunal que demonstram a sua inocência. A sua crença através do método da demonstração da corte satisfaz 1-4 [as quatro condições da teoria de Nozick], vamos supor, porém o método baseado na fé não as satisfaz. Se o seu filho fosse culpado, ele ainda acreditaria que o filho é inocente, por causa da sua fé (NOZICK, 1981, p.180).

No exemplo acima, dois ou mais métodos produzem simultaneamente a mesma crença. Se avaliarmos a situação epistêmica do infeliz pai de acordo com um dos métodos, deveríamos dizer que o pai sabe que o filho é inocente; se avaliarmos a situação de acordo com o outro, diríamos que ele não sabe. Como determinar qual das duas alternativas é a correta?

Para resolver o impasse, Nozick sugere que devemos comparar os veredictos dos métodos em circunstâncias contrafactuais, e verificar qual dos métodos prevaleceria em casos de conflito. Um método  $m_2$  prevalece sobre um  $m_1$  se e somente se quando o resultado das aplicações de  $m_2$  e  $m_1$  são distintas, o sujeito S forma a sua crença de acordo com o método  $m_2$ .

Aplicando a solução de Nozick ao exemplo do tribunal, parece claro que a teoria diz que o pai não sabe que o seu filho é inocente. Embora o método do tribunal seja sensível à verdade, o método da fé não é, e esse prevalece sobre aquele: caso a corte julgasse que o filho

---

<sup>7</sup> Não trato aqui de verificar se é possível que dois métodos diferentes e independentes possam gerar simultaneamente uma mesma crença. Nozick considera este um fenômeno possível, e alterou a sua teoria em razão dele, e, por isto, a questão deve ser aqui mencionada.

é culpado, ainda assim o pai acreditaria, por causa da sua fé inabalável, que o seu filho é inocente.

Fazendo novamente as alterações necessárias, uma versão mais desenvolvida da teoria afirma que S sabe que p por meio de um método m se, e se somente se:

- **Verdade** - é verdade que p;
- **Crença** - S acredita que p com base em um método m;
- **Sensibilidade** - se fosse falso que p, então S não acreditaria que p com base em m;
- **Aderência** - se fosse verdadeiro que p, então S acreditaria que p com base em m;
- **Métodos múltiplos** - não existe nenhum método M', tal que S crê que p com base em m', m' não satisfaz a condições da sensibilidade e/ou aderência e m' prevalece sobre m.

Simbolicamente, as condições da aderência e sensibilidade serão representadas assim<sup>8</sup>:

- **Sensibilidade** -  $\neg p \Box \rightarrow \neg \beta_m p$ ;
- **Aderência** -  $p \Box \rightarrow \beta_m p$ .

## 2.2 COMO NOZICK INDIVIDUA OS MÉTODOS

Conforme vimos na seção anterior, Nozick julgou necessário relativizar os condicionais da sensibilidade e aderências aos métodos empregados. É preciso esclarecer o conceito de *método* e torná-lo, na medida do possível, um pouco mais preciso. A tarefa é essencial, tendo em vista que, dependendo da caracterização do conceito, diferentes juízos sobre presença ou ausência de conhecimento serão produzidos.

*Grosso modo*, o método de formação de uma crença é o modo ou maneira como uma crença foi formada. Fundamentalmente, as caracterizações dos métodos podem variar segundo três dimensões: localização, acessibilidade e generalidade<sup>9</sup>. Irei primeiro explicar em que consistem cada uma dessas três dimensões, em seguida tratarei da sugestão nozickiana e logo depois de como ela se classifica segundo o esquema aqui proposto.

---

<sup>8</sup> Por *p-mundo(s)* entendo o(s) mundo(s) em que p é verdadeiro.

<sup>9</sup> A diferença entre as três dimensões não é apontada por Nozick. Uma esquematização parecida é sugerida por Rabinowitz (2016).

Métodos podem estar localizados interna ou externamente. Métodos possuem localização interna se fatores referentes a propriedades mentais intrínsecas do sujeito (por exemplo, a ocorrência de certas aparências perceptuais com certos conteúdos) não fazem parte do critério de identidade dos métodos. Métodos possuem localização externa se fatores que não dizem respeito às propriedades mentais intrínsecas do sujeito fazem parte do critério de identidade dos métodos.

Podemos definir a localização dos métodos por meio do conceito auxiliar de *duplos mentais*. S e S' são duplos mentais relativamente à uma proposição *p* se e somente se possuem as mesmas propriedades mentais intrínsecas relevantes<sup>10</sup>. Assim, vamos supor que os duplos mentais S e S' creem que *p*, e que *p* = *eu tenho mãos*. Se S parece estar percebendo suas mãos, então S' também o está; se S possui uma certa evidência E, de natureza perceptual, para a crença de que *p*, então S' também possui a evidência E para a crença de que *p*.

Dois sujeitos em circunstâncias diferentes podem ser duplos mentais, desde que as referidas diferenças sejam meramente externas. Por exemplo, duplos mentais em ambientes radicalmente diferentes, como um sujeito em mundo "normal" e sua contraparte em um cenário cético, compartilham precisamente os mesmos métodos, quando estes são localizados internamente<sup>11</sup>. É irrelevante, para a identificação dos métodos, quando localizados internamente, que a crença de um seja causada pelo objeto externo percebido, e a crença do outro seja produzida, por exemplo, por um gênio maligno. O que conta para a identificação de um método é como as coisas são "internamente".

A acessibilidade de um método diz respeito ao conhecimento (potencial ou atual) que um sujeito tem de que está empregando um certo método, quando o utiliza. Um método *m* é acessível para S somente se: se S utiliza um método *m*, então ele sabe (ou está em posição de saber) que está utilizando o método *m*.

Por fim, métodos podem ser caracterizados de acordo com diferentes níveis ou graus de generalidade. Por exemplo, eu tenho agora a crença de que existe um computador em minha frente. Podemos caracterizar método de formação da crença de forma *muito* ampla. De forma extremamente genérica, poderíamos dizer que o método é simplesmente a percepção.

---

<sup>10</sup> Não é implicado pelo que digo que S e S' sejam sujeitos diferentes. Eles podem ser o mesmo sujeito em situações diferentes.

<sup>11</sup> É importante notar que externistas de conteúdo mental não aceitariam essa maneira de entender a localização interna dos métodos. Por conseguinte, esta classificação precisaria ser reformulada, caso viessemos a adotar alguma forma de externismo de conteúdo mental.

Um pouco menos, mas ainda bem genericamente, poderíamos dizer que o método é simplesmente percepção visual. Em um nível de generalidade muito mais estrito, poder-se-ia dizer que o método empregado é *formação de crenças visuais com base em aparências perceptuais de tal forma*.

A generalidade é gradativa, podendo ir de um extremo (generalidade ampla) a um outro extremo (generalidade restrita), ou ocupar qualquer uma das posições intermediárias. A maneira como um autor identifica ou individua os métodos pode estar mais próxima de um dos extremos, ou do outro. Vejamos agora o que Nozick entende por *métodos*, e logo em seguida como sua proposta pode ser classificada de acordo com o esquema que apresentamos. Em uma significativa passagem, ele afirma o seguinte:

Uma pessoa pode empregar um método (no meu sentido) sem proceder metodicamente, e sem conhecimento ou consciência do método que está utilizando. Normalmente, um método terá um resultado final na experiência na qual a crença é baseada, tal como ocorre com a experiência visual, e então: (a) nenhum método sem este resultado será o mesmo método, e (b) qualquer método experiencialmente idêntico, "o mesmo de um ponto de vista interno", contará como o mesmo método. Baseando nossas crenças em nossas experiências, eu e a pessoa flutuando no tanque estamos utilizando, para estes propósitos, o mesmo método. (NOZICK, 1981, p.184)

O que se sugere, nessas poucas linhas dedicadas ao assunto, é que os métodos são caracterizados internamente, de acordo com a experiência do sujeito<sup>12</sup>. Assim, se em duas circunstâncias diferentes, digamos as circunstâncias  $\alpha$  e  $\beta$ , um sujeito S possui experiências relevantes diferentes, então o sujeito S utiliza métodos diferentes nas circunstâncias  $\alpha$  e  $\beta$ . De forma semelhante, se em duas circunstâncias diferentes, digamos as circunstâncias  $\gamma$  e  $\phi$ , um sujeito S não possui as mesmas experiências de um ponto de vista interno, então S não utiliza o mesmo método nas circunstâncias  $\gamma$  e  $\phi$ .

Conforme sugere Nozick, ainda que não obviamente com estas mesmas palavras, duplos mentais compartilham os mesmos métodos, pois duplos mentais têm a mesma “vida mental” (pelo menos nos aspectos relevantes), ainda que possam estar situados em ambientes radicalmente diferentes. Na caracterização proposta por Nozick, fatores referentes ao ambiente ou circunstâncias do sujeito *não* fazem parte do critério de identidade de um método.

---

<sup>12</sup> Na passagem citada é sugerido implicitamente pelo advérbio *normalmente* que existem métodos sem qualquer correlato experiencial. Contudo, como Nozick nada diz sobre como deveriam ser entendidos os métodos nestas circunstâncias atípicas, optei por não tratar desta questão.

Há duas maneiras de compreender o par interno/externo; a maneira como Nozick individua os métodos é interna em um sentido, e externa em outro. Ela é interna, no sentido de que os métodos são determinados por eventos ou estados internos à mente do sujeito; ela é externa, no sentido de que a ocorrência de um determinado método não é necessariamente acessível, isto é, um sujeito pode utilizar um método sem saber que o está utilizando. Em outros termos, os métodos, segundo Nozick, não são acessíveis (a segunda dimensão de que tratamos).

No que diz respeito à generalidade, Nozick individua os métodos muito restritamente. Os métodos são individuados de acordo com a experiência (ou de acordo com outros fatos mentais relevantes) que os sujeitos possuem em determinadas circunstâncias específicas. Tais fatores *identificam* o método utilizado, o que sugere que os métodos são caracterizados segundo um grau muito restrito de generalidade.

### 2.3 A SEMÂNTICA DOS CONTRAFACTUAIS

De acordo com Nozick, não é necessária uma teoria precisa da semântica dos condicionais contrafactuais para os seus propósitos estritamente epistemológicos. Podemos nos satisfazer com um entendimento meramente intuitivo de como funcionam os contrafactuais e esperar que, talvez em um futuro próximo, venhamos a ganhar maior compreensão sobre o assunto. Assim, ele escreve o seguinte:

Uma vez que nós não especificamos uma teoria precisa dos subjuntivos, ou especificarmos precisamente como identificar um método e quando ele deve ser mantido fixo, existe uma certa flutuação [*leeway*] em nossa explicação. Pode ser que esta flutuação possibilite à nossa explicação lidar com estes exemplos e outros casos, ao utilizar as noções constituintes da teoria de forma imprecisa e intuitiva. Isto não é uma objeção, mas uma razão para pensar que as noções podem ser especificadas mais precisamente para lidar com os casos - uma condição da sua especificação é que ela trate corretamente dos casos- desde que a discussão destes casos não flutue inconsistentemente, primeiro levando à uma direção, depois à outra (NOZICK, 1981, p.38)

Apesar das alegações do autor de que um entendimento meramente intuitivo dos contrafactuais é suficiente para seus propósitos, é útil referirmo-nos brevemente a algumas teorias sobre a semântica dos contrafactuais. Duas razões sustentam esta necessidade. Em primeiro lugar, algumas das mais importantes e conhecidas teorias acerca dos contrafactuais, propostas

por Lewis (2005) e Stalnaker (1981), impactam diretamente em questões epistemológicas<sup>13</sup>. Em segundo lugar, o próprio Nozick sugere, ainda que com certa hesitação, que utilizará os conceitos de *mundos possíveis* e a semântica lewisiana para representar e interpretar os condicionais subjuntivos, sem querer comprometer-se com uma explicação ou análise da semântica dos condicionais subjuntivos em termos de mundos possíveis<sup>14</sup>. Assim, por exemplo, ele afirma:

Eu não pretendo endossar qualquer explicação dos subjuntivos em termos de mundos possíveis, nem estou comprometido com este tipo de explicação. Entretanto, eu às vezes a utilizarei, quando ilustrar alguns pontos de forma especialmente clara. (NOZICK, 1981, p.186)

O que a passagem acima citada afirma é que, embora Nozick não endosse uma análise dos contrafactuais em termos de mundos possíveis *a la* David Lewis, ele a utilizará quando for conveniente para *representar* o funcionamento dos condicionais subjuntivos, como uma espécie de guia ou critério de decisão. Assim, independentemente do fato de Nozick endossar ou não as semânticas dos subjuntivos de Stalnaker e Lewis, ele frequentemente as utiliza. Como as teorias de Lewis e Stalnaker são as mais influentes explicações dos contrafactuais, e referida tanto por Nozick como pela literatura subsequente sobre a teoria rastreamento, é imprescindível explicar em que elas consistem, e como elas se conectam diretamente com alguns tópicos epistemológicos relevantes.

### 2.3.1 As propostas de Lewis e Stalnaker

Os fundamentos de uma análise dos condicionais contrafactuais em termos de mundos possíveis devem-se ao trabalho de Robert Stalnaker. Ele propõe um modelo e uma explicação dos contrafactuais, que seria posteriormente desenvolvida por outros, notadamente David Lewis.

Stalnaker assinala que é pertinente distinguir entre dois problemas relativos a condicionais em geral: pragmático e lógico. O primeiro diz respeito aos critérios de decisão que empregamos para julgar acerca da verdade ou falsidade de condicionais. O segundo problema diz

---

<sup>13</sup> A interface semântica/epistemologia se tornará mais evidente quando tratarmos do problema das verdades necessárias para a teoria do rastreamento e da condição de aderência.

<sup>14</sup> Nozick não parece ser completamente coerente nessa questão. Apesar das afirmações de que não está comprometido com uma teoria específica dos condicionais subjuntivos, em algumas passagens Nozick parece fazer o contrário. Veremos isto quando tratarmos da questão das verdades necessárias e dos condicionais verdade-verdade.

respeito aos significado e propriedades formais dos condicionais, bem como as suas respectivas condições de verdade. Uma resposta ao problema pragmático não é uma resposta ao problema lógico. Contudo, uma resposta ao problema pragmático pode servir como guia relativamente satisfatório para a identificação das suas respectivas condições de verdade. As questões relativas a essa distinção foram assim resumidas por Stalnaker:

Eventualmente, eu irei defender uma hipótese sobre as condições de verdades para enunciados que possuem a forma condicional, porém eu começarei perguntando uma questão mais prática: como se avalia um enunciado condicional? Como se decide se acreditamos ou não acreditamos que ele é verdadeiro? Uma resposta a este tipo questão não será um conjunto de condições de verdade, mas servirá como um auxílio heurístico na busca por tal conjunto. (STALNAKER, 1981, p.42)

Conforme vimos, Stalnaker sugere uma resposta ao problema pragmático para condicionais, e logo em seguida sugere uma resposta ao problema lógico. No que diz respeito aos condicionais contrafactuais especificamente, o critério ou guia proposto por Stalnaker é o seguinte:

Primeiro, adicione o antecedente (hipoteticamente) ao seu estoque de crenças; segundo, faça quaisquer ajustes necessários para manter a consistência (sem modificar a crença hipotética no antecedente); finalmente, considere se, após a mudança, o consequente é verdadeiro. (STALNAKER, 1981, p.44)

Vou denominar o procedimento acima de *procedimento pragmático*. Segundo Stalnaker, o procedimento pragmático nos indica um caminho para especificar as condições de verdade de contrafactuais. Para fazer a transição, ele sugere introduzir o conceito de *mundo possível*, pois “um mundo possível é o análogo ontológico de um estoque de crenças hipotéticas” (Stalnaker, 1981, p.45). Poderíamos também afirmar, embora Stalnaker não faça desta forma, que o mundo possível minimamente diferente do mundo atual em que o antecedente do contrafactual é verdadeiro é o “análogo ontológico” do estoque total de crenças do sujeito, após estas serem devidamente ajustadas.

Stalnaker propõe que um condicional contrafactual é verdadeiro no mundo  $w$  somente se o consequente é verdadeiro em um conjunto relevante de mundos que diferem minimamente de  $w$  em que o antecedente também é verdadeiro ou não existe nenhum mundo em que é verdadeiro o antecedente<sup>15</sup>.

*Grosso modo*, podemos resumir a ideia central da explicação dos contrafactuais em termos da proximidade de mundos possíveis da seguinte forma:

---

<sup>15</sup> Para um contrafactual da forma  $p \square \rightarrow q$ , direi que  $p$  é o antecedente e  $q$  o consequente. Consideremos por exemplo o seguinte contrafactual: *se chovesse, a rua estaria molhada*. O antecedente e o consequente do contrafactual são, respectivamente, *está chovendo* e *a rua está molhada*.

$p \Box \rightarrow q$  em um mundo @ se e somente se todos os p-mundos mais próximos a @ são q-mundos.

Segundo Stalnaker, contrafactuais são "enunciados sobre mundos contrafactuais particulares" (Stalnaker, 1981, p.46). Esses mundos possíveis particulares, entretanto, devem estar submetidos a uma certa restrição: devem ser mundos tais que diferem minimamente do mundo atual, tanto quanto for compatível com a verdade do antecedente. De acordo com Stalnaker, existe *apenas um* mundo possível em que é verdadeiro o antecedente do contrafactual e que difere minimamente do mundo atual. Considerando isto, ele propõe as seguintes condições de verdade para os contrafactuais, para quaisquer proposições  $p$  e  $q$ :

**CS** -  $p \Box \rightarrow q$  em  $w$  se e somente se (i)  $q$  é verdadeiro no p-mundo que difere minimamente do mundo  $w$  ou (ii) não existe nenhum p-mundo.

O segundo disjuncto de CS foi introduzido justamente para lidar com as situações nas quais o antecedente do contrafactual é impossível. Quando é impossível que  $p$ , não existe nenhum p-mundo, *a fortiori* nenhum p-mundo que difere minimamente de  $w$ . Assim, de acordo com a proposta de Stalnaker, todos os condicionais contrafactuais com antecedentes impossíveis, os chamados *contrapossíveis*, são verdadeiros, o que expressamos por meio do seguinte princípio, para quaisquer proposições  $p$  e  $q$ :

**TC** -  $\neg \Diamond p \rightarrow (p \Box \rightarrow q)$ .

Stalnaker também sustenta três princípios, dos quais derivaremos consequências importantes, conforme veremos mais tarde. São os seguintes, para quaisquer mundos  $w$  e  $w'$ :

- **Centramento** - nenhum mundo é tão próximo a  $w$  como o próprio  $w$ ;
- **Limite** - para qualquer mundo possível  $w$  e qualquer sentença logicamente consistente  $p$ , existe algum mundo  $w'$ , tal que  $w'$  é um p-mundo que difere minimamente de  $w$ ;
- **Unicidade** - para qualquer mundo possível  $w$  e qualquer sentença logicamente consistente  $p$ , existe no máximo um p-mundo que difere minimamente de  $w$ .

Diremos que um mundo  $w$  difere menos de um mundo  $w'$  quanto mais próximo  $w$  é de  $w'$ . O conjunto dos mundos possíveis pode ser ordenado em relação ao mundo  $w$  (ou a qualquer outro mundo), tal que alguns mundos estão mais distantes ou mais próximos ao mundo  $w'$ . Proximidade é uma relação de similaridade entre mundos: quanto mais similares são dois

mundos sob determinado aspecto, mais próximos eles são. O mundo que difere minimamente do mundo  $w$  é justamente o mundo mais similar (i.e., mais próximo). Stalnaker considera que para determinar a verdade de um condicional contrafactual devemos considerar um único mundo possível (se houver): o mundo mais próximo.

Contudo, existem autores que julgam que o conjunto de mundos possíveis relevantes deve ser estendido, e ir além do mundo mais próximo. Em sua influente obra, Lewis (2013) propôs uma teoria semântica dos contrafactuais em termos de mundos possíveis que, embora inspirada na teoria de Stalnaker, é diferente dela em aspectos que serão importantes para nós. Mais especificamente, Lewis rejeita suas teses importantes, os princípios da unicidade e do limite.

Primeiramente, o princípio do limite. Consideremos a seguinte sentença: *se eu tivesse menos de 1,70m, então eu não teria sido atingido pelo objeto*. É esta sentença verdadeira no mundo  $w$ ? Digamos que no mundo  $w$  eu possua 1,80m. Qual será o mundo mais próximo a  $w$  que eu possuo menos de 1,70m? Talvez o mundo  $w$  em que eu possuo 1,69? Aparentemente não, pois existe um outro mundo, digamos o mundo  $w'$ , em que eu possuo 1,699m. Será esse último então o mundo mais próximo? Também não, pois existe ainda outro mundo mais próximo ao mundo  $w$ , digamos o mundo  $w''$ , em que eu possuo 1,6999m. Pelas mesmas razões, também não poderá ser  $w''$  o mundo mais próximo, e assim sucessivamente. Podemos recuar indefinidamente, sem jamais encontrar um ponto de parada. Não parece existir, para essa situação, um único mundo que seja simplesmente "o mundo que difere minimamente do mundo atual".

Passemos a examinar o princípio de unicidade. Consideremos agora outra sentença: *se Bizet e Verdi fossem compatriotas, então eles teriam se conhecido*. Qual é o mundo que difere menos do mundo atual, tal que o antecedente do condicional é verdadeiro? Deveríamos considerar o mundo em que ambos são franceses, ou o mundo em que ambos são italianos? Não parece haver qualquer critério para decidir entre as duas possibilidades, nem qualquer motivo para afirmar que um deles é mais (ou menos) similar ao mundo atual.

Para lidar com problemas como esses, e não se comprometer com os princípios de unicidade e do limite, Lewis propõe uma alteração na semântica de Stalnaker. Segundo a semântica proposta por Lewis, para quaisquer  $p$  e  $q$ :

CL -  $p \Box \rightarrow q$  é verdadeiro em um mundo  $w$  se e somente se (i) existe algum  $(p \rightarrow q)$ -mundo mais próximo a  $w$  que qualquer  $(p \rightarrow \neg q)$ -mundo; ou (ii) não existe nenhum  $p$ -mundo.

Note-se que a condição (i) acima é compatível com a existência de "empate" de graus de similaridade. Dois ou mais  $p$ -mundos  $w$  e  $w'$  podem ser os mundos mais próximos ao mundo  $w$ , e ainda assim a condição (i) ser satisfeita, desde que  $w$  e  $w'$  sejam mais próximo que qualquer outro  $(p \rightarrow \neg q)$ -mundo. Tampouco o conjunto dos mundos mais próximos a  $w$  precisa ser finito. Pode não existir um mundo que seja o mundo mais próximo que qualquer outro mundo (porque sempre existe um mundo mais próximo ainda), e ainda assim a condição (i) ser satisfeita.

Lewis faz algumas outras observações, que devemos aqui salientar. Segundo o autor, a negação de  $p \Box \rightarrow q$ , i.e.  $\neg (p \Box \rightarrow q)$ , não é equivalente a  $p \Box \rightarrow \neg q$ . De acordo com a sua proposta, um condicional da forma  $p \Box \rightarrow q$  é verdadeiro quando não poderia ser o caso que:  $p$  é verdadeiro, e ainda assim  $q$  não é o caso. De forma semelhante, um condicional da forma  $p \Box \rightarrow q$  é falso quando poderia acontecer de  $p$  ser verdadeira e  $q$  ser falsa.

## 2.4 O PROBLEMA DE GETTIER

O poder explicativo é a principal virtude da teoria do rastreamento, e a teoria fornece explicações com grande elegância e simplicidade para difíceis problemas filosóficos. Entre as alegadas soluções geradas pela teoria nozickiana, destaca-se a resposta ao problema de Gettier.

Em um curto artigo com não mais de 3 páginas, Gettier (1963) produziu contraexemplos à *análise tradicional do conhecimento*. Desde então, foram muitas as tentativas, muitas delas não tão bem-sucedidas, de identificar o que há de equivocado com esta análise e como ela deve ser corrigida, ou eventualmente substituída por uma proposta radicalmente diferente. Como é bem conhecido, este segue sendo um dos tópicos mais controversos e complicados da agenda epistemológica contemporânea, mesmo passados mais de cinquenta anos desde o artigo de Gettier.

Segundo a análise tradicional do conhecimento, para qualquer proposição  $p$ , um sujeito  $S$  sabe que  $p$  se e somente:

- **Verdade** - é verdade que  $p$ ;
- **Crença** -  $S$  acredita que  $p$ ;
- **Justificação** - a crença de  $S$  de que  $p$  está justificada.

Os contraexemplos de Gettier dizem respeito a determinadas situações nas quais alegadamente as cláusulas da crença, verdade e justificação são satisfeitas, e, no entanto, o sujeito epistêmico não possui conhecimento. Por conseguinte, crença, verdade e justificação não podem ser suficientes para a posse de conhecimento.

A ausência de conhecimento em casos Gettier, segundo a maioria das interpretações, é explicada pela intervenção de algum tipo de sorte ou coincidência. Ainda que o sujeito  $S$  creia justificada e verdadeiramente que  $P$ , é uma mera coincidência que ele tenha, por assim dizer, "atingido" a verdade.

Gettier afirma que os contraexemplos apresentados por ele dependem de duas suposições, ambas muito plausíveis. A primeira toma a seguinte forma:

- **Crença justificada falsa (CJF)** - é possível  $S$  estar justificado ao crer em uma proposição falsa.

O segundo princípio, de cuja versão relativa ao conhecimento trataremos extensamente mais tarde, diz o seguinte:

- **Fecho dedutivo da justificação (FDJ)** - para quaisquer  $p$  e  $q$ : se  $S$  está justificado em crer que  $p$ , e  $S$  deduz validamente que  $q$  a partir de  $p$ , vindo a aceitar que  $q$  em razão de  $p$ , então  $S$  está justificado em crer que  $q$ .

Feitas estas observações, examinemos um dos contraexemplos sugerido por Gettier, que denominaremos *caso das dez moedas*. Smith e Jones estão competindo por uma vaga de emprego, e estão juntos, na sala de espera, aguardando a chamada para a entrevista. Smith ouve da secretária, logo antes de ser chamado, que Jones obterá a vaga disputada, e que a entrevista é uma mera formalidade. Assim, ele forma a crença justificada de que: (1) Jones conquistará a vaga de emprego. Agora vamos supor que Smith viu Jones, por algum motivo qualquer, retirar algumas moedas do seu bolso, e que Smith contou exatamente dez moedas. Com

base nisto, ele forma a crença, também justificada, de que: (2) Jones possui dez moedas em seu bolso. Com base em (1) e (2) ele infere validamente que:

(3) a pessoa selecionada para a vaga de emprego possui dez moedas em seu bolso.

Agora, vamos supor que é verdade que:

(4) Smith tem dez moedas no bolso e foi selecionado para a vaga de emprego.

Assim, ao contrário do que Smith pensa, é ele mesmo - não seu concorrente Jones - que foi selecionado para a vaga e, ao mesmo tempo, Smith possui dez moedas em seu bolso. Consequentemente, é verdadeira a crença de que (3).

Intuitivamente, diríamos que a crença em (3) é justificada. Smith parece possuir boas razões, no caso as crenças também justificadas em (1) e (2), para a sua crença em (3). Em razão de FDJ, e da justificação de (1) e (2), a crença de que (3) também é justificada. Temos assim que Jones acredita justificadamente que (3).

A proposição (3), além de justificada, também é verdadeira. A crença de (3) é verdadeira, uma vez que (4) é verdadeira. Consequentemente, Smith crê verdadeira e justificadamente que (3). Contudo, não diríamos que Smith sabe que (3) - parece ter sido uma mera coincidência que ele tenha uma crença verdadeira nestas circunstâncias<sup>16</sup>.

Em seu artigo, Gettier também apresenta um segundo contraexemplo à análise tradicional do conhecimento, que denominaremos *caso do carro Ford*. Digamos que Smith possui excelente evidência, suficiente para a justificação, para a crença de que:

(5) Jones possui um carro Ford.

A evidência para (5) pode ser o testemunho de Jones, o fato de que ele foi visto dirigindo o carro, de que outros também afirmaram que Jones possui um Ford, e assim por diante. A crença de Smith de que (5) é justificada. De (5) ele infere validamente que:

---

<sup>16</sup> Conforme apontado corretamente por Tiegue Vieira Rodrigues, a proposição *a pessoa selecionada para a vaga de emprego possui dez moedas em seu bolso* é ambígua, uma vez que a descrição definida *a pessoa selecionada para a vaga de emprego* é também ambígua, e pode ser usada tanto atributivamente como referencialmente. No entanto, podemos modificar o caso original proposto por Gettier, de modo a evitar a ambiguidade. Basta substituir (3) por *alguém simultaneamente foi selecionado para a vaga de emprego e possui dez moedas no bolso*, sendo que esta última poderia ser parafraseada como *existe um x, tal que x vai conquistar a vaga de emprego e x possui dez moedas no bolso*. Agradeço ao Tiegue por ter feito a observação sobre o uso referencial e atributivo da descrição definida.

(6) Jones possui um carro Ford ou Brown está em Barcelona.

Digamos que Smith não possua nenhuma outra evidência para (6) além da própria (5).

Vamos supor agora que é falso que (5), isto é, que Jones não possui um Ford. Jones fingiu, aliás muito competidamente, ser dono de um carro Ford a fim de parecer ter mais riquezas do que efetivamente possui. Todavia, por uma estranha coincidência, Brown realmente está em Barcelona, ainda que Smith não tenha quaisquer informações a respeito. Portanto, é verdade que Jones possuiu um carro Ford ou Brown está em Barcelona.

O caso do carro Ford segue exatamente a mesma estrutura de o caso das dez moedas. A crença de Smith de que (5) é justificada, ainda que falsa. Smith infere validamente (6) de (5). Em razão de FDJ, e da justificação de (5), a crença de (6) é justificada. Contudo, a crença em (6) não é um caso de conhecimento. Smith teve sorte de formar uma crença verdadeira, e ele parece ter formado uma crença verdadeira em razão de uma mera coincidência, qual seja, a coincidência de que Jones fingiu possuir um carro Ford e Brown está em Barcelona.

Como reconhece quase unanimemente toda a literatura epistemológica pós-Gettier, o núcleo dos contraexemplos é a intervenção de algum tipo de sorte ou coincidência. Em ambos os casos, o sujeito cognoscente simplesmente *teve sorte*. A grande lição que os epistemólogos tiraram destes casos é que a cláusula da justificação deve ser complementada, ou eventualmente substituída, por alguma outra condição que exclua a sorte epistêmica. O tão famoso problema de Gettier consiste justamente na pergunta pela natureza dessa cláusula.

Ambos os princípios mencionados por Gettier, FDJ e CJF, são necessários a fim de que os contraexemplos possam funcionar como desejado. Em ambas as situações, são realizadas inferências a partir de crenças justificadas e falsas. No primeiro, é falsa a crença de que Jones conquistará a vaga de emprego; no segundo, é falsa a crença de que Jones é o dono de um carro de Ford. Todavia, em cada caso a justificação da crença-alvo (as crenças em (3) e (6)) é inferencial, e depende da justificação das premissas de cada raciocínio. Uma vez aceito que as crenças nas premissas de cada raciocínio são justificadas, e que é verdadeiro o princípio de fecho dedutivo da justificação, segue-se que as crenças nas respectivas conclusões também são justificadas.

Há duas características importantes nos contraexemplos apresentados originalmente por Gettier, que não pertencem a todos os contraexemplos que foram denominados *casos Gettier* na literatura sobre o assunto. Em primeiro lugar, as crenças gettierizadas dos exemplos

que tratamos são inferenciais<sup>17</sup>. Em segundo lugar, as crenças gettierizadas são deduzidas de crenças justificadas e falsas.

Conforme a literatura sobre o assunto progrediu, foram descobertos novos contra-exemplos, que não eram baseados em crenças falsas, nem eram inferenciais. Consequentemente, esses exemplos não dependem de FDJ nem de CJF. Um desses é apresentado por Goldman (1976), e aqui o denominaremos *o caso dos celeiros*. Suponhamos que Barney está viajando de carro por uma estrada rural. Ao longo da viagem, ele observa diversas estruturas que, tanto quanto pode perceber, parecem típicos celeiros. Ele forma a crença perceptual correspondente:

(7) aquele objeto é um celeiro.

Vamos supor que a crença de que (7) seja verdadeira, e que Barney observou um celeiro naquele local. Além do mais, as suas faculdades sensoriais funcionam perfeitamente, e não há qualquer razão para Barney suspeitar de que não existe um celeiro no local avistado. Assim, Barney possui uma crença verdadeira e justificada em (7), satisfazendo a análise tradicional do conhecimento.

Todavia, existe uma complicação. A estrada rural, pela qual Barney se desloca, está repleta de determinadas estruturas que, embora parecessem celeiros mesmo para o mais competente observador, são apenas réplicas ou fachadas (digamos que os falsos celeiros não contêm nada na parte de trás, não visível a partir da estrada), que visavam enganar os viajantes. Entre as tantas fachadas e réplicas muito bem construídas, esconde-se um único celeiro genuíno, precisamente o que foi avistado por ele. Barney simplesmente *teve sorte* de ter avistado, em meio a tantas réplicas enganadoras que o conduziriam ao erro, o único celeiro genuíno em toda a região.

O caso dos celeiros é significativamente diferente dos exemplos originais de Gettier. Primeiramente, no exemplo em questão a crença-alvo é perceptual e direta, sem que haja o envolvimento de algum tipo de raciocínio. Não é possível, portanto, alegar que a explicação do problema reside na falsidade das premissas de uma *inferência*.

Além do mais, ao contrário dos outros contraexemplos, os fatos que tornam a crença verdadeira e a sua causa imediata são idênticos: a presença, no ambiente, de um celeiro genuíno.

---

<sup>17</sup> Dizemos que uma crença em uma proposição alvo  $p$  está gettierizadas quando a crença é verdadeira por acidente e justificada, portanto, quando a crença é verdadeira, justificada e não é um caso de conhecimento.

ino. Nos casos Gettier originais, existe um divórcio entre o que torna a crença verdadeira e as causas da crença, o que está ausente no caso dos celeiros. Embora o sujeito cognoscente deste último esteja conectado aos fatos e à verdade da crença (afinal, a crença de que aquele objeto é um celeiro é causada pelo próprio celeiro), alguma medida de sorte relevante intervém na geração da crença verdadeira: ele teve sorte de deparar-se com o único celeiro genuíno em um ambiente dominado por réplicas indistinguíveis de celeiros.

Como Nozick explica os casos Gettier? Segundo o diagnóstico proposto por ele, a raiz do problema está na violação da condição de sensibilidade. Vejamos como isto acontece em cada uma das situações descritas.

No caso das dez moedas, a crença de Smith não é sensível à verdade da proposição alvo. Se fosse falso que o escolhido para a vaga de emprego possui dez moedas no bolso, ainda assim Smith acreditaria que o escolhido possui dez moedas no bolso. Ele presumidamente acreditaria ainda que fosse falso, porque se o escolhido não tivesse dez moedas no bolso, ainda assim Smith teria ouvido que Jones foi o escolhido, teria visto Jones contar as dez moedas e realizado precisamente a mesma inferência. Situação idêntica se repete no caso do carro Ford: ainda que a proposição alvo fosse falsa, o sujeito teria acreditado que ela é verdadeira, pois presumidamente ele ainda teria acreditado que Nogot possui um carro Ford, e feito a inferência correspondente.

O princípio de sensibilidade não é satisfeito no caso dos celeiros: é falso que se não existisse um celeiro no local avistado por ele, então Barney não acreditaria que aquele objeto é um celeiro, pois claramente Barney poderia ter acreditado que aquele objeto é um celeiro, se fosse falso que aquele objeto é um celeiro (por exemplo, se tivesse diante de uma das muitas réplicas).

Assim como nos demais exemplos, o sujeito cognoscente obteve uma crença verdadeira por simples sorte ou coincidência. Ele teve sorte de estar diante do único celeiro genuíno na região, quando poderia muito facilmente estar diante de uma das muitas réplicas. A teoria do rastreamento parece feita sob medida para lidar com o problema de Gettier e excluir a sorte epistêmica.

## 2.5 O PROBLEMA DA LOTERIA

A teoria do rastreamento também se propõe a explicar por que julgamos, em certas situações, que denominaremos de *situações lotéricas*, que o sujeito não possui conhecimento. Ao mesmo tempo, a teoria parece também explicar por que o sujeito cognoscente possui conhecimento em circunstâncias similares, ainda que possua uma evidência probabilisticamente mais fraca.

Para tratar do caso, consideremos duas situações distintas: loteria<sub>1</sub> e loteria<sub>2</sub>. Vejamos o caso loteria<sub>1</sub>. S comprou um bilhete *i* para uma loteria justa, que contém muitos bilhetes, todos com as mesmas chances de serem sorteados. O sujeito S, em razão da elevadíssima probabilidade de que seu bilhete *não* seja sorteado, forma a crença de que PL, quando *PL = o bilhete i não foi sorteado*. Por hipótese, PL é verdadeira, e o bilhete *i* não foi sorteado.

Comparemos agora o caso acima com o seguinte, o caso loteria<sub>2</sub>. S comprou um bilhete *i* para uma loteria justa, que contém muitos bilhetes, todos com as mesmas chances de serem sorteados. Após a realização do sorteio, é divulgado o resultado: o bilhete de S não foi sorteado. Após ler no jornal local, que em linhas gerais é confiável e não costuma errar a respeito desse tipo de assunto, S forma a crença de que seu bilhete não foi sorteado.

Em loteria<sub>1</sub>, diríamos que S não sabe que o seu bilhete não foi sorteado antes de verificar por algum outro meio (ninguém sabe que um bilhete de loteria não foi sorteado, simplesmente com base nas probabilidades envolvidas). Em loteria<sub>2</sub>, diríamos que o sujeito possui conhecimento. Parece se tratar de um caso normal de conhecimento testemunhal.

Quando comparamos as duas situações, notamos uma característica impressionante: a probabilidade de o bilhete de S não ser sorteado é muito maior do que a probabilidade do jornal não ter cometido algum erro ao relatar o resultado do sorteio. Contudo, tal disparidade não parece se refletir em nossos juízos sobre conhecimento<sup>18</sup>.

---

<sup>18</sup> Este ponto acerca dos casos loteria<sub>1</sub> e loteria<sub>2</sub>, de que as respectivas probabilidades envolvidas não parecem se refletir em nossos juízos sobre presença de conhecimento, foi apontada por DeRose(1996). Agradeço ao professor Rodrigo Martins Borges por ter sugerido a referência.

O que explica esta *assimetria* em nossas atribuições de conhecimento e ignorância em loteria<sub>1</sub> e loteria<sub>2</sub>? Por que dizemos que S não sabe no primeiro caso, mas sabe no segundo, apesar da probabilidade de erro da parte de S ser muito menor no primeiro do que no segundo? O diagnóstico, mais uma vez, recorre ao princípio de sensibilidade. S não sabe que o seu bilhete não foi sorteado no primeiro caso porque se a sua crença fosse falsa, isto é, se o bilhete tivesse sido sorteado, ele ainda assim acreditaria falsamente que o bilhete não foi sorteado com base no mesmo método. Presumidamente, ele continuaria acreditando falsamente que o bilhete não foi sorteado, porque a sua base evidencial, que consiste exclusivamente de considerações probabilísticas, seria a mesma em ambas as circunstâncias.

O mesmo princípio da sensibilidade, que, segundo Nozick, explica por que S não tem conhecimento no primeiro caso, explica também por que ele possui conhecimento no segundo, supondo que a condição de aderência também é satisfeita. Se fosse falso que S não ganhou na loteria, isto é, se S tivesse ganhado na loteria, então S não acreditaria que não ganhou na loteria, pois se S tivesse ganhado na loteria, então o jornal teria relatado tal fato, o que levaria S a não formar a crença de que não ganhou na loteria. O princípio de sensibilidade parece adequado para explicar a diferença entre as duas situações, assim como produz veredictos corretos de um ponto de vista intuitivo.

## 2.6 A RESPOSTA DE NOZICK AO CETICISMO

Nós pensamos saber várias coisas. Eu penso, por exemplo, saber que estou agora escrevendo em meu computador. Contudo, eu *realmente* sei que estou aqui agora escrevendo em meu computador? Eu poderia estar agora sonhando em minha cama que estou escrevendo em meu computador com as mesmas impressões e experiências perceptuais que tenho agora (com precisamente estas mesmas imagens, lembranças, etc.). Eu *sei* que essa hipótese não descreve precisamente a minha situação no momento atual?

Pensemos um pouco mais sobre o assunto. Se eu estivesse em minha cama sonhando que estou em frente ao meu computador, então tudo me pareceria exatamente como me parece agora. Se eu estivesse sonhando que estou escrevendo em meu computador, teria exatamente a mesma sequência de imagens e sensações que tenho agora. Assim, argumenta o cético, eu não tenho como afastar ou eliminar a hipótese cética de que estou apenas sonhando que estou

em frente ao meu computador, e desta forma vir a saber que a hipótese cética não é verdadeira (afinal, *como* eu poderia saber algo deste tipo?). Como não sei que não estou sonhando que estou escrevendo em meu computador (quando não existe nenhum computador), continua o argumento cético, então não sei que estou aqui agora escrevendo em meu computador. O argumento pode ser repetido para qualquer presumido objeto de conhecimento perceptual. O resultado é profundamente desconcertante e paradoxal, e, no entanto, parece difícil, até mesmo impossível, escapar dele.

Tipicamente, o cético procura impugnar de maneira sistemática nossas alegações de conhecimento em um determinado campo específico (p.ex., sobre o mundo exterior) mediante a apresentação de hipótese, a denominada *hipótese cética*, incompatível com a proposição supostamente objeto de conhecimento. Como o sujeito cognoscente não é capaz de eliminar ou afastar a hipótese levantada, argumenta o cético, então ele não sabe o que originalmente julgava saber. Dada a aparente facilidade de levantar, para qualquer alegação de conhecimento, uma hipótese incompatível, mas que não se consegue afastar, o ceticismo tem se mostrado, ao longo da história da filosofia, de difícil refutação, e um problema perene, para o qual voltam-se com regularidade os filósofos.

É costume reconstruir o ataque cético por meio de um argumento simples, composto por três premissas. O argumento, denominado de *argumento da ignorância* (AI) por DeRose (1995), tem a seguinte forma, onde  $p$  é uma proposição ordinária (p.ex., *eu possuo mãos*), e  $hc$  uma hipótese cética qualquer incompatível com  $p$ :

1. se eu sei que  $p$ , então eu sei que  $\neg hc$ ;
2. eu não sei que  $\neg hc$ ;
3. logo, eu não sei que  $p$ .

O argumento é plenamente generalizável, aplicando-se a qualquer proposição ordinária  $p$ . O argumento também é válido, tendo a simples forma de um *modus tollens*. Se quisermos evitar a conclusão cética, e aceitarmos que o argumento é válido, como parece sê-lo, só nos resta recusar alguma das premissas do argumento. Qual das duas deve ser rejeitada? Estamos diante de argumento válido, cujas premissas são extremamente plausíveis e, no entanto, a conclusão é desconcertante.

O problema suscitado por AI é particularmente interessante quando contemplado sob outra perspectiva. Podemos compreender o ceticismo como um paradoxo, formado por um

conjunto de três sentenças, individualmente plausíveis e intuitivas, porém conjuntamente inconsistentes. O problema sob esta perspectiva consiste em identificar como podemos escapar do referido paradoxo. As três sentenças são as seguintes, para quaisquer proposições ordinárias  $p$  e quaisquer hipóteses céticas  $hc$ :

1. eu sei que  $p$ ;
2. se eu sei que  $p$ , então eu sei que  $\neg hc$ <sup>19</sup>;
3. eu não sei que  $\neg hc$ .

Denominarei de *problema do ceticismo* a seguinte questão: qual das três sentenças acima devemos rejeitar? Naturalmente, a resposta deve ser motivada e não arbitrária, além de, obviamente, coerente e intelectualmente satisfatória. Existem três reações possíveis ao argumento<sup>20</sup>:

1. rejeitar que eu sei que  $p$ ;
2. rejeitar que se eu sei  $p$ , então eu sei que  $\neg hc$ ;
3. rejeitar que eu não sei que  $\neg hc$ .

Conforme veremos a seguir, Nozick optou pela segunda alternativa.

### 2.6.1 A resposta de Nozick

Nozick propõe a rejeição de que se  $S$  sabe que  $p$ , então  $S$  sabe que  $\neg hc$ . A proposição é falsa, segundo Nozick, visto ser possível que o antecedente do condicional seja verdadeiro, e o seu conseqüente falso. Trata-se de uma consequência direta da teoria do rastreamento, e permite, ao menos em princípio, bloquear o argumento cético<sup>21</sup>.

Nozick tentou, de maneira bastante engenhosa, mostrar que é possível compatibilizar as alegações ordinárias de conhecimento - por exemplo, de que estou aqui agora escrevendo em meu computador - com a ignorância a respeito da verdade das hipóteses céticas. Por meio de tal estratégia, ele argumenta que, apesar da nossa incapacidade de afastar as hipóteses in-

<sup>19</sup> Esta é uma apresentação simplificada do paradoxo. Conforme veremos, esta segunda sentença deve ser modificada, pois obviamente não sabemos todas as consequências lógicas de tudo aquilo em que acreditamos.

<sup>20</sup> Não trato nesse momento de algumas propostas, tais como a teoria contextualista, que sugere que as três sentenças podem ser verdadeiras, desde que asseridas em diferentes contextos conversacionais.

<sup>21</sup> Eu suponho que, ao menos implicitamente, é atribuído algum tipo de necessidade da proposição.

compatíveis levantadas pelo cético, ainda podemos saber muito daquilo que normalmente julgamos saber.

Suponhamos que S se encontra em um mundo mais ou menos como o senso comum o representa (o que deve ser, ao menos, possível),  $p$  seja uma proposição ordinária e que  $hc$  seja uma hipótese cética. Neste caso, a crença de que  $p$  é sensível à verdade. Seja  $p$ , por exemplo, *eu possuo mãos*. Nas circunstâncias ou situações muito semelhantes à atual em que S não possui mãos (o que seria o caso se, por exemplo, perdesse as mãos em um acidente, tivesse nascido sem mãos, e assim por diante) S *não* acreditaria que possui mãos. A crença de S de que  $p$  varia conforme a verdade ou falsidade de que  $p$ : se fosse falso que  $p$ , então S não acreditaria que  $p$ ; se fosse verdadeiro que  $p$ , então S acreditaria que  $p$ .

No que diz respeito a crença em  $\neg hc$ , a sua situação de S é nitidamente diferente. Seja  $hc$ , por exemplo, *S está sendo iludido por um Demônio Maligno, que faz com que ele tenha experiências de mãos, quando não existe nenhuma mão*. É falso que se fosse falso que  $\neg hc$  (i.e., verdadeiro que  $hc$ ), então S não acreditaria que  $\neg hc$ . Com efeito, ocorre precisamente o contrário: se fosse verdadeiro que  $hc$ , então S acreditaria que  $\neg hc$ . Se um Demônio Maligno produzisse em S precisamente as mesmas experiências e sensações que ele possui no mundo atual, então ele acreditaria, tal como acredita no mundo atual, que não há um Demônio Maligno (ou qualquer outra entidade do tipo) o enganando a respeito da existência das suas mãos. Portanto, S não sabe que  $\neg hc$ , de acordo com a teoria de Nozick, uma vez que a crença em  $\neg hc$  não satisfaz uma das condições necessárias do conhecimento, qual seja, a condição da sensibilidade. Consequentemente, é *possível* saber que  $p$  e não saber  $\neg hc$ . Assim, é falso que se eu sei que  $p$  então eu sei que  $\neg hc$ , uma vez que é possível que o antecedente do condicional seja verdadeiro e o seu conseqüente falso.

Nozick não foi o primeiro a sugerir que a recusa da segunda premissa do paradoxo é a chave para uma resposta ao ceticismo. A este respeito, ele foi precedido por Dretske (1970), e é da recusa desta segunda premissa por Dretske que trataremos a seguir.

### 2.6.2 Princípios de fecho do conhecimento

Para aprofundar a discussão da resposta de Nozick ao ceticismo, e explicar alguns princípios subjacentes envolvidos, é importante esclarecer alguns conceitos. Um operador

sentencial é uma expressão que, quando anexada a uma sentença qualquer, produz outra sentença. Assim, por exemplo, são operadores sentenciais as seguintes expressões: *é verdade que*, *é fato que*, *S acredita que*, *S sabe que*, e assim por diante.

*Fechamento* é originalmente um termo técnico da matemática. *Grosso modo*, dizemos que os objetos de um conjunto qualquer  $A$  são fechados por uma relação  $R$  qualquer se e somente se, para quaisquer objetos  $x$  e  $y$ :

$$(x \in A \ \& \ xRy) \rightarrow y \in A.$$

As sentenças às quais se aplicam verdadeiramente um operador sentencial também formam um conjunto. Isto permite combinar os dois conceitos que vimos anteriormente: operador e fechamento. Assim, dizemos que um operador sentencial  $O$  é fechado por uma relação  $R$  se e somente se<sup>22</sup>:

$$(Op \ \& \ pRq) \rightarrow Oq.$$

Eis alguns exemplos:

- O operador sentencial *é verdade que* é fechado pela relação de implicação lógica. Se é verdade que  $p$ , e  $p$  implica logicamente  $q$ , então é verdade que  $q$ .
- O operador sentencial *S acredita que* não é fechado pela relação de implicação lógica. Não é verdade que se  $S$  acredita que  $p$ , e  $p$  implica logicamente  $q$ , então  $S$  acredita que  $q$ .
- O operador *é fato que* é fechado pela relação de implicação lógica. Se é fato que  $p$ , e  $p$  implica logicamente  $q$ , então é fato que  $q$ .

Pois bem, o que diremos do operador *S sabe que*? É o operador *S sabe que* (ou, mais sucintamente, o conhecimento) fechado pela relação de implicação lógica? Uma vez que dispomos da ideia de um operador sentencial fechado por uma relação, podemos formular diferentes princípios de fecho do conhecimento, que dizem ser o conhecimento fechado (ou aberto) mediante uma certa relação.

A tese de que o conhecimento é fechado pela relação de implicação lógica é o seguinte, para quaisquer proposições  $p$  e  $q$ :

---

<sup>22</sup>  $Op$  é sentença produzida pela operação de  $O$  na sentença  $p$ . A expressão  $pRq$  se lê como *p está na relação R com q*.

**Fechamento epistêmico pela implicação lógica (FI)** - Necessariamente: se S sabe que  $p$ , e  $p$  implica  $q$ , então S sabe que  $q$ .

FI é claramente falso. Nós não somos logicamente oniscientes.  $P$  pode implicar  $q$ , e ainda assim não saber que  $q$ , porque S não sabe que  $p$  implica  $q$ .

Consideremos, então, o seguinte princípio, que visa suprir a lacuna das relações desconhecidas de implicação lógica, para quaisquer proposições  $p$  e  $q$ :

**Fechamento epistêmico pela implicação conhecida (FIC)** - Necessariamente: se S sabe que  $p$ , S sabe que  $p$  implica  $q$ , então S sabe que  $q$ .

Este último é muito mais plausível que FI, e talvez seja FIC que verdadeiramente opera na argumentação cética. É este último princípio que é atacado diretamente tanto por Dretske como por Nozick. Contudo, mesmo este pode ser excessivamente forte, e parece invalidado por detalhes menores, alheios ao espírito da discussão. Afinal, um sujeito S pode saber que  $p$ , e saber que  $p$  implica  $q$ , mas, por assim dizer, não "juntar os pontos", e não crer que  $q$  por causa de  $p$ .

A alteração mínima que acomoda o problema menor que afeta FIC - a ausência da crença em  $q$  por causa da crença de que  $p$  - é resolvida por meio do princípio do fechamento do conhecimento pela inferência válida, para quaisquer proposições  $p$  e  $q$ :

**Fechamento epistêmico pela inferência válida (FIV)** - Necessariamente: se S sabe que  $p$ , e S deduz  $q$  validamente de  $p$ , então S sabe que  $q$ .

A lacuna presente em FIC é sanada por FIV, uma vez que se um sujeito S infere validamente  $q$  de  $p$ , então

1. S acredita que  $q$ ;
2. S acredita que  $q$  por causa de  $p$ ;
3.  $p$  implica logicamente  $q$ .

Conforme vemos, FIV captura um dos *insights* que parece motivar a discussão em torno dos princípios de fecho do conhecimento, qual seja, de que o conhecimento é infalivelmente transmitido pela inferência válida.

FIV autoriza a seguinte inferência, onde  $p$  é uma proposição de senso comum qualquer, e  $hc$  é uma hipótese cética incompatível com  $p$ :

1. S sabe que p;
2. S infere validamente  $\neg hc$  de p;
3. Logo, S sabe que  $\neg hc$ .

FIV também implica que uma vez que você saiba que p, e que p implique q, você sempre está em posição de saber que q, bastando, para tanto, apenas realizar a inferência válida disponível. Sempre que você realiza uma inferência válida a partir de premissas, que são elas mesmas objetos de conhecimento, você sabe que é verdadeira a conclusão do argumento. A rejeição de FIV bloqueia a argumentação cética naquele que parece ser um dos seus pontos mais fortes: a inferência de que S não sabe que p a partir de S não sabe que  $\neg hc$ .

É comum que os autores que tratam de princípios de fecho do conhecimento não distingam entre diferentes espécies ou variedades do princípio que expusemos: FI, FIC e FIV. Em todo o caso, o que parece estar em jogo na disputa é melhor capturado por FIV, uma vez que os princípios FI e FIC são falsificados por razões aparentemente triviais. Consequentemente, quando trato de princípios de fecho epistêmico, estou normalmente me referindo a FIV.

Aqueles que rejeitam FIV não afirmam que nunca (ou mesmo que apenas raramente) o conhecimento é transmitido pela inferência válida. Eles não condenam a inferência válida, o que faria a sua tese absurda. Eles apenas rejeitam a universalidade da inferência válida como modo de estender ou ampliar o conhecimento. Conforme veremos no capítulo 2.8, o próprio Nozick possui uma teoria de como adquirimos conhecimento através da inferência. Os autores que rejeitam FIV normalmente afirmam uma tese muito mais fraca, e mesmo assim ainda surpreendente: é possível S saber que p, p implicar logicamente q e S *ainda assim* não poder saber que q por meio de uma inferência válida de q a partir de p<sup>23</sup>.

Nozick não recuou diante da consequência de que devemos rejeitar o princípio de fecho do conhecimento. Ele tomou este resultado como uma qualidade da sua teoria, talvez a única maneira de explicar como podemos preservar nossas alegações ordinárias de conhecimento. A argumentação cética, segundo esta perspectiva, depende de um princípio falso tacitamente assumido, ainda que extremamente plausível. Contudo, é possível mostrar como e por que o referido princípio é falso, o que decorre diretamente da teoria do rastreamento.

---

<sup>23</sup> Dretske (1970) ressaltou o mesmo ponto em uma linguagem ligeiramente diferente. Dretske distinguiu entre operadores penetrantes, semi-penetrantes e não-penetrantes. Na linguagem dretskeana podemos dizer que o conhecimento é um operador semi-penetrante.

### 2.6.2.1 Reformulando o paradoxo do ceticismo

Uma vez cientes da distinção entre os diferentes tipos de princípio de fecho, podemos refinar a versão crua do paradoxo do ceticismo que apresentamos previamente. A sentença *se S sabe que p, então S sabe que  $\neg hc$*  é claramente falsa, e por razões aparentemente triviais: S pode não acreditar que  $\neg hc$ , não saber que p implica  $\neg hc$  ou não inferir  $\neg hc$  de p. Assim, parece que devemos reformar a sentença 2 do paradoxo do ceticismo, a fim de tornar o problema mais preciso.

A breve explicação que fizemos a respeito dos princípios de fecho do conhecimento parece ir diretamente ao ponto. O que é *prima facie* implausível não é que S saiba que p e não saiba que  $\neg hc$ , mas que S saiba que p e não possa saber, por meio de uma inferência válida, que  $\neg hc$ . Tal constatação sugere o seguinte refinamento do paradoxo do ceticismo:

1. Eu sei que p;
2. Se eu sei que p, e infiro validamente  $\neg hc$  a partir de p, então eu sei que  $\neg hc$ ;
3. Eu não sei que  $\neg hc$  ( ao inferir validamente  $\neg hc$  a partir de p).

Esta segunda formulação do paradoxo do ceticismo evita os problemas relativamente triviais que parecem afetar a formulação inicial. E mais importante ainda, a segunda formulação está mais explicitamente conectada com a breve explicação anterior dos princípios de fecho do conhecimento. Com efeito, a segunda sentença do paradoxo do ceticismo reformulado é meramente uma instância do princípio de fecho epistêmico pela inferência válida.

Como a segunda sentença do paradoxo do ceticismo reformulado é uma instância de FIV, a rejeição da segunda premissa deve ser acompanhada da rejeição de FIV. Nozick, assim como Dretske anteriormente, aceitou tal consequência. Ele pensa que FIV (assim como outras versões refinadas de princípios de fecho do conhecimento) é falso, apesar da sua plausibilidade inicial.

A recusa de FIV por parte de Nozick não é arbitrária ou gratuita. Evidentemente não devemos recusar, sem maiores explicações e sem alguma teoria subjacente, um princípio que possui grande plausibilidade e força intuitiva. Contudo, não é este o caso. Nozick possui uma teoria, motivada por considerações independentes da sua resposta ao ceticismo, que explica a

falsidade da segunda premissa do paradoxo do ceticismo, tanto em suas versões preliminares, como nas versões reformuladas.

## 2.7 A CONDIÇÃO DE ADERÊNCIA

A maior parte das discussões relativas à teoria epistemológica de Nozick se concentram no princípio de sensibilidade e no engenhoso tratamento do problema cético. É verdade que o princípio de sensibilidade desempenha um papel mais fundamental na construção da teoria. Contudo, o rastreamento contém uma cláusula adicional - a chamada *condição da aderência*- e uma explicação para o conhecimento inferencial.

A aderência é um condicional subjuntivo adicional que, quando acoplado ao princípio de sensibilidade, permite a solução de alguns aparentes problemas para a teoria. Assim, de acordo com Nozick, o princípio de sensibilidade é apenas uma condição *necessária*, porém não suficiente, para o conhecimento. Consequentemente, não é legítimo objetar à sensibilidade alegando que existem crenças sensíveis que não são casos de conhecimento, sem que se verifique se a aderência é satisfeita nesses presumidos contraexemplos.

Consideremos o seguinte caso, foi apresentado originalmente por Harman (1968). Imaginemos que o ditador de um país foi assassinado. O jornal oficial, por um lapso dos órgãos de censura ainda perplexos, noticia corretamente que o ditador foi assassinado. Pouco tempo depois, a censura manda recolher as edições que veiculavam o assassinato, e as substitui por outra que o nega. A segunda edição, com a falsa notícia, ganha ampla circulação, e a primeira é logo suprimida da maioria dos lugares. Quase todos passam a acreditar falsamente que o ditador está vivo. Suponhamos que existe uma determinada pessoa, digamos um sujeito S, que teve a sorte de encontrar apenas a primeira edição do jornal, e não ouviu falar da segunda edição, como quase todos os habitantes do país. A crença de S é verdadeira, justificada e, no entanto, ao menos segundo Nozick, não diríamos que S sabe que o ditador está morto.

Parece que S simplesmente teve sorte de obter apenas a primeira edição do jornal, o qual, também por acidente ou sorte, permitiu a veiculação da verdade sobre o assunto. O relevante sobre este caso para a teoria de Nozick é que a crença de S parece satisfazer o princípio de sensibilidade. Se o ditador não tivesse sido assassinado, então a primeira edição não teria veiculado a informação de que ele foi assassinado, e assim S não teria acreditado que o ditador foi assassinado.

A aplicação do princípio de aderência ao caso do ditador permite que o veredicto correto, novamente segundo Nozick, seja gerado. É falso que se o ditador tivesse sido assassinado, então S teria acreditado que o ditador foi assassinado. Existem mundos muito próximos no qual o ditador é assassinado, por exemplo o mundo em que S recebe a segunda edição do jornal, ao invés da primeira, e no qual S não acredita que ele foi assassinado. Isto é suficiente, também segundo Nozick, para mostrar que o princípio de aderência não é satisfeito no caso em questão.

O princípio de aderência também contribui para o tratamento de outros contraexemplos importantes, para os quais o princípio de sensibilidade não oferece nenhuma resposta adequada.

Consideremos o seguinte caso, que foi apresentado por Nozick (1981). Uma lesão no cérebro de S causa a crença irracional em S de que ele possui uma lesão no cérebro. A crença simplesmente "surge" na mente, de forma aparentemente espontânea, sem que S saiba como ou porquê. A crença de S de que possui um cérebro satisfaz o princípio da sensibilidade: se ele não tivesse uma lesão no cérebro, então ele não teria a crença de que possui uma lesão no cérebro.

Assim, como nos casos das verdades logicamente necessárias, o princípio de sensibilidade não nos fornece meios de distinguir entre ignorância e conhecimento. A resposta de Nozick ao problema recorre à aderência. A crença de S de que possui uma lesão no cérebro não é aderente, pois "se a lesão cerebral tivesse sido ligeiramente diferente, a crença não teria sido formada ainda que se utilizasse a mesma rota" (NOZICK, 1981, p.190). Em outros termos, existe um mundo *w* muito próximo ao mundo atual, tal que em *w* o sujeito S forma a crença com base no mesmo método, e S não acredita que possui uma lesão cerebral<sup>24</sup>.

### 2.7.1 Aderência e verdades logicamente necessárias

A razão mais importante para uma introdução da condição de aderência na teoria é a insuficiência da condição de sensibilidade para tratar do conhecimento de verdades necessá-

---

<sup>24</sup> Note-se que com uma pequena modificação no exemplo, a crença de S de que possui uma lesão no cérebro passa a ser também aderente. Basta imaginarmos que a lesão, devido ao local específico em que ela ocorre, normalmente causa nos afetados por ela a crença de que possuem uma lesão cerebral. Não é claro qual seria o veredicto de Nozick a respeito deste novo caso.

rias. O princípio de sensibilidade ou bem é inaplicável a quaisquer verdades logicamente necessárias, ou bem ele é trivialmente satisfeito<sup>25</sup>. Em quaisquer um dos casos, devemos procurar alguma cláusula extra para a posse de conhecimento.

Seja  $p$  uma verdade necessária qualquer, isto é, uma proposição qualquer tal que  $\Box p$ . O condicional da sensibilidade aplicado a  $p$  diz que  $\neg p \Box \rightarrow \neg \beta p$ . Ora, de acordo com a interdefinibilidade das noções de possibilidade e necessidade, para quaisquer  $p$  e  $q$ , é o caso que  $\Box p \longleftrightarrow \neg \Diamond \neg p$ . Por conseguinte, quando  $p$  é uma verdade necessária, o condicional da sensibilidade contém um antecedente impossível. Pois bem, como avaliar condicionais subjuntivos com antecedentes impossíveis, os chamados *contrapossíveis*? Como avaliar o que seria o caso, se o impossível acontecesse? De acordo com o princípio da trivialidade dos contrapossíveis que vimos anteriormente:

$$\neg \Diamond p \rightarrow (p \Box \rightarrow q).$$

A interdefinibilidade de necessidade e possibilidade diz que  $\Box p$  é equivalente a  $\neg \Diamond \neg p$ . Por conseguinte, o condicional da sensibilidade, quando aplicado a crenças em verdades necessárias, é um contrapossível. Em razão da trivialidade dos contrapossíveis, é verdade, para qualquer proposição  $p$ , que:

$$\Box p \rightarrow (\neg p \Box \rightarrow \neg \beta p).$$

O princípio acima diz que toda crença em uma verdade necessária satisfaz o princípio de sensibilidade, independentemente de como a crença foi formada. Todavia, toda crença em uma verdadeira necessária é ela própria verdadeira. Portanto, na ausência de uma condição adicional, basta acreditar em uma proposição necessariamente verdadeira para saber que a proposição é verdadeira. Assim, para qualquer verdade necessária  $p$ :

$$\beta p \rightarrow Kp.$$

Em suma, desaparece a diferença entre crença em verdades necessárias e conhecimento de verdades necessárias<sup>26</sup>.

---

<sup>25</sup> Nesta seção, trato do problema das verdades logicamente necessárias, embora penso ser possível apresentar problema semelhante para as verdades nomologicamente necessárias. Por brevidade, e a menos seja indicado em contrário, daqui por diante sempre me refiro à necessidade lógica.

### 2.7.2 A resposta de Nozick ao problema das verdades necessárias

No que diz respeito ao problema da aplicação do princípio de sensibilidade às crenças em verdades necessárias, Nozick afirma que:

Uma vez que a condição 3 não entra em jogo para a verdade necessária P, a explicação da teoria para estas sentenças, quando verdadeiramente acreditadas, reduz-se a 4: se P fosse verdadeiro e S fosse usar o método M para formar crença a respeito de P, então S acreditaria, via M, que P (NOZICK, 1981, p.186).

É quase inevitável a insatisfação com a ideia sugerida por Nozick. Esperamos que uma análise do conceito de conhecimento seja igualmente aplicável a todos tipos de conhecimento, seja este conhecimento de verdades necessárias ou de verdades contingentes, sob pena de não ser uma *análise* do conceito de conhecimento. Se Nozick aceitasse que condicionais subjuntivos com antecedentes impossíveis são necessariamente verdadeiros, ele poderia ainda argumentar que, nos casos relevantes de crenças em verdades necessárias que não são conhecimento, o princípio de aderência sempre falha. Tal alternativa seria menos obscura e mais coerente.

De qualquer forma, segundo Nozick, o princípio de aderência evita a trivialização do conhecimento de verdades necessárias. Suponhamos que um sujeito S acredita verdadeiramente em uma determinada proposição matemática, digamos a proposição  $p$ . S acredita que  $p$  simplesmente porque um amigo de S, que é ignorante em matemática, disse a ele que  $p$ , e S acredita, por razões puramente emocionais e sem qualquer base evidencial, em que tudo o que seu amigo diz.

De acordo com Nozick, S poderia muito facilmente, com pequenas alterações nas circunstâncias, e preservando o mesmo método, não ter acreditado que  $p$ . Existe um mundo muito próximo onde o amigo de S não afirma que  $p$  (afinal, ele é um ignorante sobre o assunto), mais precisamente, ele afirma que  $\neg p$ , e S não acredita que  $p$ , pois acredita em  $\neg p$  (suponho que ele não forme as duas crenças contraditórias ao mesmo tempo). S simplesmente teve *sorte*

---

<sup>26</sup> Alguém poderia pensar que as verdades necessárias são tais que basta compreender o conteúdo da proposição para estar em posição de saber que ela é verdadeira. Não seria necessário justificar a asserção de que crenças em verdades necessárias deve ser distinto de conhecimento de verdades necessárias? Mesmo que este venha a ser o caso para certas verdades necessárias, sem dúvida esta não é o caso para **toda** proposição necessariamente verdadeira. Podemos compreender sentenças verdadeiras da lógica ou da matemática sem saber que elas são verdadeiras, pois não vemos como elas seguem logicamente de outras proposições. Agradeço ao professor Rodrigo Borges por ter apontado esta possível objeção.

de estar em uma das situações em que seu amigo, sendo um ignorante sobre o assunto, asseriu sentenças matemáticas verdadeiras. Trata-se de um caso interessante, pois ele mostra que é possível ter uma crença acidentalmente verdadeira em uma verdade necessária. A sorte epistêmica é compatível com a crença em verdades necessárias, como demonstra o exemplo.

### 2.7.3 Condição de aderência e os contrafactuais verdade-verdade

Uma segunda questão importante para o tratamento da condição da aderência diz respeito aos contrafactuais com antecedentes e consequentes verdadeiros (os *contrafactuais verdade-verdade*, como os denominaremos).

Digamos que  $p$  e  $q$  sejam verdadeiros em um mundo possível  $w$ . Vimos anteriormente o seguinte princípio:

**Centramento** - não existe nenhum mundo tão próximo a um mundo  $w$  como o próprio mundo  $w$ .

Em razão do princípio de centramento, se  $p$  é verdadeiro em  $w$ , então próprio  $w$  é o mundo mais próximo a  $w$ . Portanto, temos que  $p \Box \rightarrow q$  é verdadeiro em  $w$  se  $p \rightarrow q$  em  $w$ . Assim, para quaisquer  $p$  e  $q$ :

$$(p \& q) \rightarrow (p \Box \rightarrow q).$$

As verdades do antecedente e do consequente garantem a verdade do condicional subjuntivo, de acordo com Stalnaker e Lewis (que vimos na seção 2.3). Por conseguinte, quando é verdade que  $p$  e  $S$  acredita que  $p$ , a condição de aderência é satisfeita. Dessa forma, para quaisquer  $p$  e  $q$ ,

$$(\beta p \& p) \rightarrow (p \Box \rightarrow \beta p).$$

Tal resultado torna o princípio de aderência supérfluo para uma análise do conhecimento, visto que não distingue entre casos de crença verdadeira *simpliciter* e casos de conhecimento. Se isto for correto, então não há qualquer razão para introduzir a cláusula. A única maneira de evitar esta consequência é rejeitar a trivialidade dos condicionais verdade-verdade. A esse respeito, Nozick escreve o seguinte:

Não apenas  $p$  é verdade e  $S$  acredita em  $p$ , mas se fosse verdade que  $p$ , então  $S$  acreditaria em  $p$ . Compare: não apenas o fóton foi emitido e foi para a esquerda, mas (era verdade no momento que): se fosse emitido, iria para a esquerda. As verdades do antecedente e do consequente não são suficientes para a verdade dos subjuntivos; 4 diz mais que 1 e 2. De acordo com a sugestão hesitantemente feita acima, 4 é verdade se não apenas ele acredita que  $p$ , mas se nos mundos “próximos” em que  $p$  é verdadeiro, ele acredita que  $p$  (NOZICK, 1981, p.176).

Nozick afirma que a crença de um sujeito  $S$  de que  $p$  satisfaz a condição da aderência somente se  $S$  acredita que  $p$  em um determinado conjunto de  $p$ -mundos mais extenso que o conjunto formado *unicamente* pelo mundo atual. Se existem mundos possíveis próximos ao mundo atual em que é verdade que  $p$  e ainda assim  $S$  não acredita que  $p$ , então a condição de aderência não é satisfeita, mesmo que  $S$  acredite verdadeiramente que  $p$  no mundo atual. Assim, pelo menos de acordo com Nozick, a condição de aderência não é trivialmente satisfeita toda vez que um sujeito crê verdadeiramente.

Naturalmente, esperávamos que Nozick viesse a fornecer um tratamento uniforme dos condicionais subjuntivos, e que deveria aplicar-se igualmente para as condições da aderência e da sensibilidade. Apesar da afirmação nozickiana de que não fornecerá uma análise dos condicionais subjuntivos, ele parece *justamente* fazer isto quando trata do princípio de aderência e dos condicionais verdade-verdade.

Não penso que seja inviável adaptar o tratamento dado por Nozick ao princípio de aderência, de tal forma que seria também aplicável ao princípio de sensibilidade. Uma possibilidade é a seguinte: a crença de  $S$  de que  $p$  no mundo  $w$  é sensível se e somente se  $S$  não acredita que  $p$  em todos os  $\neg p$ -mundos próximos ao mundo  $w$  (não apenas no único  $\neg p$ -mundo *mais* próximo, se houver um  $\neg p$ -mundo mais próximo).

Podemos subdividir o conjunto de mundos possíveis em que é o caso  $p$  em dois outros grandes conjuntos: o conjunto dos mundos possíveis próximos (i.e., não muito diferentes) em que  $p$ , e o conjunto dos mundos possíveis não-próximos (i.e. muito diferentes) em que  $p$ . A aderência é satisfeita se nos mundos possíveis próximos, porém diferentes do mundo atual, em que é verdade  $p$  e  $S$  utiliza o mesmo método, também é verdade que  $S$  acredita que  $p$ .

Assim, proponho a seguinte regra para interpretar a condição de aderência: a crença de um sujeito  $S$  de que  $p$  é aderente no mundo  $w$  se e somente se nos mundos próximos a  $w$  em que é verdade que  $p$  e  $S$  utiliza o mesmo método,  $S$  acredita que  $p$ . A regra vai ao encontro ao que Nozick diz sobre o assunto, embora não seja explicitamente formulada por ele, e é a partir dela que interpretaremos a condição de aderência.

## 2.8 A TEORIA DE NOZICK DO CONHECIMENTO INFERENCIAL

Além de propor uma análise do conceito de conhecimento em geral, Nozick também propôs, paralelamente, uma abordagem do conhecimento inferencial. Embora esta seja comumente negligenciada na literatura sobre o assunto, ela é parte importante da teoria nozickiana, e deve ser aqui examinada<sup>27</sup>. Nozick rejeita FIV, sem obviamente rejeitar que a inferência válida seja ampliativa do nosso conhecimento. Ele também possui uma teoria que explica como e quando adquirimos conhecimento inferencial, e por que não é possível adquirir conhecimento inferencial de que  $\neg hc$  (uma hipótese cética qualquer) a partir de uma proposição do senso comum (p.ex., eu possuo mãos).

Para um filósofo adepto de uma análise do conhecimento em termos de contrafactuais, Nozick, não surpreendentemente, também propôs uma análise contrafactual para o conhecimento inferencial. Segundo ele, um sujeito *S* sabe que *q* por meio de uma inferência a partir de *p* se e somente se:

1. É verdade que *q*;
2. *S* sabe que *p*;
3.  $\neg q \square \rightarrow \neg \beta q$  via inferência de *p*;
4.  $q \square \rightarrow \beta q$  via inferência de *p*.

A terceira e a quarta condição do conhecimento inferencial dizem, respectivamente, o seguinte: se fosse falso que *q* (a conclusão da inferência), então *S* não acreditaria que *q* a partir de uma inferência de *p*; se fosse verdadeiro que *q*, então *S* acreditaria que *q* a partir de uma inferência de *p*.

Uma das consequências mais importantes da análise do conhecimento inferencial é que ela não permite que um sujeito *S* saiba que  $\neg hc$  inferindo que *p*, quando *hc* é uma típica hipótese cética e *p* é uma proposição ordinária objeto de conhecimento. Se fosse falso que

---

<sup>27</sup> Uma exceção é Baumann (2012), que faz uma excelente análise do tópico.

$\neg hc$ , i.e., verdadeiro que  $hc$ , então  $S$  acreditaria que  $p$  (supondo que  $S$  considerasse a proposição  $p$ ), pois se fosse verdadeiro que  $hc$ , então  $S$  teria as mesmas evidências e razões que no mundo atual que o levam a acreditar que  $p$ , e a inferir  $\neg hc$  de  $p$ . Em outros termos, a teoria de Nozick do conhecimento inferencial não autoriza o seguinte raciocínio, mesmo quando  $Kp$ :

1.  $p$ ;
2. Se  $p$ , então  $\neg hc$ ;
3. Logo,  $\neg hc$ .

A inferência acima não transmite conhecimento de que  $p$  para  $\neg hc$ , de acordo com a teoria do conhecimento inferencial que vimos acima. Assim, a teoria de Nozick do conhecimento inferencial bloqueia a proposta de que  $S$  poderia saber que  $\neg hc$ , a partir de uma inferência válida de  $p$ , quando a crença de que  $p$  é um caso de conhecimento.

### 3 OBJEÇÕES À TEORIA DO RASTREAMENTO

#### 3.1 A TEORIA É MUITO RESTRITIVA

Nos capítulos anteriores foi explicada a teoria do rastreamento. Também examinamos as alegadas soluções que a teoria produz, notadamente a sua criativa resposta ao problema céptico. Contudo, a teoria, desde o seu surgimento, tem sido alvo de pesadas críticas, incluindo-se aqui uma grande quantidade de contraexemplos. A seguir, examinaremos as objeções que sugerem ser a teoria muito restritiva, isto é, que classifica casos de conhecimento como casos de ignorância.

##### 3.1.1 Indução

Um dos problemas mais graves que a teoria do rastreamento enfrenta diz respeito ao conhecimento indutivo<sup>28</sup>. Foi Jonathan Vogel quem primeiro chamou atenção para a aparente incompatibilidade da teoria com o conhecimento indutivo. Eis um dos seus contraexemplos:

*Urânio.* Roger coloca um pedaço de urânio em uma placa fotográfica, e descobre que a placa se tornou embaçada. Ele repete o experimento muitas vezes. Roger sabe, por indução, que a placa recentemente exposta está embaçada, mesmo antes de inspeciona-la (VOGEL, 2012, p. 131)

Nós diríamos- ceticismo sobre indução à parte - que Roger sabe que a placa recém colocada está embaçada. Contudo, argumenta Vogel, a crença de Roger não é sensível. Se a placa não estivesse embaçada - supondo tudo o mais constante - Roger acreditaria falsamente que a placa está embaçada.

Um outro exemplo, também apresentado por Vogel (2012), é o caso que denominaremos de *caso do copo de gelo*. Em um dia quente de verão, eu esqueço um copo com cubos de gelo na sacada do meu apartamento, ficando exposto ao sol e calor. Horas depois, eu lembro que esqueci o copo na sacada do meu apartamento, e concluo que o gelo contido no copo já está derretido. Mesmo antes de observar com os meus próprios olhos o copo com derre-

---

<sup>28</sup> Exposições detalhadas do problema podem ser encontradas em Vogel (1987), Vogel (2012), Sosa (1999 a) e Sosa (1999b).

tido, não diríamos que eu *sei* que o gelo contido no copo derreteu, após horas de exposição ao sol e calor?

Assim como ocorre no exemplo da placa embaçada, o princípio de sensibilidade também parece falhar no caso do copo de gelo. Se, por alguma razão, a água contida no copo continuasse sob a forma sólida, mesmo após intenso sol e calor, eu ainda continuaria acreditando que ela está na forma líquida. O princípio de sensibilidade não parece aplicar-se, novamente, a um caso de conhecimento indutivo.

Por fim, mencionemos mais um exemplo conhecido, que foi introduzido por Sosa (1999b), e que denominaremos de *caso da lixeira*. Do último andar do meu prédio, através de um compartimento, eu joga um saco de lixo, o qual atravessa um duto, cai diretamente na lixeira do prédio, e é recolhido por ela. Este mecanismo funciona muito bem e com regularidade. Suponhamos que eu acabei de jogar o lixo através do compartimento, e que tudo funcionou como sempre tem funcionado.

Eu *sei* que o lixo caiu na lixeira, logo após jogá-lo pelo compartimento? Aparentemente sim, especialmente se o mecanismo tem funcionado com regularidade, e não há nada de anormal ou estranho na situação. Todavia, o princípio de sensibilidade parece falhar no caso. Se o lixo não fosse recolhido pela lixeira, eu ainda acreditaria que ele foi recolhido. A minha crença não rastreia a verdade. Por conseguinte, se é verdadeira a teoria do rastreamento, então eu não sei que o lixo foi recolhido pela lixeira.

### 3.1.2 A resposta de Nozick

Nozick estava ciente de que a teoria do rastreamento parece ter dificuldades em acomodar o conhecimento indutivo. Contudo, ele pensou ser possível tratar da questão, e produzir uma explicação do conhecimento indutivo compatível com a sua teoria.

Digamos que um sujeito S acredita que  $p$  com base em uma evidência  $e$ . A crença de S de que  $p$  é produzida por uma inferência indutiva a partir de  $e$ . De acordo com Nozick, se  $p$  fosse falsa, então a falsidade de  $p$  teria sido refletida no conjunto das evidências que S utiliza ao formar a sua respectiva crença, e ele não teria formado a crença de que  $p$ .

Nozick coloca da seguinte maneira: A seguinte condição (denomine-a de "I" de inferência) especifica quando a inferência de  $q$  de  $p$  (que é conhecido) produz conhecimento de que  $q$ :

I: se  $q$  fosse falso,  $S$  não acreditaria que  $p$  (ou não inferiria  $q$  de  $p$ )  
(NOZICK, 1981, p.231)

Isto parece sugerir a seguinte regra para o conhecimento indutivo:

**(IND)** se  $S$  infere indutivamente que  $p$  a partir de uma evidência  $q$ , e  $S$  sabe que  $p$  em razão da inferência, então: se fosse falso que  $p$ , então  $S$  não acreditaria que  $p$  com base em  $q$  ou não  $S$  não acredita que  $q$ .

Em outra passagem importante, Nozick explica o conhecimento indutivo da seguinte forma:

Nós sustentamos uma crença sobre o futuro na base de alguns fatos passados e recentes. Nós sabemos que  $p$ ? Apenas se nossa crença de que  $p$  rastreia o fato de que  $p$ . Isto é, apenas se: se  $p$  não fosse verdadeiro, então isto teria sido refletido em alguns fatos passados e recentes, e eles não teriam acontecido (NOZICK, 1981, p.222)

Vamos agora mostrar como a teoria do conhecimento inferencial de Nozick poderia ser aplicado a casos de conhecimento indutivo. Retomando o exemplo de Vogel, é verdade que  $S$  sabe que os cubos de gelo derreteram somente se:

(9) se os cubos de gelo não tivessem derretido, então  $S$  não teria obtido as evidências que atualmente possui para a crença de que os cubos de gelo derreteram (p.ex., evidências sobre o tempo em que os cubos ficaram expostos ao sol, a temperatura do dia, o comportamento da água em estado sólido quando submetida ao calor intenso, etc.) e fatos passados e recentes teriam sido diferentes (p.ex., o tempo em que os cubos ficaram expostos ao sol teria sido diferente, a temperatura do dia teria sido diferente, o comportamento da água em estado sólido quando submetida ao calor intenso teria sido diferente, e assim por diante), e  $S$  não teria acreditado que os cubos de gelo derreteram a partir das evidências que atualmente possui.

Contrafactuais em que o tempo do antecedente é posterior ao tempo do consequente são os denominados *contrafactuais reversos*. A resposta de Nozick ao conhecimento claramente utiliza ou depende de um contrafactual reverso.

### 3.1.3 A crítica de Vogel

Vogel faz duas críticas à solução proposta por Nozick. Na primeira delas, Vogel trata de como interpretamos o critério de similaridade, e alega que contrafactuais como (9) são indeterminados ou, na melhor das hipóteses, falsos. Para entender a argumentação de Vogel, devemos nos colocar algumas questões. Que tipo de alterações deveriam ser realizadas em mundos próximos em que o gelo não derreteu? Devemos considerar que o comportamento da água em estado sólido quando submetida ao calor é geralmente diferente nestes mundos? Ou devemos considerar que tudo o mais é constante, e que apenas aconteceu uma exceção, isto é, que apenas *este* cubo de gelo *nestas* circunstâncias não derreteu, e que tal acontecimento não se repetiria?

Um mundo  $W_1$  onde apenas *esta* exceção às leis naturais aconteceu é mais ou menos similar ao mundo atual que outro mundo  $W_2$  em que os fatos relevantes para o exemplo são diferentes, em que a história pregressa do mundo é fixada, porém as leis naturais não sofrem exceções? Evidentemente, precisaríamos aqui de um critério de similaridade. Qual dos dois fatores devemos preservar, a semelhança no que diz respeito à história pregressa do mundo (até o momento em que aquele copo de gelo foi exposto ao sol) ou a inviolabilidade das leis naturais? Vogel alega que não existe uma resposta clara para estas questões, e, em virtude de tal obscuridade, não é evidente como devemos aplicar o critério de similaridade a condicionais do tipo (9)<sup>29</sup>.

Além do problema da indeterminação, Vogel também sugere que contrafactuais reversos como (9), se não são indeterminados, são claramente falsos. Ele alega que o condicional (9) é incompatível com outro condicional muito plausível:

(10) se estes cubos de gelo não tivessem derretido, então S ficaria surpreso (i.e., não esperaria que estes cubos não derretessem).

---

<sup>29</sup> Vogel argumenta que contrafactuais reversos como (9) são falsos ou indeterminados. Penso que é muito mais plausível que condicionais como (9) sejam considerados falsos, e que isto concorda com nossos juízos intuitivos sobre o assunto. Quando ouvimos alguém enunciar um condicional como (9), nós não pensamos que o que a pessoa enunciou não é verdadeiro nem falso: nós imediatamente pensamos que ela disse algo *falso*.

Contudo, se a explicação de Nozick para o conhecimento indutivo estiver correta, deve ser verdade que se os cubos de gelo não tivessem derretido, então S teria tido sinais ou indícios de que isto iria acontecer, o que parece sugerir que:

(11) se os cubos de gelo não tivessem derretido, então não S ficaria surpreso.

Também parece ser o caso que se a interpretação de Nozick for correta, o não-derretimento dos cubos deveria ter sido acompanhado por sinais de que o cubo de gelo não iria derreter e pelo não-derretimento de cubos de gelo (e de outros objetos semelhantes) em situações desse tipo. Isto sugere, por sua vez, que:

(12) se estes cubos de gelo não tivessem derretido, então estes cubos de gelo não seriam diferentes dos outros cubos de gelo.

Vogel alega que também é verdadeiro o seguinte condicional, incompatível com (12):

(13) se estes cubos de gelo não tivessem derretido, então estes cubos de gelo seriam muitos diferentes dos outros cubos de gelo.

Como é possível que os condicionais (10) e (11) sejam ambos verdadeiros? Como é possível, por sua vez, que os condicionais (12) e (13) também sejam verdadeiros? Note-se que os respectivos condicionais possuem os mesmos antecedentes, porém os consequentes de ambos são contraditórios um do outro<sup>30</sup>. Isto sugere que as crenças indutivas colocam um problema grave para o conhecimento indutivo, segundo a proposta de Nozick.

#### 3.1.4 Reversão de contrafactuais

---

<sup>30</sup> Uma maneira de tentar responder a essa objeção é alegar que os condicionais (10) e (11) podem ser simultaneamente verdadeiros, dada a dependência dos contrafactuais do contexto conversacional. O defensor desta réplica precisaria identificar a diferença contextual relevante, e mostrar como ela opera para produzir uma aparência falsa de incompatibilidade entre as duas sentenças. De qualquer forma, este tipo de manobra tende a fazer com que a posse de conhecimento dependa de sutis variações contextuais, o que não é pretendido de Nozick. Uma combinação de uma versão mitigada do princípio de sensibilidade com uma teoria contextualista da atribuição de conhecimento pode ser encontrada em DeRose (1995).

Imaginemos a seguinte situação<sup>31</sup>. Você vê seu amigo próximo à sacada de um prédio muito alto. Imediatamente, você pensa: *se meu amigo se jogasse pela janela, então ele se machucaria*. Este parece um excelente exemplo de contrafactual verdadeiro.

Suponhamos agora que, diante desta situação, e tendo em mente a ausência de quaisquer disposições suicidas do seu amigo, alguém pensasse da seguinte forma: não é o caso que se meu amigo se jogasse pela janela, então ele teria se machucado, pois se meu amigo se jogasse pela janela, teria existido previamente uma proteção fora do prédio.

O argumento em questão parece profundamente defeituoso. Quando raciocinamos sobre contrafactuais, não fazemos este tipo de inferência. Quando um sujeito S alega que um contrafactual é falso (ou verdadeiro) em razão de um contrafactual reverso, da forma como faz o sujeito a respeito do seu amigo, diremos que ele *raciocina inversamente*. A admissão deste tipo de procedimento gera resultados profundamente contraintuitivos, tal como a crença de que meu amigo não se machucaria se ele se jogasse de uma janela de um prédio alto.

Pois bem, o que agora quero salientar é que o argumento de Nozick em favor da sensibilidade das crenças indutivas é precisamente um raciocínio inverso, e análogo ao raciocínio do sujeito que conclui que seu amigo não se machucaria, caso ele se jogasse pela janela. Segundo a proposta de Nozick, se o gelo não tivesse derretido, então fatos passados e recentes relevantes (como o comportamento passado da água em estado sólido quando submetida ao calor intenso, e outras coisas do gênero) teriam sido diferentes, e, por isto, S não teria obtido as evidências de que o gelo derreteu. Consequentemente, se o gelo não tivesse derretido, então S não teria acreditado que o gelo derreteu. Se esta caracterização da argumentação de Nozick está correta, então o argumento do autor tem a seguinte forma:

- Se fosse falso que p, então fatos passados e recentes relevantes teriam sido diferentes;
- Se fatos passados e recentes relevantes tivessem sido diferentes, então S não teria acreditado que p com base em e;
- Logo, se fosse falso que p, então S não teria acreditado que p com base em e.

O argumento de Nozick exemplifica o mesmo tipo de defeito que constatamos no argumento do sujeito que concluiu que seu amigo não se machucaria, caso ele se jogasse pela janela. Contudo, tal como vimos, raciocínios deste tipo são extremamente problemáticos. Um ponto importante a salientar é que o argumento de Nozick assume implicitamente a transitivi-

---

<sup>31</sup> Exemplo extraído de Khoo (2016).

dade dos condicionais contrafactuais. Todavia, condicionais contrafactuais não são transitivos, e o raciocínio por silogismo hipotético, válido para o condicional material, não pode ser transposto para condicionais contrafactuais<sup>32</sup>.

Recapitulando, podemos extrair da crítica Vogel três objeções fundamentais à solução de Nozick para o problema das crenças indutivas:

- Não temos um critério claro de similaridade que nos permita decidir se contrafactuais reversos como (9) são verdadeiros ou falsos.
- A solução nozickiana é incompatível com certos contrafactuais, que nos parecem verdadeiros.
- Nozick defende a sensibilidade por meio de um raciocínio questionável.

Estas três considerações mostram que a resposta Nozick não resolve o problema das crenças indutivas.

### 3.1.5 Um último recurso: a resposta de Kelly Becker

Diante do que vimos, parece haver um último recurso: alegar que não temos conhecimento indutivo nos casos de que tratamos, porém que isto não é tão desastroso como parece à primeira vista. Esta é a posição defendida por Becker (2007). Segundo o autor, o fato de que a teoria do rastreamento sugere que o sujeito não possui conhecimento nos exemplos que vimos não deve ser tomado como um defeito ou problema para a teoria: nós *realmente* não temos conhecimento indutivo nesses casos, e o fato do princípio de sensibilidade gerar tal resultado conta *em seu favor*. Referindo-se ao caso da lixeira, ele afirma: “ nós podemos plausivelmente dizer isto: S não sabe que o saco de lixo logo alcançará o chão; ao invés disso, ele sabe que muito provavelmente ele chegará lá” (Becker, 2007, p. 55).

---

<sup>32</sup> Mais detalhes acerca da falha do silogismo hipotético para condicional contrafactual podem ser encontrados em Bennet (2003). Um exemplo de falha do silogismo hipotético para condicionais contrafactuais é relatado por Bennet (2003, p. 160): se você tivesse caminhado sobre o gelo, enquanto apoia todo o seu peso em alguém que não se encontra sobre o gelo, então você teria caminhado sobre o gelo. Se você tivesse caminhado sobre o gelo, então o gelo teria partido. Logo, se você tivesse caminhado sobre o gelo, apoiando todo o seu peso em alguém que não se encontra sobre o gelo, então o gelo teria partido.

Diante da previsível recusa de alguns em aceitar a sua tese, ele apresenta o seguinte argumento, sugerindo que a sua tese é mais amplamente compartilhada do que parece à primeira vista:

Esta pode parecer uma alegação drástica, mas é apoiada até mesmo por livros introdutórios de lógica, pois quase todos eles definem um argumento indutivo como aquele que torna a conclusão provavelmente verdadeira. Isto, eu sugiro, é razão suficiente para negar que S sabe que o saco lixo logo alcançará o chão, e que o policial sabe que o bandido errou o alvo [ aqui Vogel refere-se a outro contraexemplo], ao mesmo tempo em que se preserva o conhecimento indutivo (BECKER, 2007, p. 55).

Da minha parte, não vejo por que livros introdutórios de lógica, com todos os problemas e simplificações que são encontradas em manuais introdutórios, deveriam servir de autoridade para guiar nossos juízos acerca da natureza das inferências. De qualquer modo, vamos aceitar que a definição aludida por ele está correta. O problema é que afirmar que argumentos indutivos tornam a suas conclusões provavelmente verdadeiras não é o mesmo que afirmar que (nem implica que):

(i) a conclusão de um argumento indutivo é sempre uma sentença da forma *provavelmente p*;

(ii) argumentos indutivos não geram conhecimento de que são verdadeiras as suas respectivas conclusões.

É difícil ver como a tese de Becker poderia ser apoiada pela evidência apontada por ele.

Além do mais, existem outras duas dificuldades com o tratamento dado por Becker às crenças formadas por meio da inferência indutiva. Substituir o conhecimento indutivo de que *p* pelo conhecimento de que *provavelmente p* não preserva ou salva o conhecimento indutivo. Afinal de contas, o conhecimento indutivo não é idêntico ao conhecimento de que são verdadeiras proposições que dizem respeito a probabilidades. Na melhor das hipóteses, apenas alguns casos, uma fração pequena da totalidade do nosso estoque de conhecimento indutivo, diz respeito a probabilidades. Se este é o caso, então a proposta de Becker não é muito distinta de outras formas de ceticismo quanto ao conhecimento indutivo. Eu não poderia saber, por exemplo, que o sol nascerá amanhã, independentemente do fato que eu poderia saber *outras coisas*, por exemplo, que é provável que o sol nascerá amanhã.

Pois bem, se a proposta de Becker leva ao ceticismo sobre indução (exceto quando trata de proposições sobre probabilidades), então ela é dificilmente conciliável com as motivações originais da teoria do rastreamento. Uma das principais razões em favor da teoria do rastreamento é a sua alegada capacidade de responder ao ceticismo filosófico, em especial a sua variante cartesiana. Nozick concede bastante ao ceticismo, devemos reconhecer, porém ele pensa ser possível reter grande parcela das nossas alegações ordinárias de conhecimento. Todavia, se Becker estiver certo, então a teoria do rastreamento nos coloca na situação desconfortável de ter de rejeitar uma ampla parcela das nossas alegações ordinárias de conhecimento indutivo, o que é uma forma de ceticismo sobre a indução. Para uma teoria motivada pela preocupação cética, este não parece um resultado aceitável, nem uma opção viável.

Ciente das resistências que a resposta anterior poderia suscitar, Becker sugeriu que existe uma maneira alternativa de conciliar sensibilidade e conhecimento indutivo. Ele coloca da seguinte forma:

Entretanto, alguns irão pensar que recusar o conhecimento de que algum evento específico irá ocorrer nesses casos é ir longe demais, talvez especialmente no que diz respeito ao caso do nascimento do sol. Para responder a essas reclamações, nós podemos tentar rejeitar o dilema em questão (a sensibilidade é falsa ou conhecimento indutivo não existe). Para fazer isto, nós precisamos esclarecer o método de formação da crença. Nós dissemos que S acredita que o sol irá nascer com base no fato de que ele nunca experimentou um dia em que isto nunca aconteceu. Talvez nós devêssemos acrescentar: e não existe nenhuma evidência em contrário. Certamente, se existisse tal evidência, S a utilizaria (esta é a base da réplica de Nozick). Em geral, é da própria natureza da inferência indutiva que o método não apenas considera toda evidência positiva, mas, ao menos implicitamente, considera possíveis contraevidências. O método de S, portanto, consiste em inferir que o sol irá nascer baseando-se em toda evidência, positiva e negativa. Se o sol não fosse nascer amanhã, então, usando *este* método, S não acreditaria que o sol irá nascer amanhã (assumindo que a evidência da alteração da rotação da terra teria sido "refletida nos fatos", conforme diz Nozick) (BECKER, 2007, p. 55-56).

O problema da proposta é que a solução de Becker não difere essencialmente da proposta de Nozick, conforme o próprio Becker parece admitir. A exigência de que a falsidade da proposição alvo seja refletida no passado é precisamente a solução proposta por Nozick, e emprega os mesmos contrafactuais reversos, que mostramos ser problemáticos. Na medida em que a solução nozickiana não se mostrou, em última instância, uma alternativa viável, também não deve ser esta proposta por Becker.

### 3.1.6 O conhecimento de segunda-ordem

Segundo alguns autores, a teoria do rastreamento não é compatível com certas formas de conhecimento de segunda ordem<sup>33</sup>. De acordo com esta objeção, aqui denominada de *objeção do conhecimento de segunda ordem* (OSE), um sujeito S não pode saber, a respeito de si mesmo, que é verdade que *eu não creio falsamente que p*, onde *p* é uma proposição qualquer que o sujeito crê verdadeiramente. Como deve ser possível, ao menos em princípio, ter conhecimento deste gênero, alega-se que a teoria do rastreamento é muito restritiva, isto é, classifica certos casos de conhecimento como casos de ignorância.

Suponhamos que um sujeito S acredita que ele próprio não acredita falsamente que *p*. Simbolicamente, representamos a proposição alvo por:

$$\neg(\beta p \ \& \ \neg p).$$

Vamos supor, contrariamente ao que queremos demonstrar, que S sabe que *p*. Neste caso, a crença de que  $\neg(\beta p \ \& \ \neg p)$  satisfaz o princípio de sensibilidade. Portanto, deve ser verdade que:

$$\neg\neg(\beta p \ \& \ \neg p) \ \Box \rightarrow \ \neg \beta \neg(\beta p \ \& \ \neg p).$$

Pela eliminação da dupla negação no antecedente do condicional, temos que:

$$(\beta p \ \& \ \neg p) \ \Box \rightarrow \ \neg \beta \neg(\beta p \ \& \ \neg p).$$

A proposição acima afirma que se S acreditasse falsamente que *p*, então ele não acreditaria que não acredita falsamente que *p*.

Jonathan Vogel pensa que se um sujeito S acreditasse falsamente que *p*, então S acreditaria que *não* acredita falsamente que *p*. Se referindo a um de seus exemplos, no qual um sujeito S acredita que *Omar possui sapatos novos*, Vogel afirma:

Do modo como as coisas são, você acredita que a sua crença de que Omar possui sapatos novos não é falsa. Mas e se fosse? Se de alguma forma a sua crença de que Omar possui sapatos novos fosse falsa, você ainda assim acreditaria que a sua crença é verdadeira, não-falsa. É difícil conceber você não acreditando que algo que você acredita é verdadeiro, independente do que quer que passe em sua mente. Então, se a sua crença de que Omar possui sapatos novos fosse falsa, você ainda assim acreditaria que a sua crença é verdadeira, não-falsa (VOGEL, 2000, p.610).

---

<sup>33</sup> Esta objeção pode ser encontrada em Sosa (1999a), Sosa (1999b), Vogel (2000) e Vogel (2007).

Assim, o que justifica a alegação de Vogel é que, para qualquer crença de um sujeito S de que p, seja ela verdadeira ou falsa, S normalmente acredita que não acredita falsamente que p. Com efeito, Vogel parece ter razão: é difícil imaginar um sujeito acreditando que p e, ao mesmo tempo, não acreditando (quando considera se crê ou não crê verdadeiramente) que a sua crença de que p não é verdadeira.

Alguns autores rejeitam o argumento acima de Vogel porque julgam que não há razão para supor que a crença de primeira ordem deva vir acompanhada pela crença de segunda ordem de que a crença de primeira ordem não é falsa. Assim, por exemplo, Joe Salerno afirma:

Em primeiro lugar, não é óbvio que acreditar que p implica ter a crença de segunda ordem de que não se está enganado ao acreditar que p. Isto implica que crianças pequenas e outros pensadores não reflexivos tem crenças sobre suas próprias crenças. Mais diretamente ao ponto, nenhuma contradição segue da assunção de que existe um pensador que, por qualquer razão que seja, é capaz de formar apenas crenças de primeira ordem (i.e., crenças que não possuem o conceito de crença como parte do seu conteúdo). Então nós temos razão para começar a suspeitar de (\*) [ (\*) é a proposição  $\beta p \Box \rightarrow (\neg \beta \neg (\beta p \& \neg p))$  ] (SALERNO, 2010, p.76)

Penso que a semântica dos contrafactuais que vimos anteriormente fornece meios de responder à crítica de Salerno. O autor pode estar certo ao afirmar que a existência de uma crença de primeira ordem não implica logicamente a existência de uma crença de segunda ordem de que a crença de primeira ordem não é falsa. Contudo, isto dificilmente é relevante para o caso; afinal de contas, mundos em que S é uma criança ou não possui qualquer crença de segunda ordem não fazem parte dos  $(\beta \& \neg p)$ -mundos mais próximos a @. Os mundos relevantes para verificar a satisfação do condicional de sensibilidade são apenas aqueles em que o sujeito é reflexivo, e considera se a crença de primeira ordem de que p é verdadeira ou falsa.

A objeção do conhecimento de segunda-ordem é bastante convincente à primeira vista, e muitos pensaram que ela constitui uma refutação da teoria de Nozick. Apesar da força inicial da objeção, autores como Salerno (2010) e Becker (2007) julgaram que ela não é decisiva, e ser possível respondê-la. A natureza da resposta de cada um dos autores, e por que penso serem ambas inadequadas, é o que veremos a seguir.

### 3.1.6.1 As respostas de Becker e Salerno

Segundo Salerno (2010) e Becker (2007), Vogel realiza uma formalização incorreta da sentença *eu não acredito falsamente que p*, e faz isto explorando uma certa ambiguidade da sentença. De acordo com Salerno, " $\neg (Bp \ \& \ \neg p)$  não é melhor representação formal do pensamento de que eu não acredito falsamente que p" (Salerno, 2010, p.77).

Salerno também faz questão de distinguir duas leituras possíveis da sentença em questão:

Nós começamos notando que "eu não acredito falsamente que p" (ou "eu não me engano ao acreditar que p") é ambíguo. Em uma leitura, a alegação de

(+) eu não acredito falsamente que p

é equivalente a

(1) eu acredito que p e p não é falso.

Mais formalmente,

$Bp \ \& \ \neg \neg p$ .

Ou equivalentemente,

$Bp \ \& \ p$ .

Esta leitura tem a implicação trivial de que eu acredito que p. A segunda leitura não tem esta implicação. Consequentemente, pode ser argumentado que não acredito falsamente que p, se eu não acredito que p.

Na segunda leitura, a alegação de que

(-) eu não acredito falsamente que p

é equivalente a

(2) Não é caso o que ambos: acredito que p e não-p.

Formalmente,

$\neg (Bp \ \& \ \neg p)$ .

A leitura (2) não tem a implicação de que eu acredito que p. Minha falha em acreditar que p é suficiente para a verdade de (2). (SALERNO, 2010, p.77).

Segundo Salerno (2010) e Becker (2007), crenças em proposições da forma  $(\beta p \ \& \ p)$  podem satisfazer o princípio de sensibilidade. Isto mostraria que é possível ter conhecimento de segunda ordem<sup>34</sup>.

---

<sup>34</sup> Rogel Esteves de Oliveria me sugeriu que a crença de S em  $(\beta p \ \& \ p)$  também não é sensível à verdade. Embora simpatize com a sugestão, penso que existe uma questão a ser

Após notar a aparente ambiguidade da sentença *eu não acredito falsamente que p*, Salerno explica porque opta pela primeira interpretação:

A minha crença de que a minha crença de que p não é falsa ocorre em um contexto em que eu acredito p. O pensamento "minha crença de que p não é falso" é sobre a minha crença de que p. Consequentemente, "a minha crença de que p não é falsa" é melhor representada por uma expressão que implica que eu acredito que p.  $Bp \ \& \ p$  é a melhor leitura (SALERNO, 2010, p.78).

É importante atentar para uma distinção importante. Devemos distinguir a sentença:

**Crença de segunda ordem (CSE)** - eu não creio falsamente que p.

De outra sentença superficialmente semelhante:

**Crença de segunda ordem\* (CSE\*)** - a minha crença de que p não é falsa.

E também da sentença:

**Crença de segunda ordem\*\* (CSE\*\*)**: eu não estou enganado ao crer que p.

A expressão *a minha crença de que p* é uma descrição definida, presente em CSE\*, porém ausente em CSE. Sem entrar em detalhes no tópico da análise das descrições definidas, gostaria apenas de salientar que a implicação sugerida por Salerno, a implicação de *tenho uma crença de que p* por *a minha crença de que p não é falsa*, é apoiada por algumas das mais importantes teorias sobre o assunto, tal como a análise russeliana das descrições definidas.

Conforme vimos no trecho supracitado, os autores confundem duas sentenças superficialmente semelhantes, e a distinção deveria ser feita, a fim de verificarmos se a formalização proposta para *eu não acredito falsamente que p* é realmente correta.

Em seu artigo, Salerno também apresenta outra razão em favor da sua interpretação da sentença *eu não acredito falsamente que p*. Segundo ele, a asserção da sentença *eu não acredito falsamente que p* cumpre uma certa função no discurso ordinário, qual seja, expressar a confiança epistêmica de que não se está enganado:

A alegação de que eu não estou enganado em pensar que p deve estar dizendo algo que pode servir como expressão da minha confiança epistêmica. Assim, a alegação de que eu não estou enganado está dizendo, entre outras coisas, que eu

---

respondida aqui antes de adotarmos a proposta de Rogel: o  $\neg (\beta p \ \& \ p)$ -mundo mais próximo ao mundo do exemplo é um mundo em que  $(\neg \beta p \ \& \ p)$ ,  $(\neg \beta p \ \& \ \neg p)$  ou  $(\beta p \ \& \ \neg p)$ ? Sem uma resposta a esta pergunta, é difícil avaliar se o princípio de sensibilidade é satisfeito. A ausência de uma resposta clara deve ser tomada como evidência contrária à proposta de Salerno.

acredito que p. Se eu estou dizendo algo que não implica que eu acredito que p, então eu não estou dizendo algo que expressa minha confiança epistêmica de que p. A leitura que parece mais apropriada é a leitura (1),  $(Bp \ \& \ p)$  (SALERNO, 2010, p.78).

Salerno e Becker afirmam que um sujeito pode saber que  $\beta p \ \& \ p$ , o que mostraria ser possível o conhecimento de segunda de ordem de que não se está enganado. O problema da solução proposta é que, ao aceitarem que um sujeito não pode saber que é verdade, a respeito de si mesmo, que *não é o caso que simultaneamente eu acredito que p e p é falso*, os autores comprometem-se com um conjunto muito implausível de afirmações.

Este problema é mais claro no texto de Becker (2007), porém também está presente em Salerno (2010). Ao sumarizar a sua resposta à objeção de Vogel, Becker afirma explicitamente que "o resultado de tudo isto é que alguém pode ter o conhecimento de segunda ordem de que possui uma crença que é verdadeira e que possui uma crença que não é falsa, mas não que não possui uma crença que é falsa"(Becker, 2007, p.60).

No trecho acima citado, Becker *explicitamente* faz três afirmações, que representamos abaixo:

- $K(\beta p \ \& \ p)$ ;
- $K(\beta p \ \& \ \neg p)$ ;
- $\neg K\neg (\beta p \ \& \ \neg p)$ .

Ora, estas três proposições não formam um conjunto coerente, pelo menos quando um sujeito S está ciente das relações de implicação lógica e faz as inferências relevantes. É extremamente surpreendente que alguém possa saber que a sua crença de que p não é falsa (ou que é verdadeira), mas não possa saber que não tem a crença falsa de que p.  $\beta p \ \& \ q$  implica logicamente  $\neg (\beta p \ \& \ \neg q)$ , e um sujeito que tem conhecimento de que  $p \ \& \ p$ , e que esteja ciente da relação de implicação relevante, deve estar em posição de saber, por meio de uma inferência válida, que  $\neg (\beta p \ \& \ \neg q)$ . Ainda que o princípio do fecho do conhecimento pela implicação conhecida pudesse falhar em outras circunstâncias, esta não parece uma situação em que isto seria plausível.

Conforme vemos, a tentativa de resposta ao problema do conhecimento de segunda ordem proposta por Becker não é bem-sucedida. Contudo, existe uma segunda tentativa de responder ao problema, que é o que veremos.

Pois bem, vamos supor que Becker e Salerno estão corretos quanto a formalização da sentença da *eu não acredito falsamente que p*<sup>35</sup>. O que dizer da seguinte sentença, sugerida por Salerno?

Conforme o próprio Salerno reconhece em trecho supracitado, a sentença CSE\*\* não se presta a formalização proposta para CSE. Uma vez que CSE\*\* seja formalizada como  $\neg(\beta p \ \& \ \neg p)$ , é difícil ver como a crença em CSE\*\* poderia ser um caso de conhecimento. Todavia, conforme mostramos antes, um sujeito S que sabe  $\beta p \ \& \ p$ , também deve poder saber  $\neg(\beta p \ \& \ \neg p)$ . Se um sujeito S sabe que é verdadeira, a respeito de si mesmo, *eu não creio falsamente que p*, então ele deve estar em posição de saber, a respeito de si mesmo, que é verdade que *não é o caso que ambos: eu creio que p e é falso que p*. A questão relativa a formalização correta de *eu não creio falsamente que p* passa ao lado do problema que colocamos.

Até o momento, ao tratarmos da OSE, não fizemos qualquer referência aos métodos empregados na formação da crença. Contudo, os princípios de sensibilidade e aderência foram relativizados aos métodos na teoria de Nozick, e devemos averiguar se a consideração dos métodos indica uma resposta possível à objeção. Uma solução deste gênero é sugerida em Salerno (2010).

Existe um argumento que mostra que o recurso aos métodos de formação não melhora a situação do defensor do princípio de sensibilidade. Suponhamos que S acredita, em @, com base em um método m, que *não é o caso que ambos: eu acredito que p e é falso que p*. A aplicação do princípio de sensibilidade, com a devida referência aos métodos, ficaria assim:

$$(\beta p \ \& \ \neg p) \ \Box \rightarrow \ \neg \beta_m (\neg (\beta p \ \& \ \neg p)).$$

---

<sup>35</sup> Agradeço ao Luís Fernando Munaretti Rosa por ter apontado que eu não mostrei ser equivocada a formalização proposta por Becker e Salerno. Todavia, permanece o problema que apontei aqui: um sujeito saber verdade que  $(\beta p \ \& \ \neg p)$ , deve estar em posição de saber que  $\neg(\beta p \ \& \ \neg q)$ .

Examinemos agora o antecedente do condicional acima. Seja  $w$  o mundo mais próximo a  $@$  em que o antecedente do condicional é verdadeiro, isto é, o mundo mais próximo ao mundo  $@$  em que  $S$  acredita falsamente que  $p$ .

Existem três possibilidades a respeito do método  $m$  no mundo  $w$ <sup>36</sup>:

1.  $m$  recomenda a crença de que  $(\beta p \ \& \ \neg p)$ ;
2.  $m$  recomenda a suspensão de juízo a respeito de  $(\beta p \ \& \ \neg p)$ ;
3.  $m$  recomenda  $\neg (\beta p \ \& \ \neg p)$ ;

Ora, nem a primeira alternativa, nem a segunda, são compatíveis com um sujeito minimamente coerente, o que supomos ser o caso em  $@$  e, por consequência, em  $w$ . Se 1 fosse o caso, então sujeito  $S$  acreditaria que  $p$  e acreditaria que acredita falsamente que  $p$ , um tipo de irracionalidade que julgamos incompatível com a situação. Se 2 fosse o caso, então o sujeito  $S$  consideraria se é o caso que acredita falsamente e ainda assim teria dúvidas a respeito de  $\beta p \ \& \ \neg p$ , o que também seria considerado um caso grave de irracionalidade. Por exclusão, resta apenas a terceira opção. Contudo, se esta última for verdadeira, então é verdade que:

$$(\beta p \ \& \ \neg p) \ \Box \rightarrow \beta_m \neg (\beta p \ \& \ \neg p).$$

Ora, esta última afirma que se  $S$  acreditasse falsamente que  $p$ , então  $S$  acreditaria com base em  $m$  que não acredita falsamente  $p$ . Em outros termos, se a terceira alternativa é correta, então a crença de que *não é o caso que ambos: eu creio que  $p$  e  $p$  é falso  $p$*  não é sensível à verdade. Como as três alternativas esgotam as possibilidades relevantes, temos a condição da sensibilidade relativizada aos métodos não é satisfeita para qualquer crença da forma *não é o caso que ambos: eu creio que  $p$  e  $p$  é falso  $p$* . Contudo, se  $S$  não pode saber, a respeito de si mesmo, que *não é o caso que ambos: eu creio que  $p$  e  $p$  é falso  $p$* , então  $S$  não deve poder saber que não crê falsamente que  $p$ . Conforme verificamos, a relativização aos métodos de formação de crença não salva a teoria de Nozick da OSE. Como a sensibilidade deve ser relativizada aos métodos de formação de crença, sob pena de incorrer em outros problemas, te-

---

<sup>36</sup> Rodrigo Borges me sugeriu que existe uma quarta alternativa: o método  $m$  não recomenda nenhuma atitude a respeito de  $(\beta p \ \& \ \neg p)$ . O problema é que o sujeito  $S$  acredita em  $@$  que  $\neg (\beta p \ \& \ \neg p)$ . Presumidamente, o mesmo método  $m$  também deveria recomendar alguma atitude doxástico a respeito de  $(\beta p \ \& \ \neg p)$  em  $w$ .

mos que o princípio de sensibilidade se mostra incompatível com o conhecimento de que não se crê falsamente.

Note-se que a objeção OSE não afirma que sempre, ou mesmo regularmente, um sujeito sabe que não crê falsamente, ou que toda vez que crê verdadeiramente, ele sabe que não crê falsamente. Tudo o que a objeção afirma é que, ao menos em algumas circunstâncias, um sujeito S sabe que não crê falsamente. A alternativa é uma forma de ceticismo quanto ao conhecimento de segunda-ordem. Naturalmente, a força da objeção depende de julgamentos independentes acerca da legitimidade dessa forma de ceticismo.

### 3.1.7 Disjunções

Conforme salientou Vogel (2012), a teoria do rastreamento parece ter dificuldades em explicar conhecimento de disjunções. Conforme veremos por meio de alguns exemplos discutidos logo a seguir, a teoria parece ser muito restritiva nestes casos, e fornece juízos e classificações incorretas.

#### 3.1.7.1 As testemunhas corroborantes

Consideremos o seguinte exemplo proposto por Jonathan Vogel:

*Corroboração.* Hercule Poirot está interrogando duas testemunhas, que foram mantidas isoladas uma da outra. Ambas testemunhas dizem que o bandido era homem. Seja  $W_1$  = 'testemunha 1 está dizendo a verdade' e seja  $W_2$  = 'testemunha 2 está dizendo a verdade'. Poirot sabe tanto  $W_1$  como  $W_2$ , porque as testemunhas corroboram umas às outras (VOGEL, 2012, p.125).

Note-se, antes de prosseguir, que nós supomos que se  $W_1$  fosse falso, então a testemunha 1 discordaria da testemunha 2, e que se  $W_2$  fosse falso, então a testemunha 2 discordaria da testemunha 1. A razão de aceitarmos estes condicionais subjuntivos é que as testemunhas do exemplo são independentes entre si, e a verdade ou falsidade do que uma testemunha diria não afetaria a verdade/falsidade do que a outra diria. Assim, não podemos assumir que se  $W_1$  fosse falso, então  $W_2$  seria falso, ou que se  $W_2$  fosse falso, então  $W_1$  seria falso.

Assim, S sabe que  $W_1$  e S sabe que  $W_2$ . Se  $W_1$  fosse falsa, então S não acreditaria que  $W_1$ , pois S notaria que  $W_1$  e  $W_2$  estão em discordância, e não acreditaria que  $W_1$ . O mesmo parece acontecer no que diz respeito a  $W_2$ . Se  $W_2$  fosse falsa, então S notaria que  $W_1$  e  $W_2$  discordam entre si, e por isso mesmo não acreditaria em  $W_2$ .

Até aqui, a teoria parece fornecer os veredictos corretos, concordando com nossos juízos intuitivos sobre a ausência ou presença de conhecimento. Pois bem, consideremos agora a crença de que ao menos uma das testemunhas está dizendo a verdade, isto é, que  $W_1 \vee W_2$ .

O princípio da sensibilidade aplicado ao caso diz que:

$$\neg (W_1 \vee W_2) \square \rightarrow \neg \beta (W_1 \vee W_2).$$

Ora, temos aqui mais uma equivalência lógica:

$$\neg (W_1 \vee W_2) \longleftrightarrow \neg W_1 \ \& \ \neg W_2.$$

Substituindo as equivalências lógicas no antecedente do condicional subjuntivo:

$$(\neg W_1 \ \& \ \neg W_2) \square \rightarrow \neg \beta (W_1 \vee W_2).$$

Este último condicional, equivalente ao princípio da sensibilidade aplicado à crença na disjunção, parece claramente falso. Se ambas as testemunhas não estivessem dizendo a verdade, então elas concordariam entre si, e Poirot acreditaria que ao menos uma delas está dizendo a verdade. Todavia, tal resultado é muito surpreendente: como eu posso saber que uma das testemunhas diz a verdade sem poder saber que, ao menos, *alguma* delas está dizendo a verdade? Esta parece uma combinação implausível e paradoxal de conhecimento e ignorância. De acordo com a teoria do rastreamento, S não sabe que  $W_1 \vee W_2$ ; contudo, a teoria deveria dizer que ele *sabe* que  $W_1 \vee W_2$ .

### 3.1.8 O reator

Consideremos agora o seguinte contraexemplo, semelhante ao caso das testemunhas corroborantes, também proposto por Vogel (2012). Homer trabalha em uma usina nuclear, e uma das suas funções é monitorar o funcionamento do reator nuclear. Um painel contém três

lâmpadas que funcionam como indicadores do estado do reator:  $L_1$ ,  $L_2$  e  $L_3$ . A função dos sinalizadores é indicar se o reator está em funcionamento: uma lâmpada ligada é um sinal de que o reator está funcionando; uma lâmpada desligada indica que o reator não está funcionando. Se apenas um dos sinalizadores está desligado, Homer observa o funcionamento dos outros dois, que estão ligados, e forma a crença de acordo com os dois últimos.

Suponhamos que o reator está em funcionamento. Por hipótese,  $L_2$  indica falsamente que o reator da usina não está funcionando, e  $L_1$  e  $L_3$  indicam corretamente que o reator está funcionando. Seja  $P = L_1$  indica corretamente o estado do reator da usina. Segundo Vogel Homer sabe que  $P$ . Todavia, a crença de que  $P$  não é sensível. Se fosse falso que  $P$ , então  $L_1$  concordaria com  $L_2$ , o que levaria Homer a acreditar (falsamente) que  $L_1$  está indicando corretamente o estado do reator. A teoria do rastreamento parece entrar em conflito com uma instância óbvia de conhecimento.

Apesar do contraexemplo de Vogel ter alguma plausibilidade, penso que pode haver uma resistência legítima por um defensor do princípio de sensibilidade. É tão claro assim que Homer sabe que  $P$ ? Para mostrar que não é evidente ou claro que Homer sabe que  $P$ , façamos algumas observações.

Sejam  $P_1 = L_1$  indica que o reator está ligado,  $P_2 = L_2$  indica que o reator está desligado e  $P_3 = L_3$  indica que o reator está ligado. Uma vez que  $P_1$  é o caso,  $P_2$  é contraevidência para  $P$ . Um sujeito epistêmico não deve simplesmente dar de ombros a uma contraevidência simplesmente porque ela é em menor número que as evidências favoráveis (são dois indicadores favoráveis contra um desfavorável, por assim dizer). Pode haver algum conflito de intuições aqui. De qualquer forma, há espaço para pensar que Homer não sabe  $P$ , o que gera dúvidas quanto à eficácia do contraexemplo por Vogel<sup>37</sup>.

Consideremos um caso análogo, que deveria ser julgado (de maneira também análoga) ao exemplo do reator. Três pessoas independentes entre si testemunham um assassinato.  $T_1$  e  $T_2$  dizem que  $P = o$  assassino é um homem, e  $T_3$  diz que o assassino não é um homem, isto é,  $\neg P$ . Suponhamos que o assassino seja realmente um homem. Um sujeito  $S$  sabendo dos três testemunhos, conclui que a testemunha 1 está dizendo a verdade porque  $T_1$  e  $T_2$  concordam entre si.

---

<sup>37</sup> Note-se como o contraexemplo do reator é dissimilar ao contraexemplo das testemunhas corroborantes. No caso das testemunhas concordantes, não havia qualquer contraevidência.

Da mesma forma, a teoria de Nozick fornece o mesmo resultado: S não sabe que T1 está dizendo a verdade, pois se isto fosse falso, T1 concordaria com T3 e S acreditaria falsamente que T1 está dizendo a verdade. Todavia, assim como no caso do reator, penso que a teoria de Nozick fornece o veredicto correto: deveríamos dizer que S não sabe que T1 está dizendo a verdade, visto que S está simplesmente *ignorando* uma contraevidência importante para o caso. Mesmo que a teoria do rastreamento venha a dispensar a justificação para a posse do conhecimento, algo que Nozick não faz explicitamente, a teoria não pode atribuir conhecimento a um sujeito que ignora uma contraevidência importante<sup>38</sup>.

### 3.1.9 Os contraexemplos de Peter Klein

Em Klein (1987) também encontramos alguns importantes contraexemplos ao princípio de sensibilidade. Klein inicia sua crítica à teoria de Nozick apresentando um exemplo, ao estilo de Gettier, onde a teoria parece fornecer o veredicto ou resultado correto. Eu acredito que o tanque de gasolina do meu carro está vazio porque o leitor de combustível assim o indica. Por hipótese, o tanque está vazio. Contudo, sem que eu o saiba, o fio que conecta o tanque ao leitor de combustível foi cortado por uma pedra. Quando o leitor não está conectado ao tanque, o leitor indica que o tanque está vazio, não importa quanto combustível esteja contido no tanque. Uma vez que a pedra cortou a conexão entre tanque e leitor, o leitor diria que o tanque está vazio, não importa quanto combustível estivesse contido no tanque.

O exemplo acima é um típico caso Gettier. O leitor de combustível aponta corretamente a quantidade de combustível no tanque por pura sorte ou coincidência; se fosse falso que o tanque está vazio, o sujeito continuaria acreditando que o tanque está vazio, dado que o leitor continuaria indicando (falsamente) que o tanque está vazio.

Consideremos agora o seguinte exemplo, também sugerido por Klein. Eu acredito no mundo *w* que o tanque de gasolina do meu carro está vazio porque o leitor de combustível assim o indica. Por hipótese, a minha crença é verdadeira, i.e., o tanque está vazio. O leitor de combustível funciona normalmente, e não há qualquer objeto de qualquer tipo interferindo na

---

<sup>38</sup> Nosso juízo acerca do conhecimento/ignorância de S de que T1 está dizendo a verdade mudaria se houvesse um número astronomicamente alto de testemunhas em favor de P, e apenas uma única testemunha isolada contrária, que talvez possa ser legitimamente ignorada? Para quem julgar que esta poderia ser uma resposta plausível, gostaria de lembrar que se este fosse o caso, então a crença não seria mais sensível.

conexão entre o tanque e o leitor. Quando o leitor não está conectado ao tanque, o leitor indica que o tanque está vazio, não importa quanto combustível esteja contido no tanque. A mesma pedra, que cortou a ligação entre o leitor e o tanque no primeiro exemplo, tem uma trajetória, e não acerta o fio. A pedra *quase* corta a conexão entre tanque e leitor, porém não o faz.

Digamos que exista um mundo possível  $w'$ , muito próximo ao mundo atual  $w$ , tal que o tanque está com mais combustível, e o peso gerado pelo combustível adicional faz com que a trajetória pedra desta vez corte a conexão entre leitor e tanque, o que, por sua vez, faz com que o leitor indique que o tanque está vazio, quando o tanque não está. No mundo  $w'$ , o sujeito  $S$  acredita falsamente que o leitor está vazio.

Segundo Klein, diríamos que eu sei que o tanque está vazio, embora a teoria do rastreamento pareça indicar o contrário, uma vez que existe um mundo  $w'$  muito próximo ao mundo atual  $w$ , tal que em  $w'$  eu acredito falsamente que o tanque está vazio por meio do mesmo método. O princípio de sensibilidade não é satisfeito. A teoria de Nozick é muito restritiva, e classifica um caso de conhecimento como casos de ignorância.

Klein sugere que podemos abstrair de TANQUE II uma certa receita ou procedimento para produzir contraexemplos a teoria do rastreamento. Klein descreve o esquema da seguinte forma: considere um processo de dois estágios que resulta na crença de que  $q$  e faça a saída (*output*) do primeiro estágio idêntica à entrada (*input*) do segundo estágio que resulta na crença que  $q$ . Agora modifique o mundo atual, de tal sorte que existe causa uma "esperando" para produzir a mesma saída (*output*) do primeiro estágio mesmo quando a entrada é diferente. Assim, ainda que a entrada do primeiro estágio fosse diferente,  $S$  ainda acreditaria que  $q$  (KLEIN, 1987, p.272)

Vamos aplicar a explicação de Klein ao segundo exemplo do tanque de combustível. No mundo atual  $w$  temos a seguinte cadeia causal: o tanque está vazio  $\mapsto$  o leitor de combustível indica que o tanque está vazio  $\mapsto$   $S$  forma a crença verdadeira de que o tanque está vazio. No mundo  $w'$ , temos a seguinte cadeia causal: o tanque está com mais gasolina  $\mapsto$  o combustível excedente contido no tanque altera a trajetória de uma pedra  $\mapsto$  pedra corta a conexão entre tanque e leitor  $\mapsto$  o leitor indica que o tanque está vazio  $\mapsto$   $S$  acredita falsamente que o tanque está vazio.

Vamos dizer que  $S$  e  $S'$  são, respectivamente, os sujeitos nos mundos  $w$  e  $w'$ . Klein diz que  $S$  e  $S'$  utilizam o mesmo método, de acordo com o critério nozickiano de individuação de métodos. Internamente, e para todos os aspectos relevantes,  $S$  e  $S'$  são indistinguíveis: as diferenças entre  $S$  e  $S'$  não residem nos seus estados ou no conteúdo dos seus respectivos estados mentais.

O sucesso do contraexemplo de Klein depende da atribuição correta de conhecimento no caso. Klein parece confiante de que, no segundo exemplo do tanque de combustível, o sujeito S sabe que o tanque está vazio. Admitindo tal pressuposto, parece que temos um contraexemplo de considerável força.

### 3.1.10 Margens de erro

Williamson (2000) argumentou muito plausivelmente que o conhecimento requer certa margem de erro, e que isto cria imensas dificuldades para o princípio de sensibilidade. Consideremos o caso a seguir. Um sujeito S observa, em condições ambientais e cognitivas apropriadas (boa luminosidade, distância, aparelho visual funcionando bem, etc.) uma pessoa R e julga que R possui mais de 1,50m de altura. O sujeito R, por hipótese, é uma pessoa muito alta e possui 2 metros de altura. Mesmo que S não discrimine entre pessoas com 1,50m e pessoas com um pouco menos de 1,50m (ou digamos 1,49m), diríamos que S, ao observar R, sabe que R possui mais de 1,50 m.

O fato de S frequentemente pensar que pessoas com altura levemente inferior a 1,50m possuem mais de 1,50m não parece impedir que S saiba, diante de uma pessoa R de 2m, que R possui mais de 1,50m. Deve ser possível que S saiba disso, e qualquer teoria que negasse tal possibilidade deve ser considerada profundamente contraintuitiva.

Seja  $p = R$  possui mais de 1,50m de altura. Se S sabe que  $p$ , então  $\neg p \square \rightarrow \neg \beta p$ . Isto é verdadeiro em @ quando:

(14)  $\neg \beta p$  no  $\neg p$ -mundo mais próximo a @.

Ora, o  $\neg p$ -mundo mais próximos a @ é um mundos onde R não tem mais que 1,50m, mais precisamente, mundos onde R possui *exatamente* 1,50m. Todavia, S não discrimina, por hipótese, entre um sujeito que possui 1,50m e um sujeito que possui mais de 1,50m, e se estivesse de um sujeito R de *exatamente* 1,50m de altura, ele pensaria que R possui mais de 1,50m. Assim, acreditaria que  $p$  no  $\neg p$ -mundo mais próximos a @, que significa que o prin-

cípio (14) não é satisfeito<sup>39</sup>. Como consequência, S não sabe que p, o que constitui um problema sério para a teoria do rastreamento.

### 3.1.10.1 A resposta de Becker ao problema da margem de erro

Becker (2007) propõe que uma vez considerado o método envolvido no caso, verifica-se que S sabe que P, contrariamente ao que sugere Williamson. Recordemos, a este respeito, que Nozick propôs que os métodos são distinguidos internamente, de acordo as experiências que o sujeito tem (pelo menos, quando a crença envolve algum tipo de experiência). Becker afirma que a adoção do critério nozickiano de individuação de métodos poderia produzir uma resposta satisfatória a objeção de Williamson.

Seja  $w_1$  o mundo em que o sujeito S observa R (uma pessoa de 2m), e forma a crença de que p. O método que S utiliza para formar a crença de que P em  $w_1$  é o método  $m_1$ . Ora, os  $\neg p$ -mundos em que R não possui mais que 1,50m não são mundos onde S utiliza o mesmo método  $m_1$ , visto que se R não tivesse mais que 1,50m, então R teria uma aparência *muito* diferente da aparência que tem em  $w_1$ , e S utilizaria outros métodos. O fato de que S se engana a respeito de p nos mundos em que R tem um pouco menos de 1,50m não é suficiente para impugnar o conhecimento que S tem de que p no mundo  $w_1$  utilizando o método  $m_1$ .

Não penso que a solução proposta Becker ao problema é satisfatória, e julgo que, no final das contas, quando lida mais atentamente, ela apenas confirma a objeção original de Williamson. Vamos supor que uma certa aparência perceptual seja mantida nos mundos contrafactuais relevantes para a satisfação de (14) (no caso, o mundo  $w_2$ ). Fazendo as devidas alterações, o condicional da sensibilidade é satisfeito quando:

(15)  $\neg \beta p$  com base no método  $m_1$  no mundo  $w_2$ .

Todavia, conforme vimos, o mundo  $w_2$  deve ser um mundo tal que o sujeito R não possui mais de 1,50m (afinal,  $w_2$  é um mundo em que a proposição alvo é falsa). Ao mesmo tempo, Becker alega que aparência de R para S deve ser fixada (pelo menos, em linhas gerais). Combinando as duas teses, temos  $w_2$  deve ser um mundo tal que R tem a mesma aparência para S (a aparência de uma pessoa de aproximadamente 2 metros de altura) e, ao mes-

---

mo tempo, R não possui mais de 1,50m. Todavia, se este é o caso, então  $w_2$  deve ser um mundo tal que S acredita falsamente que R possui mais de 1,50m, pois ele baseia sua crença na aparência de R (incluída no método  $m_1$ , segundo Becker). O princípio (15) não é satisfeito. Portanto, a crença de S não é sensível à verdade, e S não sabe que p.

Além do problema acima, a solução Becker está submetida às mesmas dificuldades que afligem a concepção nozickiana demasiadamente restrita e interna de individuação dos métodos. Na medida em que a proposta de Becker depende de tal maneira de entender os métodos, ele só é viável na medida em que também o é a proposta de Nozick. Conforme veremos mais tarde (capítulo 3.4), a especificação puramente interna dos métodos defendida por Nozick é inviável.

## 3.2 A TEORIA É MUITO PERMISSIVA

Até o momento, tratamos apenas de contraexemplos que procuram mostrar que a teoria do rastreamento é muito restritiva. A seguir, discutiremos as objeções que procuram mostrar que a teoria é muito permissiva, isto é, que classifica certos casos de ignorância como casos de conhecimento. As objeções do tipo permissivo tendem a exigir mais do objetor que as objeções do tipo de restritivo. A razão é simples, ainda que frequentemente ignorada na literatura sobre o assunto. Para que as objeções do primeiro tipo sejam cogentes é suficiente mostrar que uma das condições da teoria (geralmente o princípio de sensibilidade) não é satisfeita em casos de conhecimento; para que as objeções do segundo tipo sejam cogentes é necessário que **todas** as condições da teoria sejam satisfeitas em algum caso de ignorância. Não devemos nos esquecer deste detalhe, especialmente quando uma alternativa possível ao defensor da teoria do rastreamento é fazer uso de alguma condição adicional, quando uma das condições da teoria parece muito facilmente satisfeita em casos de ignorância (é o que veremos no problema das verdades necessárias, conforme trataremos mais tarde).

### 3.2.1 Conjunções

Anteriormente, vimos que a teoria do rastreamento parece enfrentar problemas relativos ao conhecimento de disjunções. Quando analisamos como a teoria trata do conhecimento de disjunções, verificamos a teoria se mostrou muito restritiva, classificando casos óbvios de conhecimento como casos de ignorância. No entanto, a teoria parece ser muito permissiva no que diz respeito ao conhecimento de conjunções, classificando casos óbvios de ignorância como casos de conhecimento. No que diz respeito às disjunções, a teoria torna o conhecimento muito difícil; no que diz respeito às conjunções, a teoria torna o conhecimento muito fácil.

Suponhamos que S acredita na conjunção  $p \ \& \ q$ . A crença de S é sensível? O princípio de sensibilidade aplicado a  $p \ \& \ q$  diz:

**Sensibilidade de conjunções** -  $\neg(p \ \& \ q) \ \Box \rightarrow \neg\beta_m(p \ \& \ q)$

Substituindo  $\neg(p \ \& \ q)$  pela logicamente equivalente  $\neg p \vee \neg q$ , temos que:

**Sensibilidade de conjunções\*** -  $(\neg p \vee \neg q) \ \Box \rightarrow \neg\beta_m(p \ \& \ q)$

Assim, o condicional da sensibilidade para conjunções é um condicional contrafactual com antecedente disjuntivo. Segundo Vogel, contrafactuais disjuntivos (contrafactuais com antecedentes disjuntivos) são bastante problemáticos.

Vogel aponta dois tratamentos possíveis para contrafactuais disjuntivos: resolução forte e resolução fraca. De acordo com a resolução fraca:

**Resolução fraca** -  $((p \vee q) \ \Box \rightarrow r) \ \longleftrightarrow \text{o } (p \vee q)\text{-mundo mais próximo é um } r\text{-mundo.}$

De acordo com a resolução forte:

**Resolução forte** -  $((p \vee q) \ \Box \rightarrow r) \ \longleftrightarrow ((p \ \Box \rightarrow r) \ \& \ (q \ \Box \rightarrow r)).$

Vejam os primeiros a resolução fraca. Suponhamos que S acredita que  $p \ \& \ q$  em  $w$  e que a crença de S de que  $p$  é sensível a verdade de  $p$ , isto é, que  $\neg p \ \Box \rightarrow \neg\beta p$ . Diremos que  $w'$  é o  $\neg p$ - mundo mais próximo a  $w$ . Em  $w'$ , S não acredita que  $p$ . Presumidamente, em  $w'$  também S não acredita que  $p \ \& \ q$ .

Vamos supor ainda que o  $\neg p$ -mundo mais próximo a  $w$  é mais próximo a  $w$  que *qualquer*  $\neg q$ -mundo. Por conseguinte, o  $\neg(p \vee q)$ -mundo mais próximo a  $w$  é o  $\neg p$ -mundo mais

próximo, a saber, o próprio  $w$ . Assim, para que a crença de  $S$  de que  $p \vee q$  seja sensível a verdade é suficiente que  $S$  não acredite que  $p \ \& \ q$  em  $w$ . Ora, conforme vimos,  $S$  não acredita que  $p \ \& \ q$  em  $w$ . Portanto, a crença de  $S$  de que  $p \ \& \ q$  é sensível à verdade.

Consideremos o seguinte caso, que é muito semelhante a um exemplo proposto por Vogel (2012). Um sujeito  $S$  acredita que está tomando um chá gelado. A sua crença de que está tomando um chá gelado satisfaz os princípios de sensibilidade e aderência.  $S$  também acredita, a respeito de si mesmo, que: (16) estou tomando um chá gelado e o planeta terra está aquecendo.

O mundo mais próximo ao mundo atual em que a conjunção (16) é falsa é o mundo  $W_1$  em que o primeiro conjunto (*estou tomando um chá gelado*) é falso e o segundo (*a terra está aquecendo*) é verdadeiro. No mundo  $W_1$  eu não acredito que é verdadeira a conjunção porque eu não acredito no primeiro conjunto. Parece que eu simplesmente posso saber que é verdadeira a conjunção simplesmente por saber ou rastrear o primeiro conjunto, desde que os mundos em que o segundo conjunto é falso sejam muito mais distantes do mundo atual que os mundos em que o primeiro conjunto é falso. O exemplo acima pode ser generalizado: o segundo conjunto pode ser qualquer proposição desde que satisfeitas as condições que apresentamos anteriormente.

Aparentemente, a resolução fraca não parece ser apropriada para uma explicação do conhecimento de conjunções por meio da teoria do rastreamento. Uma vez que a resolução fraca não parece adequada, devemos agora examinar os méritos da resolução forte. Vejamos o seguinte caso proposto por Vogel. Asgar, um fã de ópera, assiste a um concerto de Maria Callas. Asgar reconhece o tom de voz e a aparência de Callas, e sabe que  $C$ : *Maria Callas está cantando*. Presumivelmente, se Maria Callas não estivesse cantando, então Asgar não acreditaria que Maria Callas está cantando.

Agora, consideremos a seguinte proposição:

**I** - uma impostora extremamente parecida com Maria Callas está cantando no lugar de Maria Callas.

A resolução forte, quando combinada com a teoria do rastreamento, diz que  $S$  não sabe que  $C \ \& \ \neg I$ . Para saber que  $C \ \& \ \neg I$  Asgar deveria rastrear tanto  $C$  como  $\neg I$ . Todavia, ele não rastreia  $\neg I$ : se fosse falso que  $\neg I$  (i.e., se fosse verdadeiro que  $I$ ), então Asgar ainda assim acreditaria que  $I$ , uma vez que Asgar não seria capaz de distinguir Maria Callas de uma im-

postora extremamente parecida com ela em todos os aspectos relevantes. O problema, segundo Vogel, é que  $C$  e  $(C \ \& \ \neg I)$  são logicamente equivalentes. Todavia, isto implica que um sujeito  $S$  pode saber que  $p$ , saber que  $p \leftrightarrow q$ , e ainda assim não saber que  $q$  (mesmo quando realiza a inferência correspondente). Vogel considera isto um sério problema para a teoria de Nozick, e que inviabiliza a resolução forte: "ou seja, Asgar não sabe  $(C \ \& \ \neg I)$  ainda que ele saiba a logicamente equivalente  $C$ . Algo deu terrivelmente errado" (VOGEL, 2012, p.130).

### 3.2.2 Klein sobre rastreamento e verdades necessárias

Em artigo recente, Klein (NO PRELO) afirmou que a teoria do rastreamento possui dificuldades para lidar com certas crenças em verdades necessárias. Embora Klein não afirme que a teoria do rastreamento trivialize o conhecimento de verdades necessárias, tal que toda crença em verdades necessárias é um caso de conhecimento, ele apresenta um contraexemplo no qual a teoria claramente produz o resultado incorreto.

Suponhamos um sujeito  $S$  e uma proposição necessária  $p$ . O sujeito  $S$  acredita que  $p$  porque ele é um sujeito extremamente crédulo, tal que ele acredita em toda e qualquer proposição que ele considera. Como  $p$  é uma verdade necessária, ou bem a condição da sensibilidade pode ser ignorada, ou bem ela é trivialmente satisfeita. Resta, então, averiguar se a condição de aderência é satisfeita.

O problema é que a aderência é claramente satisfeita no caso proposto por Klein. O método de  $S$  para formar uma crença a respeito de uma proposição qualquer envolve considerar a proposição alvo e formar a crença de que ela é verdadeira simplesmente porque a entendeu em sua mente. Um método deste tipo - e não parece haver qualquer razão pela qual isto não poderia ser um método de formação de crenças - automaticamente satisfaz a condição de aderência para qualquer proposição necessária  $p$ . Como sensibilidade e aderência são satisfeitas, o sujeito  $S$  sabe que  $p$ . Ora, isto é absurdo. Ninguém adquire conhecimento desta forma.

### 3.2.3 Conhecimento de segunda ordem e verdades necessárias

Agora irei introduzir a seguinte variação do problema das verdades necessárias para a teoria do rastreamento. Um sujeito S acredita que p sem saber que p, porque a sua crença de que p não rastreia a verdade. Temos assim que  $\neg Kp$ . Ao mesmo tempo, p é uma verdade necessária. Suponhamos ainda que S acredita que acredita verdadeiramente que p, isto é, acredita que  $(p \ \& \ \beta p)$ .

Será o caso que  $K(p \ \& \ \beta p)$ ? Para averiguar se S tem conhecimento, de acordo com a teoria de Nozick, devemos verificar se a crença na proposição alvo é aderente e sensível. Em outros termos, devemos considerar as seguintes proposições:

$$(17) \ \neg(p \ \& \ \beta p) \ \Box \rightarrow \neg \beta(p \ \& \ \beta p);$$

$$(18) \ (p \ \& \ \beta p) \ \Box \rightarrow \beta(p \ \& \ \beta p).$$

Vejamos primeiro a proposição  $\neg(p \ \& \ \beta p) \ \Box \rightarrow \neg \beta(p \ \& \ \beta p)$ . A proposição  $\neg(p \ \& \ \beta p) \ \Box \rightarrow \neg \beta(p \ \& \ \beta p)$  é verdadeira em um mundo qualquer w somente se é verdade que  $\neg \beta(p \ \& \ \beta p)$  no  $\neg(p \ \& \ \beta p)$ -mundo mais próximo a w. Seja w' o  $\neg(p \ \& \ \beta p)$ -mundo mais próximo a w. Ora, como é o caso  $\Box p$ , i.e., p é verdadeira em todos os mundos possíveis, então deve ser o caso que p em w'. Se p e  $\neg(p \ \& \ \beta p)$  são verdadeiros em w', então deve ser o caso que  $\neg \beta p$  em w'. Ora, assim como S não acredita que p em w', também deve ser o caso que S não acredita que  $(p \ \& \ \beta p)$  em w'. Em outros termos, é verdade que  $\neg \beta(p \ \& \ \beta p)$  em w'. Portanto, é verdade que  $\neg(p \ \& \ \beta p) \ \Box \rightarrow \neg \beta(p \ \& \ \beta p)$ .

Vejamos agora a proposição  $(p \ \& \ \beta p) \ \Box \rightarrow \beta(p \ \& \ \beta p)$ . Qualquer que seja o método envolvido no caso em questão, ele exige que o sujeito considere ou entretenha a proposição  $(p \ \& \ \beta p)$ . Ora, parece claro que se S considerasse ou entretivesse a proposição  $(p \ \& \ \beta p)$ , a fim de decidir se ela é verdadeira ou falsa, e S acreditasse que p, então S acreditaria que  $(p \ \& \ \beta p)$ . Portanto, a proposição  $(p \ \& \ \beta p) \ \Box \rightarrow \beta(p \ \& \ \beta p)$ , instância do princípio de aderência, é satisfeita. Ora, vimos que (17) e (18) são satisfeitas. Ao mesmo tempo, o sujeito também tem uma crença verdadeira. Portanto, é o caso que

$$(19) \ K(p \ \& \ \beta p).$$

Ora, S não sabe que p. Combinando isto com  $\neg K_s p$  e (19), temos que

(20)  $\neg K p \ \& \ K(p \ \& \ \beta p)$ .

Ora, a proposição logo acima é extremamente implausível. Se alguém sabe que acredita verdadeiramente que p, então deve saber que é verdade que p. Todavia, é precisamente este resultado que é implicado pela teoria de Nozick. Além do mais, não apenas S sabe que acredita verdadeiramente que p, e não sabe que p, como também o conhecimento de S de que acredita verdadeiramente que p não depende do método utilizado. O resultado é absurdo.

A este respeito, é interessante fazer a conexão com outra objeção de que tratamos anteriormente: a objeção do conhecimento de segunda ordem. Existe um certo tipo de conhecimento de segunda ordem, que ordinariamente supomos ser possível, cuja existência é incompatível com o princípio de sensibilidade. Entretanto, quando a proposição de primeira ordem é ela própria necessária, a objeção do conhecimento de segunda ordem torna-se uma objeção permissiva: o princípio de sensibilidade torna muito fácil e abundante um certo tipo de conhecimento de segunda ordem, contrariamente ao que deveria ser o caso.

### 3.3 VIOLAÇÕES IMPLAUSÍVEIS DO FECHO EPISTÊMICO

Além das objeções permissivas e restritivas, existem outras objeções importantes que devemos examinar. Estas não alegam, propriamente falando, que a teoria do rastreamento é muito permissiva ou restritiva. Elas pretendem mostrar que a teoria do rastreamento produz certas combinações implausíveis de conhecimentos e ignorância, que violam princípios de fecho epistêmico, tais como FIV. A principal objeção deste tipo é o contraexemplo do celeiro vermelho, do qual trataremos a seguir.

#### 3.3.1 O celeiro vermelho

Uma objeção bastante poderosa à teoria foi apresentada em Kripke (2011, p.186). Trata-se de uma variação do caso original dos celeiros. Vejamos do que trata o caso.

Barney está viajando de carro por uma estrada rural. Ao longo da viagem, ele observa diversas estruturas que parecem típicos celeiros. Ele observa um destes objetos e forma a crença perceptual correspondente:

(p) aquele objeto é um celeiro.

Vamos supor que a crença seja verdadeira, e que Barney observou um celeiro naquele local. Além do mais, as suas faculdades sensoriais funcionam perfeitamente, e não há qualquer razão para Barney suspeitar de que não existe um celeiro no local avistado. Todavia, existe uma complicação. A estrada rural, pela qual Barney se desloca, está repleta de determinadas estruturas que, embora parecessem celeiros mesmo para o mais competente observador, são apenas réplicas ou fachadas (digamos que os falsos celeiros não contêm nada na parte de trás, não visível a partir da estrada), que visavam enganar os viajantes. Entre as tantas fachadas e réplicas muito bem construídas, esconde-se um único celeiro genuíno, precisamente o que foi avistado por ele.

Barney simplesmente *teve sorte* de ter avistado, em meio tantas réplicas enganadoras que o conduziram ao erro, o único celeiro genuíno em toda a região. Nós diríamos que Barney não sabe que o objeto avistado é um celeiro. Agora acrescentemos mais algumas características ao ambiente. Por alguma razão qualquer, todas as réplicas são de uma mesma cor (digamos, marrom), o celeiro avistado por ele é de uma outra cor, por exemplo, vermelho, e ninguém construiria uma réplica de celeiro vermelho (não importa o motivo). Além de formar a crença perceptual de que está diante de um celeiro, Barney também nota que o objeto possui uma cor específica, qual seja vermelho, e forma crença correspondente de que:

(q) aquele objeto é um celeiro vermelho

A crença de que q parece satisfazer o princípio de sensibilidade. Se não existisse um celeiro vermelho, então Barney não acreditaria que existe um celeiro vermelho. Com efeito, existe um mundo muito próximo em que Barney acreditaria falsamente que está diante de um celeiro (porque está diante de uma réplica), mas este é um mundo onde a réplica *não é vermelha*. Também parece que a condição de aderência é satisfeita. Nos mundos próximos em que é verdade que existe um celeiro vermelho, ele acredita que existe um celeiro vermelho. Supondo mantida a inexistência de réplicas de celeiro vermelho em mundos próximos, tal como é estipulado pelo exemplo, é difícil ver como a teoria poderia dar o veredicto de que Barney não sabe que q.

O que há de interessante no exemplo é que, de acordo com a teoria do rastreamento, S sabe que existe um celeiro vermelho e simultaneamente não sabe que existe um celeiro *simpliciter*. Tal situação constitui uma violação do princípio de fecho epistêmico. Todavia, ao

contrário de outras situações nas quais a falha do princípio tem alguma plausibilidade intuitiva, o que se verifica nesta situação é precisamente o contrário. Mais especificamente, o exemplo do celeiro vermelho parece mostrar uma falha implausível, e sem nenhum apoio intuitivo ou independente da teoria do rastreamento, do princípio de fecho do conhecimento sob eliminação da conjunção. Para verificar como isto se dá, basta notar que a sentença  $q$  pode ser parafraseada como uma conjunção:

(**r**) aquele objeto é um celeiro e aquele objeto é vermelho.

Ora, a sentença  $r$  tem a forma de uma conjunção, e o primeiro conjunto de  $r$  é justamente a sentença  $p$ . Ora,  $p$  é dedutível de  $r$  pela regra de eliminação da conjunção. Por conseguinte, temos uma falha do princípio de fecho do conhecimento pela eliminação da conjunção.

É interessante pensar o contraexemplo do celeiro vermelho à luz do problema geral que as crenças em conjunções colocam para a teoria do rastreamento. Com efeito, o caso do celeiro vermelho, quando analisado apropriadamente, mostra que a teoria do rastreamento é permissiva no que diz respeito ao conhecimento de conjunções. Se Barney não tem ciência da conexão entre a vermelhidão e autenticidade dos celeiros (i.e., que todos objetos que parecem celeiros são réplicas, com exceção dos vermelhos), então Barney *não* sabe que a objeto avisado é um celeiro vermelho, apesar da teoria do rastreamento dizer o contrário. Desta forma, o caso do celeiro vermelho serve de evidência para a crítica mais geral que havíamos proposto à teoria do rastreamento, qual seja, que ela é muito permissiva no que diz respeito ao conhecimento de conjunções. O resultado disso tudo é que pode ser *muito mais fácil* conhecer uma proposição conjuntiva do que conhecer qualquer um dos seus conjuntos separadamente.

### 3.3.2 Contrafactuais disjuntivos e o caso do celeiro vermelho

O contraexemplo do celeiro vermelho pretende mostrar que a teoria do rastreamento diz que sujeito  $S$  sabe que é verdadeira uma proposição conjuntiva, mas não sabe que são verdadeiros ambos os conjuntos. Curiosamente, esse exemplo é semelhante a outro contraexemplo de que tratamos anteriormente: o exemplo do chá gelado (seção 3.2.1). O resultado, entretanto, é distinto, e a resolução forte nos diz que o sujeito cognoscente possui conhecimento no contraexemplo do celeiro vermelho.

Consideremos a seguinte proposição conjuntiva, onde  $p = \textit{aquele objeto é um celeiro}$  e  $q = \textit{aquele objeto é vermelho}$ :  $p \ \& \ q$ . *Aquele* refere-se ao objeto que é apontado por Barney. A crença em  $p \ \& \ q$  é sensível somente se  $\neg(p \ \& \ q) \ \Box \rightarrow \neg \beta(p \ \& \ q)$ . Substituindo  $\neg(p \ \& \ q)$  por  $\neg p \vee \neg q$ , temos que  $p \ \& \ q$  é sensível somente se:

$$(\neg p \vee \neg q) \ \Box \rightarrow \neg \beta(p \ \& \ q).$$

O condicional acima é um contrafactual disjuntivo (contrafactual em que o antecedente é uma disjunção). Agora, vamos considerar como uma das resoluções dos contrafactuais disjuntivos: forte e fraca. A resolução forte dos contrafactuais disjuntivos, aplicada ao exemplo, diz que  $(\neg p \vee \neg q) \ \Box \rightarrow \neg \beta(p \ \& \ q)$  somente se  $(\neg p \ \Box \rightarrow \neg \beta(p \ \& \ q)) \ \& \ (\neg q \ \Box \rightarrow \neg \beta(p \ \& \ q))$ . Ora, é verdadeiro que  $\neg p \ \Box \rightarrow \neg \beta(p \ \& \ q)$ . Se o objeto não fosse um celeiro, então Barney não teria as impressões sensoriais de vermelho e, presumidamente, não acreditaria que objeto é um celeiro e é vermelho. Também é verdadeiro que  $\neg q \ \Box \rightarrow \neg \beta(p \ \& \ q)$ . Se o objeto não fosse vermelho, então Barney não teria aparências perceptuais de um celeiro vermelho (pois não existem réplicas de celeiros vermelhos). Consequentemente, se adotarmos a resolução forte para os contrafactuais disjuntivos, deveríamos aceitar que Barney sabe que objeto é um celeiro vermelho, ao mesmo tempo em que não sabe que objeto é um celeiro.

Assim, a resolução forte dos contrafactuais disjuntivos, quando acoplada com a teoria do rastreamento, produz o resultado intuitivamente incorreto no caso do celeiro vermelho. Por estas razões adicionais, a resolução forte dos contrafactuais disjuntivos não pode adequadamente responder ao problema das conjunções.

A resolução fraca dos contrafactuais disjuntivos não é uma alternativa promissora. De acordo com a resolução fraca, Barney sabe que é verdade que  $p \ \& \ q$  somente se  $\neg \beta(p \ \& \ q)$  no  $(\neg p \vee \neg q)$ -mundo mais próximo. Seja  $w$  o  $(\neg p \vee \neg q)$ -mundo mais próximo ao mundo atual. O  $(\neg p \vee \neg q)$ -mundo mais próximo é um mundo em que se dá uma e apenas uma das seguintes três alternativas:  $\neg p \ \& \ q$ ,  $p \ \& \ \neg q$  ou  $\neg p \ \& \ \neg q$ . Se a primeira alternativa estiver correta, então o objeto apontado é vermelho, e não é um celeiro. Como não existem réplicas de celeiros vermelhos em mundos próximos, o objeto para o qual ele aponta não deve ser um objeto que tomaria por um celeiro. Presumivelmente, Barney também não acredita que o objeto é um celeiro, e não acreditaria que o objeto é um celeiro e é vermelho. Se a segunda alternativa estiver correta, então o objeto é um celeiro genuíno não-vermelho em  $w$ . Presumivelmente,

Barney não acredita em  $w$  que está diante um celeiro vermelho, dado que ele não está diante de um objeto que parece vermelho. Por fim, vamos supor que  $w$  é um  $(\neg p \ \& \ \neg q)$ -mundo. Se este é o caso, então Barney está diante de um objeto que não é vermelho, e não é um celeiro, e Barney não pensaria que o objeto é um celeiro e vermelho, uma vez que presumivelmente o objeto não parece vermelho. Como as três alternativas esgotam as possibilidades quanto ao  $(\neg p \vee \neg q)$ -mundo mais próximo, e Barney não acredita que está vendo um celeiro vermelho em todas elas, segue que Barney não acredita que está diante de um celeiro vermelho no  $(\neg p \vee \neg q)$ -mundo mais próximo. Portanto, a resolução fraca dos contrafactuais disjuntivos, quando aplicada ao caso do celeiro vermelho, implica que a crença de Barney é sensível à verdade.

### 3.3.3 A resposta de Adams e Clarke à objeção de Kripke

Embora o contraexemplo do celeiro vermelho tenha uma força considerável, nem todos autores pensam que ele seja uma objeção definitiva e irrefutável. Adams e Clarke (2005) pensam que é possível responder às objeções de Kripke mantendo-se fiel à teoria de Nozick. Nas palavras dos autores, “este parece ser um caso em que alguém sabe que lá está um celeiro vermelho, mas que não sabe que lá está um celeiro, apenas se a pessoa perde de vista o método pelo qual se adquire conhecimento” (Adams & Clarke, 2005, p.214).

Pois bem, como a consideração do método permite dissipar a objeção levantada por Kripke? Segundo os autores, é o mesmo método que o sujeito utiliza para formar ambas as crenças, de que existe um celeiro e de que existe um celeiro vermelho. Novamente, segundo os autores:

Peg [o nome da personagem na descrição do exemplo fornecida por Adams e Clarke] está usando a aparência de celeiro vermelho para detectar tanto a informação de que existe algo vermelho, como de que existe um celeiro. Dado que não podem existir réplicas enganosas de celeiros vermelhos (de acordo com a estipulação de Kripke), a aparência de celeiro vermelho transmite fielmente ambos os tipos de informação. Uma vez que Peg está utilizando a aparência de celeiro vermelho para formar a crença de que existe um celeiro vermelho, ela satisfaz a condição de rastreamento de Nozick, de que se não houvesse um celeiro vermelho, ela não acreditaria que existe um celeiro vermelho. E se não houvesse um celeiro vermelho presente, e Peg fosse acreditar que existe um celeiro, então ela não usaria a mesma aparência de celeiro vermelho. (ADAMS & CLARKE, 2005. p. 214-215).

Há três alegações fundamentais realizadas pelos autores. Em primeiro lugar, eles afirmam que o sujeito do exemplo (no caso, Peg) utiliza o *mesmo* método tanto para a formação

da crença de que existe um celeiro, como para a formação da crença de que existe um celeiro vermelho. Em segundo lugar, eles afirmam que o método utilizado consiste em basear-se na aparência/impressão de um celeiro vermelho. Em terceiro lugar, eles afirmam que uma vez que o método é o mesmo para as duas crenças, ambas satisfazem o princípio de sensibilidade, e não há, portanto, qualquer violação do princípio de fecho dedutivo.

A ideia dos autores tem alguma plausibilidade, e, uma vez aceita a caracterização do método proposta por eles, a solução parece funcionar. Entretanto, existem problemas com a forma como os autores entendem o método. Mais especificamente, penso que eles não identificam de maneira correta o método utilizado na formação da crença, por parte de Peg, de que está vendo um celeiro.

Um método de formação de crenças é a maneira ou processo que o sujeito utiliza para formar a crença na proposição alvo. O conceito de *método* tem a função de explicar a maneira como a crença veio a ser formada. Isto significa que o método utilizado deve ter um *papel causal* na produção da crença, e a sua aplicação/utilização deve explicar a formação da crença. Se o método utilizado deve explicar a produção da crença do sujeito, e esta é a sua função essencial, então a caracterização ou individuação do método não deve conter elementos ou fatores supérfluos de um ponto de vista explicativo e/ou causal. Por conseguinte, uma descrição do método que não contém atributos supérfluos deve ser preferida, *ceteris paribus*, a uma descrição que contém elementos supérfluos de um ponto de vista explicativo e/ou causal.

Em vista do que foi dito acima, podemos nos perguntar: a maneira como Adams e Clarke descrevem o método utilizado por Peg contém atributos supérfluos de um ponto de vista explicativo e/ou causal? Ora, a resposta é decididamente positiva. Afinal de contas, a descrição do método feita por Adams e Clarke inclui *a cor vermelha* do celeiro. Todavia, a cor vermelha do celeiro não parece ter um papel causal e/ou explicativo na produção da crença de Peg. Ela não formou a crença de que está vendo um celeiro especificamente em razão da cor avermelhada do objeto<sup>40</sup>. Tudo isto indica que a maneira como Adams e Clarke descreveram o método de ambas as crenças de Peg é incorreta. Por conseguinte, eles não produziram uma resposta satisfatória ao contraexemplo de Kripke.

---

<sup>40</sup> Luís Rosa me sugeriu que, como ambas as crenças de Peg surgiram como resultado da mesma experiência, elas são geradas pelo mesmo método. O problema é que identificar o método com a experiência individual do sujeito do sujeito, com precisamente com as características que possui, remete ao problema da caracterização interna e estrita do método, que vimos ser insustentável.

### 3.3.4 O cachorro *daschund*

Suponhamos que S está diante de um cachorro da raça *daschund*, e forma a crença de que:

**(p)** Estou diante de um *daschund*<sup>41</sup>.

Digamos ainda que todas as quatro condições da teoria de Nozick são satisfeitas, e que, portanto, S sabe que p. O sujeito S não confundiria, em situações muito semelhantes, qualquer outro animal com um *daschund*. Consideremos agora a seguinte proposição:

**(q)** Estou diante de um cachorro.

Suponhamos agora que, se não existisse um *daschund*, existiria um outro animal, digamos uma ovelha, que seria confundida por S com um cachorro. Digamos que existem muitas ovelhas pela região, e que S confunde regularmente ovelhas com cachorros (talvez ele pense que ovelhas são cachorros da raça *sheepdog*). Nestas circunstâncias, parece que S não sabe que existe um cachorro de acordo com a teoria de Nozick, uma vez que sua crença não é sensível à verdade de que p: se fosse falso que p, então S continuaria acreditando (falsamente) que p.

Ora, isto parece paradoxal. Como uma pessoa pode saber que está diante de um *daschund*, saber que *daschunds* são um tipo de cachorro, e ainda assim não saber que está diante de um cachorro (ao menos, quando faz a inferência correspondente)? Novamente, parece que temos mais uma falha contraintuitiva do princípio de fecho.

A resposta de Adams e Clarke também poderia ser adaptada ao contraexemplo sugerido por Goldman. S observa um cachorro *daschund*, tem experiências ou aparências perceptuais típicas de um cachorro *daschund*, e forma a crenças de que o objeto que está observando é um cachorro e é um *daschund*. O método utilizado para formar ambas as crenças, de acordo com Adams e Clarke, é precisamente o mesmo (digamos, o método  $m_1$ ). Se S não estivesse diante de um *daschund*, ele não acreditaria que está diante de um *daschund* com base no método  $m_1$ . Similarmente, se S não estivesse diante de um cachorro, então S não acreditaria que está diante de um cachorro com base no método  $m_1$ . Uma vez admitida que ambas as crenças

---

<sup>41</sup> Este exemplo, em formulação ligeiramente diferente, aparece em Goldman (1976).

são formadas com base no mesmo método, verifica-se que o caso *daschund* não é razão definitiva para rejeitar a teoria do rastreamento.

### 3.4 PROBLEMAS COM A CARACTERIZAÇÃO DO MÉTODO

Nozick, por boas razões, pensou ser necessário incluir o conceito de *método* em sua análise do conhecimento. Embora bem motivada, a sua manobra incorpora as dificuldades pertinentes ao uso e entendimento do conceito. A seguir, mostraremos a inviabilidade da concepção de método proposta por Nozick. Ele caracterizou o método internamente, que esclarecemos mediante o conceito de *duplos mentais*. Isto garante que os cenários céticos não envolvem a alteração do método empregado em relação aos cenários normais.

Todavia, a concepção internista do conceito de método coloca Nozick em sérias dificuldades. A concepção interna e estrita do método gera resultados profundamente insatisfatórios quando combinada com os princípios de sensibilidade e aderência. Tal constatação sugere que se deve procurar uma caracterização alternativa dos métodos, se quisermos ainda preservar os *insights* fundamentais da teoria de Nozick.

Um sujeito S acredita, com base em suas experiências perceptuais, e vem a saber que: (p) aqui está a minha mão. Por hipótese, crença de S de que p é sensível à verdade de que p: nos mundos próximos em que é falso que p, o sujeito S não acredita que p. Naturalmente, existem mundos distantes em que S acredita falsamente que p, por exemplo, em certos cenários céticos, mas estes mundos são irrelevantes para a satisfação das condições do conhecimento de S de que P segundo a teoria de Nozick. A condição de aderência, por hipótese, também é satisfeita.

Digamos que o método utilizado seja o método m, e que sensibilidade quando aplicada a crença de S em P seja:

$$\neg p \square \rightarrow \neg \beta_m p.$$

Utilizando da semântica dos mundos possíveis para os contrafactuais, podemos verificar que é verdade que um sujeito S sabe que p no mundo atual @ quando é verdade que no

mundo possível mais próximo em que é falso que  $p$  e  $S$  utiliza o método  $m$ ,  $S$  não acredita que  $p$ .

Seja  $w$  o mundo possível mais próximo a  $@$  em que é falso que  $p$ . Para que sensibilidade seja satisfeita,  $S$  não deve acreditar que  $p$  em  $W$ . Conforme exige Nozick, devemos manter fixo o método  $m$  em  $W$  e  $@$ . De acordo com a caracterização nozickiana do método, o sujeito  $S$  tem as mesmas aparências perceptuais relevantes em  $@$  e  $W$ . Como as experiências relevantes seriam as mesmas em  $W$  e em  $@$ , eles seriam “duplicatas internas”, e ele acredita com base nessas experiências, o sujeito  $S$  acreditaria que  $p$  em  $W$ . Portanto,  $S$  *não* sabe que  $p$  em  $@$ .

O que vimos a respeito deste caso individual, pode ser facilmente transposto para praticamente qualquer outra crença. Se a experiência perceptual relevante para a formação da crença de que  $P$  em um mundo  $@$  com base em um método  $m$  está atrelada à identidade do próprio método  $m$ , então qualquer outro mundo  $W$  em que  $S$  utilize o método  $m$  deve ser um mundo tal que ele tem as mesmas experiências perceptuais relevantes para a formação da crença de que  $p$  que  $S$  tem em  $@$ . Nestas circunstâncias, é muito difícil entender como  $S$  poderia *não* crer que  $p$  em qualquer mundo em que ele utilize o método  $m$ . Isto é evidentemente um problema, pois mundos em que  $S$  tem alguma experiência diferente, ou existe alguma diferença interna importante, devem ser levados em conta.

### 3.5 CONSIDERAÇÕES FINAIS SOBRE O PRINCÍPIO DE SENSIBILIDADE

É hora de recapitular o que vimos até aqui acerca da teoria do rastreamento, e que conclusão pudemos tirar disto tudo. As objeções apresentadas à teoria do rastreamento não são iguais em força e plausibilidade. Enquanto algumas podem ser respondidas com os recursos contidos na própria teoria de Nozick, existem outras que são consideravelmente resistentes, e, mesmo após exame e reflexão cuidadosa, mantêm a sua força original.

Penso que existem duas objeções particularmente fortes, entre aquelas que foram aqui tratadas, que merecem algum destaque. Em primeiro lugar, existe o problema do conhecimento de crenças indutivas, que sugere ser o princípio de sensibilidade incompatível com uma vasta classe de casos de conhecimento. Em segundo lugar, existe o problema das crenças de segunda ordem a respeito de proposições necessárias de primeira ordem, e que mostra como a teoria do rastreamento nos compromete com resultados absurdos e contraintuitivos.

Como espero ter demonstrado, existem muitos problemas que afetam a teoria do rastreamento, e o balanço das razões sugere que a teoria *em sua forma original* não deve ser aceita. Contudo, isto não deve significar que nenhuma teoria semelhante, ou inspirada nela, não possa ser razoavelmente aceita. Há algo de atraente no *insight* de que o conhecimento deve levar em conta as respostas contrafactuais do sujeito a estados de coisas diferentes, algum tipo de sensibilidade (para utilizar o termo de Nozick) à verdade. A seguir, veremos uma tentativa de desenvolver as ideias de Nozick, ainda que em uma nova direção.

#### 4 A CONDIÇÃO DE SEGURANÇA

Agora irei examinar a epistemologia modal concorrente, e principal sucessora, da proposta nozickiana: a denominada *teoria da segurança*. Embora os seus diferentes defensores discordem acerca dos detalhes de como a condição deve ser formulada, a ideia básica é relativamente clara: um sujeito S sabe que  $p$  somente se S não facilmente acreditaria falsamente que  $p$ . Uma crença de que  $p$  é segura para um sujeito cognoscente S quando, ao formar a crença de que  $p$ , o sujeito S não corre risco de crer falsamente.

Vejam os alguns exemplos a fim de tornar mais concreta esta ideia. Comparemos duas situações. Eu vejo meu amigo e penso comigo: *lá está meu amigo João*. Meus órgãos sensoriais funcionam normalmente, eu sei como meu amigo se parece, estou vendo em boas condições de iluminação e distância e assim por diante. Minha crença é segura: dadas as circunstâncias em que me encontro, seria muito difícil de me enganar a este respeito. Eu não corro risco elevado de me enganar. Muitas coisas deveriam ser diferentes para que eu acreditasse que meu amigo João está lá e ainda assim João não estivesse lá.

Consideremos agora uma outra situação levemente diferente. Eu vejo meu amigo e formo a crença: *lá está meu amigo João*. Como no exemplo anterior, estou vendo o próprio João e meus órgãos sensoriais funcionam apropriadamente. Todavia, existe um diferencial. João possui um irmão gêmeo Carlos, cuja existência eu desconheço, e que circula pelo mesmo ambiente e região. Minha crença de que estou vendo João não é segura: existe um mundo ou circunstância muito próximo onde eu formo a crença de que estou vendo meu amigo João e é falso que João está lá, a saber, o mundo próximo em que estou vendo Carlos. Ao formar a crença do modo como eu a formei, eu não estou imune ao risco: eu poderia *facilmente* me enganar.

Vejam os agora em mais detalhes como os principais teóricos da segurança pensaram a esse respeito, e como Duncan Pritchard e Ernest Sosa procuraram desenvolver esta ideia central.

#### 4.1 A CONDIÇÃO DE SEGURANÇA EM SOSA

Ernest Sosa formulou o princípio de segurança da seguinte maneira<sup>42</sup>:

Digamos que a crença de S é "segura" sse: S acreditaria que p somente se fosse o caso que p. Alternativamente, a crença de que de p é segura sse: S não acreditaria que p se não fosse o caso que p; ou, melhor ainda, sse: como matéria de fato, ainda que não seja uma matéria de necessidade estrita, não facilmente S acreditaria que p sem ser o caso de que p. (SOSA, 1999a, p.142).

Em outra passagem, Sosa escreve:

Em resumo: a crença B (p) de S é sensível sse  $\neg p \square \rightarrow \neg B(p)$ , enquanto a crença de S é segura sse  $B(p) \square \rightarrow p$ . Estes não são equivalentes, uma vez que condicionais subjuntivos não se contrapõem. (SOSA, 1999a, p. 142).

Conforme vemos, Sosa formula o princípio de segurança de diferentes maneiras, e considera todas elas equivalentes. Assim, podemos considerar que um sujeito S acredita seguramente que p, segundo Sosa, se qualquer uma das seguintes condições são satisfeitas:

- S não facilmente acreditaria falsamente que p;
- S acreditaria p somente se fosse o caso de que p;
- S não acreditaria que p sem que fosse o caso de que p;
- se S acreditasse que p, então seria o caso de que p<sup>43</sup>.

Apesar das semelhanças entre sensibilidade e segurança, elas devem ser claramente distinguidas. A sensibilidade diz respeito ao que um sujeito S não acreditaria, se a proposição-alvo fosse falsa. A segurança diz respeito ao que seria verdadeiro se o sujeito S acreditasse na proposição-alvo.

Assim como foi necessário acrescentar algumas emendas à teoria da sensibilidade, a segurança também foi devidamente modificada. Para verificar a necessidade da alteração, consideremos o seguinte exemplo produzido por Sosa:

---

<sup>42</sup> Sosa abandonou o princípio de segurança em favor de uma teoria da virtude epistêmica, o que pode ser observado, por exemplo, em Sosa (2007). Obviamente, quando tratamos das ideias de Sosa nesta tese, nos referimos a fase do seu pensamento em que ainda defendia a a condição de segurança.

<sup>43</sup> Evidentemente, se pensarmos que subjuntivos com antecedente e consequente verdadeiros são trivialmente verdadeiros, então esta formulação está equivocada. Infelizmente, Sosa não fornece maiores explicações sobre o assunto.

Se eu vejo um pássaro voando, apenas por estar olhando pela janela, o que eu facilmente poderia não ter feito, eu coloco-me em condição na qual o pássaro estaria voando se eu acreditasse nisso? Obviamente não; muito facilmente eu poderia ter olhado em outra direção. Mesmo quando estou olhando na direção correta, desde que eu facilmente poderia ter olhado para outra direção; em tal caso, eu sei, enquanto olho pela janela, algo que não rastreio: que um pássaro está voando. Rastreamento simples [*outright tracking*] não pode ser uma condição necessária para o conhecimento. (SOSA, 2002, p.266).

Em função de contraexemplos deste gênero, Sosa sugere que a condição de segurança deve ser relativizada. A ideia inicial é de que nossas fontes de conhecimento, quando em operação, geram ou informam certos resultados (*deliverances*), os quais consistem em certas proposições. Assim, por exemplo, minha visão "diz" que estou em frente ao meu computador, a minha audição "diz" que alguém bateu na porta, a minha memória "diz" que eu tomei café da manhã, e assim por diante. Novamente, nas palavras de Sosa:

Uma indicação de que  $\langle p \rangle$  para um sujeito S é um "dizer" que p testemunhado por S. Diferentes fontes correlacionam-se com diferentes modos em que se pode dizer que p. Alguém pode dizer literalmente, é claro, pessoalmente ou por escrito, e S pode ter visto ou escutado. Se nós podemos acreditar em nossos olhos e ouvidos, ademais, é por que eles nos contam coisas. Nós experienciamos visualmente ou auditivamente como se p. Normalmente, nós aceitamos as indicações de nossos sentidos, a menos que nós detectemos algo estranho. (SOSA, 2002, p.268)

Quando uma certa fonte de conhecimento produziria o resultado de que p somente se p Sosa diz que a fonte de que p é um *indicador* de que p. Uma crença de que p é segura somente se baseada em um indicador de que p.

A ideia de uma fonte de conhecimento parece desempenhar o mesmo papel que o conceito de métodos desempenha na teoria nozickiana. Sosa parece sugerir, em determinadas passagens, à semelhança de Nozick, que uma indicação envolve algum tipo de aparência ou apresentação, seja ela sensorial ou intelectual. Assim, por exemplo, ele afirma:

Uma 'indicação', uma instância do processo de indicar, é um estado de coisas no qual algo parece para o sujeito ser de uma certa forma- um estado de coisas mais ou menos complexo que pode ter conteúdo tanto intelectual como sensorial. (SOSA, 2002, p.275).

Assim, segurança, em uma versão mais elaborada, pode ser definida da seguinte forma: um sujeito S crê seguramente que p com base em uma indicação I se e somente se I indicaria que p somente se p fosse o caso. Se preferirmos uma formulação alternativa equivalente, podemos dizer que S acredita seguramente que p com base na indicação I somente se S não facilmente acreditaria falsamente que p com base em I. Conforme Sosa parece sugerir, se trata

de uma espécie de "necessidade local", e uma forma mitigada de infalibilismo: nas circunstâncias em que o sujeito S se encontra, ele *não poderia* crer falsamente que p<sup>44</sup>.

#### 4.1.1 Segurança e conhecimento indutivo

Uma das vantagens da segurança *vis-a-vis* a sensibilidade parece ser a maior facilidade com que segurança trata do conhecimento indutivo. Retomemos um dos exemplos que tratamos anteriormente. Um sujeito S joga o lixo por um compartimento no seu edifício, e acredita que o lixo caiu na lixeira, conforme geralmente acontece. S acredita que o lixo caiu na lixeira em razão do fato de que, normalmente, o lixo é recolhido pela lixeira antes de cair no chão, e ela tem funcionado geralmente bem, sem apresentar defeitos. Supondo que o mecanismo da lixeira tem funcionado corretamente até então, e que não há nada de anormal ou atípico nas circunstâncias do sujeito, seria muito difícil que S acreditasse que P com base na indicação I, e ainda assim o lixo não ser recolhido. Afinal de contas, esta seria uma situação rara, incomum, um tanto *difícil* de acontecer nessas circunstâncias.

O princípio de sensibilidade exige que consideremos mundos ou circunstâncias em que o lixo não foi recolhido pela lixeira, e que examinemos então o que sujeito S acreditaria em tais circunstâncias. O problema é que estes mundos em que as crenças do sujeito seriam falsas podem ser muito diferentes - em aspectos importantes - do mundo em que se encontra atualmente o sujeito cognoscente. Para a satisfação da condição de segurança, não parece importar o que acontece, ou o que seria o caso, em mundos distantes. Os seus defensores consideram isto uma qualidade da sua teoria, e uma vantagem em relação a sensibilidade.

A condição de segurança, contrariamente à sensibilidade, não parece ter em princípio quaisquer dificuldades em explicar o conhecimento indutivo, tal como no exemplo de que tratamos aqui. Segundo Sosa, a crença de S de que o lixo foi recolhido pela lixeira com base na indicação I é segura. Seria muito difícil de acontecer que S acredite que p com base em um mecanismo que tem funcionado muito bem e com excepcional regularidade e *ainda assim* a crença de S ser falsa. Obviamente, coisas estranhas e incomuns são logicamente possíveis, e o mecanismo não é infalível; contudo, nada disto parece impedir a aquisição do conhecimento por parte do sujeito cognoscente.

---

<sup>44</sup> Neste ponto, Sosa defende um infalibilismo muito semelhante ao que é defendido por Dretske (1971), na denominada *teoria das razões conclusivas*.

Conforme vimos anteriormente, a dificuldade em acomodar o conhecimento indutivo é um dos principais problemas para o princípio de sensibilidade. Se a condição de segurança consegue explicar os mesmos problemas que a sensibilidade, mantendo algumas ideias ou *insights* centrais da teoria nozickiana, e ainda assim não parece ter grande dificuldade em acomodar o conhecimento indutivo, a teoria de segurança é preferível à sensibilidade, e parece ser uma alternativa promissora. Vejamos agora como a sensibilidade pode tratar de outros problemas que afetam a sensibilidade.

#### 4.1.2 O fecho epistêmico

Ao contrário da sensibilidade, a condição de segurança pretende preservar o fecho epistêmico. Consideremos, por exemplo, uma proposição qualquer ordinária do senso comum, a proposição  $p = \textit{estou agora escrevendo em meu computador}$ , e uma alternativa cética, a proposição  $hc = \textit{estou sonhando que estou escrevendo em meu computador, e não estou escrevendo em meu computador}$ . Vamos supor que a crença de S de que  $p$  satisfaz a condição de segurança, tal como entendida por Sosa: S acreditaria que  $p$  com base em  $I$ , nas circunstâncias em que se encontra, somente se fosse o caso que  $p$ . Vamos supor ainda que S se encontra em um mundo aproximadamente tal como o senso comum o descreve (não existem grandes ilusões massivas, a sua percepção funciona correta e confiavelmente, e assim por diante). Ao mesmo tempo, S também sabe que se está escrevendo em seu computador, então ele não está sonhando que está escrevendo em seu computador, quando não existe nenhum computador. Será o caso de que S também sabe que não é o caso de que  $\neg hc$ ? Esta última crença, de que a hipótese cética é falsa, satisfaz a condição de segurança?

Ao contrário do que acontece com o princípio de sensibilidade, a crença na negação das hipóteses céticas pode satisfazer o princípio de segurança. Para explicar como e porquê a condição de segurança evita a falha do fecho epistêmico, é pertinente recorrer ao vocabulário dos mundos possíveis. Na imensa maioria dos mundos próximos ao mundo atual em que S acredita que  $\neg hc$ , é verdadeiro que  $\neg hc$ . Apenas em mundos muito diferentes do nosso, radicalmente diferentes, S acreditaria falsamente que  $\neg hc$ , pois os mundos em que é falso que  $\neg hc$  são mundos em que é verdade que  $hc$ , e isto só aconteceria, por hipótese, em mundos *muito* distantes. Considerando as circunstâncias em que S se encontra, não aconteceria de S acreditar que  $\neg hc$  sem que seja o caso que  $\neg hc$ .

Conforme vemos, segurança não exige que a crença do sujeito de que p seja verdadeira em circunstâncias radicalmente diferentes das atuais. O que um sujeito acreditaria se ele estivesse em um cenário cético, quando ele se encontra atualmente em um cenário não cético e muito distante do mundo atual, é irrelevante para a satisfação do princípio. Isto evita, segundo seus defensores, o tipo de violação do fecho epistêmico aceita por autores como Fred Dretske e Robert Nozick.

#### 4.1.3 Segurança e a resposta ao ceticismo

Vimos na seção anterior que Sosa alega que segurança preserva o fecho epistêmico, e que, em circunstâncias normais, um sujeito S pode saber que são verdadeiras certas proposições do senso comum e ao mesmo tempo saber que são falsas as hipóteses céticas, cujas negações são implicadas por aquelas. Agora é o momento de explicar em mais detalhes a alegação dos teóricos da segurança de que sujeitos em condições normais podem *saber* que são falsas as hipóteses céticas, e como tal alegação relaciona-se diretamente com o problema do ceticismo.

Ao tratarmos do problema cético, o apresentamos sob forma de um paradoxo, e elencamos três respostas: rejeitar o fecho epistêmico, alegar que temos conhecimento da falsidade das hipóteses céticas ou simplesmente aceitar o ceticismo. Nozick e Dretske optaram pela primeira alternativa. Sosa, e alguns outros teóricos da segurança, defenderam a segunda: nós sabemos, ou pelo menos estamos em posição de saber, que são falsas as hipóteses céticas.

Naturalmente, para responder ao paradoxo do ceticismo, não é suficiente apontar ou indicar uma teoria, no caso a teoria da segurança, que mostra ser possível sabermos serem falsas as hipóteses céticas. É desejável também explicar a *aparente plausibilidade* da premissa cética, e indicar como e por que o cético está errado ao alegar que não podemos saber que são falsas as hipóteses céticas. Conforme o próprio autor coloca, "lembre o atraente requisito de que um tratamento completo do paradoxo deve explicar para nós por que o componente rejeitado pelo paradoxo parece tão plausível" (Sosa, 1999a, p.147).

Pois bem, como Sosa explica a aparente plausibilidade do argumento cético? A sua resposta é particularmente interessante: a atração ou força da premissa cética - de que não sabemos serem falsas as hipóteses céticas - é derivada de uma confusão, de uma *ilusão cogniti-*

va. Nós confundimos o princípio de sensibilidade com o princípio superficialmente similar, o princípio de segurança. Conforme ele diz:

c. Segurança e sensibilidade, sendo mutuamente contrapositivas, são facilmente confundidas, de tal sorte que é fácil confundir o requisito de segurança (para o conhecimento e a sua atribuição correta) com o requisito de sensibilidade. É fácil ignorar que subjuntivos não contrapõe.

d. Aqueles que acham as considerações do cético e sua distintiva premissa plausíveis com base na sensibilidade podem estar confundindo sensibilidade com segurança, e podem nesta base considerar como corretas as afirmações da premissa do cético. Afinal de contas, o requisito de segurança é bem apoiado pela sorte de considerações aduzidas geralmente pela oposição a Moore com base na sensibilidade. Sendo sensibilidade tão similar à segurança, tão fácil de confundir, não é surpresa que alguém consideraria a sensibilidade tão plausível, o suficiente para induzir alguém em erro e a avaliar como corretas as afirmações dessa premissa (SOSA, 1999a, p.148).

A resposta de Sosa ao paradoxo do ceticismo pretende combinar diferentes objetivos de maneira harmônica<sup>45</sup>. Ela preserva a possibilidade do conhecimento de senso comum, ao mesmo tempo que explica a atração da argumentação cética e mantém o fecho epistêmico.

## 4.2 A CONDIÇÃO DE SEGURANÇA EM PRITCHARD

Vimos na seção anterior a condição de segurança tal como Sosa a entende, e algumas razões que o levaram a sugerir tal princípio. Duncan Pritchard desenvolveu uma ideia semelhante. Como as ideias de Pritchard a este respeito são um pouco mais complicadas, devemos parar para examiná-las.

A proposta de Pritchard enquadra-se em um projeto filosófico mais amplo, denominado *epistemologia anti-sorte*. Pritchard se propõe a tomar o conceito de sorte como uma espé-

---

<sup>45</sup> Existe um problema bastante sério que afeta a proposta de solução fornecida por Sosa, e que me foi sugerido por Rodrigo Borges: na melhor das hipóteses, Sosa mostra que a possibilidade do cenário cético não impugna, por si só, o nosso conhecimento. Todavia, Sosa não responde a outra pergunta, que é justamente a questão essencial: como sabemos que nosso mundo *atual* não é um cenário cético? Diante de tal questionamento, penso que uma resposta coerente com o que Sosa escreve deveria ser uma resposta do estilo mooreano, a saber, alegando que a negação das hipóteses céticas segue de várias outras coisas que sabemos (por exemplo, certas alegações do senso comum sobre fatos ordinários). Não penso que esta seja uma resposta satisfatória, dado que não fornece nenhum argumento *independente* das alegações do senso comum em favor da posição anti-cética.

cie de âncora, da qual podemos derivar condições necessárias para o conhecimento (não coincidentemente, a condição de segurança) e iluminar diferentes problemas epistemológicos.

Pritchard (2012a) distingue entre dois tipos de epistemologia anti-sorte. A epistemologia anti-sorte *modesta* afirma que existe uma condição anti-sorte, que é necessária para a posse de conhecimento, e procura derivar uma condição necessária, porém não suficiente, para o conhecimento a partir do conceito de *sorte*. A epistemologia anti-sorte *robusta* afirma que a condição anti-sorte, conjuntamente com a crença e a verdade, é necessária e suficiente para a posse de conhecimento. Conforme veremos mais tarde, Pritchard é partidário da epistemologia anti-sorte modesta.

O projeto da epistemologia anti-sorte de Pritchard é desenvolvido em três etapas:

- Análise do conceito de sorte;
- Análise do conceito de sorte epistêmica;
- Especificação de uma condição anti-sorte para o conhecimento.

A seguir, iremos explicar como Pritchard aborda o conceito de *sorte*, e posteriormente a derivação de uma cláusula anti-sorte para o conhecimento.

#### 4.2.1 O Conceito de sorte

Embora o conceito de *sorte* seja frequentemente empregado na literatura epistemológica pós-Gettier, pouco foi feito para explicar em que consiste o conceito propriamente dito. Diferentemente das teorias mais conhecidas e tradicionais, Pritchard oferece uma análise mais detalhada do conceito. Ele propõe uma análise do conceito em termos estritamente modais, o que permite, conforme veremos mais tarde, gerar uma cláusula anti-sorte modal.

Sorte, segundo Pritchard (2005), pode ser caracterizada pela combinação de dois princípios. O primeiro destes princípios é o seguinte:

**(L1)** se um evento E é um caso de sorte no mundo atual, então E não ocorre em uma ampla classe dos mundos mais próximos ao mundo atual, nas quais as condições iniciais relevantes são as mesmas.

Pritchard entende proximidade e mundos possíveis da mesma forma que a semântica dos contrafactuais normalmente a entende: proximidade é uma relação de similaridade. Mundos possíveis podem ser ordenados em relação a um determinado mundo possível sob avaliação. A ordem reflete o grau de proximidade em relação ao mundo de consideração: quanto mais similar, mais próximo ao mundo possível em questão.

O segundo princípio, que visa introduzir a noção de relevância para um agente, é o seguinte:

**(L2)** se um evento E é um caso de sorte, então E é significativo para um agente S em consideração (ou seria significativo, se S avaliasse os fatos relevantes)

Uma vez estabelecidos estes princípios, Pritchard passa a esclarecer por que ambos parecem dar conta de nossas atribuições de sorte em situações relevantes. Consideremos algumas situações paradigmáticas de atribuição de sorte: ganhar em uma loteria, encontrar um objeto valioso e escapar de um acidente (por exemplo, de avião ou de carro). Vejamos cada um dos exemplos, e depois como a explicação de Pritchard parece gerar as atribuições corretas. Um sujeito S compra um bilhete de loteria, e, para a sua surpresa e felicidade, descobre que seu bilhete foi sorteado. Como na maioria das loterias, supomos que a loteria era justa e os bilhetes em competição eram muitos. Intuitivamente, diríamos que S teve sorte neste caso (*muita sorte*, se quisermos ser ainda mais enfáticos).

Os princípios L1 e L2 geram precisamente o resultado que esperávamos nestas circunstâncias. A imensa maioria dos mundos próximos ao mundo atual em que as condições relevantes iniciais são mantidas constantes, S não seria sorteado. Existe uma certa periferia ou cinturão de mundos ao redor do mundo atual, e apenas em uma mísera fração destes mundos o bilhete de S foi sorteado. Na maioria destes, S comprou o bilhete e ele *não* foi sorteado. Obviamente, o princípio L2 também é satisfeito, em razão da vantagem pecuniária.

Pritchard também menciona o caso de um sujeito S que, sem ter planejado encontra um objeto de alto valor aquisitivo escondido em algum lugar. O sujeito S teve sorte em encontrar o objeto escondido. Em uma grande porção dos mundos próximos em que as circunstâncias iniciais são mantidas fixas, ele não encontra nenhum objeto. Isto é algo muito difícil de acontecer nas circunstâncias do exemplo (em que ele, sem qualquer intenção ou propósito, encontra um objeto escondido de alto valor).

Consideremos um terceiro exemplo, que é um dos casos mais comuns e emblemáticos de sorte, embora não seja tratado explicitamente por Pritchard. Um sujeito S compra uma passagem de avião. Todavia, em função do trânsito congestionado, não consegue chegar a tempo de embarcar. O avião que ele iria tomar sofre um acidente e, como é comum em casos deste tipo, mata todos os passageiros. Nós diríamos que o sujeito S escapou de um acidente por sorte. Assim como nos casos anteriores, os princípios L1 e L2 também são satisfeitos. O evento E, que podemos caracterizar como *embarcar no avião*, embora não ocorra no mundo atual, ocorre em uma classe ampla dos mundos próximos em que as condições relevantes iniciais são as mesmas. Afinal, na maioria dos mundos em que as condições iniciais relevantes são mantidas constantes, o sujeito S embarca no avião e perde sua vida. Ele *quase* perdeu a vida, nós diríamos, e este *quase* indica que ele muito facilmente poderia ter perdido a vida, mantidas as demais circunstâncias constantes (em um mundo muito próximo, ele teria perdido a vida). De forma semelhante, o princípio L2 também é satisfeito, pelo evidente interesse de S na preservação da sua vida.

Pela combinação de L1 e L2, temos a seguinte formulação do conceito de *sorte*:

**Sorte** - Um evento E ocorre por sorte no mundo atual se e somente se (i) E não ocorre na maioria dos mundos próximos ao mundo atual com as mesmas condições iniciais relevantes; (ii) a ocorrência de E tem alguma significação ou importância para algum sujeito S (ou teria importância para S, se ele viesse ter consciência de que E).

Feitos os esclarecimentos sobre a concepção modal do conceito de *sorte*, podemos agora tratar do problema crucial: a relação entre sorte e conhecimento.

#### 4.2.2 Sorte Epistêmica

Agora iremos discriminar diferentes tipos de sorte, a fim de identificar as variedades benignas e malignas de sorte. Devemos isolar a sorte epistêmica propriamente dita, que se manifesta quando S acredita verdadeiramente que P por sorte ou acidente, das seguintes outras variedades:

- Sorte evidencial: S tem sorte em possuir evidência;
- Sorte de conteúdo: É uma sorte que p;

- Sorte cognitiva: S tem sorte em obter o conhecimento de que p;
- Sorte de faculdade: S tem sorte em possuir a faculdade cognitiva que permite S saber que p.

As quatro variedades acima de sorte são benignas, de um ponto de vista epistêmico. Nenhuma delas impede ou cancela a aquisição de conhecimento. Seria absurdo dizer que eu não posso saber que não fui atropelado por um carro hoje à tarde, simplesmente porque tive sorte em não ser atropelado. Da mesma forma, um detetive pode ter sorte em encontrar a evidência que indica o assassino, ou eu posso ter sorte em obter o conhecimento de qual foi o bilhete premiado na loteria (a fim poder resgatar o prêmio em dinheiro). Muitas dos conhecimentos que temos envolvem diferentes formas de sorte benigna.

Pois bem, se nenhuma das formas de accidentalidade ou sorte que vimos acima é epistemicamente maligna, então que tipo de sorte ou acidente é incompatível com a posse de conhecimento? Ora, o tipo maligno de sorte consiste na sorte em obter uma crença verdadeira, dadas as condições iniciais relevantes da situação, o que Pritchard (2005, p.146) também denomina de *sorte verística*. Dadas as considerações que fizemos, podemos agora fazer uma apresentação preliminar do conceito de sorte epistêmica:

Condição anti-sorte (CAS) - a crença verdadeira de um sujeito S de que P em um mundo w é um caso de sorte epistêmica se e somente se a crença de S de que p é falsa em uma ampla classe de mundos possíveis próximos ao mundo w em que as condições iniciais relevantes para a formação da crença são as mesmas.

Conforme veremos a seguir, CAS sugere que a cláusula responsável pela exclusão da sorte verística deve ter a forma de uma condição modal.

#### 4.2.3 O princípio de segurança

Duncan Pritchard apresenta uma versão preliminar da condição de segurança (o princípio S1), e oferece sucessivos enfraquecimentos do princípios (os princípios S2 e S3), a fim de remediar certos problemas. Posteriormente, esta condição de segurança, suficientemente refinada, será combinada em uma epistemologia da virtude. Iremos primeiro introduzir a ver-

são inicial da segurança, com a respectiva motivação, e posteriormente as sucessivas modificações e alterações.

Pritchard inicialmente caracteriza segurança da seguinte forma:

A crença verdadeira de S é segura sse na maioria dos mundos possíveis próximos em que S continua a formar sua crença a respeito da proposição alvo da mesma forma que o faz no mundo atual, a sua crença continua verdadeira (PRITCHARD, 2008, p.32).

Vamos dizer que esta versão inicial da condição segurança é o princípio S1. Existe um paralelo entre a condição de segurança e o princípio anti-sorte, pois S1 e CAS são equivalentes, desde que a forma como S veio a crer que P seja identificada com as condições iniciais relevantes para a formação da crença. Por conseguinte, o princípio de segurança é justamente o princípio que identifica a condição anti-sorte, responsável pela geração do problema de Gettier. Se aceitamos que o conhecimento exclui a intervenção da sorte epistêmica, o que é um lugar comum na epistemologia, então devemos também aceitar o princípio de segurança.

#### 4.2.3.1 Refinamentos da condição de segurança

Em razão de diversos problemas encontrados para as primeiras versões da condição de segurança, Pritchard produziu sucessivas modificações e refinamentos da primeira versão.

Consideremos uma dificuldade enfrentada por S1: o problema da loteria. Para tratar do caso, vamos retomar duas situações distintas que apresentamos anteriormente: loteria<sub>1</sub> e loteria<sub>2</sub>. A crença de S de que PL é segura em loteria<sub>1</sub>, segundo S1. O motivo é que uma vez que as probabilidades são massivamente contrárias ao bilhete ser sorteado, parece que na *maioria* dos mundos próximos em que as condições iniciais do exemplo são mantidas, o bilhete não é sorteado. Apenas em uma fração pequena dos mundos próximos relevantes a crença de S de que o bilhete não foi sorteado é falsa. Consequentemente, a crença de S de que seu bilhete não foi sorteado é segura.

Nas palavras de Pritchard,

O exemplo da loteria diz respeito a crença de alguém - anterior à escuta do resultado do sorteio e baseada nas probabilidades [*odds*] envolvidas- de que possui um bilhete perdedor para uma loteria justa, para a qual muito bilhetes foram lança-

dos. Intuitivamente, este não é um caso de conhecimento, apesar das excelentes bases estatísticas que se possui em favor da crença porque, mesmo quando a crença é verdadeira, é uma questão de sorte que ela é verdadeira. Afinal de contas, existem mundos possíveis próximos em que a crença é falsa (i.e., aqueles mundos em que o sujeito possui o bilhete vencedor). Se isto está correto, entretanto, então parece que exigir que a crença do sujeito seja verdadeira na 'maioria' dos mundos próximos, tal como SP [condiçãoS1] exige, não será suficiente, uma vez que, intuitivamente, esta condição é satisfeita no caso da loteria onde existem poucos mundos possíveis próximos em que se forma a crença falsa na proposição alvo (PRITCHARD, 2008, p.36).

Assim, a crença verdadeira S de que o bilhete i não foi sorteado, quando a probabilidade de o bilhete i não ser sorteado é extremamente alta, e S acredita que i não foi sorteado com base exclusivamente nas probabilidades envolvidas, envolve sorte epistêmica. Portanto, ou bem abandonamos a ideia de que segurança elimina sorte ou bem a crença de S de que não ganhou na loteria não envolve sorte. Para resolver estas questões, Pritchard sugere que elaboremos uma versão de segurança mais restritiva. Inicialmente, ele sugere a seguinte modificação:

(S2) A crença de S é segura se e somente se em todos (ou quase todos) os mundos possíveis próximos em que S continua a formar sua crença na proposição-alvo da mesma forma que faz no mundo atual, a crença é verdadeira (PRITCHARD, 2008, p.38).

Segundo Pritchard, a crença de que o bilhete i não foi sorteado na loteria não é verdadeira em todos (ou quase todos) os mundos possíveis em que S forma a respectiva crença. Embora a crença de S que i possa ser verdadeira *na maioria* dos mundos próximos relevantes, a crença de S de que i não é verdadeira em todos (*ou quase todos*) os mundos possíveis próximos relevantes em que S acredita que i. Afinal, existe um mundo possível próximo relevante, um mundo no qual p.ex., as bolas do sorteio caem em uma configuração diferente, no qual o sujeito S forma a crença alvo da mesma forma, e ainda assim a sua crença é falsa.

O princípio S2 se propõe a resolver o problema gerado pelo exemplo da loteria. Contudo, ao resolver um problema, parece que produzimos outro. O motivo é que S2 é vulnerável a certos contraexemplos envolvendo conhecimento indutivo. Ora, a capacidade da condição de segurança de acomodar o conhecimento indutivo é justamente uma das alegadas vantagens da segurança sobre a sensibilidade. Referindo-se ao contraexemplo da lixeira, que apresentamos antes, Pritchard escreve o seguinte:

Considere novamente, por exemplo, o 'caso da lixeira' descrito por Sosa, e que foi citado acima. Intuitivamente, nós queremos dizer que Sosa sabe que o seu saco de lixo alcançou o chão. O problema é que se segurança é entendida de forma robusta na linha de (SP\*) [a condição S2], então, foi alegado, este conhecimento parece excluído, uma vez que certamente existem alguns poucos mundos possíveis i.e., aqueles nos quais o saco fica preso ao longo do caminho, nos quais Sosa conti-

nua a formar a crença na mesma proposição e ainda assim forma uma crença falsa como resultado (PRITCHARD, 2008, p.38).

Para acomodar estas duas intuições, de que temos crença segura no caso da lixeira e crença insegura no caso da loteria, Pritchard propõe esta outra modificação da condição de segurança, que denominaremos S3:

A crença de S de que P é segura no mundo atual se e somente se na maioria dos mundos próximos [*nearby*] em que S continua a formar a sua crença a respeito da proposição-alvo da mesma forma que o faz no mundo atual, e em todos os mundos muito próximos [*very close nearby worlds*] em que S forma a sua crença a respeito da proposição-alvo da mesma forma que o faz no mundo atual, a sua crença continua sendo verdadeira (PRITCHARD, 2008, p.40).

A ideia básica de S3 é que não pode existir *um único* mundo muito próximo ao mundo atual em que S crê p e é falso que p. Ao mesmo tempo, no que diz respeito aos mundos próximos (porém não *muito* próximos) em que S acredita que p, é suficiente que seja verdadeiro que p na maior parte deles. A segunda cláusula explica por que não temos conhecimento no caso da loteria. Contudo, a combinação da primeira com a segunda cláusula não é incompatível com o conhecimento no caso da lixeira, desde que não exista um único mundo *muito* próximo em que S acredita que p e é falso que p.

Conforme vimos, o motivo pelo qual Pritchard julga que S3 dá a resposta correta ao problema da loteria é que ele considera que o mundo possível próximo em LOTERIA1, em que é falso que o bilhete foi sorteado, é um mundo possível muito próximo. A razão pela qual este mundo deve ser categorizado entre os mundos muito próximos, é que muito pouco precisaria ser mudado no mundo atual, para que a crença de que PL viesse a ser falsa. Apenas precisaria acontecer que as bolas numeradas do sorteio, supondo o mecanismo tradicional de muitas loterias, viesse a cair em uma configuração ligeiramente diferente. E *isto*, pelo menos segundo o autor, ocorre em mundos muito próximos.

Podemos extrair das considerações o seguinte condicional, que será importante mais tarde: se poucas modificações precisariam ser feitas no mundo W para que E viesse a ser o caso, então E ocorre em um mundo muito próximo ao mundo W.

Uma consequência da resposta de Pritchard ao problema da loteria é que ela implica o divórcio entre as noções de probabilidade e distância modal. Eventos muito improváveis no mundo atual, tal como vir a ganhar em uma loteria, podem ser modalmente muito próximos. Consequentemente, a distância modal não deve ser função da probabilidade de um evento ocorrer.

#### 4.2.4 A crítica de McEvoy

Em artigo recente, McEvoy (2009a) propôs uma crítica penetrante à Pritchard, e sugeriu que o autor não conseguiu solucionar apropriadamente o problema da loteria. A seguir, irei apresentar esta crítica, e tecer alguns comentários em seu favor.

McEvoy alega que a crença de que PL em loteria<sub>2</sub> não é segura, de acordo com S3 e com a métrica dos mundos possíveis propostas por Pritchard. O mesmo motivo que conduz Pritchard a dizer que a crença de que PL não é segura em loteria<sub>1</sub>, também deveria levar Pritchard a dizer que a crença de que PL em loteria<sub>2</sub> não é segura. Afinal, para que o sujeito S viesse a crer falsamente em loteria<sub>2</sub>, tudo o que precisaria acontecer seria que houvesse um deslize da parte do jornal. O evento em que S crê falsamente que ganhou na loteria porque um número foi acidentalmente trocado não parece exigir uma mudança significativa em relação ao mundo em que, por hipótese, o jornal relata corretamente o resultado do sorteio.

A réplica mais natural em defesa de S3 seria dizer que, dado que o jornal é bastante confiável e possui algum tipo de mecanismo para detecção de erros, o mundo possível próximo em que um erro passa despercebido pelos mecanismos de correção deve ser considerado suficientemente distante. Ele *não* é um mundo muito próximo.

Pritchard alega que jornais, pelo menos aqueles que merecem crédito, possuem certos mecanismos destinados a prevenção de erros deste gênero. Existem editores e revisores de texto, geralmente confiáveis, que trabalham para que tais falhas não aconteçam. Conforme coloca Pritchard:

O ponto aqui é que jornais confiáveis - o tipo de jornal que alguém pode usar para vir a saber que perdeu na loteria - possuem mecanismos de correção internos que visam prevenir tais tipos de erro de digitação ou impressão de virem a ocorrer. A equipe é altamente treinada e diligente, toda informação é corrigida duas vezes por eficientes revisores de texto, e, talvez mais significativamente, aqueles que conferem os resultados são provavelmente possuidores de bilhetes de loteria, e possuem um interesse pessoal na correção dos resultados. Em suma, eu penso que uma vez que descrevemos os detalhes do caso apropriadamente, não é muito plausível supor que os mundos possíveis onde nosso herói forma uma crença falsa na proposição alvo são próximos (PRITCHARD, 2009, p.36).

O problema do argumento de Pritchard é que ele parece exigir demais do conceito de confiabilidade, e do que esperaríamos de um jornal ou de outra fonte de informação semelhante. De acordo com sua tese, se uma fonte é confiável no que diz respeito a transmissão de

conhecimento então não existe *nenhum* mundo possível muito próximo em que a mesma fonte de informação erra (supondo as circunstâncias iniciais relevantes fixas).

A objeção de McEvoy afirma que, afinal de contas, erros deste tipo ocorrem em jornais tipicamente confiáveis, ainda que não com grande frequência (se fossem excessivamente frequentes, então os jornais não seriam mais propriamente confiáveis). Tudo o que seria necessário para que o sujeito acreditasse falsamente em loteria<sub>2</sub> seria que o jornal cometesse um erro ao relatar o resultado, e que o mecanismo de correção (confiável, porém falível) viesse a falhar (talvez por um lapso do corretor de texto do jornal). Conforme coloca o autor:

O ponto destes exemplos é que má impressão e outros erros podem e, de fato, ocorrem em jornais normalmente confiáveis. O fato de que erros de impressão ocorrem, apesar das boas qualidades dos jornalistas e revisores mostra que tais erros podem ocorrer sem exigir modificações transmundanas relativamente às qualidades dos agentes relevantes. Por que pensar que os relatos dos vencedores de sorteios de loteria seriam especialmente imunes a tais erros? A falibilidade dos funcionários dos jornais é tudo que é necessário para erros aconteçam em qualquer área - e uma vez que isto é uma constante entre o mundo atual e todos os mundos próximos, muito pouco deve ser diferente entre o mundo atual e mundos do tipo (ii) [ mundos em que o jornal comete erro ao relatar o resultado da loteria] (MCEVOY, 2009a, p.14).

Erros são encontrados diariamente em jornais confiáveis no mundo atual. Se este é o caso, não deveríamos dizer que a possibilidade erro do jornal deve ser considerada muito próxima, uma vez que casos semelhantes sob todos os aspectos relevantes ocorrem no mundo atual? Afinal de contas, é duvidoso, para dizer o mínimo, que o mundo em que o jornal comete um erro ao relatar o resultado da loteria não é muito similar ao mundo atual, se situações muito similares, pelo menos sob os aspectos relevantes (o jornal é confiável, existem mecanismos para prevenção de erros, e assim por diante), *já fazem parte do mundo atual*. Um mundo em que um jornal muito confiável, porém falível, efetivamente falha a respeito de um certo relato, parece muito próximo do mundo atual (supondo que o mundo atual seja um mundo em que o jornal não erra ao fazer o relato).

Curiosamente, uma versão anterior da condição de segurança, o princípio S1, parece responder corretamente a loteria<sub>2</sub>. Afinal, ainda que exista algum mundo possível muito próximo em que o jornal erra ao relatar o resultado, isto não parece ocorrer na maioria deles ou com grande frequência. O problema desta alternativa é que, em razão de loteria<sub>1</sub>, Pritchard se viu obrigado a rejeitar S1. Não penso haver uma saída viável, no quadro dos princípios sugeridos por Pritchard.

Além da incapacidade de S3 de responder à loteria<sub>2</sub>, gostaria de apontar para outro problema, que sugere fortemente que a teoria de Pritchard é inadequada. S3 depende de uma

classificação tripartite do conjunto de mundos possíveis: mundos muito próximos, mundos próximos *não muito* próximos e mundos distantes. Pritchard não fornece um critério sistemático a respeito de como devemos realizar tais distinções. Ao mesmo tempo, é duvidoso que tenhamos intuições estáveis a respeito de como distinguir entre mundos próximos não muito próximos e mundos muito próximos. A teoria que depende da capacidade de realizar esse tipo de discriminação fina parece assentada em uma base demasiado frágil e incerta.

A vagueza e imprecisão de uma distinção podem ser toleráveis - elas não são elimináveis de muitos conceitos - se tivermos alguns casos paradigmáticos em que a distinção pode ser feita com relativa segurança, e se conseguirmos sistematicamente fazê-la em uma ampla gama de casos. Aqui é útil fazer uma comparação, a fim de mostrar que a vagueza e a imprecisão não são suficientes para impugnar a utilização teórica de um conceito. Consideremos, por exemplo, a distinção entre mundos próximos e distantes. Embora a métrica da distância modal contenha alguma dose de vagueza e imprecisão, podemos, ao menos em certos casos, nos por de acordo quanto a como ela deve ser feita. Mundos em que sou o Presidente da República, ou em que todos os meus conhecidos são alienígenas disfarçados de seres humanos, ou mesmo mundos em que sou um bilionário, são certamente distantes. O mesmo podemos dizer a respeito dos mundos imaginados pelos céticos: o mundo em que todas as minhas experiências sensoriais são causadas por algum tipo de demônio é certamente distante. Por outro lado, nos colocamos de acordo quanto a classificação de certos mundos como próximos: são próximos os mundos em que possuo um fio de cabelo a mais, ou um centímetro a mais de altura. A distinção entre mundos próximos e distantes parece pertinente, pelo menos para uma ampla gama de casos. Infelizmente, não vemos o mesmo tipo de robustez na distinção entre mundos próximos não muito próximos e mundos muito próximos.

#### 4.2.5 Por que a incapacidade de responder ao exemplo da loteria é problemática

Conforme vimos, Pritchard defende uma epistemologia anti-sorte modesta. Se este é o caso, ele não poderia simplesmente alegar que existe alguma outra condição necessária adicional para o conhecimento, que não é satisfeita em loteria<sub>1</sub>? Tal resultado enfraqueceria o poder explicativo da condição de segurança, assim como eliminaria uma de suas vantagens *vis-à-vis* sensibilidade, e, por conseguinte, a sua força argumentativa. Em todo caso, isto não seria

suficiente para forçar a rejeição da teoria, uma vez que as outras razões em favor da teoria permaneceriam intactas.

Esta não é uma saída atraente para Pritchard. Eis como ele explica seus motivos, referindo-se ao exemplo que denominamos *loteria*<sub>1</sub>:

Um tipo de resposta pode consistir em suplementar as teorias da segurança do conhecimento com outras condições epistêmicas que podem, em conjunção com a segurança, tratar do caso. O problema com esta proposta, entretanto, é que uma vez que a intuição guia aqui é que conhecimento está faltando precisamente por causa da presença de sorte, e uma vez que, como nós notamos, (SP) [o princípio S1] é uma maneira natural de entender a condição anti-sorte, este movimento dialético não é muito atraente. Essencialmente, estaríamos dando conta da sorte epistêmica neste caso, não por um apelo para uma condição supostamente anti-sorte, porém por referência a outras condições epistêmicas, e isto obviamente diminui a atração teórica da visão (PRITCHARD, 2008, p.38)

Claramente, Pritchard considera que uma resposta aos casos de *loteria* é essencial. Sem um tratamento do problema, ele considera que a segurança perde a sua principal razão de ser: a capacidade de capturar a condição anti-sorte. Como Pritchard não consegue produzir o resultado correto *loteria*<sub>1</sub> sem produzir o resultado incorreto em *loteria*<sub>2</sub>, é difícil encontrar uma saída satisfatória. A teoria está em um beco sem saída.

Existe, contudo, um segundo problema. Pritchard julgou necessário acrescentar uma condição de habilidade à sua condição anti-sorte. Embora eu não queira me aprofundar nos detalhes a este respeito por não ser este o foco do presente trabalho, é pertinente fazer uma explicação. Ele defende o seguinte princípio, que combina a segurança da crença com a ideia de que a formação da crença verdadeira deve ser creditada, em larga medida, ao exercício das habilidades cognitivas do agente:

S sabe que P se e somente se a crença segura de S de que p é o produto das suas habilidades cognitivas relevantes (tal que o seu sucesso cognitivo é em um grau suficientemente alto creditável à sua agência cognitiva) (PRITCHARD, 2012b, p.273).

O problema é que, segundo o autor, a condição de habilidade isoladamente (sem referência à segurança da crença) não explica a ignorância do sujeito cognoscente. A epistemologia da virtude forte, que afirma que formação de uma crença verdadeira em razão do exercício de uma habilidade cognitivo é suficiente para distinguir conhecimento de crença verdadeira, não consegue explicar a ignorância do sujeito cognoscente. Conforme ele mesmo escreve:

Em primeiro lugar, note como a epistemologia da virtude forte luta com o caso da *loteria*. Afinal de contas, o sucesso cognitivo de Lottie [nome que Pritchard dá ao personagem de um exemplo semelhante à *loteria*<sub>1</sub>] parece ser explicado pelo exercício de suas habilidades cognitivas relevantes, não parece? Dito de outra forma,

se não é a habilidade cognitiva de Lottie que explica seu sucesso cognitivo, então o que o explica? Pois note que o único candidato plausível aqui é a boa sorte (epistêmica) de que seu bilhete é o bilhete perdedor. Entretanto, considerando as probabilidades envolvidas, é difícil ver como isto poderia ser considerado uma questão de sorte (PRITCHARD, 2012b, p.266).

Em vista disso, os problemas parecem ainda mais graves. Pritchard pensa – com alguma plausibilidade – que o sucesso cognitivo (formação de crença verdadeira) é, de alguma forma, explicado pelo exercício de habilidades cognitivas do agente. Se este é o caso, então a única maneira de compatibilizar a ignorância do sujeito S em loteria<sub>1</sub> com epistemologia da virtude anti-sorte seria alegar que a crença não é segura. Todavia, este não pode ser o caso, conforme vimos anteriormente. Consequentemente, se aceitarmos a epistemologia da virtude anti-sorte, e que as condições de habilidades e anti-sorte são satisfeitas, então teríamos de aceitar que em loteria<sub>1</sub> o sujeito S sabe que o bilhete i não foi sorteado, o que é simplesmente absurdo.

#### 4.2.6 Por que loteria<sub>1</sub> não é um caso de sorte epistêmica

Pritchard afirma explicitamente que a ignorância do sujeito S em situações do tipo loteria<sub>1</sub> é explicada pela sorte epistêmica. A seguir, irei sugerir que isto não pode ser correto. Todavia, como Pritchard considera essencial que a condição de segurança explique loteria<sub>1</sub>, em razão da condição de habilidade não explicar a ignorância do sujeito cognoscente. Para ilustrar a questão, e apresentar a natureza do problema, iremos considerar o seguinte caso, uma variação dos exemplos anteriores de loteria, que denominaremos de *loteria inversa*.

Um sujeito S compra um bilhete de loteria, digamos o bilhete i, a fim de receber um prêmio em dinheiro. As características da loteria são da seguinte forma: (a) o número de bilhetes em disputa é *extremamente* alto, tão elevado quanto quisermos imaginar; (b) a loteria é justa, e todos os bilhetes possuem exatamente a mesma chance de serem sorteados; (c) a pessoa que comprou o bilhete recebe o prêmio apenas se o bilhete que ele comprou não for o bilhete sorteado.

Pois bem, a loteria foi rodada, o bilhete i não foi sorteado e S recebeu o prêmio em dinheiro. Diríamos que S teve *sorte* em ganhar o prêmio? Curiosamente, o próprio Pritchard parece pensar que o evento em que um bilhete de um sujeito S não foi sorteado, em uma situação como aquela descrita no exemplo, não é um caso de sorte. Em um contexto no qual dis-

cute a capacidade de uma forma de epistemologia da virtude de lidar com problema da loteria (tal como os exemplos anteriores de loteria, de que tratamos), e que foi anteriormente citado, Pritchard escreve que "considerando as probabilidades envolvidas, é difícil ver como isto poderia ser considerado uma questão de sorte."(Pritchard, 2012b, p.266)

Eu penso que a observação de Pritchard a este respeito está correta em linhas gerais, o que sugere o seguinte: o sujeito S não teve sorte em loteria inversa. O que parece explicar isto é certa conexão entre os conceitos de *normalidade* e de *sorte*. Quando ocorre o que normalmente ocorre - e que esperávamos que acontecesse nas circunstâncias - nós relutamos em dizer que o evento em questão é um caso de sorte, ou que o sujeito beneficiado pelo evento teve sorte. Quando dizemos que alguém teve sorte, supomos que estamos tratando de eventos que são, em alguma medida, excepcionais. A implausibilidade da afirmação de que eventos regulares podem ser considerados casos de sorte é observada com maior clareza quando notamos a aparente incongruência, e desconforto, que nos trazem sentenças do tipo: *P é um caso de sorte, e P ocorre regularmente*.

Note-se como o caso da loteria inversa, e o diagnóstico que propomos para o caso, colocam a teoria de Pritchard em maus lençóis. Os mesmos motivos que nos levaram a julgar que S não teve sorte no caso da loteria inversa também deveriam, para todos os efeitos, nos levar a afirmar que em loteria<sub>1</sub> a crença verdadeira S de que o seu bilhete não foi sorteado não é verdadeira por sorte ou acidente. Não atribuímos sorte a tipos de eventos que ocorrem com excepcional regularidade, e que correspondem ao que normalmente ocorre. Se este é o caso, então a formação de uma crença verdadeira por parte de S não deve ser explicada pela sorte verística; afinal de contas, que o sujeito viesse a crer verazmente, dadas as circunstâncias em que a crença foi formada, é o que esperaríamos que acontecesse (o que *normalmente* acontece).

Assim, da mesma forma que um sujeito S não teve sorte em ter seu bilhete não sorteado, quando a probabilidade de ele vir a ser sorteado é extremamente pequena, também não teve sorte o sujeito que veio a crer verazmente que o seu bilhete não foi sorteado com base nas probabilidades envolvidas, quando a probabilidade de ele vir a ser sorteado é mínima. Qualquer que seja a razão que explica nossa relutância em atribuir conhecimento a S em Loteria<sub>1</sub> - não é minha intenção disputar essa inclinação - ela deve ser procurada em outro lugar. Visto que a condição de segurança é a cláusula anti-sorte por excelência, segundo Pritchard, e que a ignorância do sujeito não é explicada pela sorte epistêmica, parece que a condição de segu-

rança não consegue responder ao problema da loteria. Considerando que a condição da habilidade não gera um resultado melhor, segundo o próprio Pritchard, temos que a sua epistemologia da virtude anti-sorte, tal como proposta pelo autor, não tem meios de tratar adequadamente do problema da loteria.

#### 4.2.7 Segurança e verdades necessárias

O princípio S3 não corresponde à versão final da teoria da segurança para Pritchard, e mais refinamentos se fazem necessários. O motivo é que, assim como ocorreu com sensibilidade, S3 tem dificuldades em tratar do conhecimento das verdades necessárias. Ora, vamos supor que  $p$  seja uma verdade necessária. Neste caso,  $p$  é verdadeiro em todos os mundos possíveis, *a fortiori*, em todos os mundos possíveis em que S acredita que  $p$ . Temos assim, para quaisquer sujeitos cognoscentes e qualquer proposição  $p$ , que  $(\beta p \ \& \ \Box p) \rightarrow Kp$ . Em outros termos, basta acreditar em uma verdade necessária para saber que  $p$ . Isto é evidentemente absurdo.

Há várias maneiras de contornar este resultado. Em primeiro lugar, é possível tratar o princípio de segurança como apenas uma condição *necessária* para a posse de conhecimento, ao invés de tratar a crença verdadeira segura como uma condição suficiente. Se este é o caso, então pode existir alguma outra condição necessária que não é satisfeita em casos de ignorância de verdades necessárias (quando se crê na proposição). De qualquer forma, se esta é a situação, então a ideia de que conhecimento consiste em crença verdadeira segura é inviável. Note-se que, ao contrário de outras circunstâncias, a referência à base evidencial ou indicação segura não soluciona a dificuldade. Se  $p$  é uma verdade necessária, então  $p$  é verdade em todos os mundos possíveis, *a fortiori* em todos os mundos em que S acredita que  $p$  com base em uma determinada base ou indicação. Tampouco as sucessivas modificações sugeridas por Pritchard parecem evitar a trivialização das verdades necessárias. Crenças em verdades necessárias satisfazem quaisquer das três condições que vimos anteriormente cogitadas por Pritchard: S1, S2 e S3. Conforme o próprio autor admite, S3 aplica-se apenas a proposições plenamente contingentes (termo que o autor utiliza para cobrir tanto contingências lógicas como nômicas).

Todavia, parece que existem certos casos de crença acidentalmente verdadeira em verdades necessárias. O princípio S3 não é adequado para explicar por que o sujeito cognoscente não possui conhecimento nestas circunstâncias, nem por que a presumida ignorância resulta da intervenção da sorte verística. Além do mais, parece evidente que existem certas crenças em verdades necessárias que podem ser infectadas por algum tipo de sorte verística. Consideremos o seguinte exemplo proposto por Pritchard:

Mathema usa uma calculadora para descobrir o resultado da multiplicação  $12 \times 13$ . Como resultado, ele forma a crença verdadeira que  $12 \times 13 = 156$ . Sem que Mathema saiba, entretanto, a calculadora está quebrada e gera ‘respostas’ aleatórias (PRITCHARD, 2012b, p. 253).

Diante de tal constatação, Pritchard insere mais uma modificação em sua teoria. Desta vez, ele sugere que devemos considerar outras crenças do sujeito cognoscente, além da crença na proposição-alvo. De acordo com esta nova modificação, uma crença é insegura se, preservadas as condições iniciais relevantes, o modo ou maneira como o sujeito formou a crença poderia levar a formação de crenças falsas.

Pritchard sugere, desta forma, que o princípio de segurança deve ser modificado, de tal sorte a levar conta diferentes resultados doxásticos:

Notemos, entretanto, que é muito fácil perceber como se pode estender a explicação da segurança para estas proposições, mesmo se os detalhes podem ser complicados. Afinal de contas, tudo o que precisamos é falar do resultado doxástico do processo formador de crenças, qualquer que seja ele, e não focar apenas na crença na proposição-alvo. Por exemplo, se alguém forma a sua crença de  $2 + 2 = 4$  por meio de cara ou coroa, existe uma ampla classe de mundos possíveis próximos em que este processo formador de crenças gera um resultado doxástico falso (p.ex., o mundo possível em que alguém forma a crença de que  $2 + 2 = 5$ ) (PRITCHARD, 2009, p.34).

Pois bem, como deveria ser adaptada a condição de segurança, a fim de acomodar o conhecimento de verdades necessárias, e os diferentes resultados doxásticos do processo de formação de crenças? Embora Pritchard não o faça explicitamente, a seguinte modificação parece adequada

**(S4)** A crença verdadeira do sujeito S de que P formada por um modo/maneira m é segura no mundo atual sse **(i)** existe uma ampla classe de mundos possíveis em que o mesmo processo/método não gera uma crença falsa nas mesmas condições iniciais **(ii)** são verdadeiras todas as crenças formadas segundo m nos mundos possíveis muito próximos ao mundo atual nas mesmas condições iniciais.

S4 sugere uma modificação importante da segurança e aproxima muito a teoria de alguma forma de confiabilismo. Afinal, o que faz com que uma crença de que P seja insegura, segundo S4, não é o fato de que o sujeito muito facilmente poderia crer falsamente que P. O que torna uma crença insegura, segundo S4, é que o método empregado poderia muito facilmente produzir crenças falsas. Em todo caso, a modificação sugerida parece responder adequadamente aos problemas das verdades necessárias.

#### 4.3 ALGUNS CONTRAEXEMPLOS À CONDIÇÃO DE SEGURANÇA

A condição de segurança, em quaisquer de suas variantes, nos parece o desenvolvimento mais promissor das ideias originalmente traçadas por Nozick e Dretske, que procuram explicar o conhecimento, ao menos em parte, a partir de princípios subjuntivos/contrafactuais. Apesar das suas alegadas vantagens sobre a sensibilidade, o princípio de segurança não está imune a sérias objeções. Uma delas foi apresentada por Juan Comesaña. O contraexemplo é o seguinte:

FESTA DE HALLOWEEN: haverá uma festa de *Halloween* na casa de Andy, e eu fui convidado. A casa de Andy é muito difícil de encontrar, então ele contrata Judy para permanecer em uma bifurcação e direcionar as pessoas para a casa (o trabalho de Judy é dizer às pessoas que a festa é na casa situada na rua à esquerda). Sem que eu saiba, Andy não deseja que Michael vá à festa, e diz para Judy que, se ela vir Michael, então ela deve dizer a ele a mesma coisa que diz aos outros (que a casa é na rua à esquerda), mas deve imediatamente telefonar para Andy a fim que a festa seja transferida para a casa de Adam, que é na rua à direita. Eu seriamente considero fingir ser Michael, mas, no último momento, eu desisto. Quando eu alcanço a bifurcação, eu pergunto Judy onde será a festa, e ela me diz que será na rua à esquerda (COMESANA, 2005, p.397).

Suponhamos que S seja o sujeito cognoscente do exemplo. Parece que, no exemplo tal como descrito, S sabe que a festa irá ocorrer na rua à esquerda. Contudo, a crença de S não é segura, de acordo com as caracterizações prévias da condição de segurança, pois S facilmente poderia ter acreditado falsamente que a festa irá ocorrer na rua à esquerda, uma vez que existe um mundo W muito próximo ao mundo atual @, tal que em W o sujeito S acredita falsamente que a festa será na rua à esquerda com base no mesmo método que ele utiliza no mundo atual. O mundo W é justamente o mundo em que S decide ir à festa fingindo ser Michael. Como essa é uma situação de conhecimento inseguro, o princípio de segurança não pode ser uma condição necessária para o conhecimento.

Contraexemplos estruturalmente semelhantes ao de Comesaña foram propostos por outros autores. Neta e Rohrbaugh, por exemplo, apresentam o seguinte caso:

(A) eu estou bebendo um copo de água, a partir de uma garrafa recém aberta. Próximo a mim está uma pessoa feliz, que acabou de ganhar na loteria. Se essa pessoa tivesse perdido na loteria, ela teria maliciosamente poluído minha água com uma toxina sem gosto, odor ou cor. Porém, uma vez que ela ganhou na loteria, ela não faz tal coisa. No entanto, ela quase perdeu na loteria. Agora, eu bebo a água pura não adulterada e julgo, verdadeiramente e com conhecimento, que estou bebendo água pura, não adulterada (NETA, 2004, p.399-400).

No exemplo acima, existe um mundo próximo ao mundo atual, tal que neste mundo a outra pessoa perde na loteria (o que facilmente poderia ter acontecido), e eu venho a formar a crença falsa, com a mesma base, de que estou bebendo água pura. Portanto, a crença não é segura. Todavia, a crença é um caso de conhecimento, pelo menos segundo seus autores.

#### 4.3.1 Diagnosticando os contraexemplos

O que mostram as objeções a condição de segurança de que tratamos? Para responder a essa questão, é útil tentar compreender a estrutura dos contraexemplos. Para fazer isto é conveniente considerar um dos exemplos, e tentar abstrair a estrutura relevante.

Tomemos o caso da festa de *Halloween*. Sejam  $P = a \text{ festa ocorrerá na rua à esquerda}$ ,  $C = eu \text{ não decidi fingir ser Michael}$ ,  $W_1 =$  o mundo em que não fingi ser Michael, porém quase decidi fingir ser Michael e  $W_2 =$  o mundo próximo em que eu fingi ser Michel. Eis como entendo a estrutura/forma do contraexemplo:

no mundo  $W_1$  são verdadeiras as seguintes proposições:  $C$ ,  $\beta p$ ,  $P$ ,  $(\beta p \ \& \ C) \Box \rightarrow P$  e  $(\beta p \ \& \ \neg C) \Box \rightarrow \neg P$ . No mundo  $W_1$  as circunstâncias são tais que *quase*, por muito pouco, não é verdade que  $C$ . Também é o caso que  $C$  *antecede*  $P$  e  $\beta p$ . Consideremos agora um mundo próximo ao mundo  $W_1$ , digamos o mundo  $W_2$ . Em  $W_2$  são verdadeiras as seguintes proposições:  $\neg C$ ,  $\beta p$  e  $\neg P$ .  $\neg C$  explica por que é verdade em  $W_2$  que  $\beta p \ \& \ \neg P$ . Também é o caso que  $\neg C$  *antecede*  $\beta p$  em  $W_2$ .

Em razão da existência do mundo  $W_2$ , a condição de segurança não é satisfeita em  $W_1$ , pois:

1.  $W_2$  é um mundo muito próximo ao mundo  $W_1$ , em que as condições iniciais relevantes são as mesmas;
2.  $\beta p \ \& \ \neg p$  em  $W_2$ .

Cada um dos contraexemplos restritivos à condição de segurança de que tratamos instância o esquema que anteriormente apresentamos. Se eu tivesse fingido ser Michael e acreditasse que a festa será na rua à esquerda, então a festa não seria na rua à esquerda (pois a localização da festa teria mudado), o que expressamos por  $(\beta p \ \& \ C) \Box \rightarrow P$ . Ao mesmo tempo, é verdade que eu quase, por muito pouco, decidi fingir ser Michael momentos antes de formar a minha crença de que a festa será na rua à esquerda. Todavia, eu **não** fingi em  $W_1$ . A possibilidade próxima não-atualizada de eu fingir ser Michael é o que explica por que a minha crença de que a festa será na rua à esquerda não satisfaz a condição de segurança.

No caso da água contaminada, verificamos a mesma situação. Se a outra pessoa tivesse perdido a loteria, eu teria acreditado falsamente que estou bebendo água pura. Todavia, a outra pessoa *quase* perdeu na loteria antes de eu vir a formar a crença de que estou bebendo água pura. A possibilidade próxima não-atualizada da pessoa ter perdido na loteria é o que explica por que minha crença, no mundo atual em que isto não aconteceu, não é segura.

Uma vez que compreendemos a estrutura dos casos, podemos examinar algumas tentativas de respondê-los.

#### 4.3.2 A resposta de Bogardus<sup>46</sup>

Em um artigo recente, Bogardus (2012) apresentou algumas objeções ao contraexemplo de Comesaña, e também a outros, semelhantes, apresentados por Neta e Rohrbaugh (2004). De acordo com Bogardus, esses autores cometerem uma falácia, qual seja, “cada um move falaciosamente do fato de que S estava em risco epistêmico logo antes de formar a

---

<sup>46</sup> Esta seção reproduz um texto anterior meu, publicado em Rodrigues (2016).

crença para a conclusão de que a crença de S foi formada inseguramente” (Bogardus, 2012, p. 289).

Vejamos como a crítica de Bogardus aplica-se ao contraexemplo apresentado por Comesaña. De acordo com ele:

antes de eu decidir não fingir ser Michael, eu corria o risco de ganhar uma crença falsa no futuro com base no testemunho de Judy. Porém, após eu decidir não fingir ser Michael, eu passo a não correr mais risco. E, assim, no momento em que eu formei a minha crença, eu evitei o risco epistêmico, ao decidir não fingir ser Michael. Mas, então, crucialmente, eu não corro mais risco epistêmico a partir do momento em que eu eventualmente formei a crença'. (BOGARDUS, 2012, p. 297).

Existem alguns problemas importantes na objeção de Bogardus. É digno de nota que o primeiro desses problemas foi antecipado no próprio artigo de Comesaña (2005). Ao considerar as possíveis objeções que poderiam ser feitas ao contraexemplo da festa de *Halloween*, ele observa que o defensor da teoria da segurança poderia afirmar que, após S decidir não fingir ser Michael, a possibilidade próxima de erro torna-se irrelevante para a satisfação da condição da segurança. Conforme ele mesmo coloca:

similarmente, o defensor da condição de segurança para o conhecimento poderia afirmar que antes de eu não decidir não fingir ser Michael, existia a possibilidade próxima de que eu viesse a acreditar falsamente que a festa de Andy é na casa na rua à esquerda; agora, entretanto, esta não é mais uma possibilidade próxima: eu acredito segura e verdadeiramente (Comesaña, 2005, p. 399).

Quanto a essa objeção, Comesaña faz uma outra observação crucial. Ele nota que “quando consideramos se é o caso que a proposição que  $p$  é verdadeira e segura em  $t$ , nós levamos em conta se ela é verdadeira em mundos próximos que diferem do mundo atual em momentos logo antes de  $t$ ” (Comesaña, 2005, p.399).

Os mundos relevantes para a satisfação da condição de segurança da crença de que  $P$  em  $@$  incluem mundos diferentes de  $@$  quanto ao passado. Consequentemente, não podemos, por assim dizer, “manter fixo” o fato de que  $S$  não está vestido como Michael nos mundos relevantes simplesmente porque é verdade que  $S$  está vestido como Michael no momento  $T$  em que  $S$  forma a crença correspondente de que  $P$  em  $@$ .

Do fato de que  $S$  não finge ser Michael no mundo  $@$  no momento  $T$  em que  $S$  forma a crença de que  $P$  não se segue que não devemos considerar mundos possíveis alternativos em que  $S$  decide fingir ser como Michael logo antes do momento  $T$ . Evidentemente, não podemos considerar que todas as verdades em  $@$  devem ser mantidas nos mundos relevantes para a satisfação da condição de segurança, pois, se esse fosse o caso, conhecimento seria idêntico

a crença verdadeira. Dito de outra forma, do fato de que S não está vestido como Michael no mundo @ em T quando S forma a crença de que P não segue que o conjunto dos mundos possíveis relevantes para a crença de que P não inclui mundos possíveis em que S decide vestir-se como Michael antes de T. Presumivelmente, mundos com sequências de eventos anteriores a T, e ligeiramente diferentes da sequência de eventos que ocorre no mundo @, devem ser de alguma forma considerados. Sem uma razão adicional que legitime essa inferência, o que não está presente no artigo de Bogardus, a sua objeção perde muito da sua força inicial.

Vejamos o mesmo problema por outro ângulo. De acordo com Comesaña, o sujeito S não sabe que P, pois existe um mundo muito próximo ao mundo atual @, digamos o mundo W, em que S forma a crença de que P da mesma forma e nas mesmas circunstâncias relevantes em que ele forma a crença em @, e, contudo, é falso que P em w.

Assim, no mundo @:

- Em T1, S delibera se irá vestir-se como Michael.
- Em T2, S decide não se vestir como Michael.
- Em T3, S acredita que P com base no testemunho de Judy.

Por sua vez, o mundo W é o mundo em que:

- Em T1, S delibera se irá vestir-se como Michael.
- Em T2, S decide vestir-se como Michael.
- Em T3, S acredita que P com base no testemunho de Judy.

De acordo com o exemplo de Comesaña, a crença de S de que P formada no tempo T3 no mundo @ não é segura, isto é, não satisfaz a condição de segurança tal como vimos anteriormente, pois existe um mundo W muito próximo ao mundo @ tal que, em W, o sujeito S acredita falsamente que P da mesma maneira que ele acredita em @.

Segundo Bogardus (2012), o sujeito S corre o risco de formar uma crença falsa no tempo T1 no mundo @; porém, no tempo T3, quando o sujeito S forma a crença de que P, ele não corre mais o risco de formar uma crença falsa, uma vez que ele já decidiu não fingir ser o Michael em T2. De acordo com ele, é irrelevante a existência do mundo W em que S acredita falsamente que P. O fato responsável pela irrelevância do mundo W é a existência de uma determinada condição, digamos a condição C, tal que C é satisfeita em @ e C não é satisfeita em W. No contraexemplo da festa de *Halloween*, a condição C consiste em não estar disfarçado como Michael.

Não penso que a resposta é, em última análise, uma objeção convincente. O problema é que Bogardus parece considerar que, uma vez que a condição C não foi atualizada em @ no momento T2, antes do momento T3 em que S forma a crença de que P no mundo @, a existência de um mundo W em que  $\neg C$  torna-se irrelevante para a avaliação acerca da segurança da crença de que P formada em T3 no mundo @. Conforme vimos, esta inferência é problemática. Tais considerações são suficientes para rejeitar *prima facie* o argumento de Bogardus.

#### 4.3.3 Relativizar aos métodos e às circunstâncias não resolve o problema

Apesar dos problemas levantados anteriormente, existe uma réplica possível em favor do argumento de Bogardus. Uma alternativa disponível ao autor é fazer com que C esteja incluído no modo m que S emprega para formar a sua crença de que P. Se este fosse o caso, então o mundo w, em que S acredita falsamente que P, seria irrelevante para a satisfação da condição de segurança em @, uma vez que, em w, o sujeito estaria empregando outro método de formação de crenças.

A descrição e a individuação do método de formação de crenças compõem um dos problemas mais delicados que enfrentam todas as teorias que empregam o conceito de método de formação de crenças, o conhecido problema da generalidade. Dúvidas à parte sobre um critério geral da individuação de métodos ou modos de formas crenças, as nossas intuições no caso particular não favorecem a inserção de C como parte do método utilizado por S. Tanto em W como em @, o sujeito S parece formar exatamente a mesma crença da mesma maneira, qual seja, baseando-se no testemunho de que P dado por Judy. Não parece razoável que fatos como a aparência ou a vestimenta de S possam ser relevantes para determinar a maneira ou método de formação de crenças foi utilizado no caso. Novamente, sem uma razão adicional para inserir de C como parte de m, é difícil ver como a objeção de Bogardus poderia ser justificada por esta réplica.

Vejamos agora uma segunda alternativa. Conforme vimos, nenhum teórico da segurança defendeu a forma simples do princípio. Normalmente, o princípio é relativizado ao modo de formação de crenças ou às circunstâncias ou ambiente em que o sujeito se encontra<sup>47</sup>. A

---

<sup>47</sup> Tanto as circunstâncias, como o método, fazem parte das condições iniciais relevantes, tal como Duncan Pritchard as entende.

referência ao modo/método não parece adequada para tratar do problema. Todavia, será que a referência às circunstâncias ou ambiente em que o sujeito se encontra poderiam obter melhores resultados?

Uma alternativa deste tipo deveria alegar que os mundos contrafactuais em que a condição C não é satisfeita não são relevantes para a satisfação da segurança, pois não são relevantes os mundos em que a crença de S de que P deve ser verdadeira a fim de que o princípio de segurança seja satisfeito. O problema aqui é encontrar um motivo ou razão para sustentar essa tese, e encontrar princípios ou critérios que nos guiem na realização da distinção entre circunstâncias relevantes, que não admitem variação, e condições irrelevantes, que admitem variação. A ausência de critérios conhecidos para sistematicamente realizar tal distinção conta como uma razão forte contrária a tais propostas.<sup>48</sup>

Todavia, penso que existe outro problema. Consideremos o seguinte exemplo, que denominaremos de *celeiros*<sub>2</sub>. Barney está viajando de carro por uma estrada rural. Existem muitas fachadas de celeiro espalhadas pela estrada, e um único celeiro genuíno. Em um determinado local do trajeto percorrido por Barney, ele se encontra na seguinte situação. Existe uma fachada de celeiro, que Barney observaria se olhasse para o lado direito da estrada, e existe um celeiro genuíno, que Barney observaria se olhasse para o lado esquerdo da estrada. O celeiro e a fachada são indistinguíveis a partir da estrada. Barney *quase* dirige seu olhar para o lado direito, onde poderia observar uma fachada de celeiro. No entanto, ele não o faz. Milésimos de segundo após ele quase desviar o olhar para o lado direito, ele desvia para o lado esquerdo, e forma a crença verdadeira: estou vendo um celeiro.

Barney sabe que está vendo um celeiro? Gostaríamos de dizer que Barney não sabe que está vendo um celeiro em *celeiros*<sub>2</sub>. O caso é extremamente semelhante ao caos dos celeiros, com o acréscimo do detalhe referente ao desvio de direção do olhar. Sob todos os aspectos relevantes, *celeiros* e *celeiros*<sub>2</sub> são idênticos. Também são idênticos os juízos que faríamos acerca de Barney: ele não sabe que está vendo um celeiro.

Seja C1 = *Barney não desviou o olhar para o lado direito*. Note-se como a condição C1 instancia apresentamos anteriormente. No exemplo, temos que:

---

<sup>48</sup> Luís Rosa me sugeriu que talvez seja indecível se C deve fazer parte das condições iniciais (que devem ser mantidas nos mundos possíveis relevantes para a satisfação da condição de segurança). Eu simpatizo com esta posição, em vista da ausência de critérios comuns para sistematicamente decidir acerca do que incluir nas condições iniciais. Se este for o caso, então a objeção de Comesaña não se sustenta. Todavia, isto não é uma vantagem para o defensor da condição de segurança. A indecidibilidade é um problema tão ou mais grave que o contraexemplo de Comesaña.

- (i) C1 é verdadeira;
- (ii) quase que C1 não é verdadeira;
- (iii) se C1 fosse falsa, então Barney acreditaria falsamente que p;
- (iv) se C1 fosse verdadeira e Barney acreditasse que p, então seria verdade que p.

Ora, não permitir que C1 varie em seu valor de verdade nos mundos próximos relevantes para a satisfação do princípio de segurança, quando C1 satisfaz as condições i-iv, levaria à produção de um juízo incorreto em *celeiros*<sub>2</sub>. Isto mostra que o defensor da relativização às circunstâncias encontra-se em uma situação muito difícil e delicada. Até que seja dada uma resposta adequada a este desafio, penso que o prognóstico da relativização às circunstâncias não é promissor.

Qualquer tentativa de explicar por que um  $\neg$ C-mundo em que S crê falsamente não deve ser levado em consideração nos casos da festa de *Halloween* e da água contaminada, deveria explicar também por que não fazemos o mesmo em *celeiros*<sub>2</sub>. Afinal de contas, os casos são idênticos em todos os aspectos relevantes.

Na ausência de alguma diferença relevante entre *celeiros*<sub>2</sub>, o caso da festa de *Halloween* e da água contaminada, e de uma explicação por que levamos em conta o  $\neg$ C-mundo em *celeiros*<sub>2</sub>, mas não fazemos o mesmo nos demais casos, a referência às circunstâncias deve ser tomada como uma manobra *ad hoc*.

Em suma, relativizar às circunstâncias ambientais ou ao método de formação de crenças não responde aos contraexemplos de que tratamos. A relativização temporal - as crenças eram inseguras antes de C vir a ser o caso, mas depois de C vir a ser o caso, as crenças alvo são seguras - não é uma alternativa. Esta é a sugestão de Bogardus, e vimos que ela não é adequada. Diante de tal quadro, penso que há boas razões não respondidas para rejeitar o princípio de segurança.

## 5 CONCLUSÃO

Agora é o momento de recapitular o caminho percorrido até aqui. Ambas as epistemologias modais que examinei constituem importantes tentativas de responder a complexos e difíceis problemas filosóficos. Entre as duas teorias, entretanto, a teoria do rastreamento parece sucumbir mais facilmente frente a algumas objeções. A mais importante destas alega que o rastreamento não consegue acomodar satisfatoriamente o conhecimento indutivo, e não vejo como esta objeção poderia ser contornada ou respondida.

Também tratamos nesta tese de outros contraexemplos a teoria do rastreamento, e embora um defensor da teoria possa responder adequadamente a alguns deles, o conjunto acumulado das evidências sugere um quadro fortemente contrário à aceitação da teoria. O veredicto final é que a teoria deve ser rejeitada.

A condição de segurança evita muitos dos problemas da teoria do rastreamento. Em particular, o princípio de segurança não parece ter grandes problemas com conhecimento indutivo e de verdades necessárias. Segundo seus principais defensores, Duncan Pritchard e Ernest Sosa, a segurança também preserva o princípio de fecho epistêmico (como quer que venhamos a entendê-lo especificamente), e permite o conhecimento de que são falsas as hipóteses céticas. A depender de nossos juízos prévios sobre o princípio de fecho ou sobre o ceticismo, tal resultado conta pontos em favor da teoria.

Talvez o aspecto mais interessante da teoria da segurança seja como, pelo trabalho de Duncan Pritchard, ela foi combinada com uma análise do conceito de sorte epistêmica. Com efeito, de acordo com Pritchard, a segurança expressa a condição anti-sorte, a cláusula que visa excluir o tipo de sorte incompatível com a posse do conhecimento.

Existem dois problemas importantes que identifiquei ao longo deste texto, e ambos não são de mesma natureza e escopo. O primeiro deles diz respeito à versão de Pritchard para a condição de segurança. Segundo esta objeção, a teoria de Pritchard não consegue simultaneamente combinar o conhecimento indutivo e explicar por que o sujeito cognoscente em situações lotéricas não possui conhecimento. Como, segundo o próprio Pritchard, a teoria da segurança deve dar uma resposta ao problema da loteria, penso que isto constitui um problema bastante sério, e que mereceria uma resposta adequada da sua parte.

Em segundo lugar, existe uma outra objeção que apresentamos à condição de segurança. Trata-se da objeção proposta por Comesaña: o caso da festa de *Halloween*. A esta objeção somou-se outro contraexemplo estruturalmente idêntico, o que nos permite um tratamento conjunto, tal como vimos com caso da água contaminada. Esta objeção, ao contrário da primeira, pode ser dirigida tanto a versão da segurança proposta por Sosa, como à teoria da segurança de Pritchard.

Indiquei três maneiras de responder ao problema levantado por Comesaña, as quais me pareceram as alternativas mais promissoras: a relativização ao tempo, método ou circunstâncias. A relativização ao tempo foi proposta por Thomas Bogardus, e, após exame da sua sugestão, concluí que esta não é uma solução adequada. Tampouco produziram melhores resultados as relativizações ao método ou às circunstâncias ambientais. Na ausência de outras respostas ao problema levantado pelos casos da festa de *Halloween* e da água contaminada, penso que existe uma razão bastante forte para também rejeitar o princípio de segurança.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS, Fred; CLARKE, Murray. Resurrecting the tracking theories. **Australasian journal of Philosophy**, v. 83, n. 2, p. 207–221, 2005.

ARMSTRONG, David Malet. **Belief, truth and knowledge**. Nova York: CAMBRIDGE UNIVERSITY PRESS, 1973. 228 p.

BAUMANN, Peter. Nozick's defense of closure. In: BECKER, Kelly; BLACK, Tim. (Org.). **The Sensitivity Principle in Epistemology**. Cambridge, UK: CAMBRIDGE UNIVERSITY PRESS, 2012. p. 11–28.

BECKER, Kelly. **Epistemology modalized**. New York: ROUTLEDGE, 2007. 174 p.

BENNETT, Jonathan. **A philosophical guide to conditional**. Oxford: OXFORD UNIVERSITY PRESS, 2003. 387 p.

BOGARDUS, Thomas. Knowledge Under Threat. **Philosophy and Phenomenological Research**, v. 88, n. 2, p. 289–313, 2012.

COMESAÑA, Juan. Unsafe knowledge. **Synthese**, v. 146, n.3, p. 395–404, 2005.

DEROSE, Keith. Solving the skeptical problem. **The philosophical review**, v.104, n.1, p. 1–52, 1995.

DEROSE, Keith. Knowledge, Assertion and Lotteries. **Australasian Journal of Philosophy**, v.4, n.74, p. 568-580, 1996.

DRETSKE, Frederick. Conclusive reasons. **Australasian Journal of Philosophy**, v. 49, n. 1, p. 1–22, 1971.

DRETSKE, Frederick. Epistemic operators. **The journal of philosophy**, v. 67, n. 24, p. 1007–1023, 1970.

GETTIER, Edmund L. Is justified true belief knowledge? **Analysis**, v. 23, n. 6, p. 121–123, 1963.

GOLDMAN, Alvin I. Discrimination and Perceptual Knowledge. **The Journal of Philosophy**, v. 73, n. 20, p. 771–791, 1976.

GOLDMAN, Alan H. Nozick on knowledge: finding the right connection. In: LUPER-FOY, S. (Org.). **The possibility of knowledge: Nozick and his critics**. New Jersey: ROWMAN LITTLEFIELD, 1987, p. 182-196.

HARMAN, Gilbert. Knowledge, inference, and explanation. **American Philosophical Quarterly**, v. 5, n. 3, p. 164-173, 1968.

HARMAN, Gilbert. **Thought**. New Jersey: PRINCETON UNIVERSITY PRESS, 1973. 199 p.

KHOO, Justin. **Backtracking counterfactuals, revisited**. Disponível em: <<http://web.mit.edu/jkhoo/www/khoobcrfinal.pdf>>. Acesso em: 20/12/2016.

KLEIN, Peter. On behalf of the skeptic. In: LUPER-FOY, S. (Org.). **The possibility of knowledge: Nozick and his critics**. New Jersey: ROWMAN LITTLEFIELD, 1987. p. 267–281.

KLEIN, Peter. The Nature of knowledge. In: BORGES, R., DE ALMEIDA, C. e KLEIN, P. (Org.). **Explaining knowledge: new essays on the Gettier Problem**. Oxford: Oxford University Press, NO PRELO.

KRIPKE, Saul. Nozick on knowledge. In: \_\_\_\_\_. **Philosophical troubles**. Oxford: OXFORD UNIVERSITY PRESS, 2011. Cap. 7, p. 162-224.

LEWIS, David. **Counterfactuals**. Oxford: BLACKWELL PUBLISHING, 2005. 156 p.

MCEVOY, Mark. The Lottery Puzzle and Pritchard's safety Analysis of Knowledge. **Journal of Philosophical Research**, v. 34, p. 7–20, 2009.

NETA, Ram; ROHRBAUGH, Gilbert. Luminosity and the safety of knowledge. **Pacific Philosophical Quarterly**, v. 85, n. 4, p. 396–406, 2004.

NOZICK, Robert. **Philosophical explanations**. Cambridge, Massachusetts: HARVARD UNIVERSITY PRESS, 1981. 764 p.

PRITCHARD, Duncan. **Epistemic luck**. New York, USA: OXFORD UNIVERSITY PRESS, 2005. 290 p.

PRITCHARD, Duncan. Knowledge, luck, and lotteries. In: HENDRICKS, V.; PRITCHARD, D. (org.). **New Waves in epistemology**. NY: PALGRAVE MACMILLAN, 2008, p. 28-51

PRITCHARD, Duncan. Safety-Based Epistemology: whiter now? **Journal of Philosophical Research**, v. 34, p. 33-45, 2009.

PRITCHARD, Duncan. In defense of modest anti-luck epistemology. In: BECKER, Kelly; BLACK, Tim. (Org.). **The Sensitivity Principle in Epistemology**. Cambridge, UK: CAMBRIDGE UNIVERSITY PRESS, 2012a, p. 173-192.

PRITCHARD, Duncan. Anti-luck virtue epistemology. **The Journal of Philosophy**, v. 109, n. 3, p. 247-279, 2012b.

RABINOWITZ, David. **The safety condition for knowledge**. In: Internet Encyclopedia of philosophy. Disponível em: <<http://www.iep.utm.edu/safety-c/>>. Acessado em 4/07/2016.

RODRIGUES, Lucas Roisenberg. Sobre uma objeção à condição de segurança. In: NEIVA, A. e ROSA, L. (Org.). **Proceedings of the Brazilian research group in epistemology**. Porto Alegre: EDITORA FI, 2016, p. 66-82.

SALERNO, Joseph. Truth tracking and the problem of reflective knowledge. In: CAMPBELL J. et al. (Org.) **Knowledge and Skepticism**. Cambridge, Massachusetts: THE MIT PRESS, 2010, p. 73-85.

SOSA, Ernest. How must knowledge be modally related to what is known? **Philosophical Topics**, v. 26, n. 1, p. 373–384, 1999a.

SOSA, Ernest. How to defeat opposition to Moore. **Philosophical perspectives**, v. 13, p. 141-153, 1999b.

SOSA, Ernest. Tracking, competence, and knowledge. In: MOSER, Paul (Org.). **The Oxford Handbook of Epistemology**. Oxford: OXFORD UNIVERSITY PRESS, 2002, p.264-286.

SOSA, Ernest. **A Virtue Epistemology: Apt Belief and Reflective Knowledge, Volume I**. Oxford: OXFORD UNIVERSITY PRESS, 2007

STALNAKER, Robert. A Theory of Conditionals. In: HARPER, W. et al.(org.) **Ifs: Conditionals, belief, decision, chance and time**. Dordrecht: Springer Netherlands, 1981. cap.1, p. 41–55.

UNGER, Peter. An analysis of factual knowledge. **The Journal of Philosophy**, v. 65, n. 6, p. 157–170, 1968.

VOGEL, Jonathan. Tracking, closure, and inductive knowledge. In: LUPER-FOY, S. (Org.). **The possibility of knowledge: Nozick and his critics**. New Jersey: ROWMAN LITTLEFIELD, 1987, p.197-218.

VOGEL, Jonathan. Reliabilism leveled. **The Journal of Philosophy**, v. 97, n. 11, p. 602–623, 2000.

VOGEL, Jonathan. Subjunctivitis. **Philosophical Studies**, v. 134, n. 1, p. 73–88, 2007.

VOGEL, Jonathan. The enduring trouble with tracking. In: BECKER, K.; BLACK, T. (Org.). **The Sensitivity Principle in Epistemology**. Cambridge, UK: CAMBRIDGE UNIVERSITY PRESS, 2012, p. 122–151.