



**UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA
DOUTORADO EM FILOSOFIA**

CONHECIMENTO E SENSITIVIDADE

Luiz Helvécio Marques Segundo

FLORIANÓPOLIS
2020

Luiz Helvécio Marques Segundo

CONHECIMENTO E SENSITIVIDADE

Tese submetida ao Programa de Pós-Graduação
em Filosofia da Universidade Federal de Santa
Catariana para a obtenção do título de Doutor em
Filosofia

Orientador: Prof. Dr. Alexandre Meyer Luz

Florianópolis
2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Marques Segundo, Luiz Helvécio
Conhecimento e sensibilidade / Luiz Helvécio Marques
Segundo ; orientador, Alexandre Meyer Luz, 2020.
212 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro de Filosofia e Ciências Humanas, Programa
de Pós-Graduação em Filosofia, Florianópolis, 2020.

Inclui referências.

1. Filosofia. 2. Teoria do conhecimento . 3.
Sensibilidade . 4. Epistemologia da Função Apropriada . 5.
Ceticismo e fechamento . I. Meyer Luz, Alexandre. II.
Universidade Federal de Santa Catarina. Programa de Pós
Graduação em Filosofia. III. Título.

LUIZ HELVÉCIO MARQUES SEGUNDO
CONHECIMENTO E SENSITIVIDADE

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Dr. André Joffily Abath
Universidade Federal de Minas Gerais

Prof. Dr. Jonas Becker Arenhart
Universidade Federal de Santa Catarina

Prof. Dr. Jaimir Conte
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Filosofia.

Prof. Dr. Ivan Ferreira da Cunha
Coordenação do Programa de Pós-Graduação

Prof. Dr. Alexandre Meyer Luz
Orientador

Florianópolis, 2020.

Agradecimentos

Ao Prof. Alexandre Meyer, meu orientador, pela gentileza, acolhimento e, sobretudo, paciência.

Aos professores Jaimir Conte, Jonas Arenhart e André Abath por gentilmente aceitarem o convite para a banca. Suas correções, sugestões e objeções foram cruciais para a presente versão da tese. Obviamente que as imprecisões e erros remanescentes são de minha inteira responsabilidade.

Ao Departamento de Filosofia da UFSC pela acolhida e profissionalismo.

Ao SELF – *Seminário Livre de Filosofia* – por permitir que eu apresentasse e discutisse algumas de minhas ideias.

Aos amigos e colegas com quem convivi diariamente em Florianópolis por quase dois anos. Entre sérios debates e desprentiosas conversas aprendi bastante sobre filosofia, ciência e história.

Aos amigos pessoais que ajudaram a tornar mais amenos os meus dias de profunda angústia. Falar sobre música, literatura e sobre “coisas que nós não entende nada” [*sic*] foi essencial; parte considerável do sentido da minha vida repousa nessas atividades.

Eu não poderia deixar de manifestar minha gratidão a dois (dentre muitos) grandes gênios artísticos que embalaram boa parte desta empreitada, Lima Barreto e Gilberto Gil.

Por fim, minha maior dívida é para com Percy, a pessoa com quem divido minha vida. Ela acompanhou pacientemente cada passo deste período, incentivando-me, consolando-me e enchendo cada vez mais meu coração de amor.

Estou sentando com um filósofo no jardim; ele diz repetidas vezes “Sei que aquilo é uma árvore”, apontando para uma árvore que está perto de nós. Alguém chega e escuta isso, e eu lhe digo: “Este sujeito não está louco. Só estamos filosofando”.

Wittgenstein, *Sobre a Certeza*, §467.

RESUMO

O assunto principal desta tese é a natureza do conhecimento empírico. Penso que conhecimento seja, pelo menos, crença verdadeira *sensitiva*: uma crença que não seria formada ou sustentada pelo sujeito caso fosse falsa, mas formada ou sustentada caso verdadeira. Defenderei mais especificamente a tese de que a sensitividade é melhor compreendida no interior da teoria da função apropriada. De acordo com a epistemologia da função apropriada, uma crença verdadeira conta como conhecimento somente se (i) for produzida por um mecanismo funcionando apropriadamente, (ii) num ambiente adequado, (iii) tal mecanismo visar a verdade, (iv) tender a produzir mais crenças verdadeiras do que falsas, e (v) não houver anuladores. Penso que as condições (i) e (ii) só podem ser satisfatoriamente formuladas levando-se em conta a sensitividade, ao mesmo tempo que a sensitividade encontra a sua melhor formulação quando levadas em conta (i) e (ii). Grosso modo, defenderei que a sensitividade é uma questão da crença do sujeito ser responsiva à verdade por ser produto de um mecanismo cognitivo funcionando num ambiente favorável: um ambiente no qual ele não formaria tal crença caso tal crença fosse falsa, mas que, caso verdadeira, a formaria. Começarei, no Capítulo I, com o seguinte *desideratum*: oferecer uma condição anti-sorte ao conhecimento. Argumento que a sensitividade é a principal candidata à condição anti-sorte. No Capítulo II apresento a versão clássica da sensitividade e algumas de suas principais dificuldades. No Capítulo III, proponho a sensitividade como inerente à epistemologia da função apropriada e respondo às dificuldades levantadas no capítulo anterior. No capítulo IV, discuto uma condição anti-sorte rival, a segurança, e argumento que a sensitividade é superior. Por fim, no Capítulo V, trato de outro *desideratum* importante a qualquer teoria do conhecimento, o ceticismo. A adoção da sensitividade fornece uma resposta atraente ao ceticismo, ainda que amplamente considerada implausível, uma vez que parece violar um princípio epistêmico plausível, o princípio de fechamento. Argumentarei que podemos manter um princípio restrito de fechamento e que a sensitividade explica por que podemos fazer tal restrição.

Palavras-chave: conhecimento; sensitividade; função apropriada; epistemologia anti-sorte; ceticismo; fechamento.

ABSTRACT

The subject matter of this dissertation is the nature of empirical knowledge. My view is that knowledge is at least *sensitive* true belief: one that the subject wouldn't form or hold were it false, but he would form it or hold it were it true. I will argue that sensitivity is better understood when stated into the proper function framework. According to proper functionalism, a true belief is knowledge only if (i) it is formed by a cognitive mechanism functioning properly (ii) in a proper environment, and this mechanism (iii) has to aim truth, (iv) having the tendency to produce more true beliefs than false ones, and (v) no defeater being available. I think that conditions (i) and (ii) can't be adequately stated unless we take into account the sensitivity condition. The same is true for sensitivity: its better formulation is stated when we consider (i) and (ii). Roughly, I will argue that the sensitivity of one's belief is a matter of its responsiveness to truth due to the proper function of one's cognitive mechanism in a favorable environment: an environment where the subject wouldn't form or hold the belief were it false, but he would believe it were it true. I begin, in Chapter I, with the following desideratum: to provide an anti-luck condition to knowledge. I argue that sensitivity is the main candidate to such condition. In Chapter II the standard version of sensitivity and some of the main difficulties to it are presented. Then, in Chapter III, I propose to understand sensitivity as inherent to proper functionalism and reply to the difficulties previously presented. A rival of sensitivity, the safety condition, is discussed in Chapter IV. I defend sensitivity over safety. Finally, in Chapter V, another important desideratum to any theory of knowledge is taken in consideration: skepticism. The adoption of sensitivity provides an attractive way to respond to skeptical argument, but at the cost of the rejection of closure principle, a very plausible epistemic principle. I will argue that we can adopt a restrict version of closure and that this restriction is predicted by sensitivity.

Key words: knowledge; sensitivity; proper functionalism; anti-luck epistemology; skepticism; closure.

SUMÁRIO

INTRODUÇÃO	9
I- CONHECIMENTO E SORTE	19
1. O Problema de Gettier	19
2. Eliminando algumas categorias de contraexemplos	24
3. Algumas tentativas de solução	36
4. Sorte epistêmica.....	49
5. Conhecimento gettierizado?	54
II- SENSITIVIDADE	59
1. Rastreamento	59
2. Objeções influentes.....	72
III- SENSITIVIDADE, FUNÇÃO APROPRIADA E RESPOSTAS ÀS OBJEÇÕES	84
1. Os mundos que interessam	84
2. Função apropriada, micro e macro ambientes	93
3. Resposta às objeções	105
IV- MELHOR SEGURO DO QUE SENSITIVO?	140
1. O princípio de segurança	140
2. Segurança e a intuição anti-sorte	141
3. Segurança e confiabilidade.....	149
4. Sensitividade e confiabilidade	156
V- CETICISMO E FECHAMENTO	163
1. O argumento cético.....	164
2. Externismo e ceticismo.....	165
3. Contraexemplos ao fechamento e insensitividade.....	176
4. Fechamento restrito	186
5. Crença razoável em \neg BIV	189
6. Sensitividade e a força da hipótese cética	194
CODA	198
BIBLIOGRAFIA	200

INTRODUÇÃO

Um fato bastante curioso na filosofia é que os dados inegáveis com os quais começamos são sempre um tanto vagos e ambíguos. [...] O processo do sólido de filosofar, a mim me parece, consiste principalmente em passar daquelas coisas óbvias, vagas e ambíguas, das quais nos sentimos tão seguros, a algo preciso, claro e definido [...]

Bertrand Russell, *The Philosophy of Logical Atomism*.

Nada melhor que as palavras de Russell, em epígrafe, para abrir esta tese. O dado inegável que iremos examinar é o conhecimento. Sem qualquer dificuldade e em diversas situações dizemos “Fulano sabe que p ” ou “Sicrano não sabe que p ”. Por exemplo, quando vejo o amável Miguel indicar a um turista onde fica a Rua São José, não vacilo em atribuir-lhe o conhecimento da localidade da Rua São José. Contudo, ao vê-lo diagnosticar com gripe aviária esse mesmo turista, que espirrara momentos antes de pedir-lhe a informação, nego sumariamente que Miguel saiba que o turista tenha gripe aviária (ainda que tivesse de fato). Muito provavelmente diríamos que Miguel possui certa competência para formar crenças verdadeiras sobre as localidades de Ouro Preto, ao passo que não possui competência para formar crenças específicas verdadeiras sobre a saúde das pessoas. Mas se acaso nos pedissem uma caracterização “precisa, clara e definida” de que competência é essa que estamos falando, tão logo descobriríamos o quão vaga é tal noção. É uma competência intelectual, *reflexiva* como diriam os filósofos; ou uma competência animal, *pré-reflexiva*; ou quem sabe algo entre essas duas? É algo que se adquire com a prática, uma exclusividade dos humanos, ou é também um traço de outras espécies? A falta de uma resposta clara e não flutuante a essas e outras questões relacionadas é suficiente para ilustrar o quão o fenômeno do conhecimento, nosso “dado inegável”, é “vago e ambíguo”.

O meu objetivo nesta tese é dar alguns modestos passos na tentativa de compreensão do conhecimento. Mais especificamente abordarei as seguintes questões intimamente relacionadas: (i) Que propriedade (ou conjunto de propriedades) – se é que há – que quando

adicionada à crença verdadeira a torna conhecimento? (ii) Será o conhecimento um fenômeno parcialmente modal? E (iii) que tipo de resposta – se alguma – podemos oferecer ao ceticismo quanto ao mundo externo uma vez que tenhamos respondido a (i) e (ii). A maior parte desta tese tratará de (i) e (ii). Abordarei (iii) apenas no último capítulo.

Partirei de um ponto de vista externista, e que pode ser genericamente visto como um tipo de teoria confiabilista do conhecimento. Defenderei que o conhecimento de proposições empíricas é *crença verdadeira sensitiva*. Dessa perspectiva, filio-me às teorias rastreadoras que remontam pelo menos a Fred Dretske (1971), Alvin Goldman (1976) e Robert Nozick (1981). Por outro lado, sustentarei que a sensitividade é uma questão de uma crença ser produzida por um mecanismo cognitivo funcionando apropriadamente, o que me liga ao perspectivismo de virtudes de Ernest Sosa (1991), mas principalmente à epistemologia da função apropriada de Alvin Plantinga (1992). À perspectiva que emergirá nestas páginas podemos chamar “Epistemologia da Função Apropriada Sensitiva”, “EFAS” doravante.

A versão padrão da sensitividade pode ser posta da seguinte forma (NOZICK 1981, pp. 172-178). A crença de S de que α é sensitiva se e somente se:

- (i) se não fosse o caso que α , então S não acreditaria que α ; e
- (ii) se fosse o caso que α , então S acreditaria que α .

À condição (i) chamamos *variação* e à (ii) chamamos *aderência*. Assim, a condição de sensitividade ao conhecimento é composta por duas sub-condições: variação e aderência. Além disso, é preciso também especificar o modo pelo qual o sujeito forma a sua crença. Nozick (pp. 179-185) sugere a especificação de um método de formação de crenças, de modo que a satisfação de (i) e (ii) dependerá sempre de se manter o método fixo.

Nozick interpreta as condicionais contrafactuais em (i) e (ii) através da semântica dos mundos possíveis. Nesse sentido, podemos dizer que a teoria da sensitividade encara o conhecimento como um fenômeno modal. Diremos, por exemplo, que (i) é verdadeira se, e somente se, *nos mundos mais próximos em que α não é o caso, S não acredita que α* . A ideia de proximidade entre mundos é crucial. A menos que tenhamos uma maneira não arbitrária de selecionar os mundos que contam como próximos, não teremos uma maneira satisfatória de dizer se S sabe ou não que α . Uma das dificuldades enfrentadas pela sensitividade padrão é

parecer não haver uma maneira clara de se determinar o conjunto de mundos próximos ou mundos relevantes. É neste ponto que proponho vermos a sensibilidade por outro ângulo, considerando primeiro a epistemologia da função apropriada.

De acordo com a epistemologia da função apropriada (PLANTINGA 1993, pp. 3-20), uma crença é conhecimento somente se:

- (i) foi produzida por um mecanismo cognitivo funcionando apropriadamente,
- (ii) num ambiente adequado;
- (iii) esse mecanismo visa a verdade;
- (iv) tende a produzir mais crenças verdadeiras do que falsas.

As condições (i) e (ii) são melhor compreendidas quando levamos em conta a noção de plano funcional ou desenho [*design plan*] de um mecanismo cognitivo. Parte de nossos mecanismos cognitivos – *e.g.*, a percepção – foram moldados evolutivamente (ou mesmo por Deus, não importa) para produzir representações sobre o nosso entorno. Nesse sentido, o mecanismo funciona bem ou mal de acordo com o seu plano funcional; esperamos que nossa percepção produza representações um tanto acuradas do nosso entorno. Mas o funcionamento apropriado de um mecanismo depende também do ambiente no qual ele esteja operando. Uma águia, por mais que tenha visão acuradíssima, ainda assim pode não formar uma representação boa o bastante de seu entorno se estiver no completo breu da noite. Por isso, o funcionamento a apropriado depende bastante do ambiente no qual o mecanismo opera. As condições (iii) e (iv), por sua vez, garantem que o mecanismo não seja confiável por acaso.

Foquemo-nos, porém, no ambiente. A completa escuridão em um descampado não torna o ambiente favorável para a formação de crenças visuais em humanos. Suponha que S veja um vulto na escuridão e acredite que seja um cachorro. Mas S não sabe precisar o tamanho do vulto, seus contornos, etc. Se fosse um cachorro, seria um grande golpe de sorte S ter acertado. A sua crença teria sido verdadeira acidentalmente, por pura sorte. O normal a se pensar é que S muito provavelmente errara, que formara uma crença falsa. Num ambiente desfavorável, S tem a tendência de formar mais crenças falsas do que verdadeiras. Assim, um ambiente favorável será aquele em que S tenderá a formar mais crenças verdadeiras. Isto é, se sua visão estiver funcionando apropriadamente, (a) quando for o caso de haver um cachorro em sua frente, S acreditará que há um cachorro em sua frente; e (b) se não for o caso de haver,

S não acreditará haver tal cachorro. Em outras palavras, um ambiente favorável é aquele em que a crença do sujeito é sensitiva. Ou para usar a metáfora nozickiana, um ambiente em que a crença do sujeito rastreia a verdade.

Uma das principais atrações das teorias rastreadoras, principalmente a de Nozick, é a sua maneira de tratar o ceticismo quanto ao mundo externo (ou ceticismo radical). Frente a sucessivos fracassos dos epistemólogos na tentativa de lidar com o problema cético, a teoria rastreadora é iluminante: ao mesmo tempo em que reconhecia a extrema dificuldade de se refutar certas hipóteses e cenários céticos – *e.g.* o do gênio maligno cartesiano, o do cérebro numa cuba, etc. –, explicava por que temos conhecimento do mundo externo. O diagnóstico é o seguinte: crenças empíricas sobre o mundo externo são geralmente sensitivas, ao passo que crenças em negações de hipóteses céticas são insensitivas. Assim, sabemos várias proposições sobre o mundo externo, mas não sabemos que não estamos num cenário cético apenas acreditando falsamente naquilo que acreditamos. Isso tudo a despeito de reconhecermos que uma proposição sobre o mundo externo, *e.g.* ⟨Tenho mãos⟩, implica a negação da proposição ⟨Sou um cérebro numa cuba sendo enganado para acreditar que tenho mãos⟩. Esse resultado parece invalidar o seguinte princípio epistêmico:

(PF) Se sei que α , e sei que α implica β , então sei que β .

Contudo, se PF for inválido, como explicar a ampliação do nosso estoque de conhecimento através do raciocínio dedutivo? Por exemplo, se sei que misturar antidepressivos com álcool provoca alterações psíquicas, e vejo João tomar antidepressivos com álcool, é razoável dizer que sei que João terá alterações psíquicas. Contudo, se PF é inválido, o que explica esse caso? Muitos epistemólogos pensam ser essa uma razão derradeira para se abandonar a teoria da sensitividade.

Novamente proponho que vejamos a sensitividade por outro ângulo. Ao invés de apenas pensar naquilo que viola o fechamento, pensemos também naquilo que não viola. A minha sugestão é que a sensitividade é capaz de oferecer não apenas uma explicação dos casos em que há ampliação de conhecimento via dedução, como também explica a falha de ampliação via dedução. Há um conjunto uniforme de proposições insensitivas – no qual as hipóteses céticas radicais se incluem – para as quais não deveríamos esperar ter conhecimento. Não obstante, podemos ter crenças razoáveis acerca delas.

No **Capítulo I** começarei por preparar o terreno para a introdução da sensibilidade na investigação epistemológica. Nenhum lugar seria melhor do que o problema de Gettier. Se há alguma grande lição a ser tirada do problema de Gettier, é sem dúvida a de que conhecimento e sorte são incompatíveis. Acredito que não podemos começar a tentar investigar a natureza do conhecimento sem antes reconhecer esse dado. E não podemos desenvolver qualquer teoria satisfatória caso não sejamos capazes de encontrar uma condição anti-sorte que não seja *ad hoc*. Argumentarei que dentre os casos à la Gettier, apenas casos com a mesma estrutura do famoso caso dos celeiros falsos é que apresentam dificuldades sérias a uma teoria do conhecimento. Esses casos evidenciam que o principal problema com a crença verdadeira justificada adquirida por sorte está no ambiente. Isso, por sua vez, sugere o tipo de condição anti-sorte que deveríamos esperar. Penso que a sensibilidade é uma boa (provavelmente a melhor) candidata.

O **Capítulo II** é em geral bastante expositivo: apresento a versão padrão da condição de sensibilidade e algumas das objeções que considero centrais a ela. Há basicamente dois tipos de objeções: as que consideram a sensibilidade demasiado estrita, e as que a consideram demasiado permissiva. No primeiro tipo, levantam-se contraexemplos que apresentam casos putativos de conhecimento que são erroneamente classificados como falha de conhecimento pela sensibilidade. No segundo, levantam-se contraexemplos que apresentam casos putativos de falha de conhecimento que são erroneamente classificados como conhecimento pela sensibilidade. As respostas a essas objeções são adiadas até o próximo capítulo.

No **Capítulo III** esboço uma maneira de se determinar o conjunto de mundos possíveis próximos para avaliar a contrafactual da sensibilidade. Essa maneira leva em conta principalmente as condições ambientais presentes no momento em que o sujeito faz uso de seus mecanismos cognitivos. Os mundos próximos ou mundos relevantes são aqueles em que as condições ambientais são aproximadamente as mesmas que a do mundo atual (o mundo em que o sujeito supostamente possui conhecimento) e o sujeito faz uso do mesmo mecanismo cognitivo. Essas restrições constituem aquilo a que chamamos de ambiente favorável. A sensibilidade é vista, por conseguinte, como um desenvolvimento natural da epistemologia da função apropriada. É aqui que a EFAS surge. Por fim, respondo às objeções apresentadas no capítulo precedente.

O **Capítulo IV** é voltado para a discussão da condição anti-sorte rival da sensibilidade, a condição de segurança. Grosso modo, a crença de S de que α é segura se e somente se: se S acreditasse em α , α seria o caso. Alternativamente, se nos mundos possíveis mais próximos em que S acredita que α , α é o caso. A discussão será bastante esquemática. Considerarei dois importantes argumentos a favor da segurança. O primeiro se concentra na segurança como condição anti-sorte e o segundo na segurança como captando nossas intuições sobre a confiabilidade. Argumentarei que ambos falham no mesmo ponto: não são capazes de restringir o conjunto de mundos próximos adequadamente. Por outro lado, a condição de aderência é capaz de fazer o trabalho.

Ao contrário da ortodoxia, não penso que a teoria do rastreamento baseada na sensibilidade tenha resultados tão negativos no que diz respeito ao fechamento epistêmico. Mas isso não quer dizer que eu endosse ingenuamente algum princípio de fechamento *simpliciter*. Há já dificuldades conhecidas na bibliografia que *prima facie* apresentam problemas para o fechamento: são os casos das proposições lotéricas e da mula disfarçada. Argumento então, no **Capítulo V**, que juntamente com esses tipos de proposições, proposições descrevendo cenários céticos formam uma classe especial de proposições insensitivas para as quais o fechamento falha. As proposições desse conjunto podem no máximo ser razoavelmente acreditadas. Por outro lado, o conjunto formado pelas nossas crenças de senso comum e científicas – na medida em que são conhecimento – é fechado sob implicação conhecida. Assim, temos uma razão para formular um princípio de fechamento restrito a crenças sensitivas. Dada a sensibilidade, esse é o máximo que podemos almejar. Penso que o desafio cético, mesmo que modestamente respondido, nos mostra de maneira dramática os limites do conhecimento humano. E nesse caso, parte desse limite envolve proposições muito gerais sobre o nosso próprio conhecimento.

Eu disse que meu ponto de partida é o externismo. Posto da maneira mais genérica possível, o externismo é a tese de que o conhecimento (ou qualquer status epistêmico positivo de uma crença) depende de condições externas ao sujeito e que em geral não são diretamente acessíveis em primeira pessoa. E sendo esse o meu ponto de partida, não oferecerei qualquer

argumento contra o internismo ou mesmo qualquer argumento a favor do externismo.¹ Esse é um debate interminável e para o qual não deveríamos esperar consenso (aliás, não deveríamos esperar consenso no que quer que seja em filosofia no tocante a questões tão fundamentais!). Não tenho o que adicionar a esse debate, e minha preferência pelo externismo se deve, sobretudo, pela rejeição de certos pressupostos internistas que aparentemente tornavam a investigação sobre o conhecimento extremamente difícil e pareciam cada vez mais empurrar a epistemologia para águas profundas e obscuras.² Por exemplo, considere o problema da indução. Quando a reflexão sobre o problema da indução parecia já não ter mais para onde ir, principalmente porque os filósofos temiam a maldição da circularidade, o externismo foi capaz de jogar nova luz sobre argumentos auto-apoiadores (cf. van CLEVE 1984) e sobre justificação circular da confiabilidade de certas fontes crença (cf. ALSTON 1986). Ou considere o rumo obscuro que tomaram as discussões sobre a estrutura do conhecimento depois do ataque ao mito do dado feito por Wilfrid Sellars (1956) e que de algum modo culminaram nas posições radicais como as de Richard Rorty (1979). As análises externistas do conhecimento perceptual (cf. GOLDMAN 1967, 1976, 1979, 1986; DRETSKE 1981; ALSTON 1993), o reconhecimento da importância dos estudos empíricos para epistemologia, etc. permitiram novas versões de fundacionismo e deu novo fôlego ao debate sobre as bases e a estrutura do conhecimento (cf. SOSA 1980; 2007; PLANTINGA 1993b; 2000). Talvez, a mais importante conquista do programa externista tenha sido o desemaranhamento da noção de justificação, reconhecendo-se pelo menos três aspectos distintos: racionalidade, direito a uma crença, e garantia (cf. GOLDMAN 1986, 1988; ALSTON 1989, 2005; PLANTINGA 1993a). Não menos importante – e, de um ponto de vista institucional, uma grande conquista – foi ter impulsionado sobremaneira o desenvolvimento da epistemologia social verídica (cf. GOLDMAN 2006).

Como já mencionado, meu interesse nesta tese está voltado completamente para o conhecimento. Sendo assim, o externismo com o qual me comprometo se restringe ao

¹ Ainda hoje o grosso desse debate pode ser encontrado em *Epistemology: Internalism and Externalism*, ed. Hilary Kornblith. Blackwell (2001).

² Veja LEITE (2005). Adam Leite argumenta que o externismo mina o projeto epistemológico tradicional, *i.e.*, aquilo que pode ser visto como o projeto cartesiano, que “coloca entre parênteses todas as nossas crenças sobre o mundo, suspendendo a nossa aptidão comum de utilizá-las em nossos esforços intelectuais, na tentativa de, a partir dessa posição privilegiada, explicar o nosso conhecimento do mundo” (p. 505).

conhecimento. Dele nada de substancial se seguirá para as teorias da justificação, a não ser que a justificação seja entendida de maneira bastante estrita, como *aquela propriedade que quando adicionada em grau suficiente à crença verdadeira a torna conhecimento*.

Pelo menos desde o início do século XX, os filósofos tenderam a pensar que o conhecimento fosse essencialmente crença verdadeira falivelmente justificada. Foi natural, portanto, pensar que a justificação – o que quer que ela fosse exatamente – era aquele elemento que adicionado à crença verdadeira a tornava conhecimento. Nesse sentido, a justificação seria aquilo que confere algum status epistêmico positivo à crença verdadeira, e que em grau suficiente a torna conhecimento. Plantinga (1993a) chamou a esse status epistêmico positivo “garantia”. Porém, ele também mostrou de maneira bastante convincente que o termo “justificação” era usado num sentido deontológico para denotar o direito do sujeito a uma crença, ou para denotar a racionalidade do sujeito em sustentar uma crença. Na acepção deontológica, ter uma crença verdadeira justificada não implica que o sujeito tenha conhecimento, uma vez que nem sempre podemos culpar um sujeito ou chamá-lo de irracional por sustentar uma crença verdadeira por sorte (ou por acidente). Nesta tese nada será dito acerca da natureza da justificação no sentido deontológico. Contudo, assumirei que em muitos casos uma crença possuir garantia para um sujeito será suficiente para que essa crença seja racional, razoável, permitida, etc. Além disso, muitas crenças razoáveis podem ter pouca ou nenhuma garantia e serem insensitivas.

Na medida em que falar de conhecimento envolver falar de “boa evidência”, ou “boa base” ou “razão conclusiva”, etc. a favor de uma crença, estarei preocupado apenas com a garantia dessa crença, *i.e.*, com aquela propriedade objetiva que em grau suficiente torna a crença verdadeira em conhecimento. Sendo assim, ter “boa evidência”, etc. para uma crença não implicará que o sujeito seja capaz de acessar introspectivamente, ou mesmo avaliar, essa evidência. Mais especificamente, ao dizermos que a crença de um sujeito é *sensitiva* ou *confiável* ou *apropriadamente formada*, etc., isso não implicará que o sujeito seja capaz de dizer que a sua crença seja sensitiva, ou confiável, ou apropriadamente formada.

Ao falar de conhecimento, estarei na maior parte falando sobre conhecimento de primeira-ordem, ou para usar a feliz expressão de Sosa, de *conhecimento animal*. Pouco, ou quase

nada, direi sobre o *conhecimento reflexivo*, ou conhecimento de ordem superior. Isso porque EFAS é uma teoria que explica como as crenças que o sujeito forma sobre seu entorno contam como conhecimento, e não como as crenças que o sujeito forma sobre as suas crenças contam como conhecimento. Do fato de termos uma teoria sobre o conhecimento de primeira ordem não se segue que essa mesma teoria tenha de dar conta do conhecimento de segunda ordem. Uma analogia pode ajudar. Por exemplo, a lógica de primeira ordem explica por que certas fórmulas de primeira ordem – fórmulas que quantificam sobre objetos – são consequência lógica de outras fórmulas de primeira ordem. Mas não explica por que certas outras fórmulas de segunda ordem – fórmulas que quantificam sobre propriedades – são consequência lógicas de outras fórmulas de segunda ordem. E assim como um lógico não exige uma teoria de segunda ordem de outro lógico só por que ele fala sobre propriedades quando desenvolve uma teoria de primeira ordem, um epistemólogo não precisa exigir que uma teoria do conhecimento de primeira ordem venha acompanhada de uma teoria do conhecimento de segunda ordem.³ Na verdade, como veremos, penso que algumas crenças de segunda ordem são insensitivas – mas especificamente crenças sobre o funcionamento adequado dos nossos mecanismos cognitivos no momento de seu funcionamento. Esse resultado, em partes, nos será útil para responder ao argumento cético.

Sob pena de ser acusado de pouca modéstia epistêmica, procurarei olhar para nossas práticas epistêmicas do ponto de vista de terceira pessoa, esquecendo-me nesses instantes que estou em grande parte investigando a mim mesmo.⁴ Assumirei a postura de um naturalista extraterrestre que desembarca na Terra a fim de estudar conceitualmente as práticas cognitivas de mais uma espécie animal – os *homo sapiens*. E por falta de tempo, energia e habilidade se concentra apenas naquelas práticas mais corriqueiras e comuns em que observa

³ É claro que a lógica de segunda ordem é uma extensão da lógica de primeira ordem, o que significa que ela usa o vocabulário da lógica de primeira ordem mais o seu vocabulário novo. Porém, não tem os mesmos resultados metateóricos. O mesmo pode ser o caso com as teorias do conhecimento. Ainda que desenvolvamos uma teoria do conhecimento de segunda ordem que seja uma extensão da teoria de primeira ordem, pode ser o caso de termos resultados substanciais diferentes.

⁴ A acusação de circularidade aqui não seria injusta. Como posso explicar que o conhecimento empírico é possível ao mesmo tempo em que pressuponho que já o tenho? Como posso falar do funcionamento apropriado de um mecanismo cognitivo sem antes depender de informação produzida por esse mecanismo? Do ponto de vista externista não há outra maneira de fazê-lo. É como explicar para alguém como é a linguagem possível; não é possível fazê-lo a não ser através da linguagem. Mas alguns insistiriam num ponto de desanalogia: ao passo que não se pode pensar na linguagem sem pressupor uma linguagem, podemos pensar no conhecimento empírico apenas pressupondo o conhecimento *a priori*. Nesse caso, só seria legítimo falar sobre o conhecimento empírico depois que já tivéssemos uma garantia *a priori* de sua existência. Mas isso é demasiado forte, e muito provavelmente inalcançável. Veja a nota 2. Veja também SOSA 1994.

os humanos a dizerem “ p ”, “Sei que p ”, “Ele não sabe que p ”, e assim por diante. E como um observador extraterrestre, suporei que esses animaizinhos têm conhecimento e que, do ponto de vista filosófico, o mais importante é explicar como esse conhecimento é possível. Tendo oferecido tal explicação, o viajante pode agora se colocar no lugar desses animais e pensar sobre como eles reagiriam se lhes fosse dito que todo o seu esforço cognitivo poderia muito bem ser produto de uma grande ilusão produzida por um gênio maligno ou um cientista maluco que lhes tirou o cérebro do crânio e os cria dentro de uma cuba.

I

CONHECIMENTO E SORTE

Um estudo do problema de Gettier levará a um aumento na sensibilidade a vários tipos de exemplos que são importantes para a avaliação crítica de qualquer análise do conhecimento.

Robert Shope, *The Analysis of Knowing*.

[Os epistemólogos], enquanto um grupo, estão confiantes de que o conhecimento é imune a casos à la Gettier. Porém, enquanto grupo, estão também pouco confiantes de que realmente compreendem por que o conhecimento é imune a tais casos.

Stephen Hetherington, *How to Know*.

Prévia

Conhecimento e sorte são incompatíveis. Qualquer teoria do conhecimento que se preze tem de fazer jus a esse dado. O principal objetivo deste capítulo é tentar explicar por que várias teorias do conhecimento falharam na tarefa de oferecer uma condição anti-sorte ao conhecimento. Começo pelo Problema de Gettier. Em seguida, apresento algumas tentativas influentes de resposta e suas principais falhas. Defenderei que as teorias contrafactuais (ou modais) do conhecimento são as melhores candidatas para lidar com o problema (ou pelo menos parte dele).

1. O Problema de Gettier

O problema de Gettier é muitas vezes visto como o problema de se definir satisfatoriamente o conhecimento. No seminal “Is Justified True Belief Knowledge” (1963), Edmund Gettier oferece seus dois famosos contraexemplos ao que ele chamou de a “definição tradicional do conhecimento” como crença verdadeira justificada (CVJ).⁵ Deu-se início então a uma extensa

⁵ Gettier entende a justificação de maneira falibilista. Contudo, a definição tradicional – uma concepção mais ou menos difundida ao longo da história da filosofia – parece não ter adotado o falibilismo. Muito pelo contrário; como argumenta Julien Dutant (2015), a concepção tradicional de conhecimento parece ser majoritariamente

discussão na tentativa de oferecer uma definição do conhecimento que não fosse susceptível a contraexemplos à la Gettier.⁶ Como é de praxe em filosofia, muita tinta foi gasta, embora pouca concordância tenha emergido.

Contudo, a caracterização acima não é exata. Linda Zagzebski (1994) nos deixa bastante claro que uma das principais lições a se aprender com o problema de Gettier é que conhecimento e sorte são incompatíveis, uma vez que casos à la Gettier são sempre casos em que a crença do sujeito é verdadeira por sorte. De modo que, para afastar contraexemplos à la Gettier, é suficiente que se bloqueie crença verdadeira por sorte. Mas daí não se segue termos oferecido uma definição adequada de conhecimento. Portanto, o problema de Gettier não se reduz ao problema da definição do conhecimento.

Podemos dividir o problema de Gettier em dois sub-problemas: (i) o problema analítico, e (ii) o problema anti-sorte.⁷ O problema analítico é o de oferecer uma definição (reduziva ou não) suficientemente informativa que seja imune a casos à la Gettier. Já o problema anti-sorte, como já diz o próprio nome, é o de oferecer uma condição anti-sorte ao conhecimento a fim de bloquear casos à la Gettier. Uma solução para o problema analítico implica uma solução ao problema anti-sorte, uma vez que uma definição adequada de conhecimento tem de ser imune a contraexemplos. Mas o contrário não vale: resolver o problema anti-sorte pode ser feito independentemente. Neste capítulo darei maior atenção ao problema anti-sorte. Embora eu pense que a EFAS (a Epistemologia da Função Apropriada Sensitiva) capte bastante bem o conhecimento empírico, não me arriscarei em defender que ela seja capaz de lidar com o problema analítico no seu todo – EFAS é no máximo uma análise do conhecimento empírico. E como veremos nos capítulos vindouros, tal análise conterà entre suas condições a sensibilidade.

1.1. Estrutura e taxonomia

infalibilista. É claro que Gettier refutou CVJ; mas CVJ não é a concepção tradicional, é a concepção falibilista adotada até então por Ayer, Chisholm e Malcolm. Sobre a concepção tradicional histórica veja também LE MORVAN (2017).

⁶ Usarei a seguinte terminologia para me referir aos contraexemplos: “casos à la Gettier” para me referir de modo geral, incluindo os dois casos originais, e “casos de Gettier” especificamente para os dois casos originais.

⁷ Cf. PRITCHARD (2016; 2015; 2012).

Ambos os casos apresentados por Gettier mostram um sujeito com crença verdadeira justificada⁸ mas que intuitivamente não possui conhecimento. Esses dois exemplos, porém, não são casos raros; na verdade, depois de sua publicação, eles se multiplicaram. Há uma fôrma para criar casos à la *Gettier*. A maneira mais comum de gerá-los é simples.⁹ Um sujeito adquire uma crença justificada que (i) por *má sorte* é falsa ou muito facilmente poderia ter sido falsa, mas que (ii) por conta de um elemento de *boa sorte* se torna verdadeira. Zabzebski chama a tal fôrma “estrutura de dupla sorte”. Diremos de uma crença verdadeira justificada que aparentemente não é conhecimento que ela foi “gettierizada”; e ao processo de tornar uma crença gettierizada chamaremos “gettierização”.

Blouw, Buckwater e Turri (2017) propõem uma taxonomia capaz de refinar a estrutura de dupla sorte: “nessa estrutura básica [...] pode haver *tipos* bastante diferentes de má e boa sorte”:¹⁰

O enquadramento [para se pensar sobre as diferenças entre os casos à la Gettier] se baseia em três elementos estruturais ao longo dos quais os casos à la Gettier variam [...]. A primeira dimensão diz respeito a se um agente forma uma crença por estar diretamente em contato com um “verificador” [*truth-maker*], ou um estado de coisas que torna a sua crença verdadeira. Em alguns casos à la Gettier, o agente percebe ou infere a existência de um verificador, mas em outros casos, não. A segunda dimensão diz respeito se o contato do agente com o verificador inicial é preservado ao longo do tempo. Em alguns casos, o verificador é ameaçado embora permaneça intacto em última instância. Em outros casos, a ameaça é bem sucedida e o verificador inicial é substituído por outro. A terceira dimensão, por fim, diz respeito ao grau de semelhança entre o verificador inicial e seu substituto. Em alguns casos ambos são altamente similares. Em outros, os dois são altamente dissimilares. Por brevidade, nos referiremos às essas três dimensões estruturais em termos de *deteção*, *ameaça* e *substituição*. (p. 243)

⁸ O termo “justificação” refere-se aqui apenas àquela propriedade que supostamente diferencia o conhecimento da mera crença verdadeira, ou em outras palavras, que se adicionada à crença verdadeira a torna conhecimento. Nos últimos anos a noção de justificação tem sido assunto de bastante controvérsia, sobretudo devido ao trabalho de Alvin Plantinga (1993a; 1993b). O termo justificação, argumenta Plantinga, oscila entre um sentido deontológico – de ter o direito de acreditar – e um sentido de aquilo que em grau suficiente torna a crença verdadeira conhecimento. Ele prefere usar o termo “garantia” para essa última propriedade. Contudo, usarei “justificação” e “garantia” quase sempre como sinônimas. Usarei também “evidência” com um sentido bastante próximo.

⁹ Cf. ZAGZEBSKI (1994; 1996; 2009).

¹⁰ Eles argumentam que essa taxonomia releva não haver algo como “o caso à la Gettier”; na verdade, pensam eles, há tipos de casos à la *Gettier* e “há grupos de casos à la *Gettier* que quase nada têm em comum um com o outro” (p.243). Ademais, eles argumentam que “a noção de um ‘caso à la *Gettier*’ deve ser abandonada e substituído por uma gama de tipos de casos mais refinados” (p. 243). Na falta de tal terminologia, porém, mantereí a expressão “caso à la *Gettier*” e a qualificarei quando necessário.

De acordo com essas três dimensões – *detecção*, *ameaça* e *substituição* – eles apresentam 5 diferentes categorias que se enquadram entre a categoria de conhecimento, numa ponta, e a categoria da ignorância, na outra.¹¹ Vejamos:

- **Conhecimento:** Detecção sem Ameaça
- **Categoria 1:** Detecção + Ameaça mal sucedida
- **Categoria 2:** Detecção + Ameaça bem sucedida + Substituição Similar
- **Categoria 3:** Detecção + Ameaça bem sucedida + Substituição Dissimilar
- **Categoria 4:** Falha de Detecção + Substituição Similar
- **Categoria 5:** Falha de Detecção + Substituição Dissimilar
- **Ignorância:** Falha de Detecção + Nenhum Reparo

Os exemplos originais de Gettier se enquadram nas categorias 4 e 5.

[MOEDAS]¹² Smith e Jones estão concorrendo a uma vaga para um emprego. Smith escuta a secretária do chefe dizer que Jones ficará com o emprego. No momento em que esperavam na recepção, Smith vê Jones contando 10 moedas e colocando-as no bolso. Smith forma então a crença de que a pessoa que vai ficar com o emprego tem 10 moedas no bolso. Contudo, após a entrevista Smith se surpreende ao ser anunciado que ele é quem ficaria com o emprego. Além disso, sem que soubesse, sua mulher tinha lhe botado dez moedas no bolso antes de sair de casa.

Smith tem uma crença verdadeira justificada que não conta como conhecimento. A sua crença falha em detectar a verdade, pois Smith não está em contato com o verificador da sua crença de que quem ficará com o emprego tem 10 moedas no bolso. O que torna essa crença verdadeira é o fato de o próprio Smith ter 10 moedas no bolso e ter ficado com o emprego, e não o fato de Jones ter dez moedas no bolso. E foi a substituição desse segundo fato pelo primeiro que fez com que a crença de Smith se tornasse verdadeira. Tal substituição é similar, uma vez que pouca coisa teve de ser alterada no cenário.

Vamos ao segundo caso.

¹¹ Allan Hazlett (2015, p. 2) classifica os casos *à la* Gettier em três tipos que acabam por corresponder a três das categorias de Turri et. al.

¹² Sempre que mencionar casos *à la* Gettier, usarei um nome em caixa alta. Assim, doravante sempre que eu usar MOEDA estarei me referindo ao caso acima. O mesmo será adotado para o restante dos casos.

[FORD] Smith vê, com frequência, seu colega Jones dirigindo um Ford. Ele forma então a crença de que Jones tem um Ford. Em adição a isso, ao se lembrar de outro colega, o Brown, Smith forma a crença disjuntiva de que ou Jones tem um Ford ou Brown está em Barcelona. Acontece que Jones não tem um Ford, o carro é emprestado, mas por um golpe de sorte Brown está passando férias em Barcelona.

A única diferença entre FORD e MOEDAS está na substituição. O fato de Jones ter um Ford – o que Smith esperava que tornasse sua crença verdadeira – é bastante diferente do de Brown estar em Barcelona – o que tornou de fato a sua crença verdadeira – de modo que num sentido relevante, a substituição é dissimilar.

Vejam um exemplo da Categoria 1:

[CELEIRO] João está dirigindo pelo interior e, sem que saiba, está passando por um distrito em que as pessoas têm a mania de construir celeiros falsos, de fachada. Esses celeiros são visualmente indistinguíveis de celeiros reais. João olha para um dos celeiros à beira da estrada e forma a crença verdadeira de que há um celeiro em sua frente. Por sorte, João está olhando para o único real daquela região.

João está em contato com o verificador de sua crença – o fato de ali estar um celeiro real é o que torna a sua crença verdadeira. O ambiente, contudo, ameaça a verdade da crença de João. Há muitos celeiros falsos naquela região; por uma questão de não ter olhado alguns metros à frente, João não formou uma crença falsa. E é claro aqui que não há qualquer substituição, uma vez que a ameaça é mal sucedida.

E como exemplo da Categoria 2:

[CANETA] Katie terminou de usar uma caneta Bic azul, a deixou em cima da mesa e foi tomar banho. Ela acredita verdadeiramente que há uma caneta azul em cima da mesa. Contudo, enquanto ela está no banho, um gatuno entra em sua casa e, nesse meio tempo, enfia a caneta no bolso. Antes de sair ele acaba deixando no lugar outra caneta Bic azul, indistinguível da caneta de Katie. (Adaptado de STARMANS & FRIEDMAN, 2012, p. 276).

Katie forma a crença de que há uma caneta azul em cima da mesa olhando diretamente para a caneta, de modo que há detecção. Contudo, quando entra o ladrão e coloca a caneta no bolso, a sua crença é ameaçada (má sorte). A ameaça é bem sucedida, pois caso o ladrão não tivesse colocado lá outra caneta, a crença de Katie se tornaria falsa. Mas quando o ladrão a substitui

por outra caneta visualmente indistinguível (boa sorte), a sua crença de que há lá uma caneta Bic azul parece se tornar verdadeira (ou pelo menos não foi tornada falsa). A substituição é claramente similar.

E como exemplo da Categoria 3:

[PAPAGAIO] Sebastião é um tabelião aposentado e se lembra de ter registrado o imóvel mais caro da cidade em nome do Sr. Silva, com 30 anos na época. Ele faz as contas de quanto tempo tem que fez esse registro e forma a crença de que o imóvel mais caro da cidade pertence a alguém com 50 anos. Contudo, após a aposentadoria de Sebastião, o Sr. Silva, por extravagância, transferiu o imóvel para o seu animal de estimação, um papagaio que na época tinha exatamente a sua idade.

Sebastião forma uma crença inicialmente verdadeira em que a detecção está envolvida. Porém, a transferência do imóvel para o papagaio torna falsa a crença, que vem a se tornar novamente verdadeira graças à idade do papagaio. O estado de coisas – um papagaio de 50 anos ser titular de um imóvel – que substituiu o estado de coisas original – o Sr. Silva ser titular do imóvel – é bastante dissimilar.

2. Eliminando algumas categorias de contraexemplos

Antes de passarmos a algumas influentes tentativas de se resolver o problema da sorte epistêmica, argumentarei que os casos que se enquadram nas categorias 2-5 não são contraexemplos genuínos; por conseguinte, apenas a Categoria 1 nos fornecerá caos à la Gettier genuínos.¹³ Minha estratégia será principalmente negativa: argumentarei que a crença final atribuída aos sujeitos (nos casos das categorias 2-5) não é adequada ao contexto.

Uma reação natural a isso seria retorquir que é supérfluo tentar eliminar alguns contraexemplos à la Gettier, uma vez que é suficiente que haja uma categoria efetiva de contraexemplos. Não importa que os casos originais não funcionem, temos o caso do cão; e se por ventura o caso do cão não funcionar, temos o do celeiro; e assim por diante. De fato, foi assim que os epistemólogos pós-Gettier tenderam a proceder.¹⁴ Mas isso pode ter-lhes turvado a vista para o aspecto relevante da gettierização. A preocupação em evitar casos espúrios

¹³ Ao contrário de Turri et. al. (2017), que pensam que necessitamos de uma terminologia mais refinada para nos referirmos aos diversos casos *à la* Gettier, penso podermos seguramente continuar usando “caso *à la* Gettier”, pois estaremos nos referindo apenas aos casos da Categoria 1.

¹⁴ E isso talvez até não fizesse mesmo sentido antes da taxonomia proposta por Turri et. al.

certamente os deixou mais longe da solução ao problema. Como argumentarei mais adiante, ao voltarmos a nossa atenção apenas para os casos da Categoria 1, teremos mais clareza sobre o papel da sorte no processo de gettierização de uma crença.

Os próprios Blouw, Buckwater e Turri reconhecem tal vantagem:

[Uma] vantagem do nosso enquadramento é que ele conecta a teorização sobre os casos à la Gettier e a sorte epistêmica. [...] Em particular, a taxonomia ajuda a clarificar as configurações candidatas de *tipos* e *graus* de sorte dos quais uma abordagem correta da relação entre conhecimento e sorte possa ser traçada. [...]

[...] Uma interpretação modal do enquadramento pode naturalmente também ajudar a moldar influentes teorias do conhecimento que apelam a noções modais como a segurança e a sensibilidade. (pp. 249-250)

2.1. Categorias 4 e 5

Usemos como exemplo MOEDAS. A crença que inicialmente atribuímos a Smith é a crença falsa de que Jones irá ficar com a vaga. E juntamente com a crença verdadeira de que Jones tem 10 moedas no bolso, Smith infere que a pessoa que vai ficar com a vaga tem 10 moedas no bolso. Se a estória parasse por aqui, a crença de Smith seria falsa; porém, o fato de Smith ter ficado com a vaga, supostamente torna a sua crença verdadeira.

Contudo, se perguntássemos a Smith se em algum momento ele pensou que Jones não ficaria com a vaga, a sua resposta seria certamente que não. Isso sugere que há algo de errado com a crença final que atribuímos a Smith. A afirmação seguinte parece deixar clara a inadequação

(S) Smith acredita que a pessoa que vai ficar com a vaga tem 10 moedas no bolso mas não acredita que Jones vai ficar com o emprego.

Dada a descrição do cenário, a afirmação acima é claramente falsa. Smith acredita que Jones é quem vai ficar com o emprego. Se a crença que os epistemólogos comumente atribuem a Smith fosse adequada, a afirmação acima teria de ser verdadeira.

Um modo de ver as coisas é usando a distinção de Keith Donnellan entre uso referencial e atributivo de descrições definidas. A descrição *A pessoa que vai ficar com a vaga tem 10 moedas no bolso* pode ser lida como dizendo que seja quem for que vai ficar com

a vaga, essa pessoa tem 10 moedas no bolso – caso em que é usada atributivamente –; ou pode ser lida como dizendo que Jones, que é a pessoa que Smith pensa que vai ficar com o emprego, tem 10 moedas no bolso – caso em que é usada referencialmente. É claro que em MOEDAS o uso adequado da descrição por parte de Smith é referencial: ele está usando a descrição – ainda que falsa – para se referir a Jones.¹⁵ Nesse caso, a crença de Smith é falsa. Apenas no caso do uso referencial é que a crença de Smith seria verdadeira, mas no cenário descrito em MOEDAS não podemos atribuir tal uso a Smith. Portanto, MOEDAS é um caso de crença falsa justificada. Por conseguinte, não é um contraexemplo a CVJ.

Consideremos uma variação do caso do cão pastor (ou da ovelha do pasto) de Chisholm (1966):

[URUBU] João olha para uma árvore e forma a crença verdadeira de que nela está pousado um urubu. Acontece que João olhou para um gavião pousado na árvore. Contudo, havia de fato um urubu pousado naquela árvore, porém, pousado atrás do gavião, longe da vista de João.

Nesse caso não há uma descrição envolvida, o que não levanta, porém, uma dificuldade para o meu argumento. A chave para resolver esse caso está na generalidade. O quão geral deveríamos considerar a crença de João? Uma primeira opção seria o mais geral possível: em algum lugar daquela árvore há um urubu pousado. Mas não é esse o caso, temos de considerar um local mais específico, a saber, na mesma região em que está o gavião (que ele identifica erroneamente como um urubu). Mas se ele identifica erroneamente o gavião como um urubu, o conteúdo de sua crença é algo como *há um urubu naquele local daquela árvore*. Ora, nesse caso a crença é falsa. Apenas no caso de a crença ser mais geral – *em algum lugar da árvore* – é que a crença de João seria verdadeira. Chamemos o local onde está o gavião de *s*. Assim, novamente, a seguinte afirmação mostra a inadequação da crença originalmente atribuída a João:

João acredita que há um urubu em algum lugar da árvore mas não acredita que há um urubu em s.

¹⁵ Como o uso referencial de uma descrição funciona exatamente como um nome próprio, podemos substituir, nesse caso, a descrição *A pessoa que tem 10 moedas no bolso* pelo nome próprio “Jones”. E agora fica claramente absurda a afirmação S: Smith acredita que Jones vai ficar com a vaga mas não acredita que Jones vai ficar com a vaga.

Essa afirmação, no contexto de URUBU, é claramente falsa. O que indica que a crença adequada a se atribuir a João é a crença mais específica de que há um urubu pousado na localidade *s* daquela árvore, que é uma crença falsa. Portanto, João tem uma crença falsa justificada, o que não conta como contra exemplo a CVJ.

Há também casos em que apenas um quantificador existencial está envolvido, o que também não levanta problemas ao meu argumento. Consideremos o caso do Sr. Nogot (LEHRER 1970):

[SR. NOGOT] Um estagiário que trabalha no escritório junto com o Sr. Nogot o vê estacionando um Ford na vaga dos funcionários. Além disso, o vê com um recibo pago do IPVA do carro. O estagiário então forma a crença de que alguém naquele escritório tem um Ford. Acontece que o Sr. Havit é quem é dono do Ford; ele tinha apenas emprestado ao Sr. Nogot.

Podemos, do mesmo modo, (i) atribuir uma crença mais geral ou (ii) uma crença mais específica ao estagiário. No caso da crença mais geral, a leitura do quantificador é a leitura usual: pelo menos um indivíduo do domínio, seja quem for, tem um Ford. Numa leitura mais restrita, à qual chamarei de “leitura pragmática”: Alguém no escritório, e esse alguém é o Sr. Nogot, possui um Ford. Essa segunda leitura, na verdade, elimina o quantificador existencial, uma vez que o estagiário tem o Sr. Nogot em mente quando forma a sua crença.¹⁶ Se reparamos bem, esse é um caso especial da distinção referencial-atributivo. Com a diferença de que estamos usando uma descrição indefinida para fazer referência ao Sr. Nogot. Portanto, temos mais um caso de crença falsa justificada.

Mas e quanto a FORD? Smith forma a crença disjuntiva *Ou Jones tem um Ford ou Brown está em Barcelona*. De que modo é inadequado atribuímos tal crença a Smith uma vez que ele efetivamente infere a proposição disjuntiva de uma proposição que ele já acreditava? Não há aqui uma inadequação na atribuição da crença, mas antes no uso da disjunção. Smith

¹⁶ O que aqui chamo de “uso pragmático” do quantificador existencial é um fenômeno bastante comum na linguagem natural. Tenho um indivíduo definido em mente e quero fazer uma afirmação sobre ele, mas por algum motivo específico não quero ser específico em minha asserção. Por exemplo, vi minha irmã roubando o jornal do vizinho, e quero deixar claro para ela que vi; chego para minha mãe, quando as duas estão juntas, e digo, “Alguém sarrupiou o jornal do vizinho hoje cedo”. Ou considere o seguinte exemplo dado por Paul Grice (1975). S diz “x está se encontrando com uma mulher esta noite”. É claro que tal sujeito não está dizendo que x foi ver sua mãe ou sua irmã ou sua esposa; antes está a sugerir que é uma amante ou uma amiga proibida. E se acaso x está com a sua mãe, ainda que seja verdade que x está com uma mulher, não diríamos que S disse algo verdadeiro. Se S dissesse à esposa de x que ele está com uma mulher, a esposa de x não se mobilizaria para arma-lhe um flagrante caso considerasse que essa mulher pudesse ser a própria mãe de x.

faz uso daquilo a que chamarei “uso espúrio” da disjunção, *i.e.*, o uso da disjunção num contexto inapropriado. Consideremos uma primeira opção. Smith tem alguma evidência de que Jones tem um Ford. Suponhamos que ele tenha formado a disjunção aleatoriamente. Parece-me um tanto estranho dizermos que Smith tenha a mínima inclinação para dizer que tem alguma evidência a favor de *Brown está em Barcelona*. É claro que ele poderia dizer que tem evidência para a crença disjuntiva, não para a segunda disjunta. Mas ele só pode ter evidência para a crença disjuntiva se tiver evidência a favor de pelo menos uma das disjuntas; o que é o caso com *Jones tem um Ford*. Contudo, no contexto no qual aquilo que está em jogo é a crença de que Jones tem um Ford, a proposição *Brown está em Barcelona* é completamente irrelevante do ponto de vista evidencial, uma vez que Smith só pretende que sua crença seja verdadeira porque pensa que a proposição sobre Jones é verdadeira. Esse uso da disjunção é apenas um truque lógico; mas truques lógicos nem sempre são adequados para situações ordinárias.

Deixe-me dizer algo positivo sobre aquilo que considero o “uso legítimo”¹⁷ da disjunção num contexto ordinário ou comum. Talvez o uso mais intuitivo da disjunção em contextos comuns seja o da disjunção exclusiva. Por exemplo, no mercado quando a mãe diz ao filho “Ou sorvete ou refrigerante”, ela não se sentiria muito satisfeita em ouvir o filho dizer, “Os dois”. Ou o filho toma sorvete, ou toma refrigerante; e sem discussão! O contexto de FORD, porém, deixa claro que o uso da disjunção não é o exclusivo. Passemos, então, ao inclusivo. Os casos mais comuns do uso da disjunção inclusiva são aqueles em que o sujeito tem evidência a favor de ambas as disjuntas e não acredita que elas sejam inconsistentes entre si. Por exemplo, chego em casa e não encontro meu gato. Sei que ele costuma ir para a casa do vizinho, e que também costuma ir caçar calangos nos muros da vizinhança. É perfeitamente legítimo, nesse caso, que eu acredite que o gato está na casa da vizinha, ou algures na vizinhança caçando. E é claro que eu percebo que ele pode estar na casa da vizinha caçando calangos. Mas será que em FORD temos um contexto similar? Obviamente que não.

¹⁷ Não estou a sugerir que haja algo de errado com a lógica clássica. O que acontece aqui é um fenômeno pragmático, no sentido de que o significado de uma sentença ou frase proferida por nós não depende apenas de sua sintaxe e semântica, mas também do contexto de sua asserção (LEVINSON 1983). Um exemplo clássico é o caso do nosso uso do “se, então” na linguagem natural. Em muitos casos a condicional material da lógica clássica capta bem a condicional da linguagem natural, mas em outros usá-la seria completamente descabido. S é presidente do clube de palavras cruzadas e decide que não abrirá a sede hoje, não importa o que aconteça. Um dos membros pergunta a S se o clube abrirá, e ele responde “Se chover, não”. Seja como for, S não abriria a sede. A sua condicional é verdadeira, muito embora seja impróprio proferi-la, uma vez que S está violando alguma máxima conversacional (cf. GRANDY & WARNER 2017).

Smith não tem evidência a favor de ambas as proposições; e se tivesse, precisaríamos ainda de um contexto que deixasse claro o porquê da disjunção (e não simplesmente a formação de uma crença conjuntiva). Poderia ser um contexto em que Jones e Brown sejam casados e Smith tenha ouvido deles que eles não tinham dinheiro o suficiente para que Jones comprasse um Ford e Brown fizesse uma viagem para Barcelona. Smith, porém, não descarta a possibilidade de ambas as coisas acontecerem, pois também ouviu deles uma conversa sobre empréstimo. Suponha que Smith acredita que é mais provável que Jones tenha comprado um Ford do que Brown ter viajado para Barcelona. No caso de Brown estar em Barcelona, e sua crença se tornar verdadeira, já não temos mais uma crença gettierizada.

Outro modo de se compreender o que há de errado como o uso que Smith faz da disjunção é notar que acreditar numa proposição envolve, além da aceitação de certo conteúdo proposicional, certa disposição para agir de acordo com aquele conteúdo proposicional. No caso original, ainda que Smith forme uma crença cujo conteúdo proposicional é disjuntivo, ele está disposto apenas a agir de acordo com a proposição \langle Jones tem um Ford \rangle . Por exemplo, ele poderia pedir uma carona para Jones; ou poderia arranhar o Ford para se vingar de Jones, etc. Mas certamente não compraria uma passagem para Barcelona para ir se encontrar com Brown, ou mesmo esperar que Brown lhe trouxesse um presente de lá, etc.¹⁸ Isso sugere que o uso que Smith faz da disjunção no caso de Gettier é espúrio.

2.2. Categorias 2 e 3

Consideremos CANETA. Qual o conteúdo da crença de Katie? Quando ela entra para o banho, a crença que ela formou foi *Aquela caneta Bic azul está em cima da mesa*. A caneta que fica em cima da mesa no final – a substituição – é outra, e não faz parte do conteúdo da crença de Katie. Ela não sustentaria ao mesmo tempo, nesse contexto, que há uma caneta Bic azul em cima da mesa *e* que não é aquela caneta que ela deixou lá antes de entrar para o banho. Chamemos a caneta original de *a* e a caneta substituta de *b*. Assim, a proposição conteúdo da crença de Katie é $\langle a$ está em cima da mesa \rangle ; o fato final, depois da substituição, é o de que *b* está em cima da mesa. É óbvio que esse fato não pode ser um verificador para a crença de Katie; teria de ser o fato de que *a* está em cima da mesa – fato que não ocorre

¹⁸ De um ponto de vista bayesiano grosseiro poderíamos nos focar no comportamento de aposta de Smith. Quanto Smith estaria disposto a apostar na proposição *Jones tem um Ford ou Brown está em Barcelona*? Certamente a mesma quantia que estaria disposto a apostar em *Jones tem um Ford*. Ou seja, a proposição *Brown está em Barcelona* não desempenha qualquer papel relevante na crença de Smith.

naquele momento. A crença de Katie é, portanto, falsa. E é por isso que ela não sabe que há uma caneta Bic em cima da mesa, não porque o gatuno trocou as canetas.

Talvez uma modificação nos salvasse o exemplo. Consideremos então

[CANETA*] Katie terminou de usar uma caneta Bic azul, deixou-a em cima da mesa e foi tomar banho. Ela acredita verdadeiramente que há uma caneta azul em cima da mesa. Contudo, enquanto ela está no banho, um gatuno entra em sua casa e, nesse meio tempo, enfia a caneta no bolso. Antes de sair ele acaba deixando no lugar outra caneta Bic azul, indistinguível da caneta de Katie. O gatuno, contudo, volta para desfazer a troca das canetas, de modo que quando Katie sai do banheiro vê a sua caneta.

Essa modificação de fato evita a objeção acima, mas a custo de se tornar um caso da categoria 1 – *ameaça mal sucedida*. Note que quando o gatuno troca as canetas da primeira vez, a crença de Katie – *a está em cima da mesa* – torna-se falsa. Porém, com a destroca das canetas a crença de Katie volta a ser verdadeira. Assim, houve apenas uma ameaça mal sucedida à verdade da crença.

Por fim, vamos a PAPAGAIO. Assim como em MOEDAS, o contexto torna inapropriada a atribuição de uma crença geral a Sebastião quando a sua crença tem como constituinte o Sr. Silva. Estaria Sebastião disposto a responder “sim” para a pergunta “Você acredita que a pessoa dona do imóvel mais caro da cidade tem mais de 50 anos?” e “não” para a pergunta “Você acredita que o Sr. Silva é o dono do imóvel mais caro da cidade?”? Certamente que não. Isso sugere que, novamente como em MOEDAS, a crença adequada a se atribuir a Sebastião é *O Sr. Silva é o atual dono do imóvel mais caro da cidade*; que, por conseguinte, é uma crença falsa justificada.

2.3. Diagnóstico

Se eu estiver correto, as categorias 2-5 não constituem contra-exemplos genuínos a CVJ. Isso porque a crença mais adequada a se atribuir ao sujeito naqueles contextos teria de ser uma crença com o seu conteúdo original ou uma crença que não seja produto de uma inferência espúria. Relembrando: se alguém vê uma ovelha no campo e, a partir disso, forma a crença verdadeira de que há ali um cão pastor, a crença adequada a se atribuir a essa pessoa é *Aquele animal para o qual estou olhando, que é um cão pastor, está no campo*, e não a crença geral *Há um cão pastor no campo*.

As categorias 4 e 5 – em que há *falha de detecção* – começam com uma crença falsa. Contudo, sequer tal crença se torna verdadeira (ainda que por sorte), como sugerido por Gettier e acatado posteriormente pela maioria dos epistemólogos. Sugiro que a impressão de que tal crença se torna verdadeira surge apenas porque caímos no truque de usar uma sentença geral para nos referirmos a um objeto específico; ou porque consideremos legítimo um passo inferencial na verdade suspeito. Já as categorias 2 e 3 – em que há *detecção + substituição* – começam com crenças verdadeiras que se tornam falsas. A diferença aqui é que continuamos a pensar que a crença é verdadeira por causa do truque. A substituição do verificador torna a crença geral verdadeira; a crença específica, por sua vez, é tornada falsa.

Podemos ver as coisas assim. Uma crença geral pode ser tornada verdadeira por verificadores diferentes. Por exemplo, a crença na proposição ⟨Há políticos corruptos no Brasil⟩ é tornada verdadeira tanto pelo fato de Paulo Maluf ser um político brasileiro, quanto pelo fato de Fernando Collor também o ser. Mas uma crença particular como a crença na proposição ⟨Paulo Maluf é um político corrupto⟩ só pode ser tornada verdadeira por um único verificador, a saber, o fato de Paulo Maluf ser um político corrupto. E é claro que se é o caso que Paulo Maluf é um político, tanto a crença particular quanto a geral são verdadeiras. Mas o contrário não vale: se a crença geral é verdadeira não se segue seja o caso que Paulo Maluf é um político corrupto; talvez apenas Collor é que o seja.

Pensemos numa proposição se tornando verdadeira por sorte. Os contraexemplos originais sugerem que a crença de Smith de que a pessoa que vai ficar com o emprego tem 10 moedas no bolso é verdadeira por sorte porque o esperado era que Jones ficasse com o emprego, não o próprio Smith. Nesse sentido, é muito fácil que uma crença geral se torne verdadeira por sorte: basta que um elemento ou conjunto de elementos satisfaça a condição exigida. Mas não é esse sentido o relevante. Quando saio de casa com a eventual crença de que há uma padaria aberta em Ouro Preto, dirigindo-me especificamente à padaria da esquina, e vejo que tal padaria está fechada, ainda que eu saiba que a padaria do outro lado da cidade esteja aberta, considerarei a minha crença inicial falsa. O sentido relevante para a minha crença de que há uma padaria aberta em Ouro Preto – quando me refiro especificamente à padaria da esquina – ser verdadeira por sorte é algo como a luz do bairro ter acabado mas a padaria ter um gerador. Nesse caso, foi sorte de a padaria da esquina estar aberta. O truque

nos casos que eliminei consiste exatamente em considerar o primeiro tipo de crença verdadeira por sorte, quando na verdade deveria considerar o segundo.

Uma reação natural à minha recusa dos contraexemplos originais de Gettier e vários outros casos como contraexemplos legítimos seria apontar para casos como CELEIRO ou CANETA* e dizer que meu argumento não se aplica a tais casos. Alguém que fizesse isso estaria parcialmente correto, muito embora estivesse ignorando que estou excluindo como casos legítimos as categorias 2-5 da taxonomia proposta por Turri *et. al.* Continuo considerando como legítimos apenas os casos da Categoria 1. Sendo assim, discutirei doravante quase que exclusivamente casos à la Gettier da Categoria 1 – *i.e.* apenas casos em que o sujeito detecta a verdade da proposição mas que por algum elemento de sorte não tem conhecimento.

2.4. Rejeitando CELEIRO como um contraexemplo genuíno?

Embora eu pense que os casos do tipo CELEIRO sejam os únicos casos à la Gettier genuinamente ameaçadores, há filósofos que, ao contrário, pensam que casos como esse são efetivamente casos de conhecimento. William Lycan, por exemplo:

[...] rejeito as intuições comumente aceitas; além de não as compartilhar, também penso que estejam erradas. Suspeito que [o sujeito em CELEIRO] sabe, a despeito dos elementos de acaso que permeiam a sua situação. (2006, p. 158)

Vale notar que ele não argumenta diretamente contra CELEIRO. Ao invés, ele ataca três outros casos famosos apresentados por Gilbert Harman (1973, pp. 142-44) – aos quais chamarei LIVRO, DITADOR e ITÁLIA – que ele pensa serem do tipo de CELEIRO. De fato esses três casos parecem se enquadrar na Categoria 1 (*detecção + ameaça mal sucedida*), *i.e.*, a categoria de CELEIRO.¹⁹ Vejamos:

[LIVRO] Um funcionário da biblioteca da faculdade vê Tom num movimento suspeito em meio às prateleiras colocando um livro debaixo do casaco e saindo em disparada. O funcionário é chamado a depor e diz que foi Tom quem surrupiou o livro. Dias depois a mãe de Tom é chamada a depor e, sem que o

¹⁹ Stephen Hetherington parece também compartilhar da posição de Lycan. Ele distingue entre casos à la Gettier *proveitosos* [*helpful*] e *perigosos* [*dangerous*] (1999; 2016). Um caso é proveitoso quando a crença verdadeira justificada do sujeito se deve parcialmente a alguma estranheza do cenário; ao passo que é perigoso quando a estranheza do cenário apenas ameaça a crença verdadeira justificada do sujeito. Hetherington classifica CELEIRO como um caso perigoso, e parece não se importar muito se os casos perigosos sejam casos à la Gettier genuínos: “mencionei o caso dos celeiros falsos [...] porque muitos epistemólogos continuam a considera-lo como um caso à la Gettier” (2016, p. 9).

funcionário ficasse sabendo, disse que Tom estava numa viagem, longe dali, e que quem tinha pego o livro era seu irmão gêmeo, Buck. (adaptado)

Nesse caso, a crença verdadeira do funcionário da biblioteca deixa de contar como conhecimento, uma vez que o depoimento da mãe – mesmo falso – seria evidência facilmente disponível que, conjuntada à evidência do funcionário, anularia a sua justificação para acreditar que Tom tivesse roubado o livro.

[DITADOR] Um ditador latino-americano é assassinado e o jornal mais respeitado do país dá a notícia. Pablito, indo em direção ao campo, compra, durante a viagem, o jornal e lê a notícia. Contudo, nesse meio tempo, as principais emissoras de TV – obrigadas pelo partido do ditador – dão a notícia falsa de que o ditador estava bem e que quem tinha morrido era seu sócia. (adaptado)

Pablito, sustenta Harman, não sabe que o ditador foi assassinado, apesar de sua crença ser verdadeira e baseada em evidência adequada. O conteúdo da notícia falsa, que grande parte das pessoas assistiu pela TV, pode ser facilmente adquirido por Pablito, de modo que quando adicionado à sua evidência original, anularia sua justificação.

[ITÁLIA] Benito leva seu amigo Plínio ao aeroporto para embarcar num voo para Itália, onde passará alguns meses. Plínio decide fazer uma brincadeira com Benito. Faz várias montagens de fotos suas para parecer que está em Machu Picchu e as envia a Benito. Contudo, envia para o e-mail que Plínio checka apenas uma vez por semana. Plínio só irá ver o e-mail de Benito uma semana mais tarde. (adaptado)

Benito não sabe, portanto, que Plínio está na Itália, embora tenha uma crença verdadeira e justificada. Isso porque as fotos que Plínio lhe enviara, que estão facilmente disponíveis, se adicionadas à evidência original de Benito, anulariam a sua justificação.

Nos três casos acima, o cenário é montado de tal forma que algum aspecto do ambiente impede que a crença verdadeira com base em evidência adequada conte como conhecimento. Harman pensa que esses aspectos fornecem ao sujeito anuladores por ele não anteriormente possuídos; anuladores que o sujeito facilmente teria acesso (LYCAN 1977, p. 117; 2006, p. 162). Por exemplo, em DITADOR a notícia falsa de que o ditador não morreu sem dúvida se espalharia muito rapidamente devido à confiabilidade em geral dos telejornais. Mais cedo ou mais tarde a notícia chegaria ao ouvido de Pablito. Em ITÁLIA, Plínio está a uma semana – ou talvez menos, caso decida checkar seus e-mails novamente, ou mesmo acesse

acidentalmente a sua conta – de ver as fotos que parecem mostrar que Benito está em Machu Picchu. Já em LIVRO, o funcionário da biblioteca de algum modo seria informado do depoimento da mãe de Tom: receberia atualizações do andamento do processo pela própria biblioteca, ou pela instituição da qual pertence a biblioteca, seus colegas de trabalho poderiam lhe contar, etc. Lycan apresenta o seguinte princípio para dar conta da intuição de que os casos acima são casos à la Gettier (p. 117):

(H) S sabe que h com base na evidência e somente se não há evidência e' tal que (a) e' potencialmente anule a justificação que S tem para h com base em e ; (b) embora S não possua e' , e' esteja prontamente acessível a S; e (c) não haja evidência ulterior, igualmente acessível a S, que anule e' , e, que por sua vez, não seja anulada.²⁰

Contudo, ele argumenta que esses três casos são, na verdade, casos de conhecimento (p. 120). O argumento é essencialmente negativo e pode ser dividido em duas partes: (i) Ele primeiro aponta que parece não haver uma linha clara que distinga entre aqueles fatos externos que minam o conhecimento e os que não minam, e conclui que até que Harman e os que partilham de sua intuição forneçam tal linha demarcatória, não temos qualquer razão para pensar que há falha de conhecimento nos casos acima; chamemos esse argumento de “argumento do ônus da prova”. (ii) Em seguida, ele aponta o que suspostamente está por trás da intuição de Harman e tenta afastá-la.

Embora eu partilhe da intuição de Harman, penso não ser de grande dificuldade oferecer a linha demarcatória exigida por Lycan, ainda que de maneira um pouco vaga. Como já dito, a principal lição a ser tirada do problema de Gettier é que sorte e conhecimento são incompatíveis. Assim, se algum fato externo ou característica do ambiente faz com que um sujeito com crença verdadeira justificada careça de conhecimento, só pode ser porque tal fato ou características existem em quantidade suficiente para tornar a crença do sujeito verdadeira por sorte.

Mas agora temos um problema. Nos três casos de Harman a crença do sujeito não é verdadeira por sorte. Em LIVRO, no momento em que o funcionário da biblioteca vê Tom surrupiar o livro, não há qualquer elemento no cenário que torne a crença do funcionário verdadeira por sorte. E o mesmo ocorre em ITÁLIA e DITADOR. É apenas num momento posterior que algum elemento externo ameaça o conhecimento dos sujeitos – ou seja, a crença

²⁰ Para darmos conta de CELEIRO é preciso afrouxar um pouco a cláusula (b), uma vez que o sujeito dificilmente obteria a informação que estava numa região cheia de celeiros falsos.

dos sujeitos é gettierizada apenas posteriormente. E se tais casos são casos à la Gettier genuínos, o são de uma maneira muito estranha: em t_1 o sujeito tem crença verdadeira justificada e sabe, mas em t_2 não sabe. O próprio Lycan menciona esse resultado contraintuitivo²¹ (p. 122). Se há algum elemento de sorte nos três casos, tem de estar em t_2 . Mas não pode ser um elemento que torne a crença do sujeito verdadeira por sorte, uma vez já vimos que não é uma crença verdadeira por sorte. Resta-nos a justificação que o sujeito tem para a crença ser mantida por sorte. Assim, se os casos de Harman são genuínos, teremos de dizer que o conhecimento é incompatível com a sorte em pelo menos dois sentidos: (i) crença verdadeira por sorte não é conhecimento, e (ii) crença verdadeira cuja justificação é mantida por sorte não é conhecimento.

Não perseguirei, contudo, a linha de raciocínio acima. Não preciso que os casos de Harman sejam casos à la Gettier genuínos (embora eu pense que sejam), ou mesmo de refinar por agora a ideia de o conhecimento ser incompatível com a sorte. Preciso apenas de um caso em que o sujeito tenha, em t_1 , uma crença verdadeira justificada e que haja algum elemento de sorte também em t_1 . CELEIRO satisfaz exatamente essas exigências. No exato momento em que o sujeito forma a crença de que há um celeiro em sua frente, diversos celeiros de fachada estão dispostos no ambiente. Embora ele tenha olhado para o único celeiro real, ele teria muito facilmente acreditado que qualquer um daqueles outros celeiros também fosse real; e em todos esses outros casos a sua crença teria sido falsa. Foi realmente muita sorte da parte do sujeito ter olhado para o único celeiro real e, por conseguinte, ter formado uma crença verdadeira de que há ali um celeiro.

CELEIRO, portanto, torna-se um tanto mais intuitivo quando o examinamos tendo em mente que conhecimento e sorte são incompatíveis. A quantidade de celeiros falsos no ambiente é suficiente para tornar verdadeira, por sorte, a crença do sujeito. Assim, o argumento do ônus de Lycan deixa de ser efetivo. Por conseguinte, não será preciso examinar a segunda parte de seu argumento geral.

²¹ Um modo de tornar os casos menos contraintuitivos seria qualificando a estranheza do ambiente como mais ou menos próxima. Por exemplo, em LIVRO, se a mãe de Tom testemunhou pouco depois de o funcionário ter formado a crença de que Tom surrupiou o livro, é mais plausível pensar que o funcionário não sabe do que se o testemunho tivesse sido dado dias depois. Talvez uma analogia ajude. Suponha que compro um computador e já logo em seu primeiro uso ele queima. Liguei-o à tomada, ele funcionou por alguns segundos, e queimou. Se alguém me pergunta “O computador funcionou?”, a resposta “Não” seria tão aceitável quanto “Funcionou por alguns segundos e queimou”. Mas se o computador tivesse sido usado por dias e só então queimado, a resposta adequada seria “Funcionou por semanas e depois queimou”.

Uma implicação de minha conclusão – a de que CELEIRO apresenta um caso razoavelmente claro de falha de conhecimento – é que Lycan simplesmente não pode estender a sua intuição acerca dos casos de Harman a CELEIRO. Mas isso não parece ir contra o que afirmei no início desta seção, que todos esses casos se enquadravam na Categoria 1? Afinal, se a categoria é uniforme naquilo que é relevante para a gettierização de uma crença, então o resultado deveria ser também uniforme. Lembremo-nos do pequeno reparo feito acima na ideia de conhecimento e sorte serem incompatíveis. Tivemos de adicionar a cláusula de que a crença justificada mantida por sorte é também incompatível com o conhecimento. Isso porque os casos de Harman, se são casos à la Gettier, não o são exatamente pelo mesmo motivo que o é CELEIRO. Sugiro que a intuição de Lycan – de que os casos de Harman são casos de conhecimento – tem algo a ver com a estranheza de o sujeito saber em t_1 mas não em t_2 . Mas ainda que concedamos plausibilidade a tal intuição, ela não tem qualquer força contra CELEIRO, uma vez que ela diz respeito à perda de justificção ou de conhecimento ao longo do tempo.

Seja como for, Lycan não é bem sucedido em afastar CELEIRO como um caso genuíno à la Gettier.

O que nos resta, portanto, do seminal artigo de Gettier é apenas uma homenagem. A partir de agora considerarei como gettierizadas apenas crenças em cenários como CELEIRO. Se me permitem uma outra homenagem, poderíamos chamar aos casos da Categoria 1 *casos Ginet/Goldman*.

3. Algumas tentativas de solução

Como já mencionado anteriormente, tratarei neste capítulo apenas do *problema anti-sorte*. Essa restrição não diminui a quantidade de tentativas de solução, contudo. Na verdade, arrisco-me a dizer que a maior parte dos filósofos que se ocupou do problema de Gettier não distinguiu entre as dimensões analítica e anti-sorte. Por essa razão trataram o problema quase sempre como um problema de definir o conhecimento focando-se mais em oferecer condições anti-Gettier ou anti-sorte do que em dizer algo informativo sobre o conhecimento que fosse além de CVJ. (Na verdade, eles pensavam que estavam dizendo algo informativo sobre o conhecimento ao oferecer condições anti-sorte). E mesmo assim grande parte dessas tentativas malogrou. Seria uma tarefa maior da que posso cumprir se eu me propusesse a fazer

um levantamento exaustivo da discussão gerada pelo problema de Gettier.²² Contudo, penso que as minhas escolhas sirvam para ilustrar a dificuldade de se oferecer uma condição anti-sorte satisfatória.

As tentativas de solução ao problema de Gettier podem ser divididas em dois grandes grupos: (i) adição de uma quarta condição à CVJ, e (ii) a rejeição e substituição de alguma condição de CVJ, geralmente a justificação.

3.1. CVJ + Condição anti-Gettier

3.1.1. *Sem bases falsas*

O tipo de solução mais simples ao problema de Gettier é impedir que a crença alvo de conhecimento tenha sido obtida ou baseada em informação falsa. Isso inicialmente funcionaria para bloquear os contraexemplos originais, que dependem de fato de bases falsas. Tão fácil quanto introduzir essa condição, é encontrar novos contraexemplos – casos à la Gettier que não dependem de qualquer inferência. O caso do cão pastor é um. Porém, como vimos, tanto os casos originais quanto o do cão pastor não são contraexemplos genuínos, uma vez que corretamente interpretados, teríamos de atribuir uma crença falsa justificada ao sujeito.

Considere então CELEIRO. Esse é um caso em que nenhuma inferência ocorre. Mas suponhamos que a seguinte condição seja verdadeira: *se S sabe perceptualmente que p, então S é capaz de obter evidência proposicional a favor de p*. Aplicado ao caso do celeiro, a evidência proposicional disponível a S pode ser algo grosso modo como: *parece-me que há um celeiro em minha frente*. Se usado como base para S inferir que há um celeiro em sua frente, claramente temos uma inferência com premissa verdadeira. Contudo, ainda temos um caso à la Gettier, uma vez que S infere a partir de uma premissa verdadeira que há um celeiro em sua frente, e ainda assim a sua crença é verdadeira por sorte.

Portanto, a estratégia *Sem Bases Falsas* não é satisfatória exatamente porque não é capaz de evitar a sorte epistêmica como bem ilustrada por CELEIRO.

3.1.2. *Não-anulabilidade*

²² Felizmente temos a excelente *The Analysis of Knowing*, de Robert Shope (1983), que cobre exaustivamente a primeira década de reação ao problema de Gettier. Para novas perspectivas veja HAZLETT [ed.] (2015); BORGES, ALMEIDA & KLEIN [eds.] (2017), e HETHERINGTON [ed.] (2018).

As teorias da não-anulabilidade são as mais influentes dentre as teorias que propõe uma quarta condição à CVJ. A ideia básica é que há certas informações que, se disponíveis²³ ao sujeito, anulariam o suposto conhecimento que o sujeito tem de determinada proposição. Se entendermos a justificação como aquela propriedade que em grau suficiente torna a crença verdadeira conhecimento, tais informações – a que chamaremos anuladores – podem ou anular completamente a justificação que o sujeito tem para acreditar na proposição alvo de conhecimento, ou pode anular num grau tal que, embora a proposição permaneça justificada, não deixa de ter justificação suficiente para tornar a crença conhecimento. Os principais defensores da abordagem da não-anulabilidade foram Keith Lehrer e Thomas Paxson (1969), Peter Klein (1971, 1976) e Marshall Swain (1974).

Um modo genérico de caracterizar um anulador é o seguinte:

Uma proposição verdadeira d é um anulador para a crença de S de que p sse a adição de d ao conjunto de evidência que S tem para p faz com que p não conte como conhecimento.

Podemos dizer a mesma coisa em termos contrafactuais: se S *acreditasse* que d , não saberia que p . Assim, temos uma quarta condição a CVJ:

(iv) Não há um anulador disponível para a crença de S de que p .

Como nos interessam aqui apenas os casos à la Gettier genuínos, vejamos como a condição (iv) nos ajuda a lidar com CELEIRO. Lembre-se que João não sabe que há um celeiro em sua frente porque ele está num ambiente lotado de celeiros falsos (sem o saber) e a sua crença verdadeira de que há ali um celeiro é verdadeira por pura sorte. Ora, considere a proposição d : {João está numa região de celeiros falsos}. João tem evidência visual para acreditar que há em sua frente um celeiro; contudo, se d fosse acrescentada à sua evidência, João deixaria de ter evidência razoável para acreditar que está defronte a um celeiro, e, portanto, não saberia tal coisa.

Mas há contraexemplos à teoria da não-anulabilidade. Consideremos o seguinte:

²³ O que quero dizer aqui com “uma proposição estar disponível ao sujeito” não implica que o sujeito efetivamente acredite na proposição ou mesmo a tenha estocada na memória. “Estar disponível” aqui quer dizer que o sujeito poderia muito facilmente ter formado a crença em tal proposição.

[LAMPPIÃO] Antônio lê um livro de referência sobre o cangaço e forma a crença verdadeira de que Lampião morreu na Grotta do Angico. O autor do livro fez um trabalho competente no levantamento das informações, de modo que o seu livro é confiável. Mas há um dado da biografia do autor que ele fez questão de esconder de todos, ele era um mentiroso compulsivo.

No caso acima, Antônio sabe que Lampião morreu na Grotta dos Angicos. O trabalho do autor não foi contaminado pela sua mania compulsiva de mentir. Contudo, tivesse Antônio adquirido essa informação, a sua justificação teria sido anulada. Sendo assim, de acordo com a teoria da anulabilidade, Antônio não sabe o local onde Lampião morreu. A teoria da anulabilidade é por isso demasiado ampla, pois conta como falha de conhecimento casos que em que claramente o sujeito tem conhecimento. Logo, não pode ser uma teoria satisfatória.

Uma maneira de evitar essa consequência indesejada é interpretar os anuladores de outro modo. Como vimos acima, o anulador tinha de ser uma proposição verdadeira disponível ao sujeito, muito embora o sujeito não precisasse ter a crença efetiva na proposição. Chamemos a tais anuladores “anuladores externos”. Por outro lado, podemos entender um anulador, *d*, como uma proposição verdadeira, efetivamente acreditada pelo sujeito, e que quando adicionada à evidência do sujeito, faz com que a sua crença deixe de contar como conhecimento. Chamemos a tais anuladores “anuladores internos”.²⁴

Se considerarmos que uma proposição só funciona como anulador no caso de ser um anulador interno, a nossa intuição inicial de que LAMPPIÃO é um caso de conhecimento é mantida, uma vez que Antônio não forma qualquer crença sobre a mitomania do autor. Contudo, perdemos a resposta ao problema de Gettier. Voltemos a CELEIRO. João está dirigindo pela região dos celeiros falsos, mas não faz a menor ideia de que esteja. Assim, embora ele pudesse facilmente ter formado a crença de que ali é uma região de celeiros falsos – ter visto no mapa, ter conversado com os moradores da região, etc. –, ele não acredita efetivamente na proposição ⟨João está numa região de celeiros falsos⟩. Portanto, João não tem um anulador para a sua crença de que há em sua frente um celeiro. Logo, a anulabilidade falha em explicar por que João não sabe que há um celeiro em sua frente.

²⁴ Bergman (2006, p. 153), *e.g.*, usa a terminologia “anuladores proposicionais” e “anuladores mentais” [*mental states defeater*] para se referir ao que estou chamando aqui de anuladores internos e externos.

Chegamos então ao seguinte. Se entendemos os anuladores como externos, embora expliquemos por que João não sabe que p , acabamos classificando casos claros de conhecimento como falha de conhecimento; mas se entendemos os anuladores como internos, embora classifiquemos corretamente os casos de conhecimento, perdemos a resposta ao problema de Gettier. Seja como for, a teoria da anulabilidade não é bem sucedida.

3.2. CV + Condição anti-Gettier

3.2.1. Teoria Causal do Conhecimento

A teoria causal do conhecimento, embora datada, merece menção aqui, principalmente por estar na gênese da epistemologia modalizada. Alvin Goldman, em seu “A causal theory of knowledge” (1967) defende que conhecimento é crença verdadeira causalmente conectada (de maneira apropriada) ao seu verificador. O que explica a falha de conhecimento nos casos de Gettier, sustenta Goldman, é a ausência de uma conexão causal apropriada entre a crença do sujeito e seu verificador.

Embora a teoria causal tenha dificuldade de explicar o conhecimento da lógica e da matemática (não é possível estar causalmente conectado a entidades abstratas);²⁵ e de não ser capaz de lidar com o nosso conhecimento de verdades gerais e fatos espaço-temporalmente distantes de nós; a sua maior dificuldade é em lidar com CELEIRO, como apontado pelo próprio Goldman, em “Discrimination and Perceptual Knowledge” (1976), que fez o papel de coveiro de sua própria teoria (p. 773). Em CELEIRO o sujeito forma a crença verdadeira de que há um celeiro em sua frente através da visão, estando ele causalmente conectado ao celeiro. Sua visão funciona perfeitamente e não há qualquer cadeia causal defectiva ou desviante. Ele, contudo, não sabe. A sua crença é verdadeira por sorte, de modo que a introdução de uma condição causal é incapaz de funcionar como elemento anti-sorte.

3.2.2. Confiabilismo e Alternativas Relevantes: rumo às contrafactuais

²⁵ Goldman não pensa que isso seja um problema, uma vez que o seu alvo é o conhecimento de proposições empíricas. Ela deixa explícita a sua opinião de que a CVJ seja adequada para o conhecimento de proposições não-empíricas (p. 357). Paul Benacerraf (1973), num influente artigo, contudo, usa a teoria causal para avançar uma conclusão cética quanto ao conhecimento da matemática (interpretado de maneira realista).

Goldman (1976), então, propõe a noção de *confiabilidade*.²⁶ S sabe perceptualmente que α somente se sua crença verdadeira foi produzida por um mecanismo confiável. A confiabilidade, por sua vez, é entendida em termos contrafactuais:

Um processo cognitivo M é confiável sse para qualquer crença verdadeira, α , produzida por M, se α não fosse o caso, M teria produzido $\neg\alpha$, ou pelo menos não teria produzido α .

De central importância à abordagem confiabilista de Goldman é a capacidade de *discriminação*. Se um sujeito faz uso de um mecanismo cognitivo confiável (nesse caso a percepção, mais especificamente), então esse mecanismo o possibilitará discriminar entre o estado de coisas que produziu a sua crença e algum estado de coisas alternativo relevante. A discriminação envolve contrafactuais: numa situação contrafactualmente relevante, o sujeito formaria uma crença diferente da que formou na situação efetiva ou atual (ou pelo menos não formaria a crença que formou na situação atual). Por exemplo, em casos paradigmáticos de conhecimento perceptivo, quando a minha crença visual de que há agora um computador em minha frente, se não fosse o caso que houvesse um computador em minha frente (suponha um cenário em que há uma TV), eu não teria formado a crença de que há um computador – eu teria formado a crença de que há uma TV.

É fácil agora ver como o confiabilismo consegue lidar com CELEIRO. Uma vez que João está num ambiente lotado de celeiros falsos, que são indistinguíveis visualmente, ele não é capaz de discriminar (visualmente) entre um celeiro falso e um celeiro real. Se ele tivesse olhado para um celeiro falso, teria acreditado estar defronte a um celeiro real. Ou seja, há pelo menos uma *alternativa relevante* – João olha para o celeiro ao lado do que olhou efetivamente – e acredita falsamente haver um celeiro em sua frente.

Mas o que é exatamente uma alternativa relevante? Vamos primeiro à *alternativa*. Suponha que S acredita que α com base nalguma evidência e . Qualquer proposição α^* incompatível com α (isto é, α^* implica $\neg\alpha$) é uma alternativa a α . Obviamente, se α^* for verdadeira, S não pode saber que α . Por exemplo, S acredita que (Aquela cobra coral é venenosa) com base na disposição das cores preta, branca e vermelha no corpo da serpente.

²⁶ Apenas em “What is a justified belief” (1979) e *Epistemology and Cognition* (1986) é que Goldman desenvolve o confiabilismo quanto a justificação, entendendo a justificação como aquilo que torna a crença verdadeira conhecimento.

Claramente que a proposição ⟨Aquela cobra não é venenosa⟩ é uma alternativa à crença de S. Considere agora a evidência de S – a disposição das cores. Se S sabe que a cobra é venenosa com base nessa evidência, então, com base nessa mesma evidência, S pode eliminar a alternativa de que a cobra não é venenosa. Contudo, tal evidência não elimina a proposição alternativa, uma vez que há cobras corais falsas que mimetizam a verdadeira. Assim, S não sabe que aquela cobra coral é venenosa.

Mas o que faz uma alternativa ser *relevante*? Começemos pelo que seria uma alternativa *irrelevante*. Há várias proposições incompatíveis com ⟨Aquela cobra coral é venenosa⟩. Por exemplo, P = ⟨Aquele objeto é um demônio mimetizando uma cobra coral⟩ ou Q = ⟨Aquele objeto é uma bomba disfarçada de cobra coral⟩. A evidência visual da qual S dispõe não é capaz de eliminar P ou Q. Mas isso não levanta qualquer problema, uma vez que ainda que S não seja capaz de eliminar P ou Q, tais alternativas são irrelevantes. Se considerássemos P como uma possibilidade séria, as coisas teriam de ser bastante diferentes de como são efetivamente – mais precisamente, teria de haver demônios que por algum motivo se disfarçam de cobra coral. Ou seja, quando uma alternativa é irrelevante, supor que ela seja verdadeira envolve mudar bastante o modo como as coisas efetivamente são. Por outro lado, uma alternativa é relevante quando supor a sua verdade não altera substancialmente as coisas como efetivamente são; isto é, essa alternativa poderia ser facilmente verdadeira. Por exemplo, S está num viveiro cheio de cobras corais venenosas e não venenosas. Ele quer pegar uma cobra que acredita não ser venenosa; calha que aquela coral de fato não é venenosa. Ele, contudo, não pode saber que aquela cobra é não venenosa, uma vez que ele poderia ter pego uma cobra venenosa como base nas mesmas observações. Note, além disso, que não precisamos fazer quaisquer alterações substanciais no mundo como as coisas são para que a alternativa seja verdadeira. Por conseguinte, S poderia muito facilmente ter acreditado numa proposição falsa.

Mas o que exatamente torna uma alternativa *relevante*? Há pelo menos duas explicações: (a) características objetivas do ambiente, e (b) elementos do contexto conversacional ou pragmático. Por exemplo, um proponente de (b) poderia dizer que fatores pragmáticos – *e.g.*, o medo de S de ser atacado por um demônio – determinarão o conjunto de alternativas. É intuitivo pensarmos que tendo tal medo, a hipótese de haver um demônio disfarçado de cobra torna-se relevante para S. Por outro lado, um proponente de (a) dirá que o

fato de não haver demônios, ou pelo menos aquela região não ser a preferidas dos demônios, torna remota a possibilidade de tal hipótese ser verdadeira. Não precisamos tomar partido quanto a isso. À partida, nada impede ter combinarmos (a) e (b).

Anterior a Goldman, Fred Dretske já tinha introduzido a ideia de alternativas relevantes (1970) com o seu famoso exemplo da mula disfarçada. E muito embora Goldman tenha pensado a confiabilidade em termos contrafactuais, não caracterizou as alternativas relevantes nesses termos. Mais tarde, em seu monumental *Philosophical Explanations* (1981), Robert Nozick publica sua teoria do rastreamento e formula as alternativas relevantes em termos contrafactuais.

A teoria do conhecimento proposta por Nozick é uma versão mais refinada de outra abordagem contrafactual, a abordagem das *razões conclusivas* (DRETSKE 1971). É interessante que o próprio Dretske não tenha imediatamente caracterizado as razões conclusivas em termos contrafactuais, embora diga explicitamente bem mais tarde ter desenvolvido uma “análise do conhecimento [a abordagem das razões conclusivas] que produz o resultado de que apenas certas alternativas (àquilo que é conhecido) são evidencialmente relevantes” (2005 [2014, p. 33]).

De acordo com Dretske (1971), S sabe que p somente se tem razões conclusivas para p . E, por sua vez, S tem razões conclusivas, R, para p somente se não fosse o caso que p , R não teria sido o caso. Tal abordagem claramente dá conta de CELEIRO. Digamos que R, a suposta razão conclusiva para João saber que há um celeiro em sua frente, é o percepto resultado de sua experiência visual diante de um celeiro. Se João sabe que há um celeiro em sua frente, então se não houvesse um celeiro em sua frente, ele também não teria formado o percepto de um celeiro através da visão. Mas acontece que há uma situação muito próxima – a que ele olha para o celeiro falso ao lado, e forma o mesmo percepto. Ou seja, há uma situação em que não é o caso que há um celeiro na frente de João e ainda assim ele tem R. Isso quer dizer que R, nesse caso, não é uma razão conclusiva. Assim, João não tem uma razão conclusiva para acreditar que há um celeiro em sua frente, e, portanto, não sabe que há.

Chegamos, por fim, à formulação de Nozick. A crença verdadeira de S de que α é conhecimento se e somente se:

- (1) Se não fosse o caso que α , S não acreditaria que α ; e

(2) Se fosse o caso que α , então S acreditaria que α .

Nozick chama a (1) “variação” e a (2) “aderência”. Uma crença é *sensitiva* quando cumpre (1) e (2); e é *insensitiva* quando não cumpre pelo menos uma delas. Uma crença sensitiva rastreia a verdade da proposição nos mundos possíveis mais próximos. Nozick usa explicitamente a semântica dos mundos possíveis²⁷ para avaliar o valor de verdade das duas condicionais contrafactuais. Uma condicional contrafactual *se α fosse o caso, então β seria o caso* é verdadeira se, e somente se, nos mundos mais próximos ao atual nos quais α é verdadeira, β também é verdadeira. Ou dito de outro modo: se os mundos- α mais próximos ao atual são também mundos- β . E essa mesma condicional será falsa se houver pelo menos um mundo- α próximo que é um mundo- $\neg\beta$. Grosso modo, tão mais próximo do mundo atual será um mundo possível, quanto menos ele diferir na quantidade de verdades, ou proposições verdadeiras.

Podemos então reescrever (1) e (2) usando o jargão dos mundos possíveis:

- (1) Nos mundos mais próximos em que α não é o caso, S não acredita que α .
- (2) Nos mundos mais próximos em que α é o caso, S acredita que α .

Nozick argumenta extensivamente a favor dessa abordagem; ele mostra como ela dá conta de vários casos à la Gettier. Mas pelas razões apresentadas acima, só nos interessa CELEIRO. E já sabemos que a abordagem do rastreamento explica porque CELEIRO é um caso de falha de conhecimento, uma vez que a abordagem das razões conclusivas já fez esse trabalho. No jargão dos mundos possíveis, R é uma razão conclusiva para a crença de S de que α se, e somente se, nos mundos possíveis mais próximos ao atual nos quais não é o caso que α , R não é o caso. Embora não equivalentes, a condição para uma razão conclusiva pode ser entendida em termos da condição de variação: basta supormos que a evidência que constitui R é a responsável por S formar a crença de que α . Assim, nos mundos próximos em que α é falsa, R não é obtida e, por conseguinte, S não forma a crença de que α .²⁸ No que concerne CELEIRO: há pelo menos um mundo próximo em que João está olhando para um celeiro falso e ainda

²⁷ Nozick usa a semântica Lewis-Stalnaker, considera hoje padrão, para as contrafactuais. Uma razão pela qual muito provavelmente Dretske não lançou mão dessa semântica é o fato dela ser bastante recente à época: Stalnaker publicou seu “A theory of conditionals” em 1968, e Lewis seu *Counterfactuals* em 1973.

²⁸ No próximo capítulo, ao discutir a indexação do método de formação de crença à sensitividade, ficará ainda mais claro que uma razão conclusiva é produto da sensitividade.

assim forma a crença de que há um celeiro em sua frente. Assim, a crença de João viola a condição de variação; por conseguinte é insensitiva (ainda que verdadeira no mundo atual).

No próximo capítulo apresentarei a teoria da sensitividade de maneira mais extensiva e apresentarei alguns dos problemas mais proeminentes que ela encontra. Por hora, o importante a termos em mente é que a sensitividade é bem sucedida em lidar com a dificuldade genuína apresentada por casos da categoria de CELEIRO. Ou seja, a sensitividade nos fornece uma condição anti-sorte plausível para o conhecimento.

3.2.3. *Virtudes e função apropriada*

Há pelo menos outras três tentativas de resposta ao problema de Gettier que podem ser vistas como desenvolvimento natural da abordagem confiabilista, embora não explicitamente do confiabilismo modal ou contrafactual. São elas: (a) conhecimento como crença apta, (b) conhecimento como crença creditável, e (iii) conhecimento como crença verdadeira garantida. Todas as três abordagens se enquadram naquilo que podemos chamar de *confiabilismo de virtude*, a perspectiva de que uma crença verdadeira só conta como conhecimento se for resultado do exercício de uma virtude cognitiva – sendo uma virtude cognitiva entendida como uma capacidade, habilidade ou disposição do sujeito para formar crenças verdadeiras e evitar crenças falsas.

a. **Crença apta**

Ernest Sosa (2007) defende que conhecimento é crença apta, isto é, uma crença verdadeira produzida pelo exercício de uma virtude intelectual. O status epistêmico positivo de uma crença depende de ela ser *acurada* – *i.e.*, atingir seu alvo –, *habilidosa* – *i.e.*, manifestar competência por parte do sujeito – e, *apta* – *i.e.* atingir seu alvo devido à habilidade do sujeito. Sosa pensa que essa tríade AAA [*Acurácia, hAbilidade e Aptidão*] é em geral usada na avaliação de quaisquer performances, e, por isso, não seria diferente com performances cognitivas – *i.e.* performances que visam crença verdadeira. Seu exemplo predileto é a performance de um arqueiro. Os disparos do arqueiro visam o alvo. Mas há disparos que acertam o alvo e ainda assim diríamos não merecer qualquer crédito; disparos em que o arqueiro acerta por pura sorte. Exigimos que o disparo além de acertar o alvo, também manifeste a habilidade do arqueiro em acertar o alvo. Mas ainda assim há disparos habilidosos que acertam o alvo e não são dignos de crédito. Por exemplo, se um vento desviasse a flecha

mas alguém tropeçasse no alvo de modo que colocasse o alvo na direção da flecha; ainda assim diríamos que o acerto foi por sorte. Por isso, para que um acerto seja digno de crédito é preciso que ele seja devido à habilidade do arqueiro.

Uma crença gettierizada não é apta – uma vez que a sua acurácia não se deve à habilidade cognitiva do sujeito (p. 97). Mas isso dá conta de CELEIRO? A resposta é não. Ao olhar para o celeiro, o sujeito forma uma crença verdadeira – atinge o alvo – e essa crença é produto da sua habilidade cognitiva – manifesta competência –, e é verdadeira porque habilidosa.

b. Crença creditável

John Greco (2003) sustenta que o crédito que o sujeito merece pela sua crença verdadeira, quando possui conhecimento, se deve ao exercício de suas habilidades cognitivas confiáveis; porém, essas habilidades têm de constituir a parte mais saliente da explicação de por que a crença é verdadeira, e nenhum outro fator pode se sobrepujar à centralidade da habilidade cognitiva na explicação crença.

Num caso à la Gettier, contudo, algum fator extraordinário entra em cena e se sobrepuja à centralidade da habilidade cognitiva do sujeito na explicação da verdade da crença. Consideremos CELEIRO novamente. Pensemos primeiro numa situação normal, em que o sujeito não está na roça dos celeiros falsos. Quando o sujeito olha para o celeiro e forma a crença verdadeira de que há um celeiro em sua frente, podemos identificar pelo menos dois *explananda*: (i) o verificador da crença, e (ii) a habilidade cognitiva do sujeito em reconhecer celeiros. Nada de especial teria a crença do sujeito a não ser que o seu aparato visual tenha um papel proeminente na formação de sua crença. E exatamente porque (ii) tem esse peso explicativo é que dizemos que o sujeito sabe que há um celeiro em sua frente.

Por sua vez, o sujeito na roça dos celeiros falsos não sabe por que o fato de haver ali vários celeiros falsos desloca a importância da habilidade cognitiva do sujeito na explicação da verdade de sua crença. Além de (i) e (ii), temos também (iii) o fato de que o sujeito poderia ter muito facilmente formado uma crença falsa. Note que (iii) não exclui (ii) do *explanans*; a crença ter sido formada visualmente ainda é parte da explicação da verdade da crença. Contudo, a parte mais proeminente da explicação é que por um golpe de sorte o sujeito olhou para o único celeiro real da redondeza.

A moral da história é que casos à la Gettier deslocam a centralidade da habilidade cognitiva na explicação da crença verdadeira. Portanto, em casos à la Gettier o sujeito não merece crédito por sua crença verdadeira, e, por conseguinte, não sabe.

Embora atraente e bastante plausível à primeira vista, a abordagem da crença creditável parece pecar por excesso. Ela deixa de fora o conhecimento por testemunho (cf. LACKEY 2007). Por exemplo, sei que o pai de Lima Barreto era alcóolatra pelo testemunho de sua biógrafa; porém, não mereço qualquer crédito por ter essa crença verdadeira. Além do mais, não há qualquer habilidade cognitiva que eu tenha que seja parte saliente da explicação de por que a minha crença é verdadeira. Greco (2007) ainda insiste que a crença testemunhal merece crédito. Mas não entrarei aqui nesse debate. Penso que parte importante de sua resposta é facilmente incorporada à teoria do rastreamento. Seja lá como for que tenhamos de remendar a explicação da verdade de uma crença testemunhal para que o sujeito que aceita um testemunho seja creditado por sua crença verdadeira, na medida em que for razoável supor que crença verdadeira creditável e sorte são incompatíveis, a sensibilidade terá de desempenhar um papel importante.²⁹

c. Função apropriada

Alvin Plantinga (1993b) sustenta que conhecimento é crença verdadeira garantida, sendo a garantia aquela propriedade que em grau suficiente torna a crença verdadeira conhecimento. Para contar como garantida, por sua vez, uma crença tem de (i) ter sido formada por um mecanismo cognitivo funcionando apropriadamente, (ii) num ambiente adequado (para o qual foi moldado [pela evolução ou por desígnio] para funcionar, (iii) que vise a crença verdadeira, e (iv) tenda a produzir mais crenças verdadeiras do que falsas.³⁰

Embora as condições (i)-(iv) sejam cumpridas em todos os casos de conhecimento, elas não são suficientes para evitar casos à la Gettier.³¹ E como era de se esperar, também não dá conta de CELEIRO. O sujeito forma a sua crença através da visão, um mecanismo confiável e que funciona perfeitamente bem. Portanto, crença verdadeira garantida é compatível com a sorte.

²⁹ Quanto à sensibilidade na epistemologia do testemunho veja GOLDBERG (2012).

³⁰ No Capítulo III apresentarei a teoria da função apropriada com mais detalhes.

³¹ Linda Zagzebski (1994) apresenta o caso dos irmãos gêmeos. Não é um caso legítimo, contudo. Como argumentei, é a atribuição correta de crença ao sujeito deveria ser de uma crença falsa.

A resposta de Plantinga (1997; 2000) foca-se no ambiente. Ele insiste que o ambiente em CELEIRO não é adequado. E para isso ele introduz uma distinção entre macro e micro ambientes. Casos à la Gettier são casos em que o micro-ambiente não é favorável para a aquisição de crença verdadeira (ainda que o mecanismo cognitivo em uso seja em geral confiável). Inicialmente Plantinga caracteriza (contrafactualmente) um micro-ambiente favorável como aquele no qual “estivesse S a formar uma crença através de E [o exercício dos poderes cognitivos que produzem a crença B] em BEM [o micro-ambiente em relação a B e E], S formaria uma crença verdadeira” (2000, p. 159). Contudo, ele abandona tal caracterização em favor de uma caracterização probabilística, da qual não trataremos aqui. A minha sugestão é que podemos caracterizar um micro-ambiente favorável (e desfavorável) usando a noção de sensibilidade. Nesse caso, um micro-ambiente será favorável quando não tornar a crença de S (adquirida através de um mecanismo confiável) insensitiva. Esse é um resultado muito bem vindo, uma vez que o problema com celeiro é exatamente a insensitividade da crença do sujeito. O argumento a favor da sensibilidade incorporada à teoria da função apropriada será desenvolvido no Capítulo III.

Note que as duas perspectivas anteriores claramente levam em conta o ambiente. Porém não dão tanta centralidade a ele quanto à teoria da função apropriada. Na medida em que a teoria da função apropriada possa ser entendida como um tipo de confiabilismo de virtudes, parece-me razoável dizer que tanto a teoria da crença apta quanto a da crença creditável sejam casos especiais dela.

É importante notar, contudo, que os epistemólogos da virtude tendem a deslocar o objeto de avaliação epistêmica da crença para o sujeito. Isso as diferencia das demais teorias confiabilistas, que geralmente se preocupam mais com a produção da crença do que com as virtudes cognitivas do sujeito. Assim, para um confiabilista processual ou para o teórico da função apropriada o agente é avaliado epistemicamente apenas secundariamente; ao passo que para os epistemólogos da virtude, os agentes é que primeiro são avaliados epistemicamente, e só secundariamente é que a crença é avaliada (cf. GRECO & REIBSAMEN 2018). Essa mudança de foco se deve pelo interesse na dimensão normativa do conhecimento, e, principalmente, pela intuição de que o conhecimento requer habilidade cognitiva. Contudo, como vimos brevemente acima, o exercício de uma habilidade cognitiva por si só não é suficiente para evitar crença verdadeira por sorte. A epistemologia da virtude, por isso, não

nos fornece uma condição anti-sorte. Ela precisa, como argumenta (e efetivamente o faz) Duncan Pritchard (2012a), ser implementada com uma condição anti-sorte.³² Feito isso, temos uma teoria do conhecimento minimamente robusta e informativa, que além de satisfazer a condição anti-sorte, parece ir na direção correta quanto a dimensão analítica.

4. Sorte epistêmica

Embora a introdução da condição de sensitividade pareça lidar adequadamente com CELEIRO, o que nos garante que ela não seja apenas uma condição *ad hoc*? A razão que tenho para pensar que a sensitividade é de fato a condição adequada é a ideia de que a sorte epistêmica é um fenômeno modal. Isto é, a crença do sujeito de que α é verdadeira por sorte somente se há situações contrafactuais em que α é falsa, o sujeito está aproximadamente na mesma situação epistêmica em que se encontrava originalmente, e ainda assim forma a crença de que α . Dito de maneira mais simples: se α fosse falsa, ainda assim S acreditaria que α . Sendo um fenômeno modal, é natural que uma condição modal seja necessária para evitar a sorte epistêmica.

4.1. Sorte

Embora bastante recorrente nas discussões filosóficas, os filósofos parecem ter ignorado as complexidades e enigmas filosóficos levantados pela noção de sorte. Apenas recentemente é que apareceu aquilo a que podemos chamar “filosofia da sorte” (cf. PRITCHARD & WHITTINGTON 2015; BRECANO-BERROCAL 2016). Além de ser interessante por si própria, uma teoria sobre a natureza da sorte pode ser iluminante a vários problemas filosóficos, *e.g.*, livre-agência e responsabilidade moral, justiça, e, o que mais nos interessa aqui, o conhecimento. Não vou me ater aqui a um exame desse debate, isso iria além das minhas capacidades e do tema desta tese. Antes, apresentarei aquilo que considero ser uma concepção filosófica minimamente razoável do conceito de sorte – a saber, a concepção modal (ou pelo menos parcialmente modal). Em seguida discutirei a sorte epistêmica a partir da concepção modal.

³² Pritchard defende uma epistemologia da virtude anti-sorte. Ele, contudo, privilegia a condição de segurança ao invés da sensitividade. Na medida em que a tese defendida nesta tese se aproxima de uma epistemologia da virtude, poderíamos chama-la de “epistemologia da virtude anti-sorte baseada na sensitividade” por falta de um nome mais elegante. Contudo, prefiro manter o não mais elegante Epistemologia da Função Apropriada Sensitiva - EFAS.

A nossa maneira de falar faz parecer que os portadores primários de sorte são os eventos. Por exemplo, foi *por sorte* que Alexander Fleming descobriu a penicilina. E foi também *por sorte* que ganhei no jogo do bicho. E muito embora jogar no bicho seja uma contravenção penal (!), eu não ter sido punido por isso não foi certamente um evento de sorte. Dessa perspectiva, um evento por si só tem a propriedade de ter ocorrido por sorte. Nesse caso, dizemos que a sorte é *simpliciter*. Mas seria bastante estranho dizer que o desabamento de um determinado jequitibá no meio da mata foi algo que ocorreu por sorte. Parece mais adequado dizer que a sorte é *relacional*: um evento ocorre por sorte em relação a algum agente com determinadas expectativas. Por exemplo, se eu estivesse na mata procurando por um tronco de jequitibá e naquele momento o jequitibá desabasse ali, faria sentido eu dizer que foi pura sorte o jequitibá ter desabado ali – a despeito de se ele estava condenado ou não a desabar. Do mesmo jeito, Fleming não esperava que suas culturas de bactérias fossem contaminadas por mofo ou que quando ele voltasse de férias suas bactérias tivessem sido mortas pelo mofo. E quanto ao bicho, é claro que tenho expectativas em ganhar, muito embora as expectativas de perda sejam muito maiores devido à baixa probabilidade de se acertar o resultado.³³ Podemos então dizer que:

(SR) Um evento E ocorre por sorte para um agente A sse (i) a expectativa de A quanto a ocorrência de E for baixa, e (ii) $\neg E$ poderia facilmente ter ocorrido.³⁴

Um evento pode ocorrer por *má* sorte ou *boa* sorte. Por exemplo, ao procurar pelo tronco de jequitibá, e tendo poucas expectativas de que eu vá encontrar um (pois eu não iria derrubar um), o desabamento do jequitibá ocorreu por boa sorte. Por outro lado, suponha que eu tenha costume de vasculhar uma região de mata onde é abundante a ocorrência de troncos de jequitibá; sempre que vou lá, encontro um. Um dia, porém, chego no lugar e não encontro qualquer tronco – horas antes um grupo de camponeses passou por ali recolhendo lenha. Nesse caso, tive a má sorte de não ter encontrado qualquer tronco. Podemos dizer então que:

³³ Pritchard (2015) defende uma concepção de sorte *simpliciter*. Não considero suas razões satisfatórias, tampouco as discutirei. A título de menção, ele é um dos poucos filósofos a defender a concepção *simpliciter*. Seja como for, *simpliciter* ou relacional, essa questão não afeta o argumento a favor da concepção modal.

³⁴ SR é apenas uma simplificação. 1) Não precisamos considerar que A tenha um grau de expectativa efetivo; a exigência é que no caso do sujeito *racionalmente* pensar sobre E, ele deveria ter um grau baixo de expectativa quanto a ocorrência de E. 2) A pode estar enganado quanto a E. Ele pode ter um grau de expectativa alto e vir a formar a crença, no fim das contas, que E não ocorreu por sorte, quando na verdade ocorreu. A nossa tendência é dizer que se A tivesse pensado cuidadosamente sobre E, não teria atribuído um alto grau de expectativa a E. 3) SR vale para agentes menos que racionais ou não-racionais. Nesse caso, temos de supor um segundo agente que assumiria a perspectiva de A.

(MSR) Um evento E ocorre por má sorte para um agente A sse (i) a expectativa de A quanto a ocorrência de $\neg E$ for alta, e (ii) E não teria facilmente ocorrido.

Podemos também chamar a má sorte de “azar”.

4.2. Conceção modal

As expressões “expectativa” e “teria facilmente ocorrido” não são completamente claras. Intuitivamente poderíamos entendê-las como expressões probabilísticas. Por exemplo, o “grau de expectativa” de um agente poderia ser a *probabilidade prévia* que o agente atribui a E ; ou a *probabilidade condicional* que o sujeito atribui a E dado algum conjunto de evidências. E “ E não teria facilmente ocorrido” poderia ser a probabilidade objetiva de E não ocorrer; ou a probabilidade objetiva de $\neg E$ ocorrer.

Um exemplo estatístico nos ajuda. Suponha que Dimitri esteja jogando roleta russa com seus amigos paranaenses. Ele coloca uma bala no revólver, gira o tambor e a engatilha. As chances de ele tomar o tiro são de $1/6 = 0,16$. O seu grau de expectativa de que não vai tomar o tiro é por isso de $5/6 = 0,83$. Quando Dimitri puxa o gatilho e toma o tiro, dizemos que ele foi sortudo (ou melhor, azarado!). Esse resultado está de acordo com SR e MSR.

Mas nem todo caso de sorte pode ser decidido levando-se em conta apenas as probabilidades. Considere Hércules, que mora na parte mais alta da cidade e tem pretensões de se mudar de lá. Ao ganhar um aumento de salário, ele vai a uma imobiliária, olha uma casa para alugar, mas não fecha negócio – faltava-lhe um fiador. Dias depois a casa que ele olhou é completamente inundada por uma enchente que tomou toda a parte baixa da cidade. Hércules teve sorte de não ser atingido pela enchente. Contudo, a probabilidade de isso acontecer era baixa. Ou seja, não é o caso de a probabilidade dele ser pego pela enchente ser alta. Logo, nem todo caso de sorte envolve a satisfação da condição (ii).

Interpretemos “ E não teria facilmente ocorrido” modalmente: não há *circunstâncias contrafactuais* relevantes em que E ocorre. Ou usando o jargão dos mundos possíveis: nos mundos possíveis mais próximos ao atual E não ocorre. E do mesmo modo com “ E teria facilmente ocorrido”: há pelo menos um mundo possível próximo ao atual em que E ocorre. A interpretação modal explica o caso de Hércules. Foi sorte Hércules não ter sido atingido pela

enchente porque há pelo menos um mundo possível próximo ao atual em que ele consegue um fiador e se muda para a casa. Ou seja, a sorte é uma questão de proximidade entre mundos.³⁵

A interpretação modal, porém, não exclui a sorte probabilística. Toda sorte probabilística é também modal, mas não o contrário. Ou seja: a sorte probabilística é um caso especial de sorte modal. Dizer então que a sorte é um fenômeno modal não é negar que nenhum caso de sorte seja probabilístico; é apenas negar que todo o caso de sorte seja probabilístico.

4.3. Sorte epistêmica verítica

Há pelo menos três tipos de sorte epistêmica (PRITCHARD 2005): (i) sorte epistêmica evidencial, que é quando o sujeito adquire a evidência a favor de sua crença por sorte; (ii) sorte epistêmica doxástica, que é quando o sujeito teve sorte em acreditar que α ; e (iii) sorte epistêmica verítica, que é quando a crença do sujeito é verdadeira por sorte. (i) e (ii) não são incompatíveis com o conhecimento. Apenas (iii) é incompatível. É o tipo de sorte envolvido nos casos à la Gettier. Doravante, usarei o termo “sorte epistêmica” para me referir apenas a (iii).

Aplicando então a nossa caracterização modal da sorte à sorte epistêmica verítica, temos:

(SEV) A crença de S de que α é verdadeira por sorte sse (i) S espera (deveria esperar) que a sua crença esteja conectada apropriadamente ao seu verificador, e (ii) a crença de S de que α teria sido facilmente falsa.

Dizer que a crença de S teria sido facilmente falsa quer dizer que há pelo menos um mundo possível próximo ao atual no qual S acredita que α e α é falsa.

Tomando emprestada de Sosa (2007) a analogia com os disparos de um arqueiro: o arqueiro visa o alvo, mira e dispara a flecha. Um disparo bem sucedido é aquele em que o arqueiro atinge o alvo não por sorte, mas devido às suas habilidades. Se um vento desviasse

³⁵ Pritchard (2015) usa o caso da loteria para ilustrar que “um evento pode ser modalmente próximo mesmo quando é probabilisticamente improvável” (p. 145). Se compro um bilhete de uma loteria com 1.000 bilhetes e acredito com base nas chances que tenho de ganhar – 1/1000 – que o meu bilhete não é o premiado, a despeito do quão improvável é eu ter comprado o bilhete premiado, ainda assim não sei que meu bilhete não é o premiado. Há um mundo bastante próximo em que o meu bilhete é o sorteado, mundo esse no qual eu ainda acreditaria que não tenho o bilhete premiado.

levemente a flecha de modo que no final ela não atingisse o alvo, mas por puro acaso o mesmo vento tivesse também arredado o alvo e a flecha o atingisse, não consideraríamos que o arqueiro tivesse sido bem sucedido. Foi por sorte que ele acertou o alvo. Do mesmo modo, ainda que a crença do sujeito atinja seu alvo – a verdade –, e ainda que seu disparo tenha sido acurado – a crença foi produzida por um mecanismo confiável –, sua crença só é conhecimento se não atingiu o alvo acidentalmente (por sorte). E dizer que o arqueiro acertou o alvo por sorte significa dizer que ele teria facilmente errado aquele disparo; ou dito de outro modo: há um mundo próximo ao mundo atual em que o arqueiro erra o alvo a despeito de ter disparado a flecha em condições similares.

4.4. Condição anti-sorte

Espero que já esteja claro a esta altura que a sorte epistêmica é um fenômeno modal: um sujeito não pode ter conhecimento de que α se ele facilmente tivesse tido a mesma crença de que α caso ela fosse falsa. A condição de sensibilidade lida com esse problema: naqueles mundos mais próximos ao mundo atual nos quais α é falsa, S não acredita que α . Ou seja: se a contrafactual da variação é verdadeira, não há mundos próximos em que α seja falsa e S acredite que α . Isso é suficiente para que a introdução da sensibilidade não seja apenas uma manobra *ad hoc*, cuja única finalidade é remendar as entediantes tentativas de evitar contraexemplos à la Gettier.

Porém, outra condição modal parece também capaz de lidar com a sorte epistêmica: a condição de *segurança* (cf. LUPER-FOY 1984; SOSA 1999; PRITCHARD 2005; WILLIAMSON 2000). Grosso modo, a crença de S de que α é segura se, e somente se, se S acreditasse que α , α seria o caso. Ou no jargão dos mundos possíveis: nos mundos mais próximos ao atual em que S acredita que α , α é o caso. Assim, se o conhecimento é pelo menos crença verdadeira *segura*, então não haverá quaisquer mundos próximos em que não seja o caso que α e ainda assim S acredite que α .

A sensibilidade e a segurança, em geral, não divergem sobre os veredictos de conhecimento e na explicação das falhas de conhecimento em casos à la Gettier. O teste de plausibilidade para ambas reside na capacidade de resolução de certos problemas adicionais e pormenores. Os defensores da segurança pensam levar vantagem sobre pelo menos três aspectos: (i) fechamento epistêmico, (ii) resposta mooriana ao ceticismo, e (ii) conhecimento

indutivo. Os proponentes das teorias rastreadoras tradicionais, com bom sabido, aceitaram de bom grado a invalidade do princípio de fechamento e a impossibilidade de se refutar uma hipótese cética como a do gênio maligno. Sendo assim, não é uma vantagem óbvia que uma teoria do conhecimento tenha alguma vantagem já de saída simplesmente por manter (i) e (ii). Quanto ao conhecimento indutivo, aí sim já podemos considerar uma vantagem. Mas argumentarei (Cap. III) que a sensitividade é capaz de lidar perfeitamente bem com o conhecimento indutivo.

Por agora, apenas suporei que a sensitividade é superior à segurança. No Capítulo IV avançarei algumas razões de por que penso que a segurança não é uma condição anti-sorte satisfatória.

5. Conhecimento gettierizado?

Até agora tenho tratado – seguindo a ortodoxia – a sorte epistêmica como um entrave ao conhecimento. Vale a pena, contudo, considerar um argumento contra essa suposição. Stephen Hetherington tem argumentado (1998; 1999; 2001) que o conhecimento é compatível com a sorte epistêmica. Nesta seção considerarei seu argumento desenvolvido em seu *How to Know* (2011).

Hetherington (pp. 80-2) apresenta um contraexemplo que pode ser resumido como se segue:

[NOAH] Noah olha para uma árvore e forma a crença de que há uma árvore em sua frente. Além disso, Noah também acredita que o conhecimento perceptual – e em particular o seu suposto conhecimento de que há uma árvore em sua frente – se dá como diz o realismo ingênuo sobre a percepção. Seja como for, Noah sabe que há uma árvore sem sua frente.

A primeira pergunta a se fazer é: de que modo a crença de Noah está gettierizada? Ou seja, como a crença de Noah é verdadeira por sorte?³⁶ De acordo com Hetherington – e também com a perspectiva que venho assumindo – uma crença é verdadeira por sorte quando a evidência (entendida em sentido amplo) de que dispõe o sujeito para uma determinada crença não está adequadamente conectada ao seu verificador; ou, de maneira mais ou menos

³⁶ Hetherington usa o termo “gettierizado” de modo a não implicar que o sujeito que tenha uma crença gettierizada não tenha conhecimento. Isso é essencial para seu argumento, uma vez que se toda crença gettierizada fosse uma falha de conhecimento, não poderia haver conhecimento gettierizado, como ele pensa que há.

equivalente, quando o verificador da crença é diferente daquilo que o sujeito deveria esperar dada a evidência disponível.³⁷ Concentremo-nos então na maneira comum (e ingênua) pela qual geralmente consideramos a evidência perceptual. Um objeto de tamanho médio – a árvore – nos é apresentado aos sentidos como tendo tais e tais propriedades – com textura tal, formas e cores tal. Tendemos a pensar que esse objeto é, no mundo externo, aproximadamente como nos foi apresentado à consciência. Contudo, a descrição científica do mundo nos diz algo diferente: várias das propriedades que a física diz que os objetos têm nos são inacessíveis aos sentidos; e várias das propriedades que atribuímos aos objetos, a física nos diz que não existem nos objetos.³⁸ Chamemos a essa última perspectiva “imagem científica do mundo”, e à primeira “imagem manifesta do mundo”, para usar os termos de Wilfrid Sellars (1963). Considerando agora que Noah sabe que há uma árvore em sua frente e que ele pensa estar de acordo com a imagem manifesta do mundo, a sua evidência para acreditar que há uma árvore em sua frente aponta para um tipo de verificador distinto do que torna a sua crença verdadeira. Isto é, o verificador da crença de Noah é mais provável de estar de acordo com a imagem científica do que com a manifesta, a despeito de ele saber que há uma árvore em sua frente. Ora, isso está de acordo com uma crença ter sido verdadeira por sorte. Portanto, a crença de Noah está gettierizada. E é ao mesmo tempo conhecimento. Nas palavras do próprio Hetherington:

Podemos dizer que a boa evidência falível de Noah é parte da sua imagem *manifesta* do mundo, ao passo que o modo como a sua crença se torna efetivamente verdadeira é em última instância melhor descrita em termos da imagem *científica*. A situação é então uma situação à la Gettier, incorporando a sorte à la Gettier, pois os fatos descritíveis apenas cientificamente de como a crença de Noah vem a ser verdadeira estão longe de ser aquilo que ele esperava *dada* apenas a sua evidência observacional ingenuamente entendida. Essa última é organizada em termos de uma imagem manifesta. E é bastante desinformada acerca da imagem científica que *podia* ser aplicada (por alguém mais que não Noah) à situação de modo a descrever acuradamente como a sua crença efetivamente se tornou verdadeira. (p. 82)

Embora engenhoso, esse argumento falha em última instância. A razão é que para funcionar, teríamos de aceitar que a evidência descrita em termos da imagem manifesta não se

³⁷ À primeira circunstância Hetherington chama “desconexão à la Gettier” e à segunda “sorte à la Gettier”. Para fins de argumento ele trata as duas de maneira equivalente.

³⁸ É famosa aqui a distinção entre qualidades primárias e secundárias feita por Galileu e também Locke (cf. LANDESMAN 2002). Essa distinção deu origem ao realismo indireto – a posição segundo a qual os objetos têm um conjunto de propriedades independentes de como pensamos que o mundo seja e que a percepção nos apresenta os objetos indiretamente sem necessariamente nos revelar as propriedades reais dos objetos; temos de ter o trabalho de inferir os objetos a partir das nossas representações mentais do mundo (cf. RUSSELL 1912).

refere àquilo que é descrito pela imagem científica. Essa perspectiva tem de estar errada, pois como Noah poderia saber que há uma árvore em sua frente uma vez que ele não se refere a ela?

Uma primeira leitura do argumento de Hetherington sugere que Noah, para saber que há uma árvore em sua frente, primeiro tem de ter feito uma inferência cujas premissas são a descrição daquilo que aparece a ele e um enunciado do realismo ingênuo. Mas essa seria uma exigência demasiado alta. Mas podemos entender de maneira mais branda: pelo menos em princípio Noah teria de ser capaz de fazer tal inferência (se, por exemplo, alguém lhe perguntasse como é que ele sabe). Mesmo essa exigência ainda me parece exagerada – ainda que Noah fosse completamente incapaz de fornecer tal argumento, ele saberia que há uma árvore em sua frente. Mas concedamos, para fins de argumentação, que as coisas funcionem como quer Hetherington.

Uma vez que Hetherington diz que a evidência ingênuo aponta para um verificador diferente do que aponta a descrição científica, temos de considerar que a evidência proposicional (ingenuamente descrita) é falsa. E se considerarmos também que o realismo ingênuo é falso, temos de considerar que Noah inferiu sua crença a partir de premissas completamente falsas – *i.e.* todas as premissas do argumento são falsas. Isso nos leva então à perspectiva de que podemos ter conhecimento a partir de informações falsas. Concedamos também que isso seja possível. A pergunta a se fazer, portanto, é a seguinte: por que é razoável atribuir conhecimento a Noah?

A minha resposta é: (i) porque ainda que estritamente falsas, as sentenças usadas por Noah referem-se à árvore e não a outro objeto – *i.e.*, Noah é bem sucedido ao fazer referência ao objeto pretendido; (ii) a descrição científica do mundo não precisa ser vista como incompatível com a descrição de senso comum – podemos tomar como legítimo o discurso de senso comum, ainda que estritamente seja falso do ponto de vista científico.

A fim de estabelecer (i), comecemos com um diálogo entre um “habitante” do mundo do senso comum – Zé – e um “habitante” do mundo científico – Galileu. Vamos supor que os habitantes do mundo científico tenham evoluído de modo que seu aparato visual funcione como um microscópio potentíssimo. Onde o habitante do senso comum vê um líquido incolor e inodoro, o habitante do mundo científico vê duas moléculas de hidrogênio ligadas a uma de

oxigênio. Ambos usam o termo “água” para se referir à mesma coisa; além de ambos terem práticas bastante semelhantes no que diz respeito à água. Zé pede um copo de água a Galileu. Galileu vai até o filtro e pega água. Galileu pede a Zé que enxugue a água que escorreu pela pia; Zé vai até lá e enxuga. Os dois se comunicam perfeitamente bem. Contudo, quando Zé diz a Galileu, “Veja aquele riacho, que água límpida e transparente”; ou quando Galileu diz a Zé, “Vê aquele ângulo formado pela ligação entre as moléculas, eles têm 104,5°”; as coisas começam a não funcionar. Inicialmente eles pensam que não estão falando sobre a mesma coisa, mas com o passar tempo pensam que, à medida em que concordam em resultados práticos, só discordam quanto às descrições que fornecem da água.

É inegável que ambos, Zé e Galileu, sabem, por exemplo, que há água no riacho. E isso só pode ser o caso, se for verdade que ambos estão a se referir à mesma coisa. Mas e quanto às propriedades fenomênicas da água? Se a descrição científica está correta, implica que a descrição de senso comum seja falsa. Sim e não. Sim, porque de um ponto de vista científico propriedades fenomênicas são qualidades secundárias; se pensarmos num mundo sem agentes cognitivos que experienciam os objetos, não faz sentido atribuir propriedades fenomênicas. E não, porque se considerarmos a relação dos mesmos agentes cognitivos com os objetos, as propriedades fenomênicas emergem, de modo que faz completo sentido dizer que sentenças como “aquela porção de água é límpida” descrevem propriedades reais (embora secundárias).

Sendo assim, para estabelecer (ii), precisamos de uma teoria que explique o discurso de senso comum. Penso ser suficiente apenas apontar como isso poderia ser feito. (A) teoria do erro: o discurso de senso comum é sistematicamente falso, muito embora não sejamos capazes de evitar usá-lo. (B) reducionismo: para cada propriedade de senso comum é possível reduzir essa propriedade a um conjunto de propriedades científicas. (C) emergentismo: embora não redutíveis a um conjunto de propriedades científicas, as propriedades secundárias são contrafactualmente dependentes das propriedades científicas.

E a razão para pensar que isso seja assim, é o fato inegável de que não há qualquer teoria do senso comum que pretenda substituir a descrição científica da realidade.³⁹ Por outro

³⁹ Não estou levando em conta aqui a opinião de relativistas ou construtivistas radicais que pensam que não há qualquer realidade independente da mente, que não há categorias naturais, e que os recortes feitos por nossas teorias é que criam as categorias científicas. Esse tipo de perspectiva já é bem conhecido pelo caráter autoderrotante (cf. BOGHOSSIAN 2006; NAGEL 1997).

lado, também não estamos dispostos a abandonar o vocabulário do senso comum por amor à precisão e ao realismo científico. Mesmo o realista científico mais sangue nos olhos continua a dizer que “não gosta de usar aquele casaco porque ele é muito quente”.

Mas como isso tudo mostra que a crença de Noah de que há uma árvore em sua frente não está gettierizada? Ora, na medida em que as descrições de senso comum e as descrições científicas apanhem o mesmo objeto, a crença de Noah de que há uma árvore em sua frente é verdadeira. E uma vez que as propriedades de senso comum tenham alguma dependência das propriedades científicas – se o objeto não tivesse tais e tais propriedades científicas, tais e tais propriedades fenomênicas não apareceriam –, a crença de Noah não é verdadeira por sorte. Pensando em termos de mundos possíveis: não há mundos próximos ao mundo atual nos quais Noah tenha adquirido a evidência que adquiriu e ela seja produto de fatos científicos distintos dos fatos do mundo atual.

Portanto, Hetherington não tem razão para pensar que haja conhecimento gettierizado. Por conseguinte, não temos razão para negar que a sorte epistêmica seja incompatível com o conhecimento.

Penso ter estabelecido o principal ponto deste capítulo: uma teoria do conhecimento satisfatória tem de oferecer uma condição anti-sorte. E uma vez que a sorte é um fenômeno modal, é de se esperar que tal condição tenha também um caráter modal. Sugeri que a sensibilidade é capaz de fazer bem o trabalho. Passemos então ao capítulo seguinte a fim de considerar a sensibilidade.

II

SENSITIVIDADE

[N]ão importa o quão convincente seja um argumento contra a sensibilidade; quando alguém a vê pela primeira vez, a sua elegância e plausibilidade intuitiva são inegáveis.

Kelly Becker e Tim Black, *The resilience of sensitivity*

Prévia

O resultado a que chegamos no Capítulo I foi o de que a condição de sensibilidade é um candidato natural a condição anti-sorte. Neste capítulo esmiuçaremos um pouco mais a noção de sensibilidade. Por fim, veremos algumas objeções influentes.

1. Rastreamento

Recordemo-nos de CELEIRO:

João está dirigindo pelo interior e, sem que saiba, está passando por um distrito em que as pessoas têm a mania de construir celeiros falsos, de fachada. Esses celeiros são indistinguíveis de celeiros reais. João olha para um dos celeiros à beira da estrada e forma a crença verdadeira de que há um celeiro em sua frente. Por sorte, João está olhando para o único real daquela região.

João não sabe que há um celeiro em sua frente. Se eu estiver correto (como argumentei no Capítulo I), esse é o único caso à la Gettier genuíno. Esse tipo de caso deixa claro que há algo de errado no ambiente em que a crença é formada. Mais especificamente, a crença de João não é capaz de rastrear ou detectar a verdade da proposição em causa. Costumamos dizer que *ainda que João estivesse olhando para um celeiro falso, continuaria acreditando que há um celeiro em sua frente*. A analogia com uma máquina aqui é óbvia. Pense num detector de metais. Esperamos que um bom detector apite sempre que um metal passe pelo seu sensor. Contudo, se o detector apita quando algum material de plástico muito rígido passa pelo seu sensor, não consideraríamos tal sensor bom. Haverá casos em que ele não rastreará a presença

de metal, uma vez que *estivesse um material plástico muito rígido ao alcance de seu sensor, ele apitaria do mesmo modo*. Isso sugere que não pode haver conhecimento sem rastreamento.

Mas como tornar precisa tal noção? Note que a explicação da falha de rastreamento é dada através de uma afirmação contrafactual “se fosse o caso que tal e tal, então seria o caso que tal e tal”. No caso do celeiro, diríamos então que João sabe que há um celeiro em sua frente somente se *não fosse o caso que há um celeiro em sua frente, ele não acreditaria que há*. Ou dito de outro modo, *se houvesse um celeiro falso em sua frente, ele não acreditaria que há um celeiro em sua frente*. E isso é também o que esperaríamos do detector de metais: caso não houvesse um metal ao alcance de seu sensor, que ele não apitasse. Chamemos a uma crença que rastreia a verdade “sensitiva”. Assim:

A crença de S de que α é sensitiva sse:

- (i) Se não fosse o caso que α , então S não acreditaria que α (variação); e
- (ii) S fosse o caso que α , então S acreditaria que α (aderência).

E diremos que a crença de S de que α é *insensitiva* sse não for sensitiva. Isto é, se ou (i) ou (ii) não forem satisfeitas. Além disso,

S sabe que α somente se a crença de S de que α for sensitiva.

Como a crença de João é insensitiva, ele não sabe que há um celeiro em sua frente, a despeito de tal crença ser verdadeira e formada por um processo confiável.

Embora Nozick use “sensitividade” de modo a cobrir (1) e (2), na literatura mais recente parte dos epistemólogos tende a se restringir a (1) apenas – a condição de variação (cf. DeROSE 1995; BECKER 2007; BLACK & MURPHY 2007; PRITCHARD 2008; BECKER & BLACK 2012). Isso não quer dizer que eles rejeitem (2) como falsa. Em geral, quando a condição de variação é cumprida e o sujeito tem uma crença verdadeira (confiável) no mundo atual, a condição de aderência é automaticamente cumprida; isso supostamente a torna uma condição supérflua.⁴⁰ Seguirei parcialmente a tendência atual em prestar maior atenção à

⁴⁰ Em casos em que a contrafactual (1) é trivialmente satisfeita, a satisfação de (2) deixa de ser automática. Nesses casos talvez seja preciso tornar explícita (2).

condição de variação. Não obstante, mantereí a aderência como condição explícita para sensibilidade.⁴¹

1.1. Contrafactuais e mundos possíveis

A maneira pela qual a maioria dos filósofos pensa nas condicionais contrafactuais é através do jargão dos mundos possíveis. Para simplificar as coisas, podemos pensar em mundos possíveis como modos em que as coisas poderiam ter sido (incluindo o modo como elas efetivamente ou atualmente são). Um mundo possível é, então, um conjunto maximal consistente de proposições.⁴² O mundo atual – que representaremos por “ $w@$ ” – é descrito por um conjunto C infinitamente grande de proposições. Um mundo possível qualquer não-atual – que representaremos por “ w_n ” – é qualquer conjunto maximal consistente de proposições diferente de C . Não importa em quantas proposições difira, contanto que difira em pelo menos uma. Assim, se oferecermos uma descrição completa da realidade que difira do mundo atual apenas no valor de verdade da proposição ⟨A seleção brasileira de futebol foi eliminada da Copa da Rússia⟩, tal descrição contará como um mundo possível. Vê-se facilmente que não há limites lógicos para se criar mundos possíveis.

Para avaliarmos o valor de verdade de uma contrafactual, precisamos ainda da ideia de proximidade entre mundos. Uma contrafactual do tipo *Se fosse o caso que α , então seria o caso que β* – que representamos por “ $\alpha \square \rightarrow \beta$ ” – é verdadeira no mundo atual quando nos mundos possíveis mais próximos ao atual em que α é verdadeira, β também é verdadeira. Ou dito de outro modo, quando os mundos- α mais próximos são também mundos- β . Uma maneira intuitiva de entender a proximidade entre mundos é a seguinte: um mundo possível w_1 será tanto mais próximo de um mundo possível w_2 quanto menos w_1 diferir em proposições verdadeiras de w_2 . Por exemplo, suponha que $w_1 = \{p, q, r, s\}$, $w_2 = \{\neg p, q, r, s\}$ e $w_3 = \{\neg p, \neg q, r, s\}$. Obviamente, uma vez que todo mundo possível é igual a si mesmo, cada mundo w_n é maximamente próximo de si mesmo. w_2 é mais próximo de w_1 do que w_3 (o que implica que w_3 é mais distante de w_1 do que w_2). Suponha agora $w_4 = \{p, q, \neg r, s\}$. Note que w_4 difere em

⁴¹ A razão ficará óbvia quando eu discutir a condição de segurança nos capítulos posteriores.

⁴² Não precisamos nos comprometer com qualquer concepção substancial sobre a natureza dos mundos possíveis: nem que são entidades concretas espaço-temporalmente desconexas; nem que são entidades abstratas, ou que são apenas combinações de estados de coisas do mundo atual. Embora eu tenha escolhido falar dos mundos possíveis como um “conjunto de proposições” – o que parece sugerir que os mundos possíveis sejam entidades abstratas – escolhi-o por facilidade. Seja como for, os mundos possíveis são pelo menos dispositivos úteis para tratar de noções modais.

apenas uma proposição de w_1 , de modo que w_4 é tão próximo de w_1 como o é w_2 . Contudo, w_4 e w_2 são mais distantes um do outro do que são de w_1 . Fica fácil agora também compreender quando uma contrafactual $\alpha \Box \rightarrow \beta$ é falsa: quando houver pelo menos um mundo possível próximo ao atual em que α for verdadeira e β for falsa. (Alternativamente: quando houver pelo menos um mundo- α que seja um mundo- $\neg\beta$).⁴³

Para facilitar a avaliação das contrafactuals podemos recorrer a diagramas parecidos com aqueles que usamos para construir modelos em lógica modal. Considere o seguinte diagrama:

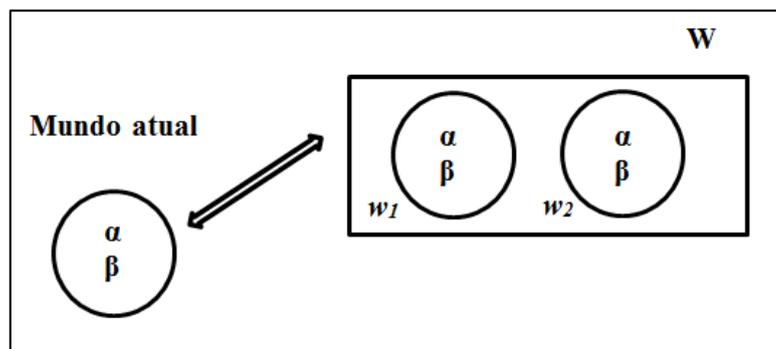


Diagrama 1

O mundo atual, $w_{@}$, é composto por apenas $\{\alpha, \beta\}$. Há apenas dois mundos próximos, w_1 e w_2 . $W = \{w_1, w_2\}$ é o conjunto de mundos próximos.⁴⁴ A condicional $\alpha \Box \rightarrow \beta$ é verdadeira em $w_{@}$ sse α é verdadeira em w_1 e β é verdadeira em w_1 , e α é verdadeira em w_2 e β é verdadeira em w_2 . Ou seja, $\alpha \Box \rightarrow \beta$ é verdadeira sse α e β são verdadeiras em todos os mundos contidos em W (o conjunto de mundos próximos). Do mesmo modo, $\alpha \Box \rightarrow \beta$ é falsa sse houver pelo menos um mundo em W em que α é verdadeira e β falsa.

Com essas noções em mente podemos acompanhar o raciocínio contrafactual na teoria do conhecimento. A sensibilidade pode agora ser posta nos seguintes termos:

⁴³ Essa é uma simplificação bastante grosseira da semântica para as contrafactuals de Lewis-Stalnaker. Há vários problemas com tal semântica, principalmente no que diz respeito à noção de proximidade entre mundos (cf. STARR 2019). Seguirei aqui a tendência da epistemologia contemporânea e simplesmente assumirei que a semântica de Lewis-Stalnaker capta razoavelmente o raciocínio contrafactual e, e que desvios mínimos são aceitáveis a fim de evitar certos resultados triviais. (cf. NOZICK 1981, p. 680-681; PLANTINGA 2000, p. 135).

⁴⁴ É claro que W poderia ser muito maior; poderia haver um conjunto bastante grande de mundos próximos ao atual. Mas para fins didáticos é melhor trabalhar com poucos.

- (i) **Variação:** nos mundos possíveis mais próximos ao atual em que é falso que α , S não acredita que α . ($\neg\alpha \Box \rightarrow \neg B\alpha$, em que “B” é o operador doxástico “S acredita que”)
- (ii) **Aderência:** nos mundos possíveis mais próximos ao atual em que é o caso que α , S acredita que α . ($\alpha \Box \rightarrow B\alpha$)

Retornemos a CELEIRO. Dissemos que João não sabe que há um celeiro em sua frente, pois *ainda que João estivesse olhando para um celeiro falso, continuaria acreditando que há um celeiro em sua frente*. Coloquemos isso em termos de mundos possíveis. No mundo atual João olha para um celeiro real – chamemos-lhe CR – alocado no lote número 1 e forma a crença de que há ali (no lote 1) um celeiro. Suponhamos que há nessa região 99 celeiros falsos além do único celeiro real (exatamente o visto por João). Vamos considerar agora alguns mundos próximos. Qualquer um dos 99 celeiros falsos poderiam ter sido construídos no lote 1; assim, há pelo menos 99 mundos possíveis⁴⁵ em que a proposição ⟨Há um celeiro no lote 1⟩ é falsa. Note que todos esses mundos diferem de modo relevante em apenas uma proposição do mundo atual. E como em todos esses mundos os celeiros falsos são visualmente indistinguíveis para João, naqueles mundos em que João olha para o lote 1 e vê um celeiro falso, ele acredita que há um celeiro em sua frente. Ora, temos então pelo menos um mundo (na verdade, 99!) em que é falso que há um celeiro no lote 1 e que João acredita que há um celeiro no lote 1. Ou seja, há pelo menos um mundo-¬CR que é também um mundo-B(CR). Portanto, é falsa a condicional *Se não fosse o caso que há um celeiro no lote 1, então João não acreditaria que há*. Note que a condicional da aderência – $CR \Box \rightarrow B(CR)$ – é satisfeita, uma vez que nos mundos-CR mais próximos (nesse caso, apenas o mundo atual) João acredita que há um celeiro em sua frente. Mas para que uma crença seja sensitiva, é preciso que seja *variante e aderente*. Por conseguinte, a crença de João não rastreia a verdade da proposição alvo. Dito de outro modo, a crença de João é *insensitiva*.

Precisamos, contudo, fazer um pequeno reparo na sensitividade. Pensemos numa pequena alteração do cenário anterior. Quando João passa pela estrada e olha em direção ao lote 1, além do celeiro, vê também um simpático caipira vendedor de cestos que costuma alertar as pessoas (quando elas dão o mínimo sinal de erro) que só no lote 1 é que há um

⁴⁵ Na verdade, há $99! = 99 * 98 * \dots * 2 * 1$ mundos próximos, um número extremamente grande. Para simplificar, vamos pensar apenas nos primeiros 99 mundos.

celeiro real. Como João parou exatamente em frente ao lote 1 e deu a entender ao caipira que estava olhando para o lugar certo, sem qualquer sombra de dúvida, o caipira continuou seu trabalho sem avisar a João. Agora é intuitivo pensarmos que se não houvesse um celeiro real no lote 1, João não acreditaria que há; uma vez que o caipira lhe teria apontado o celeiro real. Nesse caso, a crença de João parece ser sensitiva, e o diagnóstico acima torna-se equivocado.

Para evitarmos resultados desse tipo é preciso indexar o método ou processo de formação de crença pelo qual o sujeito veio a acreditar na proposição alvo. No caso original, João forma a crença de que há um celeiro no lote 1 com base na visão; e nos mundos- \neg CR próximos em que ele acredita que há um celeiro no lote 1 quando na verdade há um celeiro de fachada a sua crença é formada também com base na visão. No caso modificado, no mundo atual João forma a sua crença de que há um celeiro no lote 1 através da visão, mas nos mundos- \neg CR próximos em que ele não acredita que há um celeiro de fachada no lote 1, a sua crença não é formada através da visão e sim através do testemunho do caipira. Foi o próprio Nozick (1981, pp. 179-185) quem chamou atenção para tal modificação no seu famoso exemplo da Vó.⁴⁶ A sensitividade pode então ser assim expressa:

A crença de S de que α é sensitiva sse:

- (i) Nos mundos possíveis mais próximos ao atual em que não é o caso que α , S não forma a crença (via método usado no mundo atual) de que α . ($\neg\alpha \square \rightarrow_M \neg B\alpha$, em que “M” designa o método de formação de crença usado pelo sujeito no mundo atual)
- (ii) Nos mundos possíveis mais próximos ao atual em que é o caso que α , S forma (via M) a crença de que α .

Ou podemos dizer a mesma coisa sem fazer referência ao método usando a expressão “do mesmo modo como no mundo atual” com preferem alguns autores.⁴⁷ Doravante, suprimirei a referência ao método ao mencionar a sensitividade. Sempre que usar “nos mundos mais

⁴⁶ No caso da Vó, teríamos um caso em que intuitivamente um sujeito tem conhecimento, mas dado a sensitividade teríamos de classificar como um caso de falha de conhecimento. João sofre um acidente e sua vó vai visita-lo no hospital. Ela vê o estado do neto e forma a crença de que ele está bem. Contudo, a sua família, a fim de evitar que a vó sofresse algum mal, tinha decidido de antemão que se João não estivesse bem, eles mentiriam para ela e diriam que o médico acabara de anunciar o bom estado de João. Assim, há pelo menos um mundo próximo em que João não está bem e a sua vó acredita que ele está bem. Note que no mundo atual a crença da vó é formada pela visão, ao passo que no mundo próximo é formada pelo testemunho falso da família.

⁴⁷E.g., Duncan Pritchard (2005), John Greco (2007).

próximos em que não é o caso que α , S não acredita que α ” ($\neg\alpha \square \rightarrow \neg B\alpha$) ou “nos mundos mais próximos em que é o caso que α , S acredita que α ” ($\alpha \square \rightarrow B\alpha$), tenha o leitor em mente a indexação do método.

Uma objeção poderia ser levantada aqui contra o uso do jargão modal já desde o início: “Não é óbvio que possamos ter conhecimento de possibilidades, e ainda que tenhamos, não é óbvio o modo pelo qual adquirimos tal conhecimento; assim, tentar analisar o conhecimento lançando mão das contrafactuais é explicar o que já não é claro por algo menos claro ainda”. Mais especificamente, o objetor poderia ter o seguinte em mente. Quando dizemos algo como “Se S sabe que α , então nos mundos próximos em que α não é o caso também não é o caso que S acredita que α ”, estamos pressupondo que podemos saber algo que não é atual.⁴⁸ Uma vez que há desacordo sobre o conhecimento modal,⁴⁹ seria inútil tentar explicar o conhecimento empírico apelando à noção de mundos possíveis.

Embora interessante, não penso que tal objeção tenha de ser respondida nesta tese. Pois respondê-la seria aderir a uma estratégia paralisante de se fazer filosofia: se para cada problema filosófico que tratamos tivéssemos que responder a dificuldades prévias tão difíceis quanto a original, provavelmente nunca conseguiríamos responder ao problema original.

Mas talvez eu não esteja sendo justo com a objeção. A ideia é que a menos que eu tenha uma explicação do conhecimento modal, não adianta usar os mundos possíveis para esclarecer a natureza do conhecimento. Contudo, penso não haver essa dependência que o objetor sugere. Considere, por exemplo, o uso que o físico faz da matemática para explicar algum fenômeno. Seria injusto com o físico se só aceitássemos a sua explicação depois que ele também nos explicasse como é possível o conhecimento matemático. Há grande controvérsia filosófica sobre a natureza do conhecimento matemático, e, à partida, seja lá qual for essa explicação ele não altera a explicação do fenômeno físico. O mesmo vale para o epistemólogo. Na medida em que ele usa o jargão dos mundos possíveis como ferramenta heurística, não importa qual seja a resposta filosófica para o conhecimento modal; não deveríamos esperar que isso afetasse a explicação de como o conhecimento empírico é

⁴⁸ Essa preocupação é expressa de maneira bastante clara em HAWTHORNE (1996, p. 183): “*Suponha que o nosso discurso e pensamento modais de fato não nos comprometam com possibilita, mundos possíveis [...] Como explicar então a aparente referência a tais entidades modais em nosso discurso cotidiano [...]?* Suponha então que nosso discurso e pensamentos modais estão comprometidos com possibilita [...] Como sabemos que as tais entidades que pretendemos nos referir existem, e como sabemos como elas se parecerão?”.

⁴⁹ Para uma exposição das principais teorias sobre o conhecimento modal veja VAIDYA (2017).

possível. Afinal, seja lá qual for a natureza do conhecimento modal, todos parecem concordar que afirmações contrafactuais têm significado.

1.2. Sensitividade probabilística?

Há epistemólogos que não se sentiram satisfeitos com a sensitividade modal e acabaram por adotar uma interpretação probabilística, em termos de probabilidade condicional.⁵⁰ Grosso modo, a ideia é que a crença de S de que α é sensitiva sse (i) a probabilidade de S não acreditar que α dado que α é falsa é alta, e (ii) a probabilidade de S acreditar que α dado que α é o caso é alta (ROUSH 2005, pp. 45-46). Assim:

Sensitividade_{prob}:

- (i) $P(\neg B\alpha \mid \neg\alpha) > l$; e
- (ii) $P(B\alpha \mid \alpha) > l$, tal que $0,5 < l \leq 1$.

Roush (p. 46) estipula um valor bem alto, $l = 0.95$. É claro que poderíamos adotar um valor diferente, mas podemos deixar de lado essa discussão. Não precisamos também especificar se o proponente da abordagem probabilística tem de adotar uma interpretação subjetiva ou objetiva das probabilidades (embora eu pense que uma interpretação objetivista esteja mais de acordo com o externismo). O que me interessa aqui é apenas a ideia geral da abordagem probabilística. Penso que em geral ela não é uma concorrente à altura.

A abordagem modal e a probabilística concordarão sobre vários casos serem ou não casos de conhecimento. Tomemos um caso comum de conhecimento. Sei que estou sentado defronte ao computador digitando este texto. De acordo com a sensitividade modal, nos mundos próximos em que é falso que estou sentado em frente ao computador não acredito que estou sentado em frente ao computador. Por exemplo, no mundo próximo em que estou treinando boxe e não escrevendo este texto, não formo a crença de que estou defronte ao computador. E de acordo com a sensitividade probabilística, a probabilidade de eu não acreditar que estou sentado defronte ao computador dado que não estou sentado defronte ao computador é bastante alta.

⁵⁰ Cf. ROUSH (2005); ZALABARDO (2012). Sherrilyn Roush (e José Zalabardo concorda com ela) pensa que a abordagem probabilística é capaz de responder vários puzzles e contraexemplos (alguns dos quais serão discutidos mais adiante neste capítulo) que a abordagem modal da sensitividade não dá conta de responder.

Consideremos agora CELEIRO. De acordo com a abordagem probabilística, a crença de João de que há um celeiro no lote 1 é sensitiva se, e somente se, a probabilidade de João não acreditar que há um celeiro no lote 1 dado que não há um celeiro no lote 1 é alta. Segue-se disso que a crença é insensitiva se, e somente se, essa probabilidade condicional for baixa. Ora, lembre-se de que há 99 celeiros de fachada ao redor de João. Como os celeiros de fachada e o celeiro real poderiam estar dispostos de qualquer forma, João muito provavelmente formaria a crença de que há um celeiro no lote 1. Aparentemente, portanto, a abordagem probabilística se dá tão bem quanto a modal.

Mas apenas ao custo de pressupor a abordagem modal. Repare os advérbios “poderiam” e “formaria” logo acima; eles sutilmente denunciam a abordagem probabilística. Se abolirmos o recurso aos mundos possíveis, temos de ser consistentes e trabalhar apenas com o mundo atual. No mundo atual João olha para o celeiro do lote 1, que é o celeiro real. Se considerarmos todo o cenário, a chance de João olhar para um celeiro real é 1/100, o que torna bastante improvável que ele forme uma crença verdadeira de que há um celeiro no lote 1, embora bastante provável de que ele forme a crença de que há um celeiro no lote 1. Isso parece estar de acordo com a abordagem probabilística. Mas uma vez que só há o mundo atual e é verdade que há um celeiro no lote 1, o que explica que a baixa probabilidade de João não formar a crença? A resposta só pode ser a de que o celeiro real *poderia* estar alocado em qualquer outro lote. Ou seja, a abordagem probabilística tem de assumir que há vários arranjos possíveis para os celeiros de fachada e vários desses arranjos poderiam ter sido o caso. Nesse caso, a abordagem probabilística pressupõe a modal.

Se insistirmos, contudo, que apenas o mundo atual conta, perdemos a intuição de que João não sabe que há um celeiro no lote 1. A intuição depende do elemento de sorte; de que nesse cenário João não é capaz de discriminar entre celeiros de fachada e celeiros reais, e que, por isso, uma crença verdadeira nesse cenário é adquirida por sorte. Uma vez que João acredita numa proposição verdadeira, não faz sentido perguntar da probabilidade dele não formar essa crença dado que a proposição é falsa. Só faz sentido se perguntarmos da probabilidade dele formar essa crença nas circunstâncias em que a proposição é falsa. Mas falar de “circunstâncias” é o mesmo que falar de “nos mundos tais e tais”.

Portanto, ou a abordagem probabilística pressupõe a modal, ou tem de abandonar a intuição de que CELEIRO não é um caso genuíno de falha de conhecimento. Em qualquer dos casos ela não substitui a abordagem modal.

Além do mais, como vimos no Capítulo I, uma abordagem puramente probabilística da sorte não é adequada. Há casos em que um evento ocorre por sorte mesmo não sendo improvável, como no caso da enchente. Vimos que a sorte é primariamente um fenômeno modal. É natural, portanto, que a condição anti-sorte para o conhecimento seja também modal. A sensibilidade probabilística não precisa ser entendida como uma rival da abordagem modal; pode ser, ao invés, considerada um caso particular da abordagem modal.

1.3.Sensitividade e proximidade

Para aqueles que se incomodaram com a vagueza da noção de proximidade entre mundos na semântica das contrafactuais, não será nenhuma surpresa que se perguntem como determinar a proximidade de mundos para avaliar a sensibilidade. Dito de outro modo, qual o critério ou critérios para a determinação do conjunto de mundos possíveis relevantes para avaliação da sensibilidade? Uma sugestão inicial é dada por Fred Dretske (1971):

As circunstâncias que se supõem constantes, que são tacitamente mantidas fixas, [...] são aquelas circunstâncias vigentes na ocasião em questão [...] que são lógica e causalmente independentes do estado de coisas expresso por P [a proposição alvo]. Quando temos uma afirmação [...] contrafactual (a antecedente expressa um estado de coisas não obtido), [o conjunto de circunstâncias] incluirá aquelas circunstâncias vigentes na ocasião que são lógica e causalmente independentes do estado de coisas (ou falta desse estado) expresso pela *antecedente* da condicional. (p. 7-8)

Ou seja, alteramos apenas o valor de verdade da proposição alvo e mantemos o valor de verdade de todas aquelas proposições que são lógica e causalmente independentes da proposição alvo. De maneira mais clara: pegamos o mundo atual e alteramos primeiramente apenas a proposição alvo. Como consequência, todas aquelas proposições cujo valor de verdade depender lógica ou causalmente do valor de verdade da proposição alvo também terão seu valor de verdade alterado. Tudo o mais permanece como no mundo atual. Voltemos ao exemplo dos celeiros. No mundo atual supomos haver 100 celeiros – um real e os outros 99 de fachada. No mundo₁, a proposição ⟨Há um celeiro no lote 1⟩ torna-se falsa. Assim, alguma das proposições ⟨Há um celeiro no lote n , tal que $1 < n \leq 100$ ⟩ terá de ser falsa.

Digamos que o celeiro de fachada do lote 2 foi erigido no lote 1 e o celeiro real no lote 2. Assim, a proposição ⟨Há um celeiro de fachada no lote 2⟩ também será falsa. Suponha que no mundo atual, durante a construção do celeiro real, os operários tenham encontrado um corpo enterrado no solo. Como a construção dos celeiros de fachada não envolve a perfuração um pouco mais profunda do solo, no mundo₁ os operários não encontram o corpo. Assim, a proposição ⟨Encontraram um corpo no lote 1⟩ é falsa no mundo₁, uma vez que é causalmente dependente, em última instância, da proposição ⟨Há um celeiro no lote 1⟩, que é falsa no mundo₁.

A sugestão de Dretske parece razoável. Afinal, mantendo-se tudo como está e alterando-se apenas o valor de verdade da proposição ⟨Há um celeiro no lote 1⟩, a proposição ⟨João acredita que há um celeiro no lote 1⟩ continua verdadeira, o que produz o veredito correto sobre o caso: João não sabe que há um celeiro no lote 1. Ele usa o já famoso exemplo do termômetro para ilustrar o caso em que a sensibilidade dá o veredito correto sobre um caso de conhecimento. Sebastião tira a temperatura de seu filho usando um termômetro em perfeito estado de funcionamento. O termômetro marca 38,5° C e Sebastião forma a crença de que seu filho tem 38,5 de febre. Intuitivamente esse é um caso claro de conhecimento. Nos mundos mais próximos em que Sebastião usa o mesmo termômetro mas não é o caso que seu filho tem 38,5 (digamos que tenha 36° C), Sebastião não forma a crença de que seu filho tem 38,5 (e sim a crença correta de que tem 36). A crença de João rastreia a verdade.

Considere, contudo, o seguinte caso originalmente apresentado por Goldman (1986, p. 45):

[TERMÔMETRO] Sebastião precisa medir a temperatura de seu filho. Ele vai até sua caixa de medicamentos e retira uma bolsa com 10 termômetros. Ele pega aleatoriamente um dos termômetros, que estava em perfeito funcionamento, e o usa em seu filho. O termômetro marca 38,5° C e Sebastião acredita corretamente que seu filho tem 38,5 de febre. Porém, sem que Sebastião soubesse, dentre os 10 termômetros na bolsa, apenas aquele estava funcionando adequadamente; os outros nove estavam avariados – fosse qualquer temperatura eles marcariam 38,5.

Ao contrário do caso original, a intuição aqui é que Sebastião não sabe a temperatura de seu filho, uma vez que a sua crença em algum sentido é verdadeira por sorte; foi ao acaso que ele escolheu o único termômetro que não estava avariado. Note, porém, que a sensibilidade parece ser satisfeita: nos mundos próximos em que Sebastião usa o termômetro que usou mas

é falso que seu filho tem 38,5 (tem 36, digamos), Sebastião não forma a crença de que seu filho tem 38,5 de febre. Se a intuição de Goldman estiver correta, então não é apenas a proposição alvo (e as proposições dependentes) que importam para a determinação do conjunto de mundos próximos: há mundos em que o valor de verdade de algumas proposições lógica e causalmente independentes da proposição alvo diferem das do mundo atual e ainda assim contam como próximos o suficiente para a avaliação da contrafactual da sensibilidade.

Mas o que realmente mostra esse caso, que crença verdadeira sensível é compatível com a sorte, ou que a crença de Sebastião não é afinal de contas sensível? Se insistirmos na sugestão de Dretske, e negarmos que seja a primeira opção, então temos de afastar a intuição de Goldman de que TERMÔMETRO é um caso de falha de conhecimento. Mas se insistirmos na insensibilidade da crença de Sebastião, temos de explicar por que aumentar o conjunto de mundos próximos.

Consideremos primeiro que TERMÔMETRO não é um caso de falha de conhecimento. Uma primeira tentativa seria chamar atenção para o método de formação de crença. No mundo atual Sebastião forma a crença de que seu filho tem 38,5 de febre fazendo a leitura de um termômetro, chamemos-lhe termômetro₁. Num mundo próximo, w_2 , em que a criança tem 36° C, Sebastião forma a mesma crença com base na leitura não do termômetro₁ mas sim de um termômetro avariado, o termômetro₂. De acordo com a sensibilidade

Nos mundos mais próximos em que é falso que o filho de Sebastião tem 38,5° C, Sebastião não acredita com base na leitura do termômetro₁ que seu filho tem 38,5.

Ora, uma vez que em w_2 ele usa o termômetro₂, o fato de ele continuar acreditando que seu filho tem 38,5° C não torna falsa a contrafactual da variação. A crença de Sebastião é, portanto, sensível.

Mas como afastamos a intuição de Goldman de que TERMÔMETRO não é um caso de conhecimento? O fato de mantermos fixo o modo como Sebastião forma a sua crença ao longo dos mundos próximos não afasta a intuição de que ainda assim a crença de Sebastião é verdadeira por sorte. Talvez a seguinte especificação do método usado por Sebastião nos ajude: *pegar aleatoriamente um termômetro em meio a vários termômetros avariados (que*

marcam 38,5 seja qual for a temperatura) e medir a temperatura. Chamemos a tal método, “M_{ale}”. Vejamos:

Se não fosse o caso que o filho de Sebastião tivesse 38,5, então Sebastião não acreditaria, via M_{ale}, que seu filho tem 38,5.

Ou dito de outro modo:

Nos mundos mais próximos em que é falso que o filho de Sebastião tem 38,5°C, Sebastião não acredita via M_{ale} que seu filho tem 38,5.

Ora, é claro que usando M_{ale}, Sebastião muito provavelmente terá formado, nos mundos próximos em que seu filho não tem 38,5, a crença falsa de que seu filho tem 38,5. Sendo assim, a sua crença é insensitiva; e, portanto, não é conhecimento. Note, porém, que M_{ale} não é um método confiável: ele tende a produzir mais crenças falsas do que verdadeiras. E certamente não atribuímos conhecimento a quem quer que seja que esteja usando um método infiel. Nesse caso, então, a intuição de Goldman pode ser explicada apelando-se para a infidelidade de M_{ale} – um método que torna a crença de Sebastião insensitiva. Contudo, é bastante estranho atribuir o uso de métodos como M_{ale} a um sujeito.

Não pretendo resolver essa controvérsia agora, mas deixe-me apontar para uma solução. É inegável que Sebastião tenha tido sorte em pegar um termômetro confiável. Mas não é tão claro que disso se siga que a sua crença tenha sido verdadeira por sorte. Por exemplo, suponha que Washington seja um exímio cobrador de faltas; na entrada na área ele costuma fazer 7 gols a cada 10 cobranças. Contudo, o rendimento de Washington cai consideravelmente quando ele usa chuteiras novas. O roupeiro do time disponibiliza, antes do jogo, 3 pares de chuteiras para Washington. Sem que ele soubesse 2 eram novos. Mas Washington por sorte pega o usado. Durante o jogo, ele faz um golaço de falta. Contudo, se tivesse escolhido um dos outros dois pares de chuteiras, teria errado a cobrança. Embora tenha sido sorte Washington ter escolhido o par correto de chuteiras, está longe de ser óbvio que seu gol tenha sido um golpe de sorte. O mesmo parece valer para TERMÔMETRO. Embora Sebastião tenha pegado o termômetro confiável por sorte, está longe de ser óbvio que a sua crença de que seu filho tem 38,5 tenha sido verdadeira por sorte. Talvez a sorte envolvida

aqui não seja epistêmica, mas sim evidencial. E como mostra muito bem Pritchard (2005), a sorte evidencial é compatível com o conhecimento.

Confesso que a minha intuição hesita bastante nesse caso. Quando começo a me convencer que Sebastião tem conhecimento, subitamente a sorte envolvida na escolha do termômetro começa a me convencer de que ele não sabe. Isso talvez reflita a própria vagueza envolvida na tentativa de delimitar o conjunto de mundos relevantes. Porém, há casos claros de crença sensitiva e casos claros de crenças insensitivas. É melhor começar por eles; começar pela controvérsia não é uma boa estratégia. A esperança é que no fim tenhamos mais clareza sobre a linha que separa o conjunto de mundos relevantes dos irrelevantes – se é que há tal linha.

2. Objeções influentes

Apesar de sua plausibilidade e atração iniciais, a sensitividade enfrenta sérias objeções. Nesta seção apresentarei em algum pormenor aquelas que considero as principais objeções já existentes na literatura. Discuti-las-ei mais adiante, no Capítulo III. Há basicamente dois tipos de objeções: (i) que a sensitividade é demasiado forte (restritiva) – *i.e.*, exclui da categoria de conhecimento casos claros de conhecimento –; e (ii) que é demasiado fraca (permissiva) – *i.e.*, inclui na categoria de conhecimento casos claros de falha de conhecimento. Dito de outro modo, há conhecimento que não é crença verdadeira sensitiva, e há crença verdadeira sensitiva que não é conhecimento. Muitos autores consideram ultrajante a aparente rejeição do fechamento epistêmico por parte de Nozick, Dretske e cia. (se houver tal cia.!) e apontam a falha do fechamento epistêmico como uma objeção mortal à sensitividade. Protelarei a discussão sobre o fechamento, porém, até o Capítulo V.

2.1. Permissividade

Consideremos uma adaptação do caso do holograma do vaso apresentado por Nozick (p. 190):

[HOLOGRAMA] A taça da Copa Libertadores da América de 2013 está exposta numa redoma na sede do Clube Atlético Mineiro. Por alguma razão, o presidente do clube pediu a um engenheiro que projetasse a redoma de modo a alternar, em curtos espaços de tempo, entre a imagem da taça projetada por um holograma e o própria taça. O engenheiro projetou um sistema de detecção na base da redoma que ativa o holograma apenas quando a taça está numa posição

determinada. Reinaldo, um torcedor fanático, vai até a sede do clube, vê a imagem do holograma e acredita que está ali a taça da Libertadores.

Intuitivamente, Reinaldo não sabe que a taça está na redoma, uma vez que a sua crença foi formada pelo holograma e não pela própria taça. Contudo, a crença de Reinaldo é sensitiva: nos mundos mais próximos em que a taça não está na redoma, Reinaldo não acredita que a taça está lá. Note que nos mundos mais próximos em que não há taça, também não há o holograma, uma vez que o holograma só é ativado quando a taça está dentro da redoma. A sensitividade, portanto, permite que um caso claro de falha de conhecimento seja classificado como conhecimento.⁵¹

Embora sejam recorrentes (como veremos na seção seguinte) casos de conhecimento indutivo que parecem ser excluídos pela sensitividade, há casos em que a sensitividade parece permitir que crença baseada em evidência indutiva insuficiente conte como conhecimento:

[NEUROANATOMISTA PALEOLÍTICO] Um homem das cavernas observa que na cabeça de um ou dois animais há um cérebro. Ele supõe, sem qualquer verificação ulterior, que na sua cabeça há também um órgão desse tipo. Dado o limite de suas observações e seu conhecimento de fundo, o homem das cavernas não sabe que dentro de seu crânio há um cérebro. (VOGEL 2012, p. 148)

Uma vez que nos mundos mais próximos em que o homem das cavernas não possui um cérebro ele não pode sequer formar uma crença, e, por conseguinte, também não acredita que tenha cérebro. A sensitividade parece ser satisfeita, muito embora a base indutiva do homem das cavernas seja claramente insuficiente.

Tristan Haze (2015) apresenta dois interessantes contraexemplos. Vamos ao primeiro:

[TRIBUTARISTA] Tenho uma crença ilusória contrafactualmente bastante robusta de que o meu vizinho é um oráculo divino. Ele é de fato muito confiável e é um advogado tributarista. Ele sempre quis me dizer algo sobre direito tributário, *p*. Um dia ele me diz que *p* e eu acredito nele porque acredito que ele seja um oráculo divino. Eu nunca teria acreditado nele se ele fosse um advogado, uma vez que desconfio bastante de advogados. (p. 310)

⁵¹ Kelly Becker (2007, pp. 53-4) sugere que a adaptação de Alan Goldman (1987, p. 184) do caso do Sr. Haveit (ou Nogot) é também um caso genuíno de contraexemplo à sensitividade. Não penso que seja, pelos motivos apresentados no Capítulo 1.

Ou seja, nos mundos mais próximos em que é falso que p , o advogado (sendo confiável) não me diz que p , e, por isso, eu não acredito que p . Porém, parece claro que eu não sei p .

Eis o segundo contraexemplo:

[TRIBUTARISTA*] O meu vizinho é um tributarista [...] Por anos ele tem nutrido um plano excêntrico de me fazer acreditar na verdade de p , uma proposição sobre direito tributário, conjuntamente de outras cinco proposições falsas sobre direito tributário. A intenção dele é fazer isso de maneira contrafactualmente bem robusta. Ele pouco a pouco vai ganhando a minha confiança. Um dia, ele gentilmente me fala sobre direito tributário. Ele asseme seis proposições: p e as cinco falsas. Acredito em todas elas. (p. 310)

Ou seja, nos mundos mais próximos em que p é falsa, o advogado não teria me dito que é verdadeira, uma vez que seu plano era fazer com que eu acreditasse numa proposição verdadeira e cinco falsas. Assim, nesses mundos, também não acredito que p . A minha crença rastreia a proposição de que p , muito embora eu não saiba que p .

Mais um engenhoso contraexemplo, dessa vez oferecido por Jerome Gellman (2004, pp. 280-1):

[CORRELAÇÃO] S tem alguma evidência para acreditar que há vida em Marte. Embora a evidência seja razoável, a crença de S é falsa; não há vida em Marte. Descobertas neurocientíficas recentes mostram haver várias correlações legiformes entre crenças e estados cerebrais. Acontece de haver uma correlação legiforme entre a crença de S de que há vida em Marte e o estado cerebral E_{10} , muito embora S não faça a mínima ideia de que há. Na verdade, S tem alguma evidência para acreditar que não está no estado cerebral E_{10} . Mas S tem um hábito um tanto estranho: sempre que tem boa evidência a favor de uma proposição p qualquer, costuma formar uma proposição disjuntiva a partir de p . Assim, ela acaba por formar a seguinte crença disjuntiva: \langle Há vida em Marte ou estou no estado cerebral E_{10} \rangle .

No caso acima, a crença disjuntiva de S é verdadeira; e não verdadeira por acaso, uma vez que a proposição \langle Estou no estado cerebral E_{10} \rangle é verdadeira porque S acredita que há vida em Marte. Além do mais, a sensibilidade é satisfeita: nos mundos próximos em que \langle Há vida em Marte ou estou no estado cerebral E_{10} \rangle é falsa S não acredita que há vida em Marte. Isso porque nesses mundos a proposição S está no estado cerebral E_{10} é falsa, e se é falsa, dada a correlação legiforme, S não forma a crença de que há vida em Marte e, por conseguinte, não

introduz a disjunção para formar a crença na proposição disjuntiva em questão. Contudo, é claro que S não sabe que *há vida em Marte ou está no estado cerebral E₁₀*.

Eis uma adaptação de um exemplo de Plantinga (1993) contra o confiabilismo processual:

[LESÃO CEREBRAL] Cassius toma uma pancada muito forte na cabeça que lhe causa uma lesão cerebral. Essa lesão tem um efeito cognitivo adverso: faz com que Cassius forme a crença de que tem uma lesão cerebral. Cassius, além disso, não tem qualquer evidência adicional de que tenha uma lesão cerebral.

A crença de Cassius é verdadeira e sensitiva: nos mundos mais próximos em que Cassius não tem uma lesão cerebral, ele não forma a crença de que tenha uma lesão cerebral. Contudo, tal crença não é conhecimento. A teoria do rastreamento teria de classificar esse caso como conhecimento.

2.2. Restritividade

a. Conhecimento indutivo

Essa talvez seja uma das mais indesejadas aparentes consequências da sensitividade: crenças verdadeiras sustentadas a partir de evidência indutiva não são sensitivas. Uma vez mais, deixe-me adaptar um exemplo já famoso, dessa vez o caso do lixo de Ernest Sosa (1999):

[CALHA] Adélia é pedreira e trabalha na construção de um prédio. Ela está no décimo andar e começa a limpar o entulho acumulado. Ela dispensa o entulho por uma calha que vai até uma caçamba no térreo. Ela acha um par de botinas velhas e joga pela calha junto com o entulho. Washington, seu companheiro de trabalho, pergunta se ela sabe onde estão as suas botinas (as que ela tinha acabado de deitar fora pela calha!). Adélia diz que estão na caçamba de entulho no térreo.

A crença de Adélia que as botinas de Washington estão na caçamba é perfeitamente razoável. Dificilmente negaríamos conhecimento a ela. Contudo, se seguirmos a sensitividade é o que temos de fazer. Há pelo menos um mundo possível próximo em que a proposição ⟨a botina não chega até a caçamba⟩ é falsa – um mundo no qual ela fica presa num parafuso dentro da calha, digamos – e que Adélia acredita que a botina atingiu a caçamba. Na verdade, de acordo com a sensitividade Adélia sequer sabe se todo o entulho que ela despejou atingiu a caçamba. Ainda bem que os pedreiros não obedecem à sensitividade!

Eis outros contraexemplos já bastante difundidos (VOGEL 1987): CUBO DE GELO e ASPIRANTE. No primeiro caso, se você deixa um cubo de gelo dentro de um copo num dia de calor (calor tropical!) e depois de alguns minutos se lembra que deixou o copo, você certamente sabe que o gelo derreteu. Contudo, há pelo menos um mundo próximo em que o gelo não derrete e você acredita, com base nas mesmas evidências, que o gelo derreteu. Portanto, a sensibilidade exclui esse caso óbvio de conhecimento indutivo. Já no caso do aspirante, pense em dois policiais num conflito com um assaltante. O policial experiente conversa calmamente com o assaltante, pedindo para que ele largasse a arma, quando percebe que o aspirante vai tentar o improvável: atirar na mão do assaltante para tirar-lhe a arma. O veterano ouve o tiro e antes mesmo de olhar para a mão do assaltante forma a crença de que o aspirante errou o tiro. O aspirante de fato errou, e o veterano sabia disso. Contudo, há um mundo possível próximo em que o aspirante acerta o tiro e o veterano acredita, com base na mesma evidência, que o aspirante errou o tiro. Novamente, a sensibilidade exclui casos óbvios de conhecimento indutivo.⁵²

Becker (2007) coloca as coisas de uma maneira bastante instrutiva:

O método particular de construção de contraexemplo aqui é engenhoso. Pegue um caso de crença em que a possibilidade não- p é tão improvável que facilmente viola a sensibilidade, mas que, na mesma medida, é plausível dizer que S sabe que p . O nosso dilema é ou abandonar a sensibilidade ou sustentar que alguns dos casos mais seguros de indução não constituem conhecimento acerca de um evento que ocorrerá. (p. 55)

b. Oscar e Dack

Temos aqui o também famoso contraexemplo de Alvin Goldman:

[DACHSHUND] Oscar vê Dack, o dachshund, e acredita que há um cachorro diante dele. Se não houvesse um cachorro diante dele, estaria lá uma hiena, que Oscar classificaria erroneamente como um cachorro. (1983, p. 84-5)

A intuição inicial é que Oscar sabe que há um cachorro diante de si. Porém, o exemplo indica que há um mundo próximo em que é falso que Dack está diante de Oscar, mas é verdade que há uma hiena diante de Oscar, que ele erroneamente identifica como um cão. Assim, nesse mundo, Oscar acredita que há um cão diante de si. A crença de Oscar não rastreia a verdade

⁵² Vogel (2012) ainda apresenta mais alguns casos de conhecimento indutivo que geram problema para a sensibilidade. Os casos seguem basicamente a estrutura dos anteriores.

da proposição ⟨Há um cachorro diante de Oscar⟩, muito embora pareça óbvio que Oscar sabe que há. A sensibilidade nos obriga a rejeitar DACHSHUND como um caso de conhecimento.

c. Margem de erro

A sensibilidade parece também excluir casos de conhecimento que envolve alguma margem de erro (WILLIAMSON 2000). Consideremos o seguinte caso:

[RAMONES] Johnny é um medidor de alturas intuitivo confiável e forma a crença verdadeira de que Joey tem *pelo menos* 1,95m de altura. A sua crença é verdadeira porque Joey tem exatamente 1,98m. Há, porém, um mundo próximo em que Joey mede 1,94m – e, portanto, é falso que Joey tem pelo menos 1,95m – e Johnny forma a crença de que Joey tem pelo menos 1,95m.

É bastante claro que Johnny sabe que Joey tem pelo menos 1,95m, muito embora a sua crença seja insensitiva.

d. Holograma novamente

George Pappas e Marshall Swain (1973) apresentaram o seguinte caso (originalmente contra a abordagem das razões conclusivas de Dretske):

[HOLOGRAMA*] S vê um copo em cima da mesa e forma a crença verdadeira de que há um copo em cima da mesa. Contudo, se não houvesse um copo, mas um holograma de um copo, S teria formado a crença de que há um copo em cima da mesa.

Ou seja, eles chamam a atenção para o fato óbvio de que S sabe que há um copo em cima da mesa, muito embora a crença de S seja insensitiva.

e. Os celeiros de Kripke

Kripke (2011) – num manuscrito de 1986, agora já publicado – popularizou um novo caso envolvendo celeiros:

[CELEIRO-K] João está passando pela região dos celeiros de fachada azuis. Há apenas um celeiro real por ali, que é vermelho. Os moradores dali tem uma lei que impede a construção de celeiros de fachada vermelhos. João olha para o celeiro vermelho e forma a crença verdadeira de que há um celeiro vermelho em sua frente.

Note que a crença de João na proposição ⟨Há um celeiro vermelho em minha frente⟩ é sensitiva: nos mundos mais próximos em que não há um celeiro vermelho, João não acredita que há, pois ao olhar para um celeiro de fachada (que por lei não pode ser vermelho) ele formará a crença de que há um celeiro azul, não vermelho. Contudo, a crença de João na proposição ⟨Há um celeiro em minha frente⟩ não é sensitiva: há mundos próximos em que João olha para um celeiro de fachada, que é azul, e forma a crença de que há um celeiro em sua frente. Esse é um resultado é estranho: como pode a crença do sujeito rastrear a verdade de que há um celeiro vermelho em sua frente, mas não a crença de que há um celeiro em sua frente? Afinal, é impossível que a primeira crença seja verdadeira e a última falsa.⁵³

f. Proposições necessariamente verdadeiras⁵⁴

Considere a seguinte proposição, ⟨Machado de Assis é humano⟩. À primeira vista, é perfeitamente razoável que eu saiba tal proposição. Afinal, nos mundos próximos em que Machado de Assis não é um humano não acredito que seja. Considere agora que tal proposição não só é verdadeira, mas necessariamente verdadeira. Em todos os mundos em que Machado de Assis existe ele é um humano. Embora seja logicamente possível que Machado de Assis tivesse sido um tamanduá ou uma jabuticabeira, não é metafisicamente possível que tivesse sido. Defensavelmente, *ser humano* é uma propriedade essencial de Machado de Assis. Ora, uma vez que é metafisicamente impossível que Machado de Assis não seja humano, nos mundos mais próximos Machado de Assis é também humano. Assim, não há qualquer mundo próximo em que eu acredite que Machado seja humano e seja falso que ele não é humano. A condicional da sensitividade é, portanto, trivialmente satisfeita.⁵⁵

Considere o Sr. Rice, que nunca soubera da existência de Machado de Assis. O Sr. Rice escuta seu vizinho chamar pelo seu gato de estimação, que é homônimo ao nosso

⁵³ CELEIRO-K também levanta um contraexemplo ao princípio de fechamento, uma vez que João sabe que há uma proposição mas não sabe a proposição logicamente implicada por ela. Como dito, o fechamento será discutido no Capítulo V.

⁵⁴ Não apresentarei aqui a dificuldade de lidar com o conhecimento *a priori* de verdades necessárias (e.g. verdades da matemática e da lógica), como fez Nozick (1981, p. 186-7). Como já mencionado algures, nesta tese interesse-me apenas pelo conhecimento *aposteriori*.

⁵⁵ Dito de outro modo: A negação de uma proposição necessariamente verdadeira é uma falsidade necessária. Assim, a condicional da sensitividade terá no seu antecedente uma proposição necessariamente falsa, de modo que em nenhum mundo possível (incluindo os mundos próximos) a antecedente é satisfeita; e, por conseguinte, a condicional nunca será falsa. No caso das necessidades lógicas, não haverá qualquer mundo logicamente possível em que a antecedente seja satisfeita; no caso de necessidades reais (metafísica e nomológica), não haverá qualquer mundo realmente possível em que a antecedente seja satisfeita, embora o será em mundos logicamente possíveis.

escritor, e forma a crença que Machado de Assis é um grande escritor brasileiro e que, por isso, também acredita que é humano. É claro que o Sr. Rice não sabe que Machado de Assis é humano, muito embora a sua crença seja sensitiva. Não há qualquer mundo próximo que o Sr. Rice acredita que Machado de Assis é humano e não é o caso que Machado não o seja.

Posto de outro modo, a sensitividade não dá conta do conhecimento de verdades necessárias *a posteriori*. Para usar outro exemplo, é óbvio que sabemos que a água é H₂O. E isso é uma verdade necessária: não há mundos metafisicamente possíveis em que algo seja água e não seja H₂O.⁵⁶ Novamente, a sensitividade é trivialmente satisfeita: não há mundos próximos em que um químico acredita que a água é H₂O e não é esse o caso. Ou pense numa proposição como ⟨Deus existe⟩. Se verdadeira, é uma proposição necessariamente verdadeira (pelo menos de acordo com o teísmo tradicional, que concebe Deus como um existente necessário). E supondo que o conhecimento *a priori* de Deus seja impossível, a sensitividade uma vez mais é trivialmente satisfeita: não há mundos próximos em que o teísta acredita que Deus existe e ele não existe.

g. Leis da natureza

Outro tipo de necessidades *a posteriori* são as leis da natureza e as generalizações legiformes. Por exemplo, para crenças em proposições como ⟨Nenhum objeto viaja mais depressa que a luz⟩, ⟨A água, em condições normais de temperatura e pressão, congela a 100° C⟩, etc., a sensitividade é também trivialmente satisfeita.

h. Conhecimento de ordem-superior

Eis mais uma contribuição de Jonathan Vogel (2000; 2012) contra a sensitividade. Vogel argumenta que a sensitividade exclui o conhecimento de ordem-superior, *i.e.*, um sujeito saber que sabe.⁵⁷

⁵⁶ Não precisamos aqui nos comprometer com qualquer teoria essencialista robusta; embora o essencialismo trivial (*e.g.*, um indivíduo tem a propriedade essencial de pertencer ao conjunto unitário [*singleton*] cujo único elemento é ele próprio) não seja suficiente para gerar o necessário *a posteriori*. Para evitarmos a acusação de abuso do necessário *a posteriori* (cf. BEEBEE & SABBARTON-LEARY 2010), é suficiente que apenas sustentemos a tese condicional: se há verdades necessárias *a posteriori*, então a sensitividade é trivialmente satisfeita.

⁵⁷ Note que a objeção aqui não é a de que a sensitividade rejeita KK, o princípio que diz que se S sabe que *p*, então S sabe que sabe que *p*. É comum que posições externistas rejeitem KK, mas isso só significa que os externistas pensam que S pode saber que *p* sem que precise saber que sabe que *p*; o externista não recusa a ideia de que às vezes o sujeito sabe que sabe. A objeção de Vogel é que a sensitividade sequer permite essa tese mais fraca.

Suponha que você vai à praia e encontra com a sua amiga Ana Maria. Você nota o seu interessante biquíni de bolinhas amarelinhas que ela disse ter acabado de comprar. Assim, seguramente podemos dizer que

I) Você sabe que Ana Maria está de biquíni novo.

E podemos seguramente dizer também que

II) Você sabe que a sua crença de que Ana Maria está de biquíni novo não é falsa.

Note que (II) é conhecimento de segunda ordem. Você sabe uma proposição que é acerca de outra proposição. Para tornar isso mais claro, chamemos “ p ” à proposição ⟨Ana Maria tem um biquíni novo⟩, e “ Bp ” à proposição ⟨S acredita que Ana Maria tem um biquíni novo⟩. Dizemos que Bp é uma crença de primeira ordem, uma crença acerca diretamente do mundo. Chamemos agora “ $K(Bp)$ ” à proposição ⟨S sabe que S acredita que Ana Maria tem um biquíni novo⟩. Como conhecimento implica crença, $K(Bp)$ é conhecimento de segunda ordem.

De acordo com a sensibilidade:

Se $K\neg(Bp \wedge \neg p)$, então $(Bp \wedge \neg p) \square \rightarrow \neg B\neg(Bp \wedge p)$

Consideremos um mundo próximo em que você acredita falsamente que Ana Maria tem um biquíni novo, *i.e.*, $Bp \wedge \neg p$. Digamos que Ana Maria tenha se confundido ao te dizer que aquele era o biquíni novo, quando na verdade o novo era um muito parecido que ela colocara na gaveta errada de seu armário. Você considera o que ela diz e forma a crença de segunda ordem de que você não acredita falsamente que Ana Maria está de biquíni novo. Ou seja, $B\neg(Bp \wedge \neg p)$. Assim, nesse mundo:

$(Bp \wedge \neg p) \wedge B\neg(Bp \wedge \neg p)$

Portanto, você não sabe (no mundo atual) que você não acredita falsamente que Ana Maria está de biquíni novo.

Esse parece não ser um resultado bem vindo à sensibilidade, uma vez que é óbvio que temos conhecimento de segunda ordem.⁵⁸

⁵⁸ Um caso semelhante é apresentado por DeRose (1995, pp. 22-3). Supondo que seja óbvio que sei que tenho mãos, também sei que não acredito falsamente que tenho mãos. Mas a crença de que *não acredito falsamente*

2.3. Restritividade e permissividade

Eis agora uma dificuldade técnica que parece mostrar ao mesmo tempo o quão a sensibilidade é demasiado restrita e permissiva: o conhecimento de conjunções. Foi Jonathan Vogel (2012) – que tem um olho clínico para detectar dificuldades com a sensibilidade! – quem a apresentou. Em linhas gerais, a ideia é que para o conhecimento de conjunções – $(\alpha \wedge \beta)$ – a sensibilidade exige a satisfação de uma contrafactual cuja antecedente é uma disjunção – $(\neg\alpha \vee \neg\beta) \Box \rightarrow \neg B(\alpha \wedge \beta)$. Essa contrafactual, por sua vez, admite duas resoluções: uma forte e a outra fraca. Se optamos pela resolução fraca, então o conhecimento de conjunções é trivializado; e se, por outro lado, optamos pela resolução forte, a sensibilidade impede o conhecimento de implicações lógicas óbvias. No primeiro caso a sensibilidade é demasiado permissiva; no segundo, demasiado restritiva.

Consideremos a seguinte contrafactual disjuntiva:

$$(CD) (\alpha \vee \beta) \Box \rightarrow \gamma$$

(CD) admite duas resoluções: (i) uma fraca e (ii) a outra forte.

(CD_{fraca}) $[(\alpha \vee \beta) \Box \rightarrow \gamma]$ sse $[(\alpha \Box \rightarrow \gamma)$, caso os mundos- α sejam mais próximos ao atual do que os mundos- β], ou $[(\beta \Box \rightarrow \gamma)$, do contrário].

Isto é, uma condicional contrafactual disjuntiva é verdadeira quando os mundos- α mais próximos são também mundos- γ , e os mundos- α são mais próximos ao atual do que os mundos- β ; ou quando os mundos- β mais próximos são também mundos- γ , e os mundos- β são mais próximos ao atual do que os mundos- α . Por exemplo, *Fosse o Atlético-MG ou fosse o Cruzeiro campeão da Copa do Brasil de 2014, o campeão teria sido um time mineiro* é verdadeira e admite a resolução fraca. Nos mundos mais próximos em que Atlético-MG e Cruzeiro disputam a final da Copa do Brasil de 2014 e é falso que o Atlético vence, o Cruzeiro é quem vence. Assim, nos mundos mais próximos em que o Atlético vence, é também verdade que um time mineiro é campeão; e nos mundos próximos em que o Cruzeiro vence, é também verdade que um time mineiro vence.

que tenho mãos é insensitiva, uma vez que há mundos próximos em que acredito falsamente que tenho mãos, muito embora eu tenha a crença de que tenho mãos. Além do mais, casos como esse parecem constituir contraexemplos ao princípio de fechamento.

Considere agora o seguinte caso:⁵⁹

Suponha que Louis esteja sentado em minha frente tomando café gelado. Vejo-o e sei que p , Louis está tomando café gelado. Louis poderia muito facilmente não ter tomado café gelado. Ele poderia ter tomado algo diferente ou mesmo não tomado nada, de modo que o mundo $\neg p$ mais próximo é muito próximo ao atual. Ora, Pascal observou que tivesse o nariz de Cleópatra sido menor, toda a face do mundo teria mudado. Seja $\neg q$ = “Cleópatra não tinha um nariz pequeno”. Se Pascal tem razão, o mundo q mais próximo é muito mais dissimilar ao atual do que o mundo $\neg p$ mais próximo. Assim, dada a resolução fraca das condicionais disjuntivas, posso rastrear $p \wedge \neg q$ simplesmente por rastrear p . Em outras palavras, na medida em que a exigência de rastreamento se aplica, posso saber que (Louis está tomando café gelado e Cleópatra não tinha um nariz pequeno) apenas por ver o que Louis está tomando. (p. 128)

Assim, se admitimos a resolução fraca, a sensibilidade considera como conhecimento um caso claro em que não há conhecimento.

Vamos à resolução forte:

$$(CD_{\text{forte}}) [(\alpha \vee \beta) \Box \rightarrow \gamma] \text{ sse } [(\alpha \Box \rightarrow \gamma) \text{ e } (\beta \Box \rightarrow \gamma)]$$

(CD_{forte}) evita o resultado acima. Mas consideremos agora a seguinte proposição, r : (Milton Nascimento está cantando). Digamos que eu saiba r , uma vez que sou um ótimo ouvinte da voz de Milton Nascimento. Assim,

$$\neg r \Box \rightarrow \neg B r$$

A seguinte consideração é crucial para o argumento de Vogel:

Considere a proposição de que, i , algum impostor que se parece e canta como [Milton Nascimento] é que está cantando e não o Milton. r implica $\neg i$, de modo que r é logicamente equivalente a $(r \wedge \neg i)$. É difícil negar que o conhecimento seja fechado sob a *equivalência* lógica conhecida. Portanto, se [sei] que r , sei que $(r \wedge \neg i)$. E, se o rastreamento é necessário para o conhecimento, então [tenho] de rastrear $(r \wedge \neg i)$. (p. 129)

Assim:

$$(\neg r \vee i) \Box \rightarrow \neg B(r \wedge \neg i)$$

⁵⁹ Modifiquei as variáveis proposicionais do exemplo original.

E de acordo com CD_{forte} , as seguintes contrafactuais têm de ser verdadeiras:

- a) $\neg r \square \rightarrow \neg B(r \wedge \neg i)$, e
- b) $i \square \rightarrow \neg B(r \wedge \neg i)$.

(a) é verdadeira, uma vez que nos mundos mais próximos em que Milton não está cantando, não formo a crença de que r . Contudo, nos mundos próximos em que há um impostor no palco, formo a crença de que r e, por conseguinte, também a crença de que $\neg i$. (b), portanto, é falsa. Logo, não sei que $(r \wedge \neg i)$.

Note que as objeções acima atacam, sem exceção, a condição de variação. A razão é simples. Primeiro, porque dado o método de formação de crenças, nos mundos próximos em que α é o caso, em geral S forma a crença de que α . Segundo, porque a condição de variação parece ser mais importante, além de gerar também contraexemplos interessantes. No capítulo seguinte veremos como a EFAS – a epistemologia da função apropriada mais a sensibilidade – responde a tais contraexemplos.

III

SENSITIVIDADE, FUNÇÃO APROPRIADA E RESPOSTAS ÀS OBJEÇÕES

O propósito geral [das nossas faculdades cognitivas] é presumivelmente produzir crenças verdadeiras em nós; dito de uma forma menos passiva, elas estão moldadas de tal forma que, ao usá-las, podemos vir a adquirir crença verdadeira.

Alvin Plantinga, *Warranted Christian Belief*.

Prévia

Vimos nos capítulos anteriores que a sensibilidade é uma boa candidata a condição anti-sorte, porém não está livre de dificuldades. Neste capítulo avançarei uma maneira de entender a condição de sensibilidade que nos leva naturalmente à teoria da função apropriada. Assim formulada, a sensibilidade não apenas é uma condição anti-sorte, mas também uma componente iluminante da natureza do conhecimento. Mostrarei, ademais, que vários dos contraexemplos apresentados no Capítulo II são contornáveis.

1. Os mundos que interessam

1.1. Os mundos próximos

Grosso modo, ao avaliar a proximidade ou similaridade entre mundos temos de levar em conta: (i) o conjunto de mundos possíveis relevantes, e (ii) a quantidade de proposições que difere de um mundo para outro. Por exemplo, suponha que queremos selecionar alguns mundos próximos em que Machado de Assis ambienta *Dom Casmurro* no interior de Minas Gerais. Obviamente, tomamos como referência o mundo atual – o mundo em que *Dom Casmurro* é ambientado no Rio de Janeiro. Temos primeiro de selecionar aqueles mundos em que a proposição $\langle \text{Dom Casmurro é ambientada no Rio de Janeiro} \rangle$. Há vários mundos em que essa proposição é falsa, mas os que nos interessam são aqueles em que essa proposição é falsa

porque a proposição $\langle Dom\ Casmurro\ \acute{e}\ ambientada\ no\ interior\ de\ Minas\ Gerais \rangle$ é verdadeira. E há também vários mundos em que *Dom Casmurro* não foi escrita por Machado de Assis, muito embora tenha sido ambientada no interior de Minas; esses mundos também não nos interessam. Assim, intuitivamente, já temos alguma ideia da extensão do conjunto de mundos que nos interessa: aqueles mundos em que Machado de Assis escreve *Dom Casmurro* e a ambienta no interior de Minas Gerais. Isso, contudo, pode nos deixar com um conjunto de mundos bastante diverso. Há mundos em que Machado de Assis escreve *Dom Casmurro* e a ambienta em Minas, mas faz isso de seu escritório em Marte. Esse mundo é intuitivamente bastante distante do atual. Por isso é importante termos o mundo atual como referência; para evitar mundos inicialmente bastante distantes, começamos com o mundo atual e vamos alterando aos poucos o valor de verdade de algumas proposições. Portanto, se queremos ter uma noção da extensão do conjunto de mundos próximos em que Machado de Assis ambienta *Dom Casmurro* no interior de Minas Gerais, temos de considerar aqueles mundos que são em tudo iguais ao mundo atual, menos o fato de Machado de Assis ter ambientado *Dom Casmurro* no Rio de Janeiro. Chamemos a esse modo de determinar a distância entre mundos de “critério *simpliciter*”.

Consideremos dois mundos possíveis:

MUNDO₁ (w_1): é igual ao mundo atual em todas as proposições, com a exceção das proposições $\langle Modric\ foi\ o\ melhor\ jogador\ da\ copa\ do\ mundo\ de\ futebol\ de\ 2018 \rangle$; e $\langle Eden\ Hazard\ não\ disputou\ a\ copa\ de\ 2018 \rangle$.

MUNDO₂ (w_2): difere do mundo atual em um número massivo de proposições depois de 1950. EUA e União Soviética se embrenham numa batalha nuclear depois do final da Segunda Guerra que faz de grande parte do globo um lugar inóspito.

Simpliciter, w_1 é mais próximo ao mundo atual do que w_2 . Mas considere que em w_2 tenha havido uma copa do mundo de futebol, e que tenha sido na Rússia, e que tenha sido quase que exatamente como a copa da Rússia do mundo atual. Mais que depressa certa intuição determinista (ou algo próximo disso) nos leva a pensar que a falta de um passado pós-guerra como o do mundo atual não levaria a uma copa do mundo na Rússia em 2018. É estranho que um mundo tão diferente no pós-guerra venha a coincidir num sub-conjunto de proposições

sobre 2018. Estranho mas não impossível! E essa estranheza é explicada pela distância de w_2 do mundo atual.

Consideremos agora a seguinte contrafactual:

[MH] Se Modric não tivesse sido o melhor jogador da copa de 2018, então Hazard teria sido.

Presumivelmente MH é verdadeira. Os mundos mais próximos em que Modric não é eleito o melhor jogador, são mundos em que Hazard é eleito o melhor jogador. De acordo com a semântica padrão para as contrafactuals, temos de considerar se os mundos- $\neg M$ são mundos-H. Sendo assim, w_1 não é relevante para a avaliação de MH, uma vez que é um mundo em que Hazard sequer disputa a copa do mundo. Contudo, w_2 é um mundo relevante, uma vez que houve a copa da Rússia e Hazard atua pela Bélgica. Note, contudo, que *simpliciter* w_1 é mais próximo que w_2 , muito embora quando consideramos MH, w_2 acaba por ser mais próximo do mundo atual do que w_1 . Isso aponta para algo que intuitivamente deveria ser óbvio desde o início: na avaliação de uma contrafactual o que conta não é a proximidade *simpliciter*, mas a proximidade *relativa* (ou qualificada). Isso é o que queremos dizer com “os mundos- α mais próximos”.

Tendo em mente a ideia de proximidade relativa, ao invés de nos atermos num mundo como um todo, seria mais adequado nos atermos em um *segmento de mundo*. Um segmento de mundo será, então, um subconjunto do mundo possível em questão que descreve um determinado estado de coisas. Por exemplo, digamos que A é o conjunto de proposições que descreve o seguinte estado de coisas do mundo atual: o período que decorreu do início da Copa do Mundo da Rússia de 2018 até o seu término. Assim, em A estão contidas proposições como ⟨CR7 marcou 4 gols⟩, ⟨O Brasil foi eliminado nas quartas de final pela Bélgica⟩, ⟨Modric foi o melhor jogador⟩, ⟨Hazard foi o segundo melhor jogador⟩, ⟨A França foi a campeã⟩, ⟨No dia 17/06 fez 30° C em Sochi⟩, etc. Não é difícil ver que, além do atual, há vários mundos possíveis que têm A como subconjunto.

Note que vários mundos vão coincidir em vários segmentos. Porém, para a avaliação da contrafactual, só nos interessa um segmento. Pegamos esse segmento e variamos o valor de verdade da proposição alvo (e daquelas cujo valor de verdade depende de tal proposição) e, eventualmente, de alguma outra proposição relevante. O mais importante é que façamos

alterações mínimas nesse segmento. O restante não importa; podemos incluir A seja em que conjunto for a fim de constituir um mundo possível. A proximidade relativa dependerá quase que exclusivamente de A. Podemos agora caracterizar a proximidade relativa assim:

Dado um segmento s do mundo atual $w@$ e um segmento s^* de um mundo possível w_1 , w_1 será tão mais próximo a $w@$ de acordo com o número de proposições que diferem em valor de verdade em s e s^* .

Ou dito de outro modo:

Dado um segmento s de $w@$ e s^* de w_n , w_n será mais próximo a $w@$ do que w_{n+1} se o número de proposições de w_n que difere em valor de verdade de $w@$ (em s e s') for menor que o número de proposições de w_{n+1} que difere em valor de verdade de $w@$ (em s e s'').

1.2. Mundos epistemicamente próximos

Chamarei de “mundo epistêmico” qualquer mundo possível formado a partir de um segmento do mundo atual em que (a) descreve um sujeito formando, sustentando ou reterendo uma crença (ou conjunto de crenças) qualquer, ou (b) que nas circunstâncias relevantes descreve um sujeito deixando de formar, sustentar ou reter uma crença. Assim, a proposição em (a) será central para o segmento do mundo atual que iremos recortar. Mais especificamente, s terá de conter as proposições que descrevem o ambiente em que a crença do sujeito é formada, sustentada ou retida. Considere o seguinte caso. Carlos vai até à sua estante, olha para a lombada dos livros e identifica que *O Triste Fim de Policarpo Quaresma* está lá. Carlos, por isso, acredita que *O Triste Fim de Policarpo Quaresma* está na estante. O segmento de mundo que nos interessa aqui, ao qual chamarei “segmento epistêmico” será mais ou menos o seguinte:

SE = {⟨Carlos acredita que *O Triste Fim de Policarpo Quaresma* está na estante⟩, ⟨*O Triste Fim de Policarpo Quaresma* está na estante⟩, ⟨Os livros na estante têm a lombada legível⟩, ⟨A sala em que a estante fica está iluminada adequadamente⟩, ⟨Carlos enxerga bem em qualquer ponto dessa sala⟩, ⟨Carlos forma a sua crença a partir da visão⟩, etc.}.

De acordo com a caracterização dada anteriormente de um mundo próximo relativo, podemos caracterizar um *mundo epistemicamente próximo*:

Dado um segmento epistêmico SE do mundo atual $w@$ e um segmento epistêmico SE* de um mundo possível w_1 , w_1 será tão epistemicamente mais próximo de $w@$ de acordo com o número de proposições que diferir em valor de verdade em SE e SE*.

Ou dito de outro modo:

Dado um segmento epistêmico SE de $w@$ e SE* de w_n , w_n será mais epistemicamente próximo a $w@$ do que w_{n+1} sse o número de proposições de w_n que difere em valor de verdade de $w@$ (em SE e SE') for menor que o número de proposições de w_{n+1} que difere em valor de verdade de $w@$ (em SE e SE'').

Diremos que SE descreve (i) as condições iniciais para a formação da crença do sujeito, (ii) o processo durante o qual a crença se formou, e (iii) o resultado final do processo, que é a crença:

- (i) *Condições iniciais*: proposições que descrevem:
 - a. o mecanismo cognitivo pelo qual o sujeito forma a crença;
 - b. o funcionamento prévio do mecanismo;
 - c. o meio em que o mecanismo está funcionando, e
 - d. o input recebido do meio.
- (ii) *Processo*: proposições que descrevem cada etapa pela qual o input passa até se tornar uma crença.
- (iii) *Resultado final*: proposição que descreve o conteúdo da crença.

1.3. Sensitividade

Ao avaliarmos agora as contrafactuais da sensitividade – $(\neg\alpha \Box \rightarrow \neg B\alpha)$ e $(\alpha \Box \rightarrow B\alpha)$ – levaremos em conta o seguinte. (A) Se os mundos- $\neg\alpha$ mais próximos em que (i) as condições iniciais são aproximadamente mantidas e (ii) as etapas do processo de formação de crença são aproximadamente as mesmas, são também mundos- $\neg B\alpha$, no caso da variação. E (B) se os mundos- α mais próximos em que (i) as condições iniciais são aproximadamente mantidas e

(ii) as etapas do processo de formação de crença são aproximadamente as mesmas, são também mundos- $\beta\alpha$, no caso da aderência. Consideremos a seguinte contrafactual:

Se Jesus não tivesse aparecido, Tomé não acreditaria em sua ressurreição.

Essa contrafactual é verdadeira se e somente se nos mundos mais próximos em que Jesus não aparece para Tomé, Tomé não acredita que Jesus tenha ressuscitado. Vamos primeiro às condições iniciais nas quais Tomé formou a crença de que Jesus ressuscitou. (a) Tomé viu um homem com determinadas características que lhe permitiu identificar Jesus; o mecanismo cognitivo usado foi, portanto, a visão. (b) Além do mais, podemos supor que naquele momento a visão de Tomé funcionava normalmente; ele não tinha tomado qualquer substância que alterasse o funcionamento de sua visão, não tinha qualquer tipo de defeito que comprometesse o processamento da informação, etc. (c) Quanto ao ambiente, estava perfeitamente iluminado e Tomé estava a uma distância curta de Jesus. (d) Supondo que era de fato Jesus quem se apresentara a Tomé, o input recebido pelo aparato visual de Tomé é proveniente do próprio do próprio Jesus. Supondo-se que nada de errado ocorreu no processamento da informação provinda do meio, o resultado final do processo é a crença de Tomé de que ele está na presença física de Jesus.

Vamos ao segmento de mundo que nos importa. Como o que queremos avaliar é se os mundos próximos em que Jesus ressuscita são mundos em que Tomé acredita que Jesus ressuscitou, não podemos supor inicialmente que Tomé forma a crença. E também não podemos supor que houve qualquer input proveniente de Jesus. Assim, o segmento que isolaremos para determinar a extensão dos mundos próximos terá de satisfazer apenas (a)-(c). E claro, temos de fazer uma pequena modificação em (c), uma vez que não podemos supor que Tomé está diante de Jesus; nesse caso Tomé está apenas em um ambiente adequadamente iluminado e a uma distância curta dos objetos físicos médios dispostos no ambiente. Podemos detalhar um pouco mais o ambiente: estavam na sala além de Jesus, os onze apóstolos restantes, que sendo pessoas honestas não tentaram pregar qualquer peça em Tomé ou fazê-lo de qualquer forma acreditar na ressurreição de Jesus. Peguemos o mundo mais próximo ao atual:

$w_1 = \{\langle \text{A visão de Tomé funciona adequadamente} \rangle, \langle \text{Há iluminação adequada} \rangle, \langle \text{Tomé está a uma distância curta dos objetos físicos médios que o rodeiam} \rangle, \langle \text{Jesus não aparece para Tomé} \rangle, \text{etc.}\}$

É claro que a proposição $\langle \text{Tomé não acredita que Jesus ressuscitou} \rangle$ está contida em w_1 . Os mundos em que tal proposição é falsa e Jesus não aparece para Tomé são mundos em que pelo menos uma destas coisas ocorre: as condições (a)-(c) não foram cumpridas ou ocorreu um erro de processamento de informação durante a formação da crença. Embora tais mundos intuitivamente nos pareçam próximos, não são próximos o suficiente para a avaliação da sensibilidade.

Um problema pode surgir agora. Considere novamente CELEIRO. Listemos grosseiramente as proposições que compõem o segmento relevante:

$w_{@} = \{\langle \text{João tem a visão funcionando adequadamente} \rangle, \langle \text{Há iluminação adequada} \rangle, \langle \text{João está a uma distância razoável do celeiro}_1 \rangle, \langle \text{João olha para o lote}_1 \text{ em que está erigido o celeiro}_1 \rangle, \langle \text{Com exceção do celeiro}_1 \text{ todos os celeiros naquela região são celeiros de fachada} \rangle, \text{etc.}\}$

Consideremos o mundo mais próximo em que é falso que haja um celeiro real no lote_1 .

$w_1: \{\langle \text{João tem a visão funcionando adequadamente} \rangle, \langle \text{Há iluminação adequada} \rangle, \langle \text{João está a uma distância razoável do lote}_1 \rangle, \langle \text{João olha para o lote}_1 \rangle, \langle \text{Não há um celeiro no lote}_1 \rangle, \langle \text{Todos os celeiros, do lote 2 ao 1000 são de fachada} \rangle, \text{etc.}\}$

Podemos dizer seguramente que a proposição $\langle \text{João acredita que há um celeiro no lote}_1 \rangle$ está contida em w_1 ? Depende. A proposição $\langle \text{Não há um celeiro no lote}_1 \rangle$ não implica $\langle \text{Há um celeiro de fachada no lote}_1 \rangle$ uma vez que a proposição $\langle \text{O lote}_1 \text{ está vazio} \rangle$ pode ser o caso. Adicionemos então a w_1 a proposição $\langle \text{O lote}_1 \text{ está vazio} \rangle$:

$w_2: \{\langle \text{João tem a visão funcionando adequadamente} \rangle, \langle \text{Há iluminação adequada} \rangle, \langle \text{João está a uma distância razoável do lote}_1 \rangle, \langle \text{João olha para o lote}_1 \rangle, \langle \text{Não há um celeiro no lote}_1 \rangle, \langle \text{O lote}_1 \text{ está vazio} \rangle, \langle \text{Todos os celeiros, do lote 2 ao 1000 são de fachada} \rangle, \text{etc.}\}$

Assim, se mundos como w_2 – mundos em que o lote está vazio seja porque o celeiro real desmoronou dias antes de João passar pela região, ou porque o lote possui uma mina de água que será preservada, etc. – constituírem o limite da extensão do conjunto de mundos próximos, a sensibilidade dará como veredito que João sabe que há um celeiro no lote₁. Mas esse resultado é contraintuitivo. Em algum sentido relevante a crença de João parece ter sido verdadeira por sorte.

Mas é claro que a proposição ⟨Há um celeiro de fachada no lote₁⟩ poderia também ser verdadeira:

w_3 : {⟨João tem a visão funcionando adequadamente⟩, ⟨Há iluminação adequada⟩, ⟨João está a uma distância razoável do lote₁⟩, ⟨João olha para o lote₁⟩, ⟨Não há um celeiro no lote₁⟩, ⟨Há um celeiro de fachada no lote₁⟩, ⟨Todos os celeiros, do lote 2 ao 1000 são de fachada⟩, etc.}.

Nesse caso, se mundos como w_3 – mundos em que há um celeiro de fachada no lote₁ porque os moradores da região construíram as fachadas aleatoriamente, ou porque os moradores quiseram esconder a nascente d'água com um celeiro de fachada, etc. – constituírem o limite da extensão do conjunto de mundos próximos, a sensibilidade dará o veredito de que João não sabe que há um celeiro no lote₁. Esse é o nosso veredito preferido.

Como decidiremos qual o conjunto de mundos contará para a avaliação das contrafactuais da sensibilidade? Note que o conjunto de mundos como w_2 é um subconjunto do conjunto de mundos como w_3 . Mas isso não é problema, uma vez que para refutar uma contrafactual basta que haja apenas um mundo próximo em que a antecedente é verdadeira e a consequente falsa. Para mantermos nosso veredito inicial de que CELEIRO não constitui um caso de conhecimento, temos de escolher w_3 como limite. Mas certamente há aqueles que desde o início desconfiavam de que CELEIRO fosse um caso de falha de conhecimento; a intuição contrária à minha – de que CELEIRO é um caso genuíno de conhecimento – não é rara. Por isso é preciso alguma razão para se escolher w_3 .

O fato de considerarmos a proposição ⟨Há um celeiro falso no lote₁⟩ determinará três conjuntos distintos de mundos possíveis próximos. (i) o conjunto de mundos em que o lote está vazio; (ii) o conjunto de mundos nos quais há outra edificação qualquer no lote₁, e (iii) o conjunto de mundos nos quais há um celeiro falso no lote₁. À partida é difícil decidir qual será

o conjunto relevante. Por isso, temos de nos focar no cenário com certo cuidado. O que torna CELEIRO um caso de falha de conhecimento é a incapacidade do sujeito em discriminar visualmente um celeiro real de um celeiro falso. Nos mundos gerados em (i) e (ii), o sujeito sabe que há um celeiro em sua frente somente se supusermos que ele é capaz de discriminar um celeiro de um edifício ou de um posto de gasolina; ou que o funcionamento adequado de sua visão o permitiria ver o lote vazio, caso estivesse. Mas isso de modo algum nos chama a atenção para a incapacidade do sujeito em discriminar um celeiro real de um celeiro falso. Por exemplo, suponha um mundo em que não há um celeiro no lote₁. Nesse mundo, Sebastião chega em frente ao lote₁, olha para ele e não forma a crença de que há ali um celeiro. O que nos garante que reconhecer a incapacidade de Sebastião em discriminar entre celeiros reais e falsos não nos dará a impressão de que Sebastião não sabe que há um celeiro no lote₁? Penso que apenas uma restrição do tipo *Sebastião não teria olhado para qualquer outro lote* ou *Sebastião ignora todo o distrito, com exceção do lote₁*. Uma analogia talvez nos ajude. Osama mora numa região bélica e ao redor de sua casa há um campo minado. Ele decide comprar um detector de minas na mão de um negociante. Assim, ele começa a traçar um caminho seguro para poder ir de sua casa até as montanhas. Com sucesso ele traça o caminho. Porém, sem que ele soubesse, o detector estava avariado: não era capaz de discriminar entre uma mina e um determinado minério que se encontrava nos limites do campo minado. Quando Osama diz para as pessoas que não pisem onde há marcações vermelhas, pois lá há minas, parece bastante razoável atribuir conhecimento a Osama. Se não tivesse uma bomba ali, ele não teria acreditado que tinha, uma vez que o detector não teria apitado. É claro que isso é assim porque o terreno limítrofe ao campo minado não interessa. Se levássemos em conta o terreno com o minério, e levássemos em conta que as minas são colocadas aleatoriamente num terreno, o detector apitaria na presença do minério e Osama acreditaria que há ali uma bomba. Contudo, como me parece claro, o caso de Osama exige apenas o terreno das minas. Portanto, a menos que tenhamos alguma restrição razoável a ser feita, como no caso de Osama, CELEIRO não tem de ser interpretado desse modo, e, por conseguinte, é um caso de falha de conhecimento.

Podemos agora estruturar CELEIRO de acordo com o esquema apresentado acima:

- (i) Condições iniciais:
 - a. Sebastião olha para o lote₁

- b. Sua visão funciona perfeitamente bem
- c. Sebastião está numa região infestada de celeiros falsos
- d. Sebastião olha para o único celeiro real
- (ii) Processo: nada de anormal acontece como o processo visual de Sebastião; nenhum tipo de erro de processamento de informações em seu cérebro, etc.
- (iii) Resultado: Sebastião acredita que há um celeiro no lote₁.

E agora fazemos o mesmo com o mundo mais próximo, w_1 , em que a proposição ⟨Há um celeiro no lote₁⟩ é falsa:

- (i) Condições iniciais:
 - a. Sebastião olha para o lote₁
 - b. Sua visão funciona perfeitamente bem
 - c. Sebastião está numa região infestada de celeiros falsos
 - d. Sebastião olha para um celeiro falso**
- (ii) Processo: nada de anormal acontece como o processo visual de Sebastião; nenhum tipo de erro de processamento de informações em seu cérebro, etc.
- (iii) Resultado: Sebastião acredita que há um celeiro no lote₁.

Parece-me claro agora de que modo temos de avaliar uma crença de acordo com a sensibilidade: a) selecionamos o seguimento do mundo atual relevante, b) isolamos suas condições iniciais, c) o funcionamento do mecanismo cognitivo, e, d) a crença resultante; em seguida, fazemos o mesmo com o mundo mais próximo em que a proposição alvo é falsa e checamos se o resultado final é o mesmo. É isso, portanto, que precisamos para avaliar o valor de verdade das contrafactuais.

2. Função apropriada, micro e macro ambientes

2.1. Função apropriada e perspectivismo de virtudes

O modo pelo qual optei em apresentar a sensibilidade nos conduz muito naturalmente a duas teorias bastante próximas: o perspectivismo de virtudes (SOSA [1985] 1991) e teoria da função apropriada, primeiramente proposta por Alvin Plantinga (1993). Ambas as teorias pertencem à família mais ampla de posições confiabilistas. Plantinga pensa que conhecimento é crença verdadeira garantida, sendo a garantia entendida como aquela propriedade que em

grau suficiente torna a crença verdadeira conhecimento.⁶⁰ A garantia, por sua vez, é entendida em termos do funcionamento apropriado ou adequado do mecanismo cognitivo responsável pela produção da crença. Assim, a crença de S de que α é garantida se, e somente se:

- a. foi produzida por um mecanismo cognitivo funcionando adequadamente (de acordo com seu plano funcional [*design plan*]);
- b. num ambiente adequado;
- c. esse mecanismo visa a crença verdadeira; e
- d. visa com sucesso a crença verdadeira.

A ideia do *funcionamento apropriado* ou adequado de um mecanismo cognitivo é simples. Considere um termômetro de mercúrio. Um termômetro está funcionando apropriadamente quando registra a temperatura da maneira como foi concebido para registrar; isto é, quando executa aquilo que está de acordo com o seu plano funcional: a coluna de mercúrio se desloca no bulbo (dentro de um tubo capilar fechado a vácuo) de acordo com a expansão do mercúrio. O mesmo podemos dizer de um mecanismo cognitivo: ele tem de funcionar de modo a produzir crenças (ou, em geral, representações da realidade).⁶¹ Se o mecanismo cognitivo apresentar algum tipo de disfunção ou desordem, a crença produzida por ele não será garantida.

Assim como um termômetro de mercúrio não funciona adequadamente em ambientes com temperaturas menores do que -40°C e maiores do 360°C , devido a seus pontos de fusão e ebulição, respectivamente, também há vários ambientes em que nossos mecanismos cognitivos não funcionam adequadamente. Por exemplo, no completo breu da noite, ainda que o funcionamento da minha visão esteja em grande parte de acordo com seu plano funcional, se eu olhar diante de mim e formar a crença de que há um gambá ali, essa crença não será garantida, uma vez que a visão humana se desenvolveu para produzir representações num

⁶⁰ Ou seja, a garantia é aquela propriedade à qual pretendemos nos referir com “justificação” em sua acepção não-deontológica.

⁶¹ A ideia de um plano funcional parece sugerir desígnio ou propósito: assim como alguém planeja o termômetro com uma determinada finalidade, também alguém teria de ter projetado os mecanismos cognitivos animal a fim de produzir representações da realidade. Mas apenas sugere. A melhor biologia de que dispomos nos conta que nossos mecanismos cognitivos executam as funções que executam devido a um processo evolutivo longo e gradual. O próprio Plantinga admite que a teoria da função apropriada seja à partida neutra quanto ao desígnio inteligente e a evolução (1993, p. 13, 198; 2000, p. 154), muito embora argumente depois que uma abordagem naturalista metafísica não dá conta de explicar adequadamente a função apropriada, mais especificamente, a confiabilidade dos nossos mecanismos cognitivos (1993, caps. 11 e 12; 2011, cap. 10).

ambiente minimamente iluminado. Por isso, o funcionamento adequado depende do *meio adequado* para o funcionamento.

Contudo, uma crença produzida por um mecanismo funcionando apropriadamente num ambiente adequado pode não ser garantida. Por exemplo, se Freud estiver correto, as crenças religiosas são formadas por *wishful thinking* e visam confortar o crente “neste mundo frio, triste e miserável em que nos encontramos” (PLANTINGA 2000, 161). Assim, embora o mecanismo formador de crenças religiosas possa funcionar apropriadamente, não produz crenças garantidas porque não *visa a crença verdadeira* (uma crença não precisa ser verdadeira para ser reconfortante!).

Mas mesmo que vise a crença verdadeira, é possível que um determinado mecanismo cognitivo não tenha sucesso em produzir crença verdadeira. Não adianta construir um termômetro com a intenção de que ele seja bom em medir a temperatura e o resultado final ser um termômetro que erra na maioria das vezes. É crucial para que uma crença tenha garantia que o mecanismo que a produziu *visa com sucesso a crença verdadeira*.

Note que as duas últimas condições (c-d) dizem respeito à confiabilidade. Um mecanismo que visa com sucesso a crença verdadeira é um mecanismo que tende a produzir mais crenças verdadeiras do que falsas. E isso está relacionado com o plano funcional relevante para as duas primeiras condições (a-b): o sucesso do mecanismo cognitivo em produzir crença verdadeira depende de seu funcionamento adequado num ambiente adequado.

Já o Sosa perspectivista pensa que uma crença verdadeira é conhecimento quando produzida por uma virtude intelectual, sendo uma virtude intelectual entendida como uma disposição do mecanismo cognitivo em atingir crenças verdadeiras e evitar crenças falsas. Mais especificamente:

Uma virtude intelectual V de um sujeito S relativa a um meio-ambiente M [é uma] disposição de S para acreditar corretamente em proposições no campo F a respeito das quais S está nas condições C no meio M. (1991, p. 140)

E o seguinte deixa ainda mais claro a afinidade com a nossa proposta:

Porque S tem certa natureza interna, I, e encontra-se em certo meio-ambiente, M, S muito provavelmente tem razão a respeito de qualquer proposição X no campo F a respeito da qual S está na condição C. S poderia ser um humano; I poderia envolver a

posse de bons olhos e de um bom sistema nervoso, incluindo um cérebro funcionando bem; M poderia incluir a superfície da Terra com as suas propriedades relevantes, no interior dos parâmetros experimentados por humanos através dos séculos, ou qualquer outro experimentado por S durante sua vida ou parte dela; F poderia ser um campo de proposições especificando as cores ou formas de um objeto diante de S até certo nível de determinação e complexidade (por exemplo, ser verde e ser quadrilátero, mas não ser verde-limão ou ser quiliágono); e C poderiam ser condições como S ver o objeto sob iluminação adequada à distância de um braço e sem obstruções. (p. 139)

É claro que o modo como proponho a avaliação da sensibilidade não implica diretamente qualquer das abordagens acima. Na verdade, penso que ela seja compatível com qualquer teoria confiabilista. Embora, no fim das contas, a explicitação das condições para a sensibilidade tenha de fazer menção ao mecanismo e ao ambiente. A minha preferência pela epistemologia da função apropriada se dá, sobretudo, pela sua capacidade de unificação. Ela engloba as principais abordagens confiabilistas ao conhecimento e à garantia como o confiabilismo processual, o perspectivismo de virtudes e as abordagens rastreadoras.

Permita-me, antes de prosseguirmos, apenas um breve comentário sobre o plano funcional e uma objeção influente. A ideia de plano funcional foi principalmente introduzida para lidar com o contraexemplo de um sujeito que tem uma lesão cerebral que faz com ele próprio forme a crença de que tem uma lesão cerebral; nesse caso, o sujeito tem uma crença verdadeira formada por um mecanismo confiável, mas não é conhecimento. Porém, uma vez que a lesão cerebral é um tipo de disfunção cognitiva, a crença carece de garantia, e, portanto, não é conhecimento. Considere, porém, o famoso caso do Homem do Pântano (SOSA 1993).⁶² Um sujeito está na beira do pântano; um forte raio de sol reflete nas águas, o atinge e em seguida uma árvore. Ao ser atingido pelo raio, o sujeito se desintegra e a árvore acaba virando uma cópia exata sua, molécula a molécula. Ao sair do pântano, esse novo sujeito, forma crenças exatamente como o sujeito anterior. Contudo, uma vez que os mecanismos cognitivos desse novo sujeito não têm uma história evolutiva, não há qualquer plano funcional. Portanto, há conhecimento sem que um mecanismo cognitivo funcione de acordo com seu plano funcional. Não penso que contraexemplos desse tipo funcionem. Primeiro, o mecanismo não precisa funcionar de acordo com o *seu* plano funcional; basta que ele

⁶² Não pretendo aqui discutir em qualquer detalhe os desdobramentos da discussão desse contraexemplo. Várias posições um tanto sofisticadas já apareceram (cf. BOYCE & MOON 2016; TOLLY 2018). A minha intenção é apenas indicar que o contraexemplo não é definitivo.

funcione de acordo com *algum* plano funcional. Por exemplo, considere o nascimento do forno micro-ondas. O técnico Percy Spencer trabalhava na construção de uma peça para radares usada na detecção de aviões durante a Segunda Guerra. As micro-ondas emitidas por essa peça derreteram um chocolate em seu bolso, e em seguida ele testou se elas eram capazes de cozinhar alimentos, o que se mostrou verdadeiro. Ele então adaptou essa tecnologia para construir um forno. Ora, é claro que não fazia parte do plano funcional dessa tecnologia estourar pipocas, muito embora ela tenha se mostrado boa para isso. O mesmo vale para um mecanismo cognitivo, ele pode ter inicialmente um determinado plano funcional (ou mesmo nenhum), mas se adaptar igualmente bem para outra função.

2.2. Micro e macro ambiente

Chegamos ao componente da avaliação da sensibilidade que penso ser o mais escorregadio. A ideia é simples: qualquer mecanismo que execute determinada função, irá executá-la num ambiente determinado; e nem todo ambiente é propício para que o mecanismo execute bem a sua função. Contudo, pouca atenção tem sido dada para a especificação do ambiente.⁶³ Felizmente, Plantinga nos dá alguma luz (1996; 1997; 2000, pp. 153-60) ao refinar a noção de ambiente introduzindo as noções de micro e macro ambiente.

Consideremos primeiro a noção de ambiente. Podemos entendê-lo amplamente como o subconjunto das nossas condições iniciais para a formação da crença (alvo de conhecimento) que descreve o lugar no qual o processo cognitivo opera ou irá operar. Mas isso é demasiado amplo; considerando que o lugar onde formo a minha crença de que há um computador em minha frente é a Via Láctea, esse conjunto poderia conter a proposição de que uma estrela acaba de explodir, o que é completamente irrelevante para a formação da minha crença e para o funcionamento das minhas faculdades cognitivas. Assim temos de restringir o ambiente para o *ambiente cognitivo*, o lugar específico em que o processo opera ou irá operar.

⁶³Alvin Goldman (1976) faz menção ao ambiente na sua relação DOE (*distance-orientation-environment*) introduzida para definir uma equivalência perceptual. “O percepto é afetado pelo estado atual do *ambiente*, e.g., pela iluminação, pela presença ou ausência de objetos interferentes, pela direção e velocidade do vento” (p. 781). Sanford Goldberg (2015, p. 124) também faz menção explícita ao ambiente ao caracterizar diferentes versões de confiabilismo: “entendamos um ‘ambiente normal’ para um processo-tipo ϕ como sendo aqueles tipos de ambiente no qual ϕ está ou estaria sendo empregado na formação da crença de um sujeito cognitivamente saudável”. Ernst Sosa (1991, vide citação acima): “[o ambiente] poderia incluir a superfície da Terra com as suas propriedades relevantes, no interior dos parâmetros experimentados por humanos através dos séculos, ou qualquer outro experimentado por S durante sua vida ou parte dela”.

Assim, o ambiente cognitivo será o subconjunto que descreve as características relevantes do lugar no momento em que o processo cognitivo opera ou irá operar.

Podemos ser ainda mais específicos e dizer que o ambiente cognitivo é formado por um macro-ambiente e vários micro-ambientes. Nas palavras de Plantinga:

Um macro-ambiente cognitivo é mais geral e mais global do que um micro-ambiente cognitivo. O nosso macro-ambiente cognitivo aqui na terra incluiria características macroscópicas como a presença e as propriedades da luz e do ar, a presença de objetos visíveis, de outros objetos detectáveis através de sistemas cognitivos do nosso tipo, de objetos nem tanto detectáveis, de regularidades da natureza, da existência e da natureza geral de outras pessoas, e assim por diante. [...] [U]m determinado macro-ambiente pode conter muitos micro-ambientes diferentes [...] (1997, p. 143; 2000, p. 158)

Podemos então dizer que um micro-ambiente é um subconjunto do macro-ambiente.

Essa distinção entre micro e macro ambiente nos ajuda a entender aquilo a que Plantinga (1997, p. 142; 2000 p. 157) chamou *falta de resolução* de certos processos de formação de crença ou faculdades cognitivas. Os nossos processos cognitivos, em geral, operam de maneira bem sucedida – *i.e.*, produzindo crença verdadeira não acidentalmente – quando estão a funcionar apropriadamente no macro-ambiente cognitivo adequado. E um macro-ambiente adequado é aquele especificado pelo plano funcional (ou que pelo menos está de acordo com algum plano funcional). Mas nem sempre nossos mecanismos cognitivos terão resolução suficiente para discriminar entre certos estados de coisas, como no caso dos celeiros falsos, muito embora estejam funcionando apropriadamente num macro-ambiente adequado. Isso ocorre por conta de um micro-ambiente: em meio a um ambiente bem iluminado, cheio de objetos de tamanho-médio, etc. há um punhado de celeiros falsos visualmente indistinguíveis do único celeiro real dali. Suponha que eu esteja no campo e precise de água a 100°C. O modo como faço para saber se ela está nessa temperatura é coloca-la no fogo e esperar até que comece a ferver. Contudo, ao seguir caminho na mata, vou subindo o terreno de modo que, sem que eu faça a mínima ideia, a pressão cai para pouco menos de 1atm. Armo acampamento e coloco água para ferver. Nesse caso, formarei uma crença falsa de que a água está a 100°C, uma vez que a pressão diminui. O método que usei é confiável no macro-ambiente (próximo ao nível do mar), muito embora perca resolução em altitude elevada. Isto

é, o micro-ambiente – a parte do terreno em que me encontro em determinada altitude – não é adequado para que eu forme crenças sobre a temperatura da água usando o método que usei.

Assim, casos de falha de conhecimento como em CELEIRO ocorrem pelo seguinte: o micro-ambiente não é favorável para o funcionamento dos processos cognitivos em questão (embora o macro-ambiente seja favorável). Esse é o diagnóstico mais geral que Plantinga oferece para o problema anti-sorte. Crença verdadeira formada por mecanismos confiáveis funcionando apropriadamente num macro-ambiente para o qual foram projetados é compatível com a sorte; para evitar a sorte, é preciso que o micro-ambiente não seja desfavorável.⁶⁴ Plantinga então adiciona uma *condição de resolução*:

(CR) *Se S sabe que α , o micro-ambiente é favorável ao processo cognitivo sendo usado para formar a crença de que α .*⁶⁵

Por sua vez, o micro-ambiente é favorável se, e somente se:

(CF) estivesse S a formar uma crença por meio do exercício da faculdade cognitiva nesse micro-ambiente, S formaria uma crença verdadeira. (p. 159).

Essa condição poderia ser entendida como a condição de segurança; posta cruamente:

(SEG) *Se S acreditasse que α , então seria o caso que α ($Bp \square \rightarrow p$);*

ou dito no jargão dos mundos possíveis: nos mundos mais próximos em que S acredita que α , α é verdadeira. Por ora, deixemos de lado a condição de segurança.⁶⁶

⁶⁴ Plantinga (1997; 2000; PLANTINGA & BOYCE 2015) menciona três casos que aparentemente constituem contraexemplos à *la Gettier* à epistemologia da função apropriada. [RELÓGIO PARADO]: S olha para seu relógio de ponteiros e forma a crença verdadeira que são 7 horas. Mas o relógio estava parado já há 24 horas com o ponteiro no 7. [IRMÃOS GÊMEOS]: S vê A saindo de sua casa e forma a crença verdadeira de que A está saindo de casa. Mas A tem um irmão gêmeo, B, que S é incapaz de distinguir visualmente. B estava em casa e se tivesse saído S acreditaria que A é que estava saindo. [VAN]: Plantinga estaciona a sua van azul bebê em local proibido. A guarda municipal, sem que Plantinga saiba, reboca a sua van que, por descuido do guincho, é completamente destruída. Nesse exato momento, também sem que Plantinga soubesse, ele ganha uma van azul bebê igualzinha num bilhete que tinha comprado para ajudar uma instituição de caridade. Enquanto tudo isso acontecia, Plantinga estava na igreja contando para um novo amigo que tinha uma van azul bebê. Não penso, contudo, que esse último caso seja um contraexemplo genuíno, pelas razões apresentadas no capítulo 1: parte do conteúdo da crença de Plantinga é *aquela* van, não uma outra van qualquer (ainda que exatamente igual). Ele, portanto, tem uma crença falsa que possui *aquela* van. Os outros dois casos, porém, são genuínos, são casos como CELEIRO.

⁶⁵ Originalmente: “(CR) Uma crença B produzida pelo exercício E dos poderes cognitivos tem garantia suficiente para ser conhecimento somente se BEM (o micro-ambiente em relação a B e E) é favorável para E.” (2000, p. 159).

Plantinga, contudo, não se mostra satisfeito com (CF), a condição de favorabilidade. Isso devido ao seguinte contraexemplo:

[NEBLINA] João passa novamente pela roça dos celeiros falsos. Dessa vez uma neblina forte cobre o terreno, deixando exposto apenas o celeiro real. Ao passar pelo celeiro ele então forma a crença verdadeira que há ali um celeiro.

A crença de João satisfaz todas as exigências para a garantia, principalmente (CF): no micro-ambiente em que ele se encontra – a roça dos celeiros falsos encoberta por neblina, exceto o celeiro real – se ele formasse visualmente a crença de que há ali um celeiro, seria verdade que há um celeiro. Dito de outro modo: nos mundos próximos em que João forma a sua crença através da visão naquele micro-ambiente, ele acredita que há ali um celeiro. Porém, claramente a sua crença é verdadeira por sorte. Assim, (CR) não exclui a sorte epistêmica.

A reação de Plantinga é a formulação de uma nova condição, dessa vez probabilística ao invés de modal. A ideia é, grosso modo, que se houver algum estado de coisas (como a existência de celeiros falsos) que introduzido no micro-ambiente favorável à formação da crença verdadeira de que p (como a crença de que há um celeiro real), a probabilidade da crença de que p ser verdadeira diminui satisfatoriamente a ponto de impedir que S saiba que p (p. 160). Thad M. Botham (2003), porém, encontra um contraexemplo para essa versão probabilística de (CF). É uma modificação do caso dos irmãos gêmeos. Suponha que S vê A, que não tem um irmão gêmeo, saindo de casa. Nesse mesmo instante, um tio de A conta uma mentira a um amigo, que está a milhares de quilômetros de distância de S: diz que B tem um irmão gêmeo, e também que esse irmão gêmeo mora com A. Se adicionarmos o estado de coisas em que o tio conta uma mentira ao micro-ambiente favorável em que S forma a crença de que A está saindo de casa, a probabilidade de a crença de S ser verdadeira diminui consideravelmente. Contudo, parece claro que S sabe que A está saindo de casa. Isso fez com que Plantinga e Boyce (2015) tentassem uma nova condição de favorabilidade, dessa vez levando em conta que o exercício dos poderes cognitivos “toma por garantido” que certos estados de coisas C ocorrem e que não há qualquer outro estado de coisas C' que impeça a ocorrência de C . Nesse caso, S toma por garantido (não num sentido internista) que não há qualquer gêmeo indistinguível nas proximidades (p. 154). Eles próprios admitem ser essa apenas uma noção intuitiva e provavelmente sujeita a contraexemplos.

⁶⁶ Não seguirei a formulação da condição de favorabilidade em termos da segurança pelo motivo de que o meu interesse é a sensibilidade. No capítulo seguinte argumentarei contra a segurança.

Todavia, não interessa aqui a (FC) probabilística. Continuarei meu foco na versão modal, porém formulada de acordo com a sensibilidade. Ei-la:

(CF_{sen}) Um micro-ambiente M é favorável para a formação da crença de que α através do exercício da capacidade cognitiva E sse (i) se α fosse falsa então S não acreditaria que α por meio de E em M, e (ii) se α fosse o caso, então S acreditaria que α por meio de E em M.

Posto cruamente,

(SEN) (i) Se α fosse falsa, então S não acreditaria que α ($\neg\alpha \square \rightarrow \neg B\alpha$); e (ii) Se fosse o caso que α , então S acreditaria que α ($\alpha \square \rightarrow B\alpha$).

Ou dito no jargão dos mundos possíveis: Nos mundos próximos em que α é falsa, S não forma a crença de que α . E nos mundos em que é o caso que α , S forma a crença de que α .

Explicitando a condição de resolução temos agora:

(CF_{SEN}) M é favorável à formação da crença de que α através do exercício da capacidade cognitiva E sse (i) nos mundos mais próximos em que α é falsa, S não forma a crença de que α através de E em M, e (ii) nos mundos mais próximos em que α é o caso, S forma a crença de que α através de E em M.

Penso agora ser possível explicar por que oscilamos entre atribuir e negar conhecimento a Sebastião em TERMÔMETRO. Começemos pela intuição de que Sebastião sabe que seu filho tem 38,5°C. Se considerarmos o micro-ambiente em que Sebastião forma a sua crença, (CF_{SEN}) é satisfeita, de modo que M é favorável ao uso daquele termômetro: já tendo escolhido o termômetro confiável, naquele ambiente, se o filho de Sebastião tivesse 37°C, Sebastião não acreditaria que ele tem 38,5°C. Porém, no que diz respeito ao macro-ambiente as coisas não são tão claras. É certo que o macro-ambiente para o termômetro confiável é favorável. E é também certo que uma vez que os termômetros restantes estavam avariados, o ambiente não é favorável nem desfavorável. Mas temos de considerar a escolha aleatória do termômetro. A crença de Sebastião não depende apenas do termômetro confiável e de sua leitura confiável do termômetro; a escolha do termômetro parece contar também. E o ambiente em que ele estava não era de modo algum favorável para a escolha do termômetro. Ainda que o macro-ambiente, então, seja favorável ao funcionamento do termômetro bom,

não é favorável à formação da crença. E uma vez que o ambiente tem de ser favorável à formação da crença através de capacidade cognitiva (e do auxílio cognitivo), não é suficiente que apenas haja um ambiente favorável ao processo cognitivo.

Uma analogia talvez nos ajude. Suponha que Rabelo tenha sofrido um acidente que dividiu o módulo da visão em seu cérebro em quatro sub-módulos. Três desses sub-módulos são disfuncionais e produzem as seguintes experiências: sempre que Rabelo vê alguém tocando violão, a pessoa é apresentada como seu amigo Dino. O restante funciona normalmente. Primeiro cenário: por algum motivo, apenas o módulo bom é operante no cérebro de Rabelo; os outros três estão inoperantes (embora os especialistas não descartem a possibilidade deles começarem a operar de uma hora para outra). Segundo cenário: os quatro módulos passam a operar repentinamente no cérebro de Rabelo; e o módulo a ser usado na produção se sua crença visual é selecionado aleatoriamente. No primeiro cenário, quando Rabelo vê Dino tocando violão e acredita na proposição correspondente, parece claro que devemos atribuir conhecimento a Rabelo. Mas no segundo cenário, quando ele vê Dino tocando violão, usa o módulo bom, mas que foi selecionado aleatoriamente no seu cérebro, já não mais atribuiríamos conhecimento a Rabelo.

2.3. Relativizando o ambiente

Considerando a favorabilidade e a distinção entre micro e macro ambiente, temos as seguintes possibilidades: (i) macro e micro ambiente favoráveis, (ii) macro-ambiente favorável e micro-ambiente desfavorável, (iii) macro-ambiente desfavorável e micro ambiente favorável, e (iv) macro e micro ambiente desfavoráveis. O caso (i) é um caso claro de conhecimento; (ii) e (iv) são casos claros de falha de conhecimento. E quanto a (iii)? Aparentemente (iii) também é um caso de falha de conhecimento, uma vez que um macro-ambiente desfavorável tende a tornar o mecanismo cognitivo infiel (relativo àquele ambiente). Isso é o que sugere o diagnóstico dado a TERMÔMETRO logo acima.

Contudo, há uma maneira na qual o macro-ambiente desfavorável deixa de sê-lo. Mas para isso precisamos forçar o limite das condições ambientais de funcionamento do mecanismo cognitivo. O primeiro passo para isso é reconhecer que não precisamos manter fixo o macro-ambiente do plano funcional original; em casos especiais podemos tanto alargar como restringir esse macro-ambiente obtendo assim um novo macro-ambiente. Um exemplo

simples. No macro-ambiente em que o aparato visual humano foi moldado para funcionar, se a visão não sofre de qualquer disfunção, os humanos são capazes de reconhecer rostos de pessoas a cerca de 10 metros sob boa iluminação. Eu não, uma vez que sou míope e astigmata. Na verdade, há vários micro-ambientes – todos aqueles em que observo algum objeto a mais de 10 metros – em que a minha visão não funciona bem. É razoável supor, por isso, que, em geral, o macro-ambiente (ou parte substancial dele) para o qual a visão humana foi em geral moldada não é favorável ao funcionamento da *minha* visão. Contudo, sou bom o suficiente para discriminar rostos de pessoas e detalhes de vários objetos a uma distância curta, cerca de 1m ou 1,5m. Nesse tipo de micro-ambiente, portanto, a minha visão funciona bem. E é bastante razoável que alguém me atribua conhecimento nessas condições. Nesse caso, podemos fazer com que o micro-ambiente em que enxergo em boas condições de iluminação objetos a cerca de 1m ou 1,5m se torne o macro-ambiente. E como consequência, a possibilidade de um novo micro-ambiente desfavorável torna-se mais escassa.

A essa altura, poder-se-ia objetar que tudo isso é manobra *ad hoc*. Tudo o que fiz serve apenas para salvaguardar arbitrariamente a ideia de que a sensibilidade é uma condição necessária ao conhecimento. Que alargo e restrinjo a extensão do conjunto de mundos próximos (ou alternativas relevantes, ou micro-ambientes favoráveis) a meu bel prazer. A essa objeção tenho duas breves respostas. Primeiro, as nossas atribuições de conhecimento intuitivas a casos claros de conhecimento não são uniformes. Se vejo uma pessoa com óculos fundo-de-garrafa lendo uma página com letras miúdas e ela está tendo sucesso, a minha tendência é dizer que ela *sabe* o que está escrito ali. Porém, se logo em seguida ela tira os óculos e começa a ler outra página, dificilmente eu diria que ela sabe o que está escrito ali; diria ao invés que ou ela está fingindo ou que já tinha decorado o que estava escrito na página antes tirar os óculos. Ou pense no caso do nosso amigo Osama. Ele adquire um detector de minas que se usado para além do terreno de sua casa não daria os resultados adequados; porém, ao vermos que ele se restringe a usá-lo apenas naquela pequena região tendemos a dizer que ela sabe onde há minas ali por perto. Intuitivamente reconhecemos que algumas restrições no ambiente nos permitem atribuir ou negar conhecimento a terceiros; e o modo pelo qual estou sugerindo tratar o ambiente capta essa intuição.

Em segundo lugar, embora eu não tenha um critério que diga explicitamente quando um conjunto de mundos aumenta ou diminui a sua extensão, podemos pensar em algumas

opções. A primeira delas é puramente objetiva: o sujeito se encontra num ambiente tal e satisfaz determinadas relações de modo que não poderia deixar de ter conhecimento. Isso se adequa bem ao invariantismo. Porém, eu não gostaria de tornar uma teoria do conhecimento baseada na sensibilidade incompatível com perspectivas variantistas, como o contextualismo⁶⁷ e a infiltração pragmática⁶⁸. O conjunto de mundos pode ser determinado por padrões epistêmicos ou conversacionais que variam de contexto para contexto, ou mesmo por interesses práticos do sujeito. Embora eu tenha simpatia por essas duas últimas abordagens, deixarei essa questão em aberto, uma vez que não pensei o suficiente sobre elas e qualquer coisa que eu diga não passará de pura especulação.

Há uma maneira muito mais clara de se restringir o ambiente para que o mecanismo (ou dispositivo) se torne confiável: a adaptação consciente do plano funcional. Por exemplo, se Osama descobre que seu detector de minas não é bom quando considerada uma área maior, mas funciona bem em áreas em que o minério X não ocorre, ele pode deliberadamente especificar que seu detector funciona bem num ambiente que não ocorre o minério X. Agora, o macro-ambiente no qual o detector funciona é bastante claro: solos como comumente são encontrados na superfície da terra, mas sem a presença de X. Em geral, a descoberta de que não somos bons em formar crença verdadeira a partir de certo mecanismo cognitivo em determinados ambientes nos faz delimitar o ambiente em que somos bons.

Deixe-me encerrar com um comentário sobre a confiabilidade. Temos falado sobre crenças serem confiáveis. Uma crença é confiável quando é produzida por um processo de formação de crença (ou mecanismo cognitivo) confiável. E a confiabilidade é a tendência desse processo em produzir mais crenças verdadeiras do que falsas. Essa tendência, contudo, é explicada fazendo-se referência ao processo sendo usado num ambiente adequado que é favorável. E o ambiente é favorável quando, cumpridos todos os requisitos anteriores, o processo ou mecanismo cognitivo rastreia a verdade da proposição alvo de conhecimento: isto é, se a proposição fosse falsa o sujeito não formaria a crença correspondente, e se fosse verdadeira, formaria. Assim, podemos dizer então que uma crença é confiável se e somente se foi formada por um mecanismo cognitivo funcionando apropriadamente num ambiente favorável. É esse o resultado que pretendo com EFAS – a *Epistemologia da Função*

⁶⁷ Cf. RYSIEW (2016).

⁶⁸ Cf. KIM & McGRATH (2019).

Apropriada Sensitiva. Pelo menos no que diz respeito ao conhecimento empírico, considero EFAS uma versão consideravelmente apurada do confiabilismo.⁶⁹

3. Resposta às objeções

Vimos no Capítulo II que há uma bateria de objeções influentes contra a sensitividade. Algumas alegando que a sensitividade é demasiado permissiva e outras que é demasiado restritiva. Nesta seção, tendo já em mente a EFAS, tentarei responder a tais objeções.

3.1. Permissividade

a. Holograma

Começemos por HOLOGRAMA. Relembremos: Reinaldo tem uma crença verdadeira e sensitiva de que a taça da Libertadores está dentro da redoma, muito embora a sua crença tenha sido parcialmente causada pelo holograma da taça e não pela taça. A primeira coisa a se notar é que embora relutemos em atribuir conhecimento a Reinaldo, a sua crença não foi verdadeira por sorte. E isso é importante, pois é exatamente pelo fato de a crença ser sensitiva que ela não é verdadeira por sorte – nos mundos próximos em que não há taça, Reinaldo não acredita que a taça da Libertadores está na redoma. Se, portanto, a crença de Reinaldo não é conhecimento, tem de ser por outro motivo que não a sorte epistêmica. Sendo assim, HOLOGRAMA mostra no máximo que crença verdadeira sensitiva não é uma condição suficiente para o conhecimento.

É claro que poderíamos insistir que há algo de errado com o micro-ambiente, afinal a nossa visão não foi moldada para funcionar num ambiente em que os objetos de tamanho médio se revezam com imagens de hologramas. Ou seja, o macro-ambiente é favorável a Reinaldo, muito embora o micro-ambiente seja desfavorável. Contudo, vimos que não haveria holograma caso não houvesse a taça. E é por isso que nesse micro-ambiente a crença de Reinaldo é sensitiva. Porém, poder-se-ia insistir ainda que há um mundo próximo em que

⁶⁹ Mesmo no que diz respeito ao confiabilismo de virtudes (em que também incluo a epistemologia da virtude anti-sorte), é possível conceber EFAS no mesmo espírito. Para um confiabilista de virtudes, crença verdadeira é conhecimento quando for produzida por uma virtude intelectual (*i.e.* uma tendência ou propensão do mecanismo cognitivo em produzir crença verdadeira) e essa virtude desempenhar um papel central na explicação de por que o sujeito formou a crença que formou. Uma vez que explicar um fenômeno é pelo menos apontar como um conjunto de condições iniciais relevantes causou o *explanandum*, fica claro que em EFAS essa exigência é satisfeita. A sensitividade é um teste para a centralidade do papel da virtude intelectual na formação da crença. Em casos de ambientes desfavoráveis a virtude intelectual é driblada, a parte proeminente da explicação se desloca para aquelas condições ambientais que fazem o mecanismo cognitivo funcionar de maneira infiel.

Maria, digamos, entra na sala e rouba a taça no exato momento em que a imagem do holograma é projetada dentro da redoma, aproveitando do pequeno intervalo entre a troca do holograma para a taça. Reinaldo vira as costas e vai embora sem perceber que no instante seguinte a redoma fica vazia. Nesse caso, é falso que a taça está na redoma. Reinaldo ainda acredita que está. Portanto, HOLOGRAMA não seria um caso de crença verdadeira sensível; e, por conseguinte, não constituiria um contraexemplo à sensibilidade.

Levemos em conta, porém, a restrição do ambiente apresentada anteriormente neste capítulo. O presidente do clube pediu ao engenheiro que projetasse o holograma de modo que de nunca projetasse uma imagem da taça sem que a taça estivesse lá dentro da redoma. Ele não fez isso com a intenção de enganar os torcedores, e muito menos calhou desse dispositivo enganar os torcedores. Assim, deliberadamente o ambiente é restringido de modo a tornar a observação de um holograma (ainda que sem conhecimento do público) um meio confiável para a detecção da taça.

Se talvez você ainda tenha alguma dúvida, vejamos um caso parecido. Suponha que numa estrada muito perigosa há um ponto cego, local constante de acidentes. O departamento de trânsito decide tomar uma medida para acabar com os acidentes. Projetam um jogo de espelhos de modo a fazer os motoristas enxergarem o veículo vindo do ponto cego. Contudo, não divulgam que há esse jogo de espelhos para evitar o abuso de motoristas imprudentes. Vindo nessa estrada, então, Bino vê a imagem virtual (projetada pelos espelhos) de um caminhão e acredita que vem um caminhão bem diante dele na outra pista. Podemos facilmente conceber um mundo em que o caminhão na outra pista não está perto do caminhão de Bino, e que por uma pequena mudança na posição de um dos espelhos (feita pelo vento, digamos) a imagem projetada é de um caminhão um pouco mais distante. Sendo assim, Bino forma a crença de que o caminhão está próximo quando na verdade está distante. Contudo, esse mundo não é próximo o bastante. O fato de o ambiente ser controlado (presumivelmente a estrutura foi testada para evitar que fatores como vento, trepidação, etc. alterassem a posição dos espelhos) torna mais difícil que isso ocorra. É por isso que atribuímos conhecimento a Bino.

Portanto, HOLOGRAMA, ao invés de ser um caso de quase conhecimento (por assim dizer), parece ser um caso de conhecimento.

Mas se ainda insistirmos que a sensibilidade não é suficiente (embora necessária), a explicação de por que HOLOGRAMA não é um caso de conhecimento tem de provir da falha do cumprimento de alguma outra condição (ou condições) que conjuntamente com a sensibilidade tornam uma crença verdadeira conhecimento. Vejamos. O mecanismo cognitivo pelo qual Reinaldo forma a sua crença está funcionando adequadamente (para o macroambiente para o qual foi moldado). De modo que só nos resta o input recebido do meio. E aqui talvez tenhamos algo. A crença de Reinaldo não está causalmente conectada ao seu verificador – nesse caso ao fato de que a taça está na redoma – mas sim à imagem do holograma. E se aprendemos a lição da teoria causal do conhecimento, a crença tem de estar numa conexão causal apropriada com seu verificador. É certo que a crença não está diretamente conectada ao verificador. Contudo, a taça é causalmente responsável pelo holograma da taça, de modo que se não houvesse taça, não haveria holograma; e se não houvesse holograma, Reinaldo não formaria a crença de que a taça da Libertadores está na redoma. Assim, a crença de Reinaldo é contrafactualmente dependente de seu verificador. Por conseguinte, nada há de errado com o input.

Feitas essas considerações, só me resta concluir que HOLOGRAMA, ao contrário da intuição inicial, é um caso de conhecimento. (E na pior das hipóteses, é um caso em que alguma condição adicional deixou de ser cumprida).

b. Anatomista Paleolítico

Vejamos o caso do ANATOMISTA PALEOLÍTICO. O homem das cavernas adquire alguma evidência para pensar que tem um cérebro dentro de seu crânio a partir da observação que um ou dois animais possuem um cérebro dentro de seus crânios. O homem das cavernas tem uma crença indutiva verdadeira, baseada em evidência insuficiente, mas que é sensível.

Notemos primeiramente que uma crença indutiva não pode contar como conhecimento se baseada em evidência insuficiente. Assim, de modo algum a crença do homem das cavernas poderia ser conhecimento. Mas daí não se segue que a sensibilidade seja permissiva. Segue-se, no máximo, que a sensibilidade não é suficiente para o conhecimento.

Segundo, talvez pudéssemos insistir que a crença do homem das cavernas não é sequer sensível. Lembre-se que grande parte das crenças por si só não rastreiam a verdade da proposição; elas rastreiam via mecanismos cognitivos (processos formadores de crenças) ou

instrumentos que auxiliam a nossa cognição. No caso do homem das cavernas, é usado o raciocínio indutivo – mais especificamente, o raciocínio por analogia. A esfera de mundos próximos que contam para a avaliação da condicional da sensibilidade consiste de mundos em que a proposição ⟨O homem das cavernas não tem cérebro⟩ é falsa mas que a proposição ⟨O homem das cavernas usa o raciocínio indutivo⟩ é verdadeira. Essa última proposição implica, contudo, que de alguma coisa (que não seu cérebro biológico) o fez raciocinar. Assim, temos de considerar um conjunto de mundos em que o homem das cavernas tem um substituto para o cérebro biológico e raciocina indutivamente a partir da observação de que há cérebros dentro do crânio de dois animais não-humanos. Poderíamos considerar mundos em que o homem das cavernas teve seu cérebro biológico substituído por um hardware (resultado de experiências de uma raça alienígena muito avançada fazendo experimentos na Terra).⁷⁰ Há, agora, pelo menos um mundo próximo em que é falso que o homem das cavernas tem um cérebro mas que ele acredita que tem um cérebro. Portanto, a crença do homem das cavernas não é sensível.

Mas isso tem uma consequência indesejável: alguém que tivesse boa evidência indutiva para acreditar que tem um cérebro, ainda assim teria uma crença insensível. Ou seja, agora temos um caso em que a sensibilidade é demasiado restritiva, uma vez que exclui casos de conhecimento. Veremos o problema da sensibilidade com o conhecimento indutivo mais a frente.

É melhor, então, mantermos que a crença do homem das cavernas é sensível, embora não dependa de um mecanismo cognitivo para rastrear a verdade da proposição ⟨O homem das cavernas tem um cérebro⟩. Essa crença pertence a uma categoria (muito restrita, suponho) de crenças sensíveis. Outra crença parecida é a crença de que existo. Simplesmente não há qualquer mundo próximo (e nesse caso, mesmo distante!) em que a proposição ⟨L. H. existe⟩ seja falsa e eu ainda acredite que existo.⁷¹ Mas o mais importante é que crenças como essas

⁷⁰Poder-se-ia objetar aqui que tais mundos não são nada próximos do mundo atual, e, que, por isso, são irrelevantes para a avaliação da condicional da sensibilidade. Contudo, como defendi anteriormente, o que nos interessa é o segmento dos mundos possíveis que especifica a situação epistêmica do sujeito – daí eu ter usado a noção de um mundo *epistemicamente próximo*.

⁷¹Note que mesmo sendo sensível, apenas a crença de que existo não é suficiente para que saiba que existo. Pode parecer estranho, mas se eu tivesse adquirido a crença de que existo num estranho livro escrito por um jornalista incompetente e que quase sempre diz falsidades, a minha crença não seria conhecimento. A evidência adequada para a minha crença de que existo é introspectiva; só através desse tipo de exercício cognitivo é que sou capaz de saber que existo.

nos garantem que não encontraremos na vizinhança de mundo próximos um caso que torne a condicional da sensibilidade falsa.

c. Tributarista

Passemos aos casos do TRIBUTARISTA, de Tristan Haze. No primeiro caso, a ideia é que nos mundos próximos mantenho a minha crença de que o meu vizinho (um advogado confiável) é um oráculo divino (é isso o que Haze pretende dizer como “ter uma crença contrafactualmente robusta”), e que meu vizinho, nesses mesmos mundos não diria que p a menos que p fosse o caso. Assim, naqueles mundos próximos em que p não é o caso, meu vizinho não me diz p , e, portanto, não formo a crença de que p .

Mantendo a intuição de Haze – não sei que p –, qual a explicação dessa falha de conhecimento? Certamente não a de que a minha crença é verdadeira por sorte. A sensibilidade evita isso. Assim, a falha tem de estar na falha do cumprimento de outras condições. O que novamente mostra que o contraexemplo é no máximo à suficiência da sensibilidade. Mas qual condição é violada? Se considerarmos o mecanismo cognitivo usado – ou algo próximo a isso – a *credulidade* (disposição para acreditar em testemunhos), nada de disfuncional parece estar envolvido. Podemos supor, para fins de argumentação, que o testemunho de oráculos divinos são bastante confiáveis. Além do mais, o advogado, por ser competente, é também uma fonte confiável. Nada há de errado com o input recebido do meio também.

O que está errado então? O que explica a falha de conhecimento nesse caso? Só nos resta a anulação da crença. Haze monta o caso de modo que *se* eu descobrisse que meu vizinho é um advogado, não confiaria em sua palavra. Assim, a proposição verdadeira (O vizinho de L. H. é um advogado tributarista) funcionaria como um anulador para a minha crença de que p . Suponhamos que eu tenha justificação proposicional (mas não doxástica)⁷² de que meu vizinho é um advogado. Isso não seria suficiente para que a minha crença de que p seja conhecimento, uma vez que não formo a crença de que o meu vizinho é um advogado. É claro que poderíamos dizer que em algum mundo próximo acabo formando tal crença. Desse modo, há pelo menos um mundo próximo em que meu vizinho diz que p , p é verdadeira, mas não acredito que p (talvez eu duvide, ou acredite que $\neg p$ seja mais provável).

⁷² Assumo que a justificação doxástica é a *justificação proposicional* mais a *relação basal*.

E não pode haver conhecimento sem crença. Dito de maneira mais simples: a proposição ⟨O vizinho de L. H. é advogado⟩ é um anulador para a minha crença de que p porque eu facilmente poderia vir a acreditar que meu vizinho é um advogado. Isso explicaria a intuição de Haze de por que TRIBUTARISTA não é um caso de conhecimento.

Mas tudo depende da tal crença “contrafactualmente robusta”. Se com isso Haze pretende que não seja fácil que eu venha a acreditar que o meu vizinho é um advogado, então os mundos mais próximos em que não confio em sua palavra deixam de ser relevantes para o presente cenário. Portanto, TRIBUTARISTA é um caso de falha de conhecimento somente se não for fácil para mim acreditar que meu vizinho é advogado. Se minha situação epistêmica não torna disponível tal informação ou qualquer evidência a seu favor, é razoável dizer que sei que p . Eis um exemplo similar. Escobar acredita que calculadoras fabricadas na China são infíaveis, de modo que nunca confiou em um cálculo complexo feito numa calculadora chinesa. Escobar sempre usa a sua calculadora, uma calculadora muito boa, de uma das maiores fabricantes mundiais de calculadoras. Terminado um de seus projetos, Escobar resolve uma função em sua calculadora e acredita que o valor é 45,8. A conta está certa. Porém, sem que Escobar soubesse, a sua calculadora foi fabricada na China (a grande empresa tem uma montadora lá, embora não divulgue). Escobar sabe ou não o resultado de seu cálculo? Se a informação de que a calculadora é chinesa estiver facilmente acessível, ele formaria a crença de que a calculadora é infíavel e, portanto, não acreditaria no resultado. E isso explica por que hesitamos em atribuir conhecimento a Escobar: ele facilmente teria deixado de acreditar no resultado de sua calculadora. Mas, por outro lado, se a fabricante mantivesse seus negócios na China de maneira bastante velada, dificilmente Escobar viria a questionar a confiabilidade de sua calculadora. E nesse caso, é razoável atribuímos conhecimento a ele.

Vamos à segunda versão, TRIBUTARISTA*. Aqui, o mesmo advogado tem um plano de fazer com que eu acredite num conjunto de 6 proposições: p (uma proposição verdadeira) e outras cinco (todas falsas). Dessa vez, passo a confiar nele; ele então assere as 6 proposições, e, por meio de sua palavra, acredito nelas. Contudo, o advogado não as afirmaria a menos que p fosse verdadeira. Assim, nos mundos próximos em que p é falsa, o advogado não assere o conjunto das 6 proposições, e, por isso, não acredito que p . A solução desse caso é mais fácil. De fato não sei que p . E a razão é que embora em algum sentido a minha crença nesse

conjunto de proposições rastreia a verdade de p ao longo dos mundos próximos, o testemunho do advogado (embora possa ser concebido de modo a ser confiável)⁷³ não visa a crença verdadeira de todo. Ainda que o testemunho do advogado cumpra o propósito de produzir uma grande quantidade de crenças verdadeiras nos receptores de seu testemunho, se considerarmos o micro-ambiente para o qual esse tipo de testemunho foi concebido, fica claro que visa propositalmente crenças falsas. Portanto, considerando o micro-ambiente no qual recebo o testemunho do advogado, não tenho conhecimento devido à disfuncionalidade desse tipo de testemunho. Uma vez mais, o contraexemplo mostra no máximo que a sensibilidade não é suficiente.

d. Correlação

Passemos a CORRELAÇÃO. Embora engenhoso à primeira vista, esse caso é análogo ao caso original de Gettier (FORD). E como já argumentei no Capítulo I, a atribuição de crenças disjuntivas artificiais é inapropriada de um ponto de vista conversacional. Sendo assim, ainda que a crença disjuntiva do sujeito seja sensitiva, a atribuição de tal crença inapropriada ao sujeito impede que ela conte como conhecimento.

Para fins de argumentação, foquemos a atenção no mecanismo cognitivo (ou no método) usado por S para formar a crença. A maneira mais geral de descrevê-lo é: *uso da regra inferencial da introdução da disjunção*. Ora, inferência é processo cognitivo dependente de crenças e, portanto, condicionalmente confiável. Isto é, tende a produzir mais crenças verdadeiras do que falsas somente se alimentado com crenças verdadeiras. E o exemplo acima é construído de tal modo que o input de S – a crença de que há vida em Marte – é uma crença falsa. Sendo assim, a sua crença final não pode ter adquirido qualquer caráter epistêmico positivo. Se pensarmos em termos do ambiente, ainda que a introdução da disjunção seja confiável no macro-ambiente, no micro-ambiente em que S faz essa inferência particular, ele faz uma inferência que parte de uma premissa falsa.

e. Lesão cerebral

Passemos ao nosso último contraexemplo relativo à permissividade, LESÃO CEREBRAL. Nos mundos próximos em que Cassius não tem a lesão cerebral, ele não forma, via lesão

⁷³ Poderíamos pensar que o advogado decide fazer com que eu acredite em 5 proposições verdadeiras e apenas uma falsa. Nesse caso, esse tipo de testemunho seria confiável, uma vez que tenderia a produzir mais crenças verdadeiras do que falsas.

cerebral, a crença de que tem uma lesão cerebral. Sendo assim, a crença de Cassius no mundo atual de que ele tem a lesão cerebral é sensitiva. Mas intuitivamente não é conhecimento. À partida, o diagnóstico é o de que a crença de Cassius não é conhecimento porque foi formada por um mecanismo disfuncional. Uma lesão, por definição, é um dano; e nesse caso esse novo mecanismo cognitivo não foi moldado (nem evolutiva, nem propositalmente) visando a crença verdadeira, embora seja um mecanismo altamente confiável.

Porém, parece haver casos em que um mecanismo que não vise primariamente a crença verdadeira funcione apropriadamente. Eis um cenário. Os médicos descobrem que a lesão cerebral X, se não tratada, leva o paciente à morte num espaço de dois anos. É extremamente difícil detectar essa lesão; não se descobriu ainda um método para detectá-la. Porém, começaram a notar que todos os acometidos pela lesão, costumavam formar a crença de que tinham uma lesão determinada. Começam então a monitorar vários pacientes com essa crença específica e percebem que também morrem depois de cerca de dois anos. Na autópsia constatou-se que todos tinha a lesão. Os médicos também perceberam que não houve qualquer caso de pessoas com tal crença específica e que não tivessem a lesão, de modo que puderam inferir da correlação forte entre a crença e a lesão, que a lesão era a causa da crença. Sendo assim, lesões cerebrais em geral não visam a crença verdadeira; mas se considerarmos essa situação específica, a crença de S de que ele tem a lesão X é um indicador confiável de que é o caso que S tem a lesão X.

Assim, se LESÃO CEREBRAL é um caso de conhecimento, então não apresenta problemas para a sensitividade. Mas se não, também não apresenta; uma vez que a falha de conhecimento é explicada pelo mecanismo disfuncional, que não precisa estar diretamente ligado à sensitividade.

3.2. Restritividade

a. Conhecimento indutivo

Começemos por CALHA. Como vimos, é extremamente implausível negarmos que Adélia saiba que as botinas de Washington estejam na caçamba de entulho. Mas será que essa crença, como sugerem Sosa, Vogel e muitos outros, é insensitiva? Penso que não.

Vamos pensar um pouco sobre crenças indutivas. Há duas maneiras pelas quais podemos sustentar uma crença indutiva: i) oferecendo explicitamente um argumento

indutivo⁷⁴ cuja conclusão é a crença em questão, e ii) formando a crença de maneira semi-inferencial. Por exemplo, quando, a partir das informações que tenho sobre o volume de chuva no período de novembro em Ouro Preto nos últimos anos, acredito que neste ano o volume será aproximadamente o mesmo, estou oferecendo um argumento indutivo explícito a favor da crença de que *Choverá em Ouro Preto, no mês de novembro deste ano, algo próximo do que choveu nos últimos anos*. Mas quando acredito, depois de ver Marcolino sofrendo uma queda brusca, que ele se machucou, embora eu tenha uma crença indutiva, não ofereci qualquer argumento ou raciocínio a favor dessa crença. Chamo a esse tipo de processo “semi-inferencial” porque ele é algo entre a formação de uma crença perceptual e uma crença inferencial. Ao ver Marcolino caindo, quase que automaticamente formei a crença de que ele se machucou, sem usar quaisquer premissas para chegar a essa conclusão. Contudo, eu não teria formado essa crença se não tivesse um conjunto de crenças de fundo – crenças sobre pessoas se machucarem em quedas bruscas, sobre partes do corpo mais sensíveis, sobre a faixa etária de pessoas que comumente se machucam em quedas, etc. Seja como for, ambos os processos têm a seguinte propriedade: *a verdade das premissas não exclui a falsidade da conclusão*. Isso é o que em parte define um argumento indutivo. Dito no jargão dos mundos possíveis: ainda que no mundo atual um argumento Γ tenha premissas e conclusão verdadeiras, há pelo menos um mundo possível em que Γ tem premissas verdadeiras e conclusão falsa.

Em CALHA, a crença indutiva de Adélia é semi-inferencial; mas para efeitos de simplificação, explicitemos as crenças de fundo mais relevantes a fim de construir uma inferência indutiva. Teríamos algo como o seguinte:

- 1) As calhas usadas em construções para a dispensa de entulho são geralmente boas: sempre que você coloca o entulho, ele cai direto na caçamba.
- 2) Joguei as botinas pela calha.
- 3) Logo, (muito provavelmente) as botinas estão na caçamba.

Nos mundos mais próximos em que as botinas estão na caçamba (assim como no mundo atual) e Adélia raciocina indutivamente, ela acredita que as botinas estão na caçamba. Porém, parece que há pelo menos um mundo próximo em que (1) e (2) são verdadeiras e (3) é falsa.

⁷⁴Considero como indutivo qualquer argumento não-dedutivo. Eis as principais formas de argumentos não-dedutivos: indução por enumeração (generalização), previsão, retrovisão (“previsão” de acontecimentos passados), silogismo estatístico, analogia, e abdução.

Nesse mundo, digamos, como Adélia acredita nas premissas, ela também acredita na conclusão. Ora, isso quer dizer que há pelo menos um mundo possível no qual p é falsa e S acredita que $p - (p \wedge Bp)$. Mas será esse mundo próximo o suficiente para tornar a contrafactual $\neg p \Box \rightarrow \neg Bp$ falsa? Sosa e Vogel pensam que sim. Penso que não.

Primeiro, considerando que não há qualquer problema com o método de formação de crenças e com o macro-ambiente, resta-nos olhar para o micro-ambiente? Haverá algum micro-ambiente desfavorável à crença de Adélia? Isto é, há algum mundo possível em que Adélia forma a sua crença indutivamente, usando aquela mesma calha que funcionava perfeitamente até o momento em que ela jogou as botinas, etc. e ainda assim é falso que as botinas não caem na caçamba? Parece que não. Os mundos mais próximos nos quais as botinas não chegam à caçamba (e ainda assim Adélia acredita que chegaram) são aqueles mundos nos quais as botinas agarram na calha. E nesses mundos a calha certamente não está nas mesmas condições que no mundo atual. Ou seja, nesses mundos Adélia não forma a sua crença a partir aproximadamente das mesmas condições iniciais; essa é uma alteração relevante no ambiente.

Mas de que modo a contrafactual “Se as botinas não tivessem chegado à caçamba, Adélia não teria acreditado que chegaram” é satisfeita? De que modo nos mundos possíveis mais próximos em que as botinas não chegam à caçamba, e Adélia, raciocinando indutivamente, não acredita que a botina está na caçamba?

Considere primeiro um mundo em que aproximadamente as mesmas condições iniciais são mantidas. A calha está funcionando adequadamente; está num ângulo adequado; não está entupida, além de dificilmente entupir; as botinas são pesadas o suficiente, etc. Nesse mesmo mundo, quando jogadas na calha as botinas, por algum motivo, encalham. Dadas essas condições, Adélia não pode evitar de formar a crença de que as botinas atingiram a caçamba. Mas dadas essas mesmas condições seria um milagre que as botinas não atingissem a caçamba! Mundos desse tipo são distantes o suficiente, de modo que não tornam falsa a contrafactual. Mas daí se segue que $\neg p \Box \rightarrow \neg Bp$ é verdadeira? Ora, tome-se uma proposição α qualquer; se não há qualquer situação que torne α falsa, ela só pode ser verdadeira. Ora, se isso vale para qualquer proposição, tem de valer também para a contrafactual em questão. É

claro que excluímos da esfera de mundos relevantes os mundos- $\neg p$; mas isso pelo menos torna a nossa contrafactual vacuamente verdadeira.⁷⁵

Mas poderíamos alterar algumas das condições iniciais (afinal, a ideia é mantê-las aproximadamente as mesmas) e ver no que dá. Suponha, por exemplo, que durante a noite o vento inclinou a calha de modo que a tirou da posição adequada para deslizar o entulho até a caçamba. Adélia chega cedo para trabalhar, e, sem notar isso, joga as botinas pela calha, que acabam por parar no meio do caminho. Essa alteração parece ser razoável, e parece que temos um mundo bastante próximo. Mas considere que seria bastante estranho que se permitisse que calhas numa construção estivessem assim tão facilmente vulneráveis a mudar de posição de modo que prejudicasse o descarte do entulho. Se leves ventos da madrugada fizessem isso com calhas, elas certamente seriam objetos pouco confiáveis e, dificilmente, seriam adotadas em construções. Se, por outro lado, tivesse havido uma tempestade na madrugada com ventos fortes o bastante para deslocar a calha, Adélia seria imprudente em jogar o que quer que fosse pela calha; ela teria uma informação relevante sobre a calha que funcionaria como anulador para a sua crença indutiva. Isso, no mínimo, a faria duvidar que a botina tenha chegando na caçamba. Ao ser perguntada sobre as botinas, ele diria algo como “Veja, as joguei pela calha, mas considerando a tempestade de ontem, é bem provável que elas tenham ficado presas no meio do caminho; vá lá na caçamba e dê uma conferida”. Isso indica que nos mundos em que é falso que as botinas chegam à caçamba e Adélia não acredita que tal são mundos em que alguma condição inicial relevante é alterada. E são exatamente essas alterações que fazem com que Adélia deixe de acreditar que as botinas atingiram a caçamba.

Isso também explica eventuais falhas de conhecimento indutivo mesmo que seja verdade que as botinas atingiram a caçamba. Suponha que Adélia não tenha reparado na tempestade e no desvio do ângulo da calha. Ele chega ao trabalho, encontra as botinas e deita-as fora pela calha. As botinas agarram no meio do caminho, mas por conta de outro vento que balança a calha, elas acabam descendo. Ela acredita que elas atingiram a caçamba, muito embora não saiba. Isso porque sua crença foi verdadeira por sorte; o ambiente deixou de ser favorável para a formação da sua crença. Nos mundos próximos em que não ocorre o segundo vento, as botinas não caem, mas ela ainda acredita que elas atingiram a caçamba.

⁷⁵No caso de $\neg p \Box \rightarrow \neg Bp$ ser vacuamente satisfeita não convencer, vale notar que a condição de aderência $\neg p \Box \rightarrow Bp$ – é satisfeita não-vacuamente: nos mundos próximos em que as botinas chegam à caçamba, Adélia acredita, via indução, que chegaram.

A ideia das alterações das condições iniciais se adequa bem a uma característica crucial do raciocínio indutivo: a *não-monotonicidade*. Dizer que um argumento é não-monotônico significa que a introdução de nova informação às premissas pode alterar a força do argumento. Por exemplo, Gil Grissom encontra a digital de Paul Millander na arma do crime e forma a crença de que é Paul Millander o assassino. Contudo, Grissom vem a saber mais tarde que as digitais de Paul Millander estavam reproduzidas massivamente em mãos de brinquedo (em que Millander usou a própria mão como molde). Essa informação enfraquece a crença de Grissom, uma vez que facilmente poderiam ter plantado aquelas digitais na cena do crime. Assim, ainda que Grissom tenha inicialmente formado a crença de que Millander é o assassino, posteriormente ele deixa de aceitar tal crença. Tudo isso a despeito de as suas premissas serem verdadeiras e também a conclusão. Levando-se em conta, portanto, que naqueles mundos em que a proposição ⟨A botina está na caçamba⟩ é falsa, e levando-se em conta que nesses mundos informações relevantes sobre a calha (ou o ambiente em geral) estariam facilmente disponíveis à Adélia, Adélia não acreditaria (não daria aceitação) que a botina está na caçamba.⁷⁶

Passemos a ASPIRANTE. É muitíssimo improvável que o aspirante tenha acertado o tiro, de modo que o veterano sabe que ele errou. Embora novamente tenhamos um caso de crença semi-inferencial, explicitemos o argumento:

- 1) É muitíssimo improvável (embora não impossível) que alguém inexperiente acerte um alvo pequeno em movimento, numa situação de grande tensão, a uma distância moderadamente longa, etc.
- 2) O aspirante atira na tentativa de acertar a mão do assaltante.
- 3) Logo, (muito provavelmente) o aspirante errou o tiro.

A intuição que nos leva a pensar que a crença de que o aspirante errou o tiro é insensitiva é a de que o mundo possível em que o aspirante acerta o tiro é próximo o suficiente para a avaliação da contrafactual. Afinal, é um mundo em que a bala toma uma trajetória um pouco diferente. Mas isso não é pouco, pois as condições iniciais também teriam de ser alteradas.

⁷⁶ Note que isso é também está de acordo com um plausível enquadramento epistêmico para o raciocínio indutivo (ou não-monotônico): o raciocínio anulável (ou revogável) (cf. POLLOCK 2008; PSILLOS 2009). A ideia é que uma crença B é *prima facie* garantida para S somente se não há anuladores disponíveis para a crença de S. Há dois tipos de anuladores: enfraquecedores [*under-cutting*] e refutantes [*rebutting*]. No primeiro caso, haverá informação disponível que torne improvável ou implausível a crença de S; no segundo, haverá informação que implique a falsidade da crença de S.

Mantendo-se as aproximadamente as mesmas condições iniciais, em que o aspirante é inexperiente e atira como atirou, seria um milagre que o aspirante acertasse. Esse mundo não é próximo o suficiente para a avaliação da contrafactual.

Para tornar claro como a sensibilidade lida com casos de conhecimento indutivo, explicitemos uma condição para o conhecimento inferencial:⁷⁷

S sabe inferencialmente que β com base no conjunto de premissas $\{\alpha_1, \dots, \alpha_n\}$ sse:

- (i) β é verdadeira;
- (ii) as premissas são verdadeiras;
- (iii) $P(\beta \mid \alpha_1, \dots, \alpha_n) > P(\beta)$;
- (iv) S infere competentemente β a partir de $\alpha_1, \dots, \alpha_n$; e
- (v) A crença na verdade das premissas rastreia a verdade da conclusão:
 - a. $\neg\beta \square \rightarrow \neg(\alpha_1, \dots, \alpha_n)$, e
 - b. $\beta \square \rightarrow B(\alpha_1, \dots, \alpha_n)$.

Note que tanto em CALHA quanto em ASPIRANTE as condições (i)-(iv) são cumpridas, mas supostamente (v.a) é violada. Mas isso apenas porque os objetores estavam a considerar ou mundos em que há fatores relevantes que fariam com que o sujeito alterasse a sua crença nas premissas; ou mundos milagrosos, em que as condições iniciais são as mesmas que no mundo atual mas o evento que decorre delas é diferente.

Vale notar que inferências indutivas são inferências de risco. Quanto mais estável é o ambiente em que formamos as crenças nas premissas, menos risco corremos de terminar com uma conclusão falsa. É por isso que, ao contrário da percepção, consideramos ser mais razoável o monitoramento das condições ambientais. O que explica a intuição de os casos acima serem problemáticos para a sensibilidade é associação do alto risco de erro em certos casos de inferência indutiva a casos em que o risco é bem menor. Em certo sentido, CALHA e ASPIRANTE são mais próximos ao conhecimento perceptual do que um caso de previsão do comportamento do mercado financeiro depois de eleições presidenciais, digamos. Neste último caso, há muitas variáveis a serem levadas em conta, há muito o que monitorar, e, por

⁷⁷ Uma condição parecida a essa, porém sem a condição (iii), pode ser encontrada no próprio texto do Nozick (1981, p. 231). Foi Peter Baumann (2012) quem chamou a atenção para o fato de pouquíssimos filósofos terem se atentado a tal condição. O próprio Baumann usará tal condição para defender uma versão nozickiano do fechamento. Abordaremos isso no Capítulo V.

isso, inferências a partir desses dados costumam mais falhar do que acertar. Por outro lado, no caso de uma inferência sobre o entulho chegar à caçamba são muito menores os fatores a serem controlados. Relativo à fabricação e montagem de calhas, um entulho de tamanho adequado agarrar lá dentro é uma possibilidade remota. Os mundos mais próximos em que isso ocorresse seriam mundos em que as pessoas teriam alguma razão para monitorar mais a montagem e fabricação das calhas.

Outra observação importante é a de que quanto mais arriscada for uma inferência indutiva, mais tendemos a não considerar sua conclusão como um caso de conhecimento. Há casos mesmo em que uma massiva quantidade de evidência não é suficiente para dizermos possuir conhecimento indutivo. Por exemplo, a crença dos cientistas de que o universo tem aproximadamente 14 bilhões de anos é bastante razoável; geralmente dizemos que Carl Sagan sabe que o universo tem aproximadamente 14 bilhões de anos. Contudo, quando passamos a considerar o risco dessa conclusão, principalmente dado a história da ciência, começamos a hesitar em dizer que tal crença é conhecimento. Muito embora ainda a consideremos razoável e provavelmente verdadeira.⁷⁸

b. Oscar e Dack

Em DACHSHUND temos um caso bastante semelhante a CELEIRO. Oscar não é capaz de distinguir uma hiena de um cão. Mas é competente em identificar um dachshund (e distingui-lo de um vira-lata). Assim, ao olhar para Dack, o dachshund, ele forma a crença de que há um cão em sua frente. Contudo, há um mundo próximo em que ao invés de um dachshund, há uma hiena em sua frente, e ele ainda assim forma a crença de que há um cão em sua frente. A crença de Oscar de que há um cão em sua frente não é sensitiva. Mas ainda assim é razoável dizermos que Oscar sabe.

⁷⁸ Esse de modo algum é um exemplo imaginário. Karl Popper (1962; 1975) foi um notável defensor da ideia de que não deveríamos esperar que nossas melhores teorias científicas fossem verdadeiras; o máximo que deveríamos racionalmente acreditar é que são mais bem corroboradas do que suas antecessoras, de modo que são mais verossimilhanes. Popper, contudo, rejeitava a indução como método de obtenção de crenças. Ainda assim reconhecia que sustentar uma crença numa hipótese empírica envolve o risco de se estar errado. Essa postura, ainda que parcialmente incompreendida, foi sustentada por renomados cientistas como Carl Sagan (1995) e Peter Medawar (1984). E podemos também considerar aqui filósofos convencidos pela meta-indução pessimista, como Larry Laudan (1981): eles atribuem crença racional aos cientistas ao sustentarem hipóteses que postulam a existência de inobserváveis, muito embora pensem haver boas razões indutivas para não considerá-las conhecimento.

Consideremos a proposição ⟨Há um cão defronte a Oscar⟩ como o conteúdo da crença de Oscar. Digamos que Oscar não seja muito bom em reparar na diferença entre raças de cães; tudo o que ele consegue fazer é notar o tamanho, a cor, etc. de maneira muito geral. Digamos ainda que Oscar nunca tenha visto uma hiena na vida (sequer saiba da existência delas), e vai para um lugar em que é bastante comum a ocorrência de hienas. Ele vê Dack, o dachshund, e acredita verdadeiramente que há um cão em sua frente. Mas facilmente poderia ter visto uma hiena e formado a crença de que havia um cão em sua frente. Nesse caso, parece-me mais razoável dizer que Oscar não sabe que há um cão em sua frente. E a razão é que Oscar está num micro-ambiente desfavorável: isto é, considerando-se aquelas condições, a sua crença não rastreia a verdade da proposição alvo. Porém, considerando-se o macro-ambiente, Oscar é um discriminador razoavelmente competente confiável de cães: em geral, ele acerta mais do que erra quando aponta para animais e diz “Aquele é um cão” ou “Aquele não é um cão”. Assim, a intuição inicial de que Oscar sabe (embora de fato não saiba) é explicada por Oscar ser em geral confiável – quando não há hienas por perto – em reconhecer cães. Uma vez mais, a analogia com detectores nos ajuda. Num ambiente em que ele dificilmente nos dá o resultado errado (ainda que reconheçamos a existência de ambientes em que ele invariavelmente seria infiel) o usamos sem problemas.

O veredito, portanto, dependerá de como vamos considerar o ambiente. Se há um micro-ambiente desfavorável, Oscar não sabe. Mas se não há tal micro-ambiente desfavorável, Oscar sabe. E o fato de Oscar saber que há um cão em sua frente num ambiente favorável não mostra que há conhecimento insensitivo. Muito pelo contrário, mostra que não há naquelas proximidades hienas o suficiente para tornar aquele ambiente propício à formação de crenças falsas ou crenças verdadeiras por sorte.

c. Holograma*

Do modo como apresentado por Pappas e Swain, HOLOGRAMA* é um tanto vago. A efetividade do contraexemplo depende de como o ambiente é concebido. A mera possibilidade de um holograma não é suficiente. Se não há um micro-ambiente desfavorável, um ambiente no qual eu facilmente teria formado a crença de que há um copo em cima da mesa quando na verdade não há – por ser apenas um holograma – HOLOGRAMA* se vale de um conjunto de mundos irrelevantes para a avaliação da contrafactual. Por outro lado, se há um micro-ambiente desfavorável em que há hologramas espalhados para fazer com que as

peessoas acreditem que objetos ali quando não há, então a minha crença é insensitiva e, por conseguinte, não sei que há um copo em cima da mesa.

Novamente, se o exemplo for um caso de conhecimento, então não temos a presença de um micro-ambiente desfavorável. E nesse caso a crença não é insensitiva. Mas nos casos em que a crença é insensitiva, há um micro-ambiente desfavorável, e, por conseguinte, não sei que há um copo em cima da mesa.

d. Margem de erro

RAMONES é um caso interessante. Parece haver muitos casos claros em que sabemos coisas com certa margem de erro. Por exemplo, preparando para uma viagem, minha mulher junta uma quantidade de bagagem e me pergunta se caberá no porta-malas do carro. Por estar acostumado a arrumar as bagagens no porta-malas, sempre que olho para uma determinada quantidade de bagagens, sei se caberão ou não no porta-malas. Eu disse que aquelas bagagens caberiam no porta-malas, e estava certo sobre isso; elas de fato couberam. Contudo, há um mundo possível no qual minha mulher resolve usar uma mala um pouco maior e que junto do restante das bagagens, faz com que elas não caibam no porta-malas. Ainda assim acredito que caberão, uma vez que visualmente elas parecem ter o mesmo tamanho que têm no mundo atual. A minha crença é insensitiva, muito embora seja intuitivo dizer que sei no mundo atual que as bagagens caberão no porta-malas.

É fácil ver que vários contraexemplos podem ser gerados com expressões próximas, como “aproximadamente”, “mais ou menos”, “por volta de”, etc. Vejamos. Joey mede 1,98m e Johnny, via seu “olhomêtro” confiável, acredita que ele tem *aproximadamente* 1,95m. Pensemos agora num mundo possível onde Joey tem 1,94m. Johnny acredita que Joey tem *aproximadamente* 1,95m. Por conseguinte, a contrafactual *Se Joey não tivesse 1,98m, então Johnny não acreditaria que Joey tem aproximadamente 1,95m* é falsa. E, portanto, a crença é insensitiva.

Qualquer concepção razoável de conhecimento tem de deixar espaço para sabermos proposições com alguma margem de erro. Se Williamson estiver correto, e esse tipo de crença for insensitiva, então o conhecimento não pode ser crença sensitiva.

A falha desse tipo de contraexemplo, penso eu, está em não dar atenção para o elemento fortemente contextual presente no uso de expressões como “pelo menos”, “aproximadamente”, etc. Primeiro, por que usar expressões que indicam margem de erro? Resposta óbvia: porque nem sempre (na maioria das vezes!) somos capazes de precisão; tudo o que nos resta é contar que uma pequena margem para mais ou para menos não afete de todo a verdade de uma determinada afirmação. Segundo, qual a margem aceitável? Resposta óbvia: só o contexto é capaz de indicar.⁷⁹

Considere uma garrafa cilíndrica completamente uniforme e transparente cuja capacidade é de um litro. Ela tem exatamente 530 ml de água. Adão olha para a garrafa e forma a crença de que há pelo menos 500 ml na garrafa. Há duas maneiras de entendermos a expressão “pelo menos”. A primeira indica que Adão sabe que há 500 ml embora não saiba se há exatamente 500ml ou um pouco mais. A segunda indica que um pouco menos que 500 ml ou um pouco mais que 500 ml seria um resultado aceitável (o que equivale a “aproximadamente”). Vamos nos focar nessa segunda leitura. Suponhamos que 50 ml para mais ou para menos seja uma margem aceitável. Assim, qualquer frase do tipo “Há aproximadamente 500 ml na garrafa” significará “Há algo entre 450 e 550 ml na garrafa”. E essa margem é devido, digamos, à capacidade Adão de perceber o meio da garrafa, dividi-la intuitivamente em intervalos de 100 ml. Desse modo, podemos restringir o conjunto de mundos próximos em que a garrafa não tem 500 ml de água, mas “aproximadamente 500 ml”, no seguinte intervalo, em que “ w_n ” denota um mundo em que a proposição ⟨Há n ml de água na garrafa⟩: $W: \{w_{450}, \dots, w_{500}, \dots, w_{550}\}$. Além disso, temos de considerar não a contrafactual “Se a garrafa não tivesse 500 ml, então Adão não acreditaria que tem aproximadamente 500 ml”, mas antes “Se a garrafa não tivesse aproximadamente 500 ml, então Adão não acreditaria que tem”. Mas agora precisamos do conjunto de mundos próximos em que é falso que há aproximadamente 500 ml na garrafa. Há dois conjuntos. $W^*: \{w_{1000}, \dots, w_{551}\}$ e $W^{**}: \{w_{449}, \dots, w_0\}$. Assim, a contrafactual em questão é verdadeira sse os mundos W^* e W^{**} são também mundos em que Adão não acredita que a garrafa contém aproximadamente 500 ml de água.

⁷⁹ Como já mencionado algures, não penso que uma teoria do conhecimento baseada na sensibilidade seja incompatível com uma abordagem contextualista. Contudo, o tipo de contextualismo que estou reivindicando para resolver o problema da margem de erro é bastante local. Depende apenas da suposição fraca e incontroversa de que vários termos e expressões da linguagem natural são sensíveis ao contexto.

Uma dificuldade óbvia aqui é a seguinte: tome-se o mundo w_{449} ou o mundo w_{551} ; ambos são mundos em que Adão acredita que há aproximadamente 500 ml na garrafa. É bastante razoável supor que 1 ml seja imperceptível a Adão. Portanto, essa manobra é vã; não evita que a crença de Adão seja insensitiva.

Tal dificuldade se deve à vagueza própria da precisão da visão de Adão para perceber o aumento ou decréscimo no intervalo de poucos ml's. É certo que eu poderia introduzir w_{551} em W , o que parece bastante razoável. Mas então teríamos um problema sorite. Se w_{551} é próximo o bastante para ser um mundo em que há *aproximadamente* 500 ml de água na garrafa, então w_{552} também é, e assim por diante, até w_{1000} . O mesmo raciocínio vale para o intervalo 449-0. Isso nos levaria ao absurdo de que seja como for, Adão acreditaria que há aproximadamente 500 ml na garrafa.

A vagueza é um fenômeno que assombra muitos filósofos. Contudo, parece-me exagerado ter de rejeitar crenças que dependem de margem de erro como insensitivas só porque margens de erro em geral são vagas. Não deveríamos exigir mais precisão de um mecanismo cognitivo (ou instrumento que auxilie a cognição) do que aquela que ele pode nos dar num determinado contexto. Se Adão é bom em perceber diferenças de 100 ml na garrafa e é também bom em encontrar o meio da garrafa, ele certamente hesitaria em dizer que há aproximadamente 500 ml na garrafa ao perceber um intervalo de 100 ml entre a superfície do líquido e o que considera ser o meio da garrafa (quando sabe que sua margem de erro é de apenas 50 ml). E embora pareça arbitrário, podemos deixar uma margem de variação para o limite de W^* e W^{**} com W . E essa variação será aceitável ou não dependo do contexto.

O mesmo se aplica a RAMONES. O contexto é que determina qual a margem de erro seria aceitável quando Johnny diz que Joey tem pelo menos (entenda-se “aproximadamente”) 1,95m. Digamos que Johnny é um agente funerário e está providenciar um esquife pra o corpo de Joey. Ele sabe que em seu estoque há esquifes com comprimento entre 1,90 e 2,00m. Nesse contexto, se Joey não tivesse aproximadamente 1,95 (dentro da variação), Johnny não teria acreditado que tem (teria acreditado que tem menos ou mais).

e. Os celeiros de Kripke

Em CELEIRO-K temos uma situação intrigante. A crença do Sr. K de que há um celeiro vermelho em sua frente é verdadeira, formada por um mecanismo confiável (funcionando

adequadamente) e sensitiva. Nos mundos próximos em que não há um celeiro vermelho, o Sr. K não acredita que há um celeiro vermelho; ele acreditaria que o celeiro fosse azul (ou caso estivesse num mundo em que o lote está vazio, acreditaria que não há celeiro algum). Contudo, a sua crença de que há um celeiro em sua frente, a despeito de ser verdadeira e formada adequadamente, não é sensitiva; portanto, o Sr. K não sabe que há um celeiro em sua frente. Ora, essa é uma situação um tanto estranha: como pode o Sr. K saber que há um celeiro vermelho em sua frente mas não saber que há um celeiro em sua frente? Afinal, a proposição (Há um celeiro vermelho defronte ao Sr. K) implica (Há um celeiro defronte ao Sr. K).⁸⁰

Esse caso não é um contraexemplo real à sensitividade, porém. O truque de Kripke está na suposição de que o Sr. K sabe que *está um celeiro vermelho em sua frente*. Se essa for uma atribuição adequada de crença ao Sr. K, então ele não sabe que há um celeiro vermelho em sua frente. Mas se, contudo, insistirmos que o Sr. K sabe alguma proposição, não poderá ser a de que há um celeiro vermelho.

Primeiro à crença do Sr. K. na proposição (Há um celeiro vermelho defronte ao Sr. K.). Do ponto de vista lógico, a crença do Sr. K. envolve pelo menos duas proposições: (O objeto defronte ao Sr. K é um celeiro) e (O objeto de frente ao Sr. K. é vermelho).⁸¹ É óbvio que ambas as proposições são verdadeiras. E se o Sr. K. sabe ambas, então ele rastreia ambas. Mas ele não rastreia ambas – rastreia apenas (O objeto defronte ao Sr. K. é vermelho). Assim, ele não sabe ambas as proposições; e, por conseguinte, não sabe que *o objeto que está em sua frente é um celeiro e é vermelho*.

Para vermos que o Sr. K. rastreia a verdade da proposição (O objeto defronte ao Sr. K. é vermelho) basta considerarmos o ambiente. A sua crença de que há um objeto vermelho defronte de si não é verdadeira por sorte. Nos mundos mais próximos em que não é o caso que há um objeto vermelho defronte de si – mundos em que ou não há um celeiro ou há um celeiro de fachada azul defronte de si – e ele olha para o lote, o Sr. K. não forma a crença de que há um objeto vermelho defronte de si; ou formaria a crença de que o lote está vazio, ou a

⁸⁰ Esse caso também constitui um contraexemplo ao fechamento.

⁸¹ Podemos, mais precisamente, atribuir como conteúdo da crença do Sr. K a seguinte descrição definida: *Há um e apenas um objeto que está defronte ao Sr. K e esse objeto é um celeiro e é vermelho*.

crença de que há um objeto azul ali. Portanto, a crença na proposição ⟨O objeto defronte ao Sr. K. é vermelho⟩ pode ser seguramente considerada conhecimento.

Para vermos que o Sr. K. não rastreia a verdade de ⟨O objeto defronte ao Sr. K. é um celeiro⟩ basta notar que ele não está num ambiente favorável à formação de crenças sobre a existência ou não de celeiros. Há pelo menos um mundo próximo em que ele olha para o lote defronte a si, vê um celeiro de fachada azul e forma a crença de que há ali um celeiro. Portanto, a crença na proposição ⟨Há um celeiro defronte ao Sr. K.⟩ não pode ser considerada conhecimento.

Se S rastreia uma conjunção, então também rastreia cada uma das conjuntas. Mas o resultado acima mostra o contrário; o Sr. K. não rastreia a verdade de ⟨O objeto defronte ao Sr. K. é um celeiro⟩. Logo, a crença conjuntiva do Sr. K. de que *há defronte a mim um celeiro vermelho* não pode contar como conhecimento. A moral disso é que Kripke não está autorizado a atribuir a crença conjuntiva ao Sr. K.

Mas se seguimos a intuição de Kripke e insistimos que o Sr. K tem conhecimento, o que ele sabe então? Atente-se para o fato de que dissemos que uma das proposições que supostamente compõe a crença do Sr. K é ⟨O objeto defronte ao Sr. K é vermelho⟩. Ora, essa proposição é rastreada pelo Sr. K. Portanto, se atribuirmos a crença de que *aquele objeto, que me parece (ou acredito ser) um celeiro, é vermelho*⁸² ao Sr. K, é bastante razoável dizer que ele sabe essa proposição. Note que já não é mais estranho dizer que o Sr. K sabe que aquilo que ele acredita ser um celeiro é vermelho mas o Sr. K não sabe que é um celeiro.

f. Proposições necessariamente verdadeiras conhecidas *a posteriori*

À partida, o fato de a sensitividade ser trivialmente satisfeita no caso de crenças em proposições necessariamente verdadeiras parece comprometer o conhecimento de proposições como ⟨Machado de Assis é um ser humano⟩.⁸³ Um sujeito epistemicamente parvo, como o Sr.

⁸² Podemos atribuir ao Sr. K um uso referencial da descrição *O objeto defronte ao Sr. K é um celeiro e vermelho*. No caso anterior, ele a usa atributivamente.

⁸³ Alguns poderiam resistir à tese de que *ser humano* é uma propriedade essencial de Machado de Assis. Alguém que leve a sério a ideia de reencarnação poderia pensar que uma vez que Machado de Assis pode reencarnar em um animal de outra espécie, e que provavelmente já tenha se encarnado em muitas outras espécies no passado, não pode ser necessário que Machado de Assis é um ser humano. Se for esse o caso, é melhor considerar a proposição ⟨A água é H₂O⟩, um exemplo clássico de necessário *a posteriori*.

Rice, poderia formar a crença de que Machado de Assis é um ser humano de maneira bastante descuidada e ainda assim rastrear a verdade da proposição.⁸⁴

Mas podemos também considerar que casos como o do Sr. Rice não são contraexemplos efetivos. A menos que a crença do Sr. Rice tenha sido formada por um mecanismo confiável (funcionando apropriadamente no ambiente para o qual foi moldado para funcionar), não é suficiente que a sua crença seja sensitiva. Conhecimento é crença verdadeira formada por um mecanismo confiável que, funcionando adequadamente no ambiente apropriado, rastreia a verdade da proposição.⁸⁵

Além do mais, como o próprio Nozick já notara (1981, pp. 186-187) a condição de aderência é satisfeita – *nos mundos mais próximos em que α é o caso, S forma a crença de que α assim como formou no mundo atual* ($\alpha \square \rightarrow B\alpha$). É claro que nos mundos mais próximos em que o Sr. Rice forma a sua crença de que Machado de Assis é humano de maneira descuidada ele continua acreditando que p . Mas isso não é problema, uma vez que o método de formação de crença continua sendo infiel.

Consideremos agora um caso em que o Sr. Rice forma a crença de que Machado de Assis é humano através de um método confiável – digamos, a sua capacidade de discriminar visualmente animais humanos de não-humanos. Note que, muito embora não haja mundos próximos em que Machado de Assis não é humano, ainda assim o Sr. Rice é bom em detectar humanos: nos mundos próximos em que o Sr. Rice estivesse olhando para um chimpanzé (e não para Machado de Assis) ele acreditaria que vê um chimpanzé e não um humano. Ou seja, ele não formaria a crença de que há um humano em sua frente se lá estivesse um chimpanzé. Em suma: quando a sensitividade é trivialmente satisfeita (o que nos garante pelo menos que não há contraexemplos), a confiabilidade do mecanismo formador crença assume a centralidade na explicação do rastreamento – o mecanismo é confiável para rastrear as propriedades que, nesse caso, são instanciadas necessariamente pelo particular. Uma vez mais, uma analogia com um dispositivo de detecção pode ser esclarecedor. Suponha que um engenheiro desenvolva um radar muito bom em detectar determinado comprimento de onda: na presença da onda ele emite um sinal, e na ausência não emite sinal algum. Quando

⁸⁴ É claro que isso só é uma objeção para aqueles que defendem teses essencialistas – que há particulares que possuem propriedades essenciais não triviais.

⁸⁵ Note que isso é mais do que indexar o método de formação de crença para a avaliação da contrafactual (como faz Nozick); estamos exigindo claramente que o método seja confiável (ou que funcione apropriadamente).

instalado numa região em que há uma fonte de radiação que emite continuamente esse comprimento de onda, o radar passa a emitir o sinal continuamente. Eis o ponto da analogia: o ambiente onde há emissão constante da radiação é como o conjunto de mundos próximos, não há como estar naquele ambiente sem que esteja também a radiação. Os ambientes normais, como o que o engenheiro testou o seu radar, são como o conjunto de mundos sem Machado de Assis – assim como se houvesse um humano o Sr. Rice acreditaria que é um humano, se houver uma fonte da tal radiação nesse ambiente, o radar emitirá o sinal.

Passemos à proposição $\langle A \text{ água é } H_2O \rangle$. De acordo com muitos filósofos, tal proposição é necessariamente verdadeira. Assim, é razoável pensar que os mundos mais próximos em que tal proposição seja falsa (mundos logicamente possíveis apenas) são distantes o suficiente para não contarem na avaliação da contrafactual da variação. Não obstante, a contrafactual “Se a água não fosse H_2O , S não acreditaria que fosse” parece bastante plausível. Como explicar esse impasse? Essa plausibilidade inicial se deve apenas à sua similaridade linguística com outras contrafactuals plausíveis? Um essencialista radical certamente diria que sim. Ele diria, “você tem uma contrafactual trivialmente verdadeira, mas que é similar a muitas outras contrafactuals robustas, como por exemplo, *Se o Brasil não estivesse quase que completamente em cima de uma placa tectônica, então teria atividade sísmica constante*”.

Mas continuo a insistir. Para evitar a penosa história da descoberta da fórmula da água, pensemos num mundo em que os químicos descubrem a fórmula molecular da água usando um espectrômetro. Uma vez que o espectrômetro é um aparelho confiável, não parece razoável dizer que *se a água não fosse H_2O , S não acreditaria que fosse?* É certo que se *aquele composto* testado no espectrômetro não fosse H_2O , S não acreditaria que fosse H_2O . É nesse sentido que o espectrômetro é confiável. E claro, se aquele composto fosse H_2O , S acreditaria que é H_2O . Assim, mesmo que a variação seja trivialmente satisfeita para proposições necessárias *a posteriori*, não podemos abrir mão dela na explicação da confiabilidade em geral do método usado para acreditar na proposição.

Vejamos uma variação da metáfora do alvo de Sosa. Suponha que uma escola primária organize um campeonato de tiro ao alvo, mas não quer que alguns pais se decepcionem com o péssimo desempenho que seus filhos têm demonstrado nos últimos meses. Eles preparam as flechas e os alvos de modo que todos os tiros que fossem desviar dos alvos acabem atraídos

para os alvos. Esse ambiente é um ambiente em que não se erram tiros. Mas daí não se segue que todos os tiros sejam igualmente bons. É possível detectar os alunos que acertariam ainda que o ambiente não fosse viciado. Esses é que dão os melhores tiros. Isso porque eles têm uma disposição interna para acertar o alvo mais do que errar. Os professores tomaram o cuidado de observar a precisão de cada aluno no momento do disparo. Assim, independentemente dos pais ficarem sabendo, os professores puderam escolher os melhores atiradores. O mesmo vale para as proposições necessárias *a posteriori*. Quando os agentes formam a crença, não podem errar; mas podem acertar ou porque acertariam de qualquer forma, ou porque acertaram devido à sua disposição para formar crenças empíricas verdadeiras. Essa disposição é confiável porque em certos ambientes em que certas crenças empíricas são falsas, ele não forma a crença correspondente (variação); e quando certas crenças empíricas são verdadeiras, ele forma a crença correspondente (aderência). Portanto, ainda que a variação seja trivialmente satisfeita, não podemos abrir mão dela na explicação da confiabilidade do processo.⁸⁶

Uma tentativa não muito atraente – e que eu próprio não considero plausível – de se evitar a trivialização da contrafactual da variação seria a de encaixar o segmento do mundo atual que descreve as condições iniciais em que formo a crença de que a água é H₂O num mundo logicamente possível (embora metafisicamente impossível) em que é falso que a água é H₂O. A intuição por trás é que nem sempre a proximidade de um mundo “epistemicamente possível” tem de respeitar a possibilidade metafísica. Dito em outras palavras, um mundo pode ser epistemicamente próximo ainda que seja metafisicamente impossível. Nesse caso, uma vez que é falso que a água é H₂O, e uma vez que uso um espectrômetro funcionando adequadamente, não formo a crença de que a água é H₂O, mas X_YZ. O critério que estamos usando aqui para considerar essa possibilidade é o seguinte: se a negação de uma proposição α não implica uma contradição, então $\neg\alpha$ é possível. Contudo, essa alternativa se mostra ininteligível em última instância. Na medida em que o termo “água” for um designador rígido – e por isso usado para denotar H₂O em todos os mundos possíveis em que houver H₂O – não faz sentido eu selecionar H₂O num mundo qualquer através do termo “água” e dizer que nesse mundo H₂O não é H₂O.

g. Leis da natureza

⁸⁶ No capítulo seguinte retornarei à ideia de que a variação é essencial para a confiabilidade dos processos e métodos cognitivos em geral.

As considerações tecidas acima se aplicam igualmente ao conhecimento das leis da natureza. Um sujeito sabe uma lei da natureza qualquer somente se a sua crença rastreia a verdade dessa proposição no conjunto de mundos próximos e, além disso, essa crença foi formada por um processo cognitivo confiável. Por exemplo, considere uma das leis de Kepler: *os planetas se movem em órbitas elípticas em torno de seu sol*. De acordo com a sensibilidade, se Kepler sabia essa proposição, então nos mundos mais próximos em que os planetas não se movessem em órbitas elípticas, e Kepler tivesse raciocinado do modo como raciocinou, ele não teria acreditado que se movem em elipses. Nesses mundos os dados astronômicos usados por Kepler não seriam os mesmos, de modo que ele teria sido levado a propor outro tipo de explicação.

É preciso salientar que, ao contrário do conhecimento de senso comum, tratar do conhecimento científico é sempre mais difícil. Embora o ceticismo quanto às proposições de senso comum seja raro entre filósofos, o ceticismo quanto a parte do conhecimento científico já não é tão incomum. O antirrealismo científico – a perspectiva de que não é possível saber se as proposições científicas sobre processos e entidades inobserváveis, incluindo as leis da natureza – é um concorrente de peso ao realismo científico – a perspectiva de que o conhecimento sobre os inobserváveis é possível. À partida, não me parece justo fazer pender a sensibilidade em favor do realismo científico, uma vez que o antirrealista poderia ser também partidário da sensibilidade (para explicar o nosso conhecimento daquilo que é observável) e usá-la para explicar por que a postura mais razoável a se tomar frente ao discurso sobre os inobserváveis das teorias científicas é a suspensão do juízo. Para citar um exemplo, consideremos a perspectiva antirrealista de que as leis da natureza têm de ser concebidas como axiomas não-lógicos de um sistema formal. Presumivelmente, haverá mais de um conjunto de axiomas não-lógicos que impliquem o mesmo conjunto de proposições. É claro que, em geral, a preferência é dada por aquele conjunto de axiomas que tornem o sistema mais simples (sem perda de força explicativa). Porém, a menos que se tenha um bom argumento que conecte a verdade à simplicidade teórica, argumenta o antirrealista, não há razão para pensar que se tenha descoberto as leis fundamentais da natureza. Pode ser até o caso de um cientista acreditar verdadeiramente numa proposição L que expressa uma lei, porém, num mundo próximo em que esse mesmo cientista tem à sua disposição os mesmos dados observacionais O , usa os mesmos critérios de simplicidade C , mas contudo é falso que

L seja uma lei da natureza, ainda assim ele acredita que L é uma lei da natureza. Portanto, qualquer crença numa lei da natureza é insensitiva.

É claro que o realista poderia replicar dizendo que se L fosse falsa, o cientista não teria O ; e, portanto, ao usar C , não teria formado a sua crença de que L . Poder-se-ia aqui recorrer ao mote realista de que “o realismo é a única filosofia que não torna o sucesso da ciência um milagre”. O realista poderia insistir que se L não fosse verdadeira, seria um milagre que O tivesse ocorrido.⁸⁷

O debate é, contudo, bem mais complexo do que isso. (Talvez seja o maior de todos os debates em filosofia da ciência). Toda essa argumentação pode ser estendida a todo o tipo de inobserváveis postulados pela ciência, e toda e qualquer alegação de que tais crenças são sensitivas ou não vai muito além do escopo desta tese.⁸⁸ Seja como for, é prudente que nos mantenhamos neutros quanto ao papel da sensitividade no debate realismo-antirrealismo na ciência.

h. Conhecimento de segunda-ordem

A moral do contraexemplo de Vogel pode ser resumida assim: o conhecimento de segunda-ordem (e de ordens superiores) é impossível, uma vez que a maior parte das crenças de segunda ordem (acredito que a minha crença não é falsa) é insensitiva.

Deixe-me, contudo, colocar as coisas de modo diferente. Se você sabe que Ana Maria tem um biquíni novo, e pensa que sabe tal coisa, é bastante natural que você forme a crença de segunda-ordem de que a sua crença de que Ana Maria tem um biquíni novo não é falsa; afinal, seria absurdo alguém dizer “Sei que p mas acredito que minha crença seja falsa”. Mas também seria igualmente razoável que você acreditasse que a sua crença de que Ana Maria tem um biquíni novo é verdadeira. Assim,

- (i) S acredita que sua crença de que p não é falsa $[B\neg(Bp \wedge \neg p)]$

⁸⁷ O realista poderia tentar bloquear o argumento antirrealista mais ou menos como se segue: dependendo do quão fundamente L for, nos mundos em que L é falsa não há condições propícias para o surgimento de vida inteligente, de modo que não existiríamos, e, portanto, não há crença alguma a ser formada sobre L . Ou seja, não há qualquer mundo próximo (uma vez que L é uma necessidade nomológica) em que L seja falsa e o cientista acredite que L . Mas o antirrealista, por sua vez, poderia se valer da noção de mundos epistemicamente possíveis (que nos permite supor a existência de aparatos cognitivos em mundos metafisicamente distantes) para formular seu argumento.

⁸⁸ Veja ROUSH (2005). Ela desenvolve a sua teoria da sensitividade principalmente para explicar a evidência científica.

e

- (ii) S acredita que sua crença de que p é verdadeira $[B(Bp \wedge p)]$

são de algum modo equivalentes.

Ora, podemos então reformular o caso dizendo que *you know that your belief that Ana Maria está de biquíni novo é verdadeira*. Assim, de acordo com a sensibilidade:

$$\text{Se } K(Bp \wedge p) \text{ então } \neg(Bp \wedge p) \Box \rightarrow \neg B(Bp \wedge p)$$

Aplicando o DeMorgan à antecedente da contrafactual, temos

$$(\neg Bp \vee \neg p) \Box \rightarrow \neg B(Bp \wedge p)$$

Temos de considerar três conjuntos distintos de mundos que tornam a antecedente da contrafactual verdadeira:

- A) Mundos- $(\neg Bp \wedge \neg p)$: são aqueles mundos nos quais você não acredita que Ana Maria tem um biquíni novo e é também falso que ela esteja de biquíni novo.
- B) Mundos- $(\neg Bp \wedge p)$: são aqueles mundos nos quais você não acredita que Ana Maria esteja de biquíni novo mas é verdade que ela está de biquíni novo.
- C) Mundos- $(Bp \wedge \neg p)$: são aqueles mundos nos quais você acredita que Ana Maria está de biquíni novo mas ela não está.

Os mundos contidos em A-C são todos mundos em que você não sabe que Ana Maria está de biquíni novo, uma vez que $(\neg Bp \vee \neg p)$ implica $\neg Kp$.⁸⁹ Os mundos A e B são todos mundos- $\neg(Bp \wedge p)$, *i.e.*, mundos em que você não forma a crença de que você acredita verdadeiramente que Ana Maria está de biquíni novo. Por exemplo, naqueles mundos nos quais é falso que Ana Maria está de biquíni novo e você não acredita que ela está de biquíni novo (via testemunho de Maria), porque Maria não diz que está de biquíni novo, você também não forma a crença de que tem uma crença verdadeira de que Ana Maria está de biquíni novo. E do mesmo modo, naqueles mundos em que é o caso que Ana Maria está de biquíni novo e você não acredita que ela esteja (via testemunho), você não forma a crença de que tem uma crença verdadeira de que Ana Maria está de biquíni novo.

⁸⁹ Suponha que $Kp \rightarrow (Bp \wedge p \wedge Jp)$. Isso implica que $Kp \rightarrow (Bp \wedge p)$. Por contraposição, $\neg(Bp \wedge p) \rightarrow \neg Kp$. Ora, $\neg(Bp \wedge p) \equiv (\neg Bp \vee \neg p)$; por conseguinte, $(\neg Bp \vee \neg p) \rightarrow \neg Kp$.

Mas agora considere os mundos-C. Esses são mundos nos quais é falso que Ana Maria está de biquíni novo mas você acredita que ela esteja. Há vários mundos desses em que você acredita que tem uma crença verdadeira de que Ana Maria está de biquíni novo. Assim, temos $\neg(Bp \wedge p) \wedge B(Bp \wedge p)$ – isto é, a contrafactual é refutada e, portanto, a sua crença de segunda-ordem não rastreia a verdade. É esse o conjunto de mundos que Vogel aponta como próximo o bastante. Mas será próximo o bastante?

Lembre-se que uma vez que você acredita falsamente que Ana Maria está de biquíni novo, e que a sua crença é formada via testemunho, os mundos-C relevantes serão mundos onde Ana Maria se engana sobre o próprio biquíni, ou mundos onde ela mente para você. Esses são de fato mundos próximos o suficiente, de modo que Vogel consegue estabelecer seu ponto. E, portanto, a sensibilidade exclui o conhecimento de segunda-ordem.

Quais são exatamente os mecanismos de formação de crenças de segunda-ordem? Serão eles realmente confiáveis? Para formamos primariamente uma crença sobre outra crença que temos já de antemão (ou mesmo sobre a ausência de alguma crença), temos invariavelmente de nos voltarmos para a nossa própria mente; isso envolve a introspecção.⁹⁰ E a introspecção é admitidamente uma fonte confiável de conhecimento e justificação. Crenças sobre nossos estados conscientes em geral, sobre a fenomenologia de certas experiências (*e.g.* a minha dor é aguda; este som é horripilante, etc.), sobre dados dos sentidos ou perceptos (tenho a experiência de algo me sendo apresentado de tal e tal modo), etc. são todas crenças formadas via introspecção e exemplos paradigmáticos de conhecimento ou crença justificada. Mas será isso suficiente para estabelecer que a crença de que acredito verdadeiramente (ou a crença de que sei) que p , quando verdadeira, é conhecimento?

A resposta é não. A introspecção me permite saber que acredito que p , mas não que a minha crença de que p é verdadeira. Se assim fosse, para qualquer dúvida que eu tivesse acerca de alguma proposição p , bastaria que eu formasse uma crença de segunda-ordem para decidir se p é verdadeira ou não. Muito pelo contrário; geralmente, se tenho dúvida quanto a

⁹⁰ Eu poderia tornar-me ciente de que acredito que o Caboclo d'Água assassinou vários pescadores nas margens dos rios e lagoas do interior de Minas Gerais através do testemunho de um terceiro. Alguém poderia me dizer “Você acabou de deixar claro que acredita que o Caboclo...”. Ou talvez eu tivesse me esquecido do Caboclo d'Água, e, lendo um diário de expedição, encontro vários relatos meus sobre ele e torno-me ciente de que acredito que o Caboclo mata pescadores. Contudo, a fonte primária das minhas crenças de segunda-ordem é sem dúvida a introspecção.

verdade de p , formo uma crença de segunda-ordem de que a minha crença de que p é duvidosa, não de que é uma crença verdadeira (ou uma crença falsa).

A introspecção é um mecanismo confiável para o conteúdo interno de nossas crenças, mas não para o conteúdo externo. Considere a minha experiência perceptual de algo me ser apresentado de tal e tal modo. Digamos que essa experiência faça com que eu forme a crença de que meu gato está agora deitado no sofá. Ao voltar minha atenção para a minha mente – *i.e.*, ao introspectar-me – torno-me ciente de que acredito que meu gato está no sofá. Ou seja, torno-me ciente de que estou fazendo uma representação do mundo externo. Mas uma representação do mundo externo não é o mundo externo, de modo que o conteúdo da minha crença de segunda-ordem é um objeto interno.⁹¹ E dificilmente eu erraria que acredito que o meu gato está no sofá. Mas acreditar que acredito que meu gato está no sofá é compatível com meu gato não estar no sofá – no caso, por exemplo, em que eu estivesse tendo uma alucinação com meu gato ou estivesse a confundi-lo com um bicho de pelúcia que está em cima do sofá. E ainda que nesses casos a minha crença de primeira-ordem – uma crença sobre algo externo – seja falsa, a minha crença de segunda-ordem – uma crença sobre algo interno – é verdadeira. Sendo assim, não importa o valor de verdade das minhas crenças de primeira-ordem (e nem se são ou não conhecimento), a introspecção gerará na esmagadora maioria das vezes uma crença de segunda-ordem verdadeira: a crença de que acredito que p – $B(Bp)$. E essa crença é sensitiva: nos mundos em que não acredito que p (mundos- $\neg Bp$), não acredito, via introspecção, que acredito que p . Isto é, os mundos- $\neg Bp$ são também mundos- $\neg B(Bp)$.

O mesmo não ocorre com a minha crença de segunda-ordem de que acredito verdadeiramente que p . Não somos capazes de discriminar introspectivamente se uma crença é verdadeira ou falsa⁹² – talvez a exceção seja a crença no *cogito*. $B(Bp \wedge p)$ é verdadeira somente se a minha crença de que p for verdadeira. Ora, via de regra, quando um sujeito sustenta uma crença falsa, ele age como se aquela crença fosse verdadeira – se ele pensasse

⁹¹ Ao dizer isso, não estou a me comprometer com a tese de que o conteúdo da percepção não é se não outra coisa que um objeto interno (ou mental), algo como um dado do sentido. Por outro lado, não me sinto inclinado a dizer que o conteúdo da percepção seja inteiramente externo. Seja como for, não faz sentido dizer que uma representação mental seja algo externo.

⁹² Imagine que o Gênio Maligno crie uma Terra Gêmea. Por puro deleite ele passa o dia a te transportar da Terra para a Terra Gêmea e vice-versa. Você é incapaz de dizer, do ponto de vista de primeira pessoa, se está na Terra ou na Terra Gêmea. Na verdade, quando você está na Terra Gêmea, você acredita que está na Terra. Começa a chover na Terra Gêmea e você acredita que está caindo água, quando na verdade cai um líquido com a fórmula molecular X_yZ . A sua crença é, portanto, falsa. Mas você acredita que tem uma crença verdadeira. Se a introspecção fosse confiável para você saber que tem uma crença verdadeira, você teria sido capaz de dizer que a sua crença de que cai água (H_2O) do céu é falsa.

que a crença é falsa, não a sustentaria. Sendo assim, nos caso em que me volto para aquelas crenças que sustento e formo crenças sobre elas, invariavelmente acredito que elas são verdadeiras. Portanto, para a grande maioria das crenças que sustento – *i.e.*, crenças de primeira-ordem –, (e para as quais não disponho de qualquer evidência anuladora), a introspecção gerará a crença – de segunda-ordem – de que tais crenças de primeira-ordem são verdadeiras.

Portanto, embora a introspecção seja confiável para produzir crenças sobre as crenças que tenho – $B(Bp)$ –, não é confiável para produzir crenças sobre a verdade das crenças que tenho – $B(Bp \wedge p)$.⁹³

Mas disso não se segue de modo algum que a sensibilidade torne impossível o conhecimento de segunda-ordem. No máximo, segue-se que não é possível saber que $Bp \wedge p$ introspectivamente. Ademais, do ponto de vista externista – a perspectiva pressuposta nesta tese – não é surpresa alguma que o conhecimento de ordem superior falhe em alguns casos. Afinal, o conhecimento não é um estado mental transparente ou luminoso. Isto é, não é o caso que se S está num estado mental K de conhecimento, S sabe que está em K .⁹⁴

Não estou completamente seguro dos desdobramentos dessa conclusão; por outro lado, também não me parece de todo implausível que possamos viver sem saber que parte substancial de nossas crenças é verdadeira. É consenso, afinal, entre os externistas, que a falha do conhecimento de segunda-ordem não implica por si só a falha do conhecimento de primeira-ordem. Nesse caso, a crença razoável (racional ou justificada) de que *a minha crença de que p é verdadeira* ou de que *sei que p* , parece perfeitamente aceitável.⁹⁵

⁹³ Dito de outra maneira, a crença $B(Bp \wedge p)$ não é uma indicação confiável de que $Bp \wedge p$. Ou ainda, $Bp \wedge p$ não é uma boa evidência para $B(Bp \wedge p)$.

⁹⁴ Veja WILLIAMSON (2000). Williamson constrói um engenhoso argumento contra a luminosidade. Seu argumento, contudo, baseia-se na condição de segurança. Penso que ele possa ser também formulado em termos de sensibilidade. A apresentação desse argumento, contudo, ocuparia mais espaço do que disponho. Seja como for, o resultado acima é suficiente para estabelecer que o conhecimento não é luminoso.

⁹⁵ Pensar que o estado mental de conhecimento tem alguma marca interna distinguível introspectivamente (*e.g.*, a clareza e distinção do cogito cartesiano) levou muitos filósofos às profundezas do ceticismo. Reconhecer que minhas crenças perceptuais, por exemplo, são subdeterminadas, torna-me ciente de que não sei dizer se meu presente estado mental E é uma percepção verídica ou uma alucinação. O que implica que não sei se minha crença é verdadeira ou não. Contudo, se ter conhecimento for apenas ter crença verdadeira formada de maneira apropriada (e que é sensitiva, por conseguinte) o fato de sermos ignorantes em segunda-ordem não impede que ainda tenhamos conhecimento.

Antes de deixar esta seção, deixe-me esboçar um diagnóstico de por que parece óbvio que $K(Bp \wedge p)$ supostamente emerge, via introspecção, de $Bp \wedge p$. Primeiro, façamos uma distinção entre transparência ou luminosidade fenomenológica e transparência ou luminosidade metafísica:

[LM] Um estado mental M é *metafisicamente luminoso* sse se S está em M , então S está em posição de saber que está em M .

Note que o que estou a chamar de *luminosidade metafísica* é nada mais do que a tese comum de que os estados mentais são luminosos ou transparentes. Por outro lado,

[LF] Um estado mental M é *fenomenologicamente luminoso* sse se S está em M , então S está em posição de saber que parece que ele está M .

LM implica LF, mas não o contrário.⁹⁶ Por exemplo, estou supostamente vendo um computador diante de mim; estar nesse estado faz com que me pareça que eu esteja tendo uma percepção verídica do computador. Contudo, se considerarmos que estados alucinatorios e estados perceptuais são fenomenologicamente indistinguíveis – *i.e.*, indistinguíveis para S no momento em que S os experiencia –, estivesse eu tendo uma alucinação de um computador diante de mim, também pareceria a mim estar num estado de percepção verídica. Por outro lado, há casos especiais em que LF implica LM. Casos em que a distinção aparência-realidade plausivelmente não se aplica, como as dores. Se estou num estado mental de dor, estou em posição de saber que parece que está doendo. Mas se parece que está doendo, está doendo.

Como vimos, *acreditar que p* é um estado mental luminoso. Se acredito que p , então estou em posição de saber que acredito que p . E quando formo introspectivamente a minha crença de que acredito que p , essa crença conta como conhecimento, uma vez que é sensitiva: (i) se eu não acreditasse que p , eu não acreditaria que acredito que p , e (ii) se eu acreditasse que p , acreditaria (via introspecção) que acredito que p . O contraexemplo de Vogel deixa claro que a sensitividade implica que o conhecimento não é luminoso.

Mas então por que pensamos que seja? Por que nos parece tão claro que podemos, por introspecção, saber que sabemos ou que acreditamos verdadeiramente numa proposição? A

⁹⁶ A tese de que nossos estados perceptuais são metafisicamente transparentes é obrigatória apenas aos realistas diretos (ou ingênuos), que sustentam que o conteúdo dos nossos estados perceptuais é o próprio objeto fora de nossa mente.

minha sugestão é que a transparência fenomenológica nos engana. Assim como do ponto de vista de primeira pessoa os estados mentais de *ver um elefante* e *ter a alucinação de um elefante* são qualitativamente iguais, os estados mentais de *acreditar que p* e *acreditar verdadeiramente que p* são qualitativamente iguais. E já que sei que acredito que p , sou levado a também dizer que sei que acredito verdadeiramente que p . Acreditar, via introspecção, que acredito que p é como acreditar que estou com dor de coluna. Mas acreditar que acredito verdadeiramente que p não. Por exemplo, alguém que tem uma coceira num membro fantasma (na perna esquerda fantasma, digamos), sabe que tem uma coceira, muito embora não esteja autorizado a dizer que tem uma perna. Do mesmo modo, alguém que acredita que p , sabe que acredita que p , mas não está autorizado a dizer que sabe que acredita verdadeiramente que p .

3.3. Permissividade e restritividade

A sensitividade parece ter um problema duplo com o conhecimento de conjunções: em alguns casos a sensitividade permite o conhecimento de conjunções que não sabemos, e em outros exclui casos óbvios de conhecimento de conjunções.

Comecemos pela conjunção ⟨Louis está bebendo café gelado e Cleópatra não tinha um nariz pequeno⟩, $p \wedge \neg q$. Se S sabe que $p \wedge \neg q$, então $\neg(p \wedge \neg q) \Box \rightarrow \neg B(p \wedge \neg q)$. Considerando a resolução fraca da contrafactual:

$(\neg p \vee q) \Box \rightarrow \neg B(p \wedge \neg q)$ sse ou $[\neg p \Box \rightarrow \neg B(p \wedge \neg q)]$ ou $[q \Box \rightarrow \neg B(p \wedge \neg q)]$, contanto que os mundos- $\neg p$ sejam suficientemente distantes dos mundos- q .

Uma vez que os mundos- $\neg p$ são suficientemente distantes dos mundos- q , pensa Vogel, basta que $[\neg p \Box \rightarrow \neg B(p \wedge \neg q)]$ seja verdadeira para que S saiba que $p \wedge \neg q$. Como naqueles mundos próximos nos quais Louis não está bebendo café gelado, S não acredita, via percepção, que Louis esteja bebendo café gelado, S sabe que $p \wedge \neg q$.

Note que Vogel diz que “posso rastrear $[p \wedge \neg q]$ simplesmente por rastrear p ” (p. 128). Lembre-se que o rastreamento é feito através do processo de formação de crença, isto é, é o processo cognitivo do sujeito que rastreia a verdade da proposição (sendo a crença a indicação confiável da verdade da proposição). E há nesse caso dois processos envolvidos: o processo visual pelo qual S forma a crença de que p e o processo inferencial pelo qual forma a crença

na conjunção de que $p \wedge \neg q$. (Ainda poderíamos introduzir o processo que produziu a crença de que q , mas não é relevante). Como a crença alvo de conhecimento é a crença de que $p \wedge \neg q$, o processo relevante para a avaliação da contrafactual principal será a inferência de p e $\neg q$ para $p \wedge \neg q$. E como sabemos, num processo condicionalmente confiável, o output só será garantido se as crenças de input o forem. Assim, S sabe que $p \wedge \neg q$ se, e somente se, S sabe que p e sabe que $\neg q$.

O exemplo de Vogel não menciona como S forma a crença de que $\neg q$. Não é razoável supormos que S rastreia $p \wedge \neg q$ sem sequer acreditar que $\neg q$. Isso nos levaria ao resultado absurdo de que para quaisquer α e β em que os mundos- $\neg\beta$ são suficientemente distantes dos mundos- $\neg\alpha$, S rastrearía uma conjunção infinita ($\alpha \wedge \beta_1 \wedge \dots \wedge \beta_n \wedge \beta_{n+1}$) só por rastrear α . Temos por isso de supor que no exemplo S forma a crença de que $\neg q$. E mais, temos de supor que S forma a crença conjuntiva, via inferência, de que $p \wedge \neg q$. Podemos, para tornar o exemplo mais forte, supor que S introduz $\neg q$ aleatoriamente. Assim, nos mundos- $\neg p$ mais próximos S não acredita que p e, por conseguinte, não forma a crença de que $p \wedge \neg q$. Dada a resolução fraca, S sabe que $p \wedge \neg q$ sem saber que $\neg q$. Mas se assim for, Vogel estaria violando uma condição necessária para o conhecimento inferencial: a condição de que S tem de saber as premissas.

Mas concedendo a Vogel o seu caso, por que supor que a resolução fraca é que tenha de ser empregada? É claro que num sentido os mundos nos quais Louis não está tomando café gelado são mais próximos ao mundo atual do que os mundos nos quais o nariz de Cleópatra é diferente (e o Império Romano não existe!). Contudo, do ponto de vista epistêmico, essa distância é irrelevante. O que conta para a nossa avaliação é se nos mundos relevantes o ambiente, os mecanismos cognitivos, etc. são aproximadamente os mesmos. A distância histórica entre o mundo atual e os mundos em que o Império Romano não existiu não afeta a avaliação de se S formaria ou não a crença alvo. Dito em outras palavras, embora os mundos onde o Império Romano não existiu sejam historicamente distantes (ou dissimilares), não são epistemicamente distantes. Portanto, o contraexemplo de Vogel não corre.

Passemos à conjunção (Milton Nascimento está cantando e não é um impostor que se parece com Milton Nascimento que está cantando), $m \wedge \neg i$. Vogel pensa que $m \wedge \neg i$ “é logicamente equivalente” a m . E uma vez que sei se que α e sei que $\alpha \equiv \beta$, então sei que β ; de modo que se sei que m , também sei que $m \wedge \neg i$. Mas dada a resolução forte

$$[(\neg m \vee i) \Box \rightarrow \neg B(m \wedge \neg i)] \text{ sse } [\neg m \Box \rightarrow \neg B(m \wedge \neg i)] \text{ e } [i \Box \rightarrow \neg B(m \wedge \neg i)]$$

não posso saber que $m \wedge \neg i$, uma vez que há mundos próximos em que há um impostor se passando pelo Milton Nascimento e S acredita que é o Milton cantando e não um impostor.

Apesar de Vogel, os defensores mais convictos da sensitividade não considerariam esse um problema grave, se é que considerariam um problema. A razão é que as análises rastreadoras mais influentes – *i.e.* a de Dretske (1970, 1971) e a de Nozick (1981) – têm como resultado a falha de rastreamento de proposições como ⟨Não sou um cérebro numa cuba sendo eletroquimicamente estimulado para acreditar que tem mãos⟩, ⟨O animal na jaula não é uma mula pintada para parecer uma zebra⟩, etc. dadas certas evidências. Assim, se dada certa evidência visual e auditiva não posso saber que a pessoa cantando não é um impostor se passando por Milton Nascimento, $\neg i$, também não posso saber qualquer proposição conjuntiva em que $\neg i$ é uma das conjuntas, e, por conseguinte, não posso saber que $\neg i \wedge m$.

Mas o que parece tornar poderoso o exemplo de Vogel é que ele diz serem m e $(m \wedge \neg i)$ logicamente equivalentes. Contudo, o que ele chama de “logicamente equivalente” não é exatamente o que chamamos de “logicamente equivalente” em lógica formal. Isso fica claro, do ponto de vista formal, quando notamos a obviedade de que α não implica $\alpha \wedge \beta$. A não ser que $\alpha \rightarrow \beta$. Porém, isso não pode tornar α e $\alpha \wedge \beta$ logicamente equivalentes, uma vez que nos levaria ao absurdo de produzir uma “equivalência lógica” a cada *modus ponens* que usássemos.

Note, porém, que no caso de m e $m \wedge \neg i$, a implicação $m \rightarrow \neg i$ é necessariamente verdadeira. Isto é, não há qualquer mundo possível no qual Milton Nascimento esteja cantando e ao mesmo tempo seja uma pessoa numericamente distinta dele, porém qualitativamente indistinguível, que esteja cantando em seu lugar. Isso faz com que não haja qualquer circunstância possível em que m seja verdadeira e $\neg i$ falsa. Mas disso não se segue que sejam logicamente equivalentes, do mesmo modo que os pares ⟨João é casado⟩ e ⟨João é não solteiro⟩, e ⟨Este objeto é completamente verde⟩ e ⟨Este objeto não é completamente vermelho⟩ não são logicamente equivalentes. Mas como o primeiro par, m e $m \wedge \neg i$ não são analiticamente equivalentes – *i.e.* não dependem da verdade analítica ⟨Todo solteiro é não-casado⟩. De modo que só pode ser como o segundo par. Esse par depende da verdade conceitual ⟨Nenhum objeto pode ao mesmo tempo ter completamente mais de uma cor⟩. E de

fato parece ser isso que acontece com m e $m \wedge \neg i$. A proposição $m \rightarrow \neg i$ é uma instância da proposição (necessariamente verdadeira) mais geral (Se um indivíduo a tem a propriedade P no tempo t e porção espacial s , então um indivíduo b , diferente de a , não tem P em s e t).

Segue-se disso que Vogel não tem razão ao pensar que o caso em questão envolva uma equivalência lógica. Portanto, ainda que valha o princípio *Se S rastreia α , $\alpha \equiv \beta$, S rastreia β* , não é esse o caso de Vogel.

Ademais, como mencionado anteriormente, temos de levar em conta como o sujeito forma a sua crença de que $m \wedge \neg i$. Novamente, o método é a inferência. Assim, se S sabe que $m \wedge \neg i$, então ele sabe que m e sabe que $\neg i$. Mas ele não sabe que $\neg i$, uma vez que naqueles mundos próximos em que há um impostor indistinguível de Milton Nascimento, ele acredita que é o Milton Nascimento quem está cantando.

É claro que um proponente do argumento de Vogel poderia insistir que S sabe que $\neg i$ porque inferiu $\neg i$ a partir das proposições conhecidas m e $m \rightarrow \neg i$. Todavia, isso dependerá do proponente da sensibilidade aceitar o princípio de fechamento epistêmico: *Se S sabe que α e sabe que $\alpha \rightarrow \beta$, então S sabe que β* . Mas é exatamente esse princípio que Dretske, Nozick, dentre outros, pensam não ser válido. Vogel, por isso, comete uma petição de princípio contra o proponente da sensibilidade, uma vez que supõe o fechamento.⁹⁷

Vimos que parte das objeções à sensibilidade, ainda que bem sucedidas, mostram no máximo que a sensibilidade não é suficiente para o conhecimento. Contudo, uma vez que entendamos a sensibilidade como inerente à confiabilidade, ou ao funcionamento apropriado dos mecanismos cognitivos do agente, a sensibilidade deixa de ser implausível frente às objeções. A minha resposta às objeções mais promissoras, ambas de Vogel, talvez sejam um pouco desapontadoras. Primeiro aceitei de bom grado uma teoria do conhecimento sem conhecimento de segunda ordem. E depois que podemos não saber proposições implicadas

⁹⁷ A discussão sobre o fechamento e a sensibilidade será adiada, como já dito antes, até o Capítulo V. Ficará claro lá a inefetividade desse tipo de argumento contra a sensibilidade.

por outras proposições que sabemos. No último capítulo tentarei tornar tais posições mais plausíveis, ou pelo menos aceitáveis. Mas primeiro, no capítulo seguinte, tratemos de discutir uma condição rival à sensibilidade, a segurança.

IV

MELHOR SEGURO DO QUE SENSITIVO?

Pelo menos na epistemologia, é melhor ser seguro do que sensitivo.
John Greco, *Better safe than sensitive.*

Prévia

Parece-me razoável aceitar que a sensitividade é um componente indispensável do conhecimento. Noções como a de confiabilidade e função apropriada são entendidas adequadamente apenas à luz da sensitividade. Ademais, quando incorporada às exigências da epistemologia da função apropriada, a sensitividade se mostra uma promissora condição anti-sorte. No entanto, alguns epistemólogos insatisfeitos com várias dificuldades enfrentadas pela sensitividade, preferiram adotar um princípio modal alternativo, o princípio de segurança. Grosso modo, uma crença é segura se, e somente se, fosse o caso que α , S acreditaria que α . Os proponentes da segurança pensam que a condição de segurança tem as seguintes vantagens sobre a sensitividade: não excluem casos de conhecimento indutivo, não violam o princípio de fechamento epistêmico, e é o princípio correto para um argumento mooriano anti-cético. Como vimos no capítulo anterior, a sensitividade dá conta do conhecimento indutivo. Quanto ao ceticismo e fechamento, deixarei tal discussão para o próximo capítulo. Meu objetivo neste capítulo é atacar diretamente o princípio de segurança. Tentarei mostrar que (i) a segurança não pode captar adequadamente as nossas intuições anti-sorte (ou anti-Gettier), e (ii) nem mesmo pode captar nossas intuições acerca da confiabilidade.⁹⁸

1. O princípio de segurança

⁹⁸ Não tratarei aqui dos vários contraexemplos, já famosos na discussão, à necessidade ou suficiência da condição de segurança ao conhecimento (cf. COMESAÑA 2005; KELP 2009; PRITCHARD 2009; NETA & ROHRBAUGH 2004; BORGADUS 2012; BROCANO-BERROCAL 2014; BORGADUS & MARXEN 2014). Ainda que seja possível contornar tais contraexemplos, como penso que muitos o foram de fato, seja por atacar a própria legitimidade do caso apresentado ou por alguma adequação ou reformulação da condição de segurança, meu argumento é independente desse resultado. Para fins de argumentação, podemos mesmo assumir que tais contraexemplos falham.

O princípio de segurança é supostamente o princípio modal rival à sensibilidade. Como primeira aproximação, uma crença é segura se, e somente se, fosse o caso que S acreditasse que α , α seria o caso.

Segurança: nos mundos possíveis mais próximos em que S acredita que α , é o caso que α . [$B\alpha \square \rightarrow \alpha$]

Ernest Sosa (1999) foi um dos principais epistemólogos a colocar a condição de segurança em circulação no mercado filosófico. Ele começa por apontar duas dificuldades supostamente fatais à sensibilidade: o conhecimento indutivo e a rejeição do fechamento epistêmico (como consequência de uma resposta ao ceticismo). Em seguida, ele introduz o princípio de segurança e argumenta que ele dá o resultado correto nos casos de conhecimento indutivo e que produz um argumento anticético sem violar o fechamento.

Timothy Williamson (2000, 2009), Duncan Pritchard (2005; 2007; 2009; 2012) e John Greco (2012) foram outros a defender a segurança contra a sensibilidade. Williamson, embora não pretenda oferecer uma análise reducionista do conhecimento, pensa que a segurança torna claras várias características estruturais do conhecimento – *e.g.* a anti-luminosidade e a factividade. Pritchard, por sua vez, sustenta que a segurança é a condição anti-sorte mais adequada ao conhecimento. Já Greco pensa que a segurança capta melhor as nossas intuições acerca da confiabilidade do que a sensibilidade.

Neste capítulo, ocupar-me-ei principalmente em rebater os argumentos de Pritchard e Greco. Contra Pritchard, argumentarei que não é caso que “a condição de segurança tenha considerável mérito sobre a sensibilidade no que diz respeito a fornecer a melhor interpretação da condição anti-sorte” (p.176). E contra Greco argumentarei que a sensibilidade capta melhor as nossas intuições de confiabilidade.

2. Segurança e a intuição anti-sorte

2.1. O dilema de Greco

Vamos começar com uma definição mais elaborada da condição de segurança oferecida por Pritchard (2007, p. 238):

(SEG) A crença de S é segura sse em aproximadamente todos (se não todos) os mundos possíveis mais próximos em que S continua a formar a sua crença

sobre a proposição alvo do mesmo modo que formou no mundo atual, a sua crença continua sendo verdadeira.

A fim de entender como a segurança supostamente se sai melhor com a intuição anti-sorte, consideramos a resposta de Pritchard a um dilema posto por Greco (2007), a que podemos chamar “o dilema de Greco”. Vamos primeiro então ao dilema de Greco.

Considere os seguintes casos:

[LOTERIA] S compra um bilhete de uma loteria justa. Ao ser informado que suas chances de ganhar são de uma em um milhão, S forma a crença de que seu bilhete não será o sorteado.

[ENTULHO] S joga um par de botinas velhas pela calha de entulho. Depois de alguns segundos S forma a crença de que a botina está na caçamba de entulho no térreo.

Embora estejamos dispostos a atribuir conhecimento a S em ENTULHO, não estamos dispostos a atribuir conhecimento a S em LOTERIA. E supostamente SEG dá o veredito correto em ambos os casos. Nos mundos mais próximos em que S joga a botina pela calha e forma a crença (do mesmo modo como formou no mundo atual) que a botina está na caçamba, é o caso que a botina está na caçamba. Os mundos em que S joga a botina pela calha, acredita que ela atingiu a caçamba, mas a botina agarra na calha – e, por conseguinte, S forma uma crença falsa – são mundos distantes o suficiente para não contarem na avaliação da contrafactual. São mundos, por exemplo, em que pelo menos uma das condições iniciais do mundo atual não são mantidas. Quanto à crença de S de que ele vai perder na loteria, ela não é segura porque há pelo menos um mundo próximo no qual o bilhete de S é sorteado, mas ainda assim ele acredita (por ser informado de sua irrisória chance) que não será sorteado. Como o sorteio é aleatório, mantidas as mesmas condições iniciais do mundo atual, há pelo menos um mundo possível próximo no qual o bilhete de S é sorteado.

Contudo, o “aproximadamente todos (se não todos) os mundos possíveis mais próximos” nos permite fazer duas leituras de SEG, uma forte e a outra fraca:

(SEG-fraca) A crença de S é segura sse em *aproximadamente todos* os mundos possíveis mais próximos em que S continua a formar a sua crença sobre a

proposição alvo do mesmo modo que formou no mundo atual, a sua crença continua sendo verdadeira.

(SEG-forte) A crença de S é segura sse em *todos* os mundos possíveis mais próximos em que S continua a formar a sua crença sobre a proposição alvo do mesmo modo que formou no mundo atual, a sua crença continua sendo verdadeira.

Note que apenas SEG-forte explica por que S não sabe que seu bilhete não será sorteado. Pois, caso considerássemos SEG-fraca, haveria o mundo em que o bilhete de S é sorteado poderia não ser levado em consideração; poderia ser o mundo excluído pelo “aproximadamente”. Nesse caso, S saberia que não ganharia na loteria. Contudo, SEG-forte dá um veredito negativo a ENTULHO. Por exemplo, um mundo em que há uma ponta de prego saliente numa das junções da calha e o cadarço da botina acaba se prendendo nela é um mundo bastante próximo. Assim, só SEG-fraca é que explica por que S sabe que a botina está na caçamba de entulho. Mas nesse caso, não mais explicamos corretamente LOTERIA.

Assim, temos o seguinte dilema. Se adotamos SEG-forte, então ENTULHO não é um caso de conhecimento. Mas se adotamos SEG-fraca, LOTERIA passa a ser um caso de conhecimento. Logo, ou ENTULHO é um caso de conhecimento ou LOTERIA é um caso de falha de conhecimento.

O dilema de Greco deixa a condição de segurança em maus lençóis. Na melhor das hipóteses, a segurança é uma condição vaga; na pior, tem de ser abandonada.

2.2. A resposta de Pritchard: segurança e a natureza da sorte

Pritchard pensa que para resolver o dilema “precisamos ter em conta que a motivação para a segurança está arraigada em certa abordagem da natureza da sorte” (2012, p. 178). Como já vimos no Capítulo 1, a sorte é um fenômeno modal. Dizer que um evento E ocorreu por sorte (no mundo atual) significa que há um conjunto de mundos bastante próximos em que E não ocorre. Por exemplo, que S acertou na loteria é um evento de sorte, uma vez que há vários mundos próximos em que S compra o mesmo bilhete, o sorteio é feito aleatoriamente, e outro bilhete que não o de S é o sorteado. Do mesmo modo, um sucesso cognitivo (crença verdadeira) ocorre por sorte no mundo atual – ao que chamamos *sorte epistêmica* – quando há

um conjunto de mundos bastante próximos ao atual nos quais há falha cognitiva (crença falsa). Por exemplo, S sente um calafrio pela manhã e forma a crença de que nas próximas horas o presidente da Venezuela será deposto. Acontece que no final da tarde tal presidente é deposto. A crença de S é verdadeira por sorte – *sorte epistêmica* – uma vez que há vários mundos bastante próximos em que S tem o calafrio e forma a sua crença, mas que o presidente não é deposto. Pritchard pensa que a maneira adequada de apreciar a sorte epistêmica como um fenômeno modal nos levará a um princípio de segurança que evita o dilema de Greco.

Em termos gerais, Pritchard pensa que o princípio de segurança exige que “um sucesso cognitivo não pudesse ter sido facilmente uma falha” (2014, p. 151). Ou seja, que nos mundos mais próximos em que o sujeito forma a crença do mesmo modo como formou no mundo atual, a sua crença não seja falsa. Mas para contornar o dilema de Greco, é preciso entender a segurança de um modo mais específico,

que envolve um cenário contínuo de risco epistêmico; tendo de um lado desse contínuo riscos epistêmicos modalmente bastante próximos que são incompatíveis com o conhecimento, e que rumam em direção a riscos epistêmicos modalmente mais distantes compatíveis com o conhecimento. (p. 152)

Essa é a chave para se entender a solução ao dilema. Digamos que S corre risco epistêmico em relação a sua crença somente se há pelo menos um mundo próximo em que S forma a sua crença do mesmo modo que formou no mundo atual e ainda assim essa crença é falsa. Assim, haverá casos em que a nossa tolerância ao risco epistêmico será baixa – quando a falha cognitiva ocorrer num mundo bastante próximo. Mas à medida que a falha cognitiva passa a ocorrer em mundos menos próximos a nossa tolerância ao risco epistêmico tende a aumentar, de modo que ao atingir distância suficiente, o risco passa a ser compatível com o conhecimento. Isso sugere que Greco, na verdade, apresenta um falso dilema: não é preciso escolher entre SEG-forte e SEG-fraca. Haverá casos – como LOTERIA – em que a nossa tolerância ao risco é baixa, e, por conseguinte, algo como SEG-forte entra em cena. Por outro lado, haverá casos – como ENTULHO – em que a nossa tolerância ao risco é maior, e, por conseguinte, algo como SEG-fraca entra em cena.

Vejamos com um pouco mais de detalhe. Consideremos primeiro LOTERIA. Suponha que eu tenha comprado o bilhete número 10 numa loteria com 100.000 bilhetes. Vamos

manter fixas as seguintes condições iniciais: (a) tenho o bilhete nº 10 e (b) a loteria é justa, e (c) considero que as minhas chances são 1/10.000. Cada número sorteado gerará um mundo possível; assim, temos um mundo possível em que ganho, e um número gigante em que perco. E todos esses mundos são igualmente próximos. Sendo assim, há um mundo bastante próximo em que ganho na loteria, muito embora forme a crença de que meu bilhete não será sorteado. Não sei, portanto, que meu bilhete não será o sorteado.

O mesmo não ocorre em ENTULHO. Estabeleçamos as condições iniciais com as quais adquiro a crença de que a botina atingiu a caçamba: (a) a calha não está entupida e nem tem qualquer avaria que obstrua a passagem de um objeto da dimensão de uma botina; (b) as botinas não tem qualquer superfície que adira às paredes da calha; (c) solto as botinas na calha do mesmo modo como jogo o entulho. Se considerarmos os mundos próximos em que (a)-(c) são mantidos e que formo a crença de que a botina chegou à caçamba de entulho, em todos esses mundos é também verdade que a botina está na caçamba. Mas também é claro que há um conjunto de mundos próximos ao atual em que solto as botinas e elas não chegam à caçamba, e ainda assim formo a crença de que chegaram. Mas será que em tais mundos as condições (a)-(c) são também cumpridas? Certamente que não. Considere um mundo em que há uma ponta de arrebite saliente em uma das junções da calha. Esse mundo não é distante, muito embora a condição (a) não seja cumprida. É, portanto, mais distante que o conjunto inicial. Em mundos como esse, a chance de as botinas ficarem presas na ponta do rebite são muito baixas. O risco epistêmico que corro, então, é bem baixo. É um risco totalmente compatível com o conhecimento. A fim de obtermos um risco epistêmico alto, teríamos de alterar bastante (a)-(c). Por exemplo: (a*) a calha está bastante suja; (b*) as botinas são feitas de uma camurça que adere facilmente a superfícies grossas; (c*) fico a uma distância longa da calha e jogo as botinas com os olhos fechados. Num mundo assim as chances de que a botina não chegue à caçamba são bastante consideráveis. Levando-se (a*)-(c*) em conta certamente corro um sério risco epistêmico. Mas nesse caso, esses mundos não são próximos o bastante do mundo atual – os mundos (a)-(c) – para que tal risco ameace. Isso nos mostra que o risco epistêmico que corro em ENTULHO é compatível com o conhecimento.

As palavras de Pritchard (2012) expressam melhor do que eu poderia colocar:

A segurança capta assim uma intuição sobre a nossa tolerância ao risco do erro. Em mundos possíveis bastante próximos somos extremamente intolerantes quando tal

risco epistêmico se aproxima, de modo que não quereríamos formar quaisquer crenças falsas [a partir da mesma base]. Em mundos possíveis bem distantes, porém, somos extremamente tolerantes a tal risco epistêmico, dada a nossa lonjura. Nesse entremeio temos uma escala descendente de intolerância epistêmica, da intolerância extrema ao risco epistêmico à extrema tolerância. Quando dizemos que uma crença é segura, estamos dizendo que o risco epistêmico foi excluído a um grau suficiente a ponto de ela estar disponível (pelo menos nessa escala) no mercado do conhecimento. Com isso em mente, temos de pensar na segurança de modo que ela exclua completamente a crença falsa nos mundos possíveis mais próximos, mas se torne crescentemente tolerante à falsidade à medida que se afasta do mundo atual. Isto é, a segurança é compatível com a crença falsa em mundos possíveis próximos na medida em que tais mundos não estão especialmente próximos. (p. 179-80)

A solução de Pritchard parece bastante razoável. O fato de Greco ter ignorado o contínuo de mundos próximos o levou muito rapidamente a pensar que a segurança é uma questão de tudo ou nada, quando na verdade deveria ser entendida em graus.

Do mesmo modo, uma crença pode ser verdadeira por sorte em maior ou menor grau. Por exemplo, na cercania dos celeiros falsos, eu ter parado defronte ao único celeiro real e ter formado a crença de que há um celeiro ali é um evento de sorte. Foi muita sorte eu ter parado diante do único celeiro real num ambiente lotado de celeiros falsos (e indistinguíveis dos celeiros reais). A minha crença verdadeira com base em evidência visual não é conhecimento, uma vez que há mundos muito próximos em que paro defronte a um celeiro falso e acredito que há ali um celeiro. Porém, caso me tivessem dado o endereço do celeiro, e eu chegasse corretamente até ele e formasse a crença de que há ali um celeiro, o elemento de sorte parece ter sido diminuído consideravelmente. Um ambiente lotado de celeiros falsos mas que tenho o endereço correto do celeiro real é um ambiente em que corro um risco epistêmico muito menor, o suficiente para que seja compatível com o conhecimento.

2.3. O dilema reformulado

Ainda que a estratégia acima tenha sido suficiente para contornar o dilema de Greco, penso que podemos apresentar uma nova versão do dilema. Desta vez, iremos considerar a maneira pela qual determinamos o conjunto de mundos próximos relevantes para a segurança. Há duas: (a) consideramos a mera crença do sujeito na proposição alvo; ou (b) consideramos a *crença verdadeira* do sujeito na proposição alvo. No caso de (a), a segurança será demasiado

concessiva – considerará como próximos mundos irrelevantes; o que, por conseguinte, excluirá casos claros de conhecimento. Por outro lado, se for (b) o caso, então a segurança será demasiado restritiva – garantirá por *default* que os mundos mais próximos em que S acredita que α (aproximadamente do mesmo modo que formou no mundo atual) são aqueles mundos nos quais α é o caso. Por conseguinte, considerará como conhecimento casos de falha de conhecimento.

Vamos à primeira ponta do dilema. Tomemos um caso simples de conhecimento perceptual. Estou agora olhando para uma garrafa de café. Se sei que há uma garrafa de café em minha mesa, então, de acordo com a segurança, nos mundos mais próximos em que acredito (aproximadamente do mesmo modo que no mundo atual) na proposição ⟨Há uma garrafa de café sobre a mesa⟩, é o caso que há uma garrafa de café sobre a mesa. Há aqui dois conjuntos de fatores que trabalharão na restrição dos mundos próximos: as condições iniciais e a minha crença de que há uma garrafa de café sobre a mesa. Ou seja, para avaliar a condicional da segurança, temos de levar em conta aqueles mundos nos quais formo a minha crença a partir de condições favoráveis de observação, com luz e distância adequadas, com meu aparato visual funcionado adequadamente, etc. Mas que razões temos para supor que nesse conjunto de mundos só há mundos nos quais a proposição ⟨Há uma garrafa de café sobre a mesa⟩ é verdadeira? É bastante razoável supor que um mundo em que não há uma garrafa de café sobre a mesa, mas, ao invés, uma edição limitada da Cachaça 51 numa embalagem que imita uma garrafa de café, contará como mundo próximo se levarmos em conta a mera crença do sujeito na proposição alvo de conhecimento. Se eu estiver correto, então há um mundo próximo – no qual há a Cachaça 51 em cima da mesa – em que acredito que há uma garrafa de café, mas a proposição alvo é falsa. Portanto, tenho de considerar que não sei que há uma garrafa de café em cima de minha mesa, o que é inaceitável.

Passemos agora à outra ponta do dilema. Há uma maneira muito simples de evitarmos o resultado acima: introduzir entre as condições iniciais a de que eu estava à distância e iluminação adequados, etc. *olhando para a garrafa de café*. Isso implica que a proposição ⟨Há uma garrafa de café sobre a mesa⟩ é verdadeira em todos os mundos selecionados pela minha crença de que há uma garrafa de café sobre a mesa e pelas condições iniciais do processo de formação dessa crença no mundo atual. Ou seja, em todos os mundos próximos

em que acredito que há uma garrafa de café sobre a mesa é também o caso que há tal garrafa. Isso é garantido por *default*.

Mas agora vem o problema. Voltemos a CELEIRO. Esse é um caso claro de falha de conhecimento. O sujeito forma uma crença verdadeira por sorte de que há um celeiro diante de si. A explicação – dada tanto por proponentes da sensibilidade quando da segurança – é que há mundos próximos em que o sujeito forma a crença de que há um celeiro diante de si quando na verdade não há. Ora, se considerarmos que dentre as condições iniciais, no mundo atual, sob as quais o sujeito veio a formar a crença na proposição ⟨Há um celeiro ali⟩ está a de que ele está à boa distância do celeiro real, e mantivermos tal condição como critério de seleção de mundos, então não haverá mundo próximo em que o sujeito acredite que há um celeiro em sua frente e seja falso que há. Por conseguinte, CELEIRO passa a contar como um caso de conhecimento. O que é novamente inaceitável.

Portanto, ou a segurança é demasiado *restritiva* – pois exclui casos claros de conhecimento perceptual –, ou é demasiado *permissiva* – pois inclui casos de sorte epistêmica como conhecimento; (o que é bastante irônico, dado as pretensões da epistemologia anti-sorte de Pritchard.)

É de bom tom que chequemos se a resposta de Pritchard ao dilema de Greco não se aplica a esse novo dilema. Pois talvez eu tenha cometido o mesmo erro que Greco, diria Pritchard, ao não levar em consideração o contínuo de mundos e o risco epistêmico. Em primeiro lugar, a estratégia não funcionará se considerarmos que é apenas a mera crença do sujeito (mais as condições iniciais) que serve como parâmetro para a determinação do conjunto de mundos próximos. De acordo com a solução de Pritchard, temos de levar em conta a nossa tolerância ao erro no caso da garrafa de café. E parece que somos nesse caso bastante tolerantes; há possibilidade de eu estar errado, mas essa possibilidade é remota – um mundo possível em que a proposição ⟨Há uma garrafa de café sobre a mesa⟩ é falsa. Sendo assim, nos mundos mais próximos em que acredito que há uma garrafa de café em cima da mesa, é o caso que há. Ora, a intuição de tolerância ao erro nos empurra para o uso da crença verdadeira como parâmetro para determinar a proximidade dos mundos. Ou seja, a solução de Pritchard nos empurra diretamente para a segunda ponta do dilema. Logo, o nosso novo dilema permanece.

3. Segurança e confiabilidade

3.1. Os bons informantes

John Greco (2012) defende que “pelo menos na epistemologia, é melhor ser seguro do que sensitivo” (p. 193). Ele pensa que a segurança capta melhor as nossas intuições sobre a confiabilidade. Seguindo o trabalho de Edward Craig (1990), Greco aceita que parte do propósito de usarmos o conceito de conhecimento diz respeito à sinalização de um “bom informante”. E um bom informante é aquele que é confiável. O próprio Greco resume de maneira precisa seu argumento:

queremos que nossos informantes sejam confiáveis (ou fiáveis) num espaço de situações contrafactuais próximas, isto é, situações que são similares de maneira relevante àquelas nas quais agimos. O outro lado dessa moeda é que as situações contrafactuais “distantes” são irrelevantes para a noção de um bom informante – não deveríamos nos preocupar se nossos informantes não seriam confiáveis (infiáveis) em mundos distantes. [...] Mas isso diz a favor da condição de segurança ao conhecimento, uma vez que a segurança é *apenas* confiabilidade num espaço de situações contrafactuais próximas. (p. 193)

A força do argumento de Greco depende da ideia de que a segurança seleciona os mundos que realmente são relevantes (próximos o suficiente), ao passo que a sensibilidade sempre seleciona mundos irrelevantes (distantes o suficiente). Ele oferece um exemplo, do próprio Graig, que supostamente dá apoio a essa ideia. Imagine que um ancestral nosso, na savana africana, esteja forrageando e queira saber se há predadores por perto. Ele vê um membro de seu grupo no alto de uma árvore e lhe pergunta se há predadores por perto. Se esse membro é um bom informante, ele será confiável: ele não apenas tem uma crença verdadeira de que não há predadores por perto; ele não teria facilmente formado uma crença falsa sobre se há predadores por perto (p. 199).

Mas não é nada claro que esse exemplo favoreça a segurança sobre a sensibilidade. Por que exatamente não poderíamos dizer que o nosso informante é confiável porque *nos mundos mais próximos em que não há predadores nas proximidades ele não forma a crença de que há?* Greco não é direto nesse ponto, ao invés ele sugere que a sensibilidade de algum modo nos leva a mundos mais distantes que o necessário:

[...] *não* é exigido que você seja sensível a fatos relevantes em mundos distantes. Suponha, por exemplo, que você seja insensível quanto a estar sendo enganado por um gênio maligno acerca da existência de predadores naquela área [...]. Isso não afeta minimamente o fato de você ser um bom informante para mim aqui e agora. Uma vez mais, o que importa é que você seja confiável na situação à qual efetivamente estamos. (p. 199)

Logo em seguida ele menciona que ser um bom informante talvez exija apenas “sensibilidade em situações contrafactuais *próximas*”. Contudo, tal “sensibilidade restrita”, argumenta Greco, é ou mal motivada levando-se em conta os sujeitos como bons informantes, ou é equivalente à segurança. Podemos resumir sua estratégia do seguinte modo. Há três versões de sensibilidade restrita: (a) a sensibilidade próxima, (b) a sensibilidade restrita forte, e (c) a sensibilidade restrita fraca. (a) e (c) não fazem jus à noção de um bom informante, ao passo que (b) é equivalente à segurança.

A maneira pela qual Greco pensa na sensibilidade restrita é bem próxima ao modo pelo qual defendi no Capítulo III que deveríamos conceber a sensibilidade. Porém, penso que Greco não viu exatamente as consequências dessa maneira de se restringir a sensibilidade. Mas vejamos primeiro cada uma das premissas de seu argumento.

Quanto à (a): *sensibilidade próxima*. Greco diz que a crença de S de que α é *sensitiva-próxima* sse no mundo mais próximo no qual α não é o caso, S não acredita que α . Note que aqui há apenas “o mundo possível mais próximo”. É óbvio que essa tentativa de restrição não funciona. Greco rapidamente nota que a sensibilidade próxima não é capaz de fazer jus à ideia de informante confiável, uma vez que queremos que um informante seja confiável ao longo de um espaço de mundos possíveis próximos. Por exemplo, tenho um quintal espaçoso em que crio galinhas. O meu vizinho tem também um quintal amplo e cria galinhas; mas, ao contrário de mim, também cria perus, que ficam confinados próximos à cerca que divide nosso quintal. Ao escutar um barulho em meu quintal, vou até lá e vejo apenas galinhas. Formo assim a crença de que não há qualquer animal do vizinho em meu quintal. Dificilmente as galinhas dele passariam para o meu quintal, muito embora, com bastante facilidade os perus passariam. Assim, no mundo mais próximo em que há um animal do vizinho em meu quintal, há um peru; por conseguinte, não formo a crença de que não há um animal do vizinho em meu quintal. Porém, há mundos próximos em que as galinhas do vizinho fogem para o meu quintal; nesse caso, como não sou capaz de distinguir minhas galinhas das dele, eu acreditaria

que não há animais do vizinho em meu quintal. O que intuitivamente mostra que não sou um bom informante sobre haver animais do vizinho em meu quintal.

Embora eu tenda a concordar com Greco quanto a isso, permita-me apenas um comentário. Podemos pensar num bom informante de maneira relativizada, do mesmo como sugeri no capítulo anterior. Na ausência de um micro-ambiente desfavorável, um informante que tende a ser ruim num ambiente mais amplo, pode acabar funcionando muito bem como um bom informante num ambiente bem delimitado (favorável). Uma analogia futebolística ajuda. Há jogadores que dificilmente são escalados; geralmente servem apenas para compor elenco. Mas imagine que um desses jogadores considerados medíocres acaba se mostrando eficiente em um único tipo de jogada – arrancada pelas laterais – num tipo particular de jogo – final de jogo em que o time adversário está todo no seu campo defensivo. Sempre que acontece isso, o técnico opta por esse jogador – que em nenhuma outra situação entraria no jogo – pois ele é bastante confiável para esse tipo de jogada: acerta muito mais que seus companheiros.

Passemos à (b): *sensitividade restrita forte*. A crença de S de que α é *sensitiva-restrita-forte* sse num espaço restrito de mundos (mundos próximos) em que α não é o caso, S não acredita que α . Greco pensa que nesse caso sensitividade e segurança são coextensivos. Considere o conjunto RS de mundos formado pelo espaço restrito de mundos- $\neg\alpha$. Suponha que todos os mundos contidos em RS são mundos- $\neg B\alpha$. Seguindo Greco (pp. 202-3), vamos usar os operadores modais \Box e \Diamond restritos a RS. Se α é verdadeira em todos os mundos em RS, então $\Box\alpha$. E se α é verdadeira em pelo menos um mundo em RS, então $\Diamond\alpha$. Uma vez que a condicional material $\neg\alpha \rightarrow \neg B\alpha$ é verdadeira em todos os mundos em RS, temos $\Box(\neg\alpha \rightarrow \neg B\alpha)$. Isso é equivalente a $\neg\Diamond(\neg\alpha \wedge B\alpha)$. Ora, essa última proposição é também equivalente a $\Box(B\alpha \rightarrow \alpha)$. Portanto, $\Box(\neg\alpha \rightarrow \neg B\alpha)$ e $\Box(B\alpha \rightarrow \alpha)$ são equivalentes em RS.⁹⁹

Dado que Greco pensa que a segurança tem uma motivação intuitiva muito maior que a sensitividade, ele pensa que “a presente tentativa de salvar as teorias da sensitividade falha”,

⁹⁹ Note que a restrição do conjunto de mundos a RS, torna a contrafactual $\neg\alpha \Box\rightarrow \neg B\alpha$ equivalente à condicional material necessitada $\Box(\neg\alpha \rightarrow \neg B\alpha)$. Ora, se em RS $\neg\Diamond(\neg\alpha \wedge B\alpha)$, segue-se que em RS a condicional $(B\alpha \rightarrow \alpha)$ é vacuamente verdadeira, uma vez que não há qualquer mundo- $B\alpha$. Portanto, em RS, as contrafactuals $(\neg\alpha \Box\rightarrow \neg B\alpha)$ e $(B\alpha \Box\rightarrow \alpha)$ são contrapositivas.

que “não encontramos qualquer motivação para a [sensitividade] que não seja equivalente à [segurança]” (p. 203).

Por fim (c): *sensitividade restrita fraca*. A crença de S de que α é *sensitiva-restrita-fraca* sse na maioria dos mundos (em RS) em que não é o caso que α , S não acredita que α . Vamos agora considerar um novo conjunto de mundos próximos, RS', que é o conjunto dos mundos- α mais próximos. O resultado da união de RS e RS' será o conjunto de mundos mais próximos que contém mundos- α e mundos- $\neg\alpha$; chamemos-lhe RS*. Estipulemos então que uma crença é *segura-fraca* sse na maioria dos mundos em RS* em que S acredita que α , α é o caso. Vejamos então se comparada a essa noção de *segurança fraca*, a *sensitividade restrita fraca* é preferível.

Podemos ser infalibilistas ou falibilistas quanto a uma fonte de informação. Se formos infalibilistas e exigirmos a *sensitividade*, temos de recorrer à *sensitividade restrita forte*, uma vez que exigiremos a *confiabilidade* ao longo de todos os mundos em RS. Mas isso, argumenta Greco, é o equivalente a exigir a *segurança*, como vimos acima. Se formos falibilistas, aí sim recorreremos à *sensitividade restrita fraca*. Nesse caso, *sensitividade fraca* e *segurança fraca* não são coextensivas: uma crença pode ser fracamente segura sem ser fracamente sensitiva. Mas qual a motivação para preferir a *sensitividade restrita fraca* nesse caso? Vale a pena citar Greco:

[A] motivação é direta: ainda que a crença de alguém seja verdadeira na maioria dos [mundos-B α] (*i.e.*, a crença de que [α] é Segura-Fraca), é melhor *também* que a sua crença não seja falsa na maioria dos [mundos- $\neg\alpha$] próximos (*i.e.*, a crença de que [α] é *também* Sensitivia-Restrta-Fraca). E como já vimos, é possível ter o primeiro bem epistêmico sem ter o segundo. Porém, isso é apenas dizer que, sendo todas as coisas iguais, é melhor ser mais seguro do que menos seguro. (p. 205)

Ou seja, Greco está a nos dizer que não precisamos recorrer a mundos próximos em que α não é caso, é suficiente que na maioria dos mundos-B α seja também o caso que α . Se houver uma razão baixa de mundos-B α que são também mundos- $\neg\alpha$, então a crença é mais segura; e se houver uma razão alta, então a crença é menos segura.¹⁰⁰ Assim, a intuição sobre uma fonte de informação confiável é melhor explicada pela *segurança* (em maior ou menor grau) ao invés da *sensitividade*.

¹⁰⁰ Seja n o número de mundos-B α que são também mundos- α e m o número de mundos-B α que são mundos- $\neg\alpha$. A razão n/m é então a razão entre os mundos- α e os mundos- $\neg\alpha$.

3.2. Resposta a Greco

Embora eu pense que Greco tenha razão quanto à sensibilidade próxima, penso que seus argumentos contra a sensibilidade restrita forte e fraca não são bons o bastante para preferirmos a segurança ao invés da sensibilidade.

A razão que ele oferece para preferirmos a segurança à sensibilidade restrita forte é que nesse caso sensibilidade e segurança são coextensivas. Mas essa não é uma razão boa o suficiente. Apenas se já temos uma preferência prévia ou uma intuição a favor da segurança é que diremos que é melhor a segurança do que a sensibilidade. É claro que Greco argumenta mais a frente que a segurança capta melhor a noção de confiabilidade, mas até esse ponto seu argumento é menos que conclusivo. Seja como for, parece-me que essa equivalência entre segurança e sensibilidade é mais favorável à segurança. Vejamos.

Ao tornar a segurança e a sensibilidade coextensivas, temos que $(\neg\alpha \square \rightarrow \neg B\alpha)$ e $(B\alpha \square \rightarrow \alpha)$ são verdadeiras em exatamente os mesmos mundos possíveis em RS. Uma vez que RS está restrito aos mundos- $\neg\alpha$, que são também mundos- $\neg B\alpha$, a condicional da segurança é trivialmente satisfeita. Isso porque não há qualquer mundo em RS que torne a conjunção $(B\alpha \wedge \neg\alpha)$ verdadeira. Ora, mas não é isso que um proponente da segurança pretende; que levados em conta os mundos mais próximos em que o sujeito não acredita numa proposição falsa (a partir de aproximadamente as mesmas condições do mundo atual), a sua crença é segura porque não há qualquer mundo em RS em que o sujeito acredite falsamente nessa proposição. Antes, ele pretende que nos mundos mais próximos em que o sujeito acredite na proposição, essa proposição seja verdadeira. E a essa altura, o proponente da sensibilidade poderia dizer o mesmo, alegando prioridade à sensibilidade. Considere o conjunto de mundos- $B\alpha$ mais próximos, RS^+ . Em RS^+ os mundos- $B\alpha$ são também mundos- α . Assim, em RS^+ também não há qualquer mundo que torne a condicional da sensibilidade falsa, uma vez que nenhum mundo em RS^+ torna verdadeira a conjunção $(\neg\alpha \wedge B\alpha)$. Por conseguinte, a sensibilidade é trivialmente satisfeita, e em RS^+ sensibilidade e segurança são equivalentes. E isso parece tão insatisfatório enquanto argumento a favor da sensibilidade assim como o argumento acima a favor da segurança.

Como mencionado acima, o proponente da segurança não quer que uma crença seja trivialmente segura. Ele quer que nos mundos mais próximos em que ele acredita na

proposição alvo seja também o caso que essa proposição é verdadeira. Nesse caso, levando-se em consideração RS – que contém apenas mundos- $\neg\alpha$ – temos de expandir tal conjunto a fim de que a segurança seja satisfeita não trivialmente, *i.e.*, que haja mundos- $B\alpha$ mais próximos que sejam também mundos- α . E aqui começam os problemas. Há duas maneiras de fazermos isso. (i) usando $B\alpha$ como parâmetro, ou (ii) usando α como parâmetro.

Vamos a (i). Consideremos RS'' , o conjunto de mundos- $B\alpha$. Como vimos na seção 2.3, há mundos- $B\alpha$ próximos o suficiente que são também mundos- $\neg\alpha$. E não é sábio por parte do defensor da segurança que considere os mundos- $(B\alpha \wedge \alpha)$ – a que podemos chamar R'^* –, uma vez torna a segurança demasiado permissiva.

Passemos a (ii). Vamos adicionar o conjunto RS' , o conjunto de mundos- α mais próximos. Dadas as condições iniciais (aproximadas) nas quais o sujeito forma a sua crença no mundo atual, e dado que a sensibilidade é satisfeita, os mundos- α serão também mundos- $B\alpha$. Por exemplo, presumo que a minha crença de que há um computador em minha frente seja sensitiva. Disso segue-se que não há qualquer mundo em RS em que seja falso que há um computador em minha frente e eu acredite que há. Por conseguinte, a minha crença é trivialmente segura. Mas considerando que nos mundos mais próximos em que há um computador em minha frente e cumpro aproximadamente as mesmas condições iniciais do mundo atual, é também o caso que acredito que há um computador em minha frente. Ou seja, não há qualquer mundo em RS' em que eu acredite que tenha um computador em minha frente e que não haja qualquer computador em minha frente. Donde se segue que a sensibilidade é trivialmente satisfeita. Levando-se em conta então RS^* – a união de RS e RS' – sensibilidade e segurança parecem de fato se contrapor, embora não de modo trivial. Mas que razões temos, contrário a Greco, para preferir a sensibilidade?

A razão é que a segurança vem de bônus. Uma vez que nem a introdução de RS'' e nem a de R'^* são bem motivadas, só nos resta RS' . E a introdução de RS' nos garante que a condição de aderência ($\alpha \square \rightarrow B\alpha$) seja cumprida: nos mundos mais próximos em que α é o caso (e S cumpre aproximadamente as mesmas condições iniciais que no mundo atual), S acredita (forma a crença de) que α . Sendo assim, considerando que o nosso conjunto de mundos próximos é RS^* – *i.e.*, a união de RS e RS' –, a maneira de avaliarmos se uma crença conta como conhecimento é se ela é sensitiva: (i) se a condição de *variação* é cumprida, e (ii) se a condição de *aderência* é cumprida. Se a variação é cumprida, então a segurança é

trivialmente satisfeita; e se a aderência é cumprida, a segurança é novamente satisfeita, uma vez que em RS^* todos os mundos- $B\alpha$ são também mundos- α .

Note, uma vez mais, que o proponente da segurança não pode tentar recorrer à estratégia de reverter a argumentação e dizer que uma vez satisfeita a segurança, satisfaz-se também a variação e a aderência. Como tenho argumentado, a maneira viável de garantirmos que a condição de segurança seja satisfeita quando o sujeito sabe uma proposição, é por considerar como parâmetro os mundos- α , e não os mundos- $B\alpha$. E Greco desapercivelmente parece fazer isso ao explicar por que a sensibilidade fraca e a segurança fraca não são extensionalmente equivalentes: “a segurança fraca [...] considera os [mundos- α] próximos tanto quanto os [mundos- $\neg\alpha$] próximos” (p. 204). Ele certamente poderia ter dito – na verdade, é o que esperaríamos que ele dissesse – algo como: “a segurança fraca considera os mundos- $B\alpha$ próximos”. Talvez Greco tenha apenas cometido um deslize; mas talvez ele tenha pensado que a segurança é satisfeita usando as condições de verdade da aderência. Se o primeiro, então o problema com os mundos- $B\alpha$ próximos que são também mundos- $\neg\alpha$ permanece; se o último, então não temos razões para pensar que a segurança seja preferível.

Sosa (1999) pensava que os proponentes da sensibilidade é que tinham se confundido, que tinham “sucumbido a uma ilusão cognitiva” pelo fato de ambas as condições serem um tanto parecidas:

A segurança e a sensibilidade, sendo contrapositivas mútuas, são facilmente confundidas, de modo que é fácil confundir a exigência correta de segurança (para o conhecimento e para sua atribuição correta) com uma condição de sensibilidade. É fácil negligenciar que essas condicionais subjuntivas não se contrapõem. (p. 148)

Mas como é evidente, ele pensava que essa ilusão se devesse ao fato das condicionais aparentemente se contraporem. Isso é bastante próximo do que diz Greco sobre a sensibilidade restrita ser equivalente à segurança. Mas vimos que Greco não tem razão. Por outro lado, não me parece injusto dizermos que foi Greco (juntamente com Sosa e Pritchard) quem sucumbiu a uma ilusão cognitiva: o fato de num espaço restrito de mundos, os mundos- α serem também mundos- $B\alpha$. Esse fato modal diz respeito primariamente à aderência, tornando a segurança parasitária. Assim, na medida em que a argumentação de Greco tiver alguma plausibilidade, terá de ser a favor da aderência e não da segurança.

Essas mesmas considerações nos permitem ver por que Greco também não é bem sucedido em argumentar que a segurança fraca é preferível à sensibilidade fraca. Ainda que ele esteja certo em dizer que as nossas intuições acerca da confiabilidade são adequadamente captadas pela noção de “ser mais seguro”, o que parece ser mais adequado é dizer que uma fonte de informação é confiável na medida em que produz crenças “mais aderentes”, uma vez que a aderência tem primazia sobre a segurança. Não penso, contudo, que a aderência seja suficiente para explicar a confiabilidade. Antes, a sensibilidade é que explica. Uma fonte de informação é confiável na medida em que produz crenças sensitivas. Isto é, quando a variação e a aderência são satisfeitas. Na seção seguinte oferecerei uma breve explicação de por que a sensibilidade capta melhor nossas intuições acerca da confiabilidade.

4. Sensitividade e confiabilidade

Penso que se a EFAS – a teoria da função apropriada tem mais a sensibilidade – tiver a mínima plausibilidade, então é também plausível pensar na confiabilidade do mesmo modo. A ideia é que a confiabilidade é explicada em termos de função apropriada, que por sua vez depende da sensibilidade.

Já vimos que a sensibilidade se sai bastante bem na explicação das nossas atribuições de conhecimento e falha de conhecimento, sobretudo naquelas falhas que envolvem sorte epistêmica. Contudo, o leitor pode ainda não estar convencido de que os mesmos elementos deem conta da nossa intuição de confiabilidade.¹⁰¹ Esforçar-me-ei, nesta seção, para tornar clara tal explicação. Seja como for, o ponto principal aqui é afastar de uma vez por todas a ideia de que é melhor ser seguro do que sensitivo.

Começarei com um caso intuitivo do como geralmente pensamos sobre a confiabilidade. Penso haver certa evidência linguística para considerarmos que a sensibilidade capta melhor as nossas intuições sobre a confiabilidade do que do que a segurança. Em seguida, ofereci um argumento mais forte; sustentarei que se há algo de importante que a condição de segurança supostamente capta, a condição de aderência faz

¹⁰¹ No caso do leitor sequer partilhar da intuição de que o conhecimento envolve confiabilidade, é muito certo de que discordemos profundamente quanto à tese aqui defendida desde o início. Nesse caso, peço apenas paciência e um esforço na tentativa de pensar como um confiabilista. É possível também que a sensibilidade explique a confiabilidade, muito embora nenhuma delas esteja envolvida na explicação do conhecimento!

exatamente o mesmo trabalho. Com esse resultado seremos capazes de inverter a acusação de ilusão cognitiva.

Vamos considerar novamente os detectores de metal. Um detector é confiável quando tende a uma taxa de sucesso relativamente maior do que a uma taxa de fracasso. E essa *tendência* ao sucesso é modal: quando há um metal ele apita, e quando não há ele não apita.¹⁰² Isso sugere que a maneira intuitiva de pensarmos sobre a confiabilidade é em termos da sensibilidade: *se não houvesse um metal, o detector não apitaria* (variação), e *se houvesse um metal, o detector apitaria* (aderência). Mas também poderíamos dizer o seguinte do detector: quando ele apita, há um metal. O proponente da segurança poderia dizer também entendemos intuitivamente a confiabilidade em termos de segurança: *se o detector apitasse, então haveria um metal*. Ou como também se diz, *o detector dificilmente teria apitado sem que fosse o caso de haver ali um metal*. Não parece, no fim das contas, que temos intuições razoáveis de ambos os lados?

Talvez então fosse melhor começar por uma resposta mais neutra, uma que a princípio não pendesse mais para a sensibilidade ou para a segurança. Dizer que um detector é confiável envolve pelo menos dizer que (i) *dificilmente aconteceria de ele apitar e não haver ali um metal*. Mas isso não pode ser tudo o que basta para a confiabilidade. Se eu estivesse prestes a investir dinheiro num detector, gostaria também de saber se ele apitará sempre (ou quase sempre) que estiver por perto um metal. Ou seja, dizer que um detector é confiável envolve também dizer que (ii) *sempre (ou quase sempre) que houver ali um metal o detector apitará*.

Traduzindo a afirmação “Dificilmente aconteceria de o detector apitar e não haver ali um metal” para o idioma dos mundos possíveis temos: *não há qualquer mundo possível próximo em que o detector apite e não haja ali um metal*. Alguns conjuntos de mundos tornam essa afirmação verdadeira. Por exemplo, seja W o conjunto de mundos próximos ao mundo atual. Digamos que $W = \{w_1, w_2\}$, isto é, há apenas dois mundos próximos ao atual. Digamos também que as proposições ⟨Há um metal por perto⟩ e ⟨O detector apita⟩ – às quais

¹⁰² Levando-se em conta o modo como entendo a confiabilidade nesta tese, um mecanismo cognitivo é confiável sse: (i) tende a produzir mais crenças verdadeiras do que falsas, (ii) num ambiente adequado para o seu funcionamento, o que implica que (iii) não há um micro-ambiente desfavorável, e (iv) o uso que se faz desse mecanismo visa a crença verdadeira.

representaremos por “ p ” e “ q ”, respectivamente – são ambas verdadeiras em W . O seguinte diagrama (Diagrama 2) ilustra a situação:

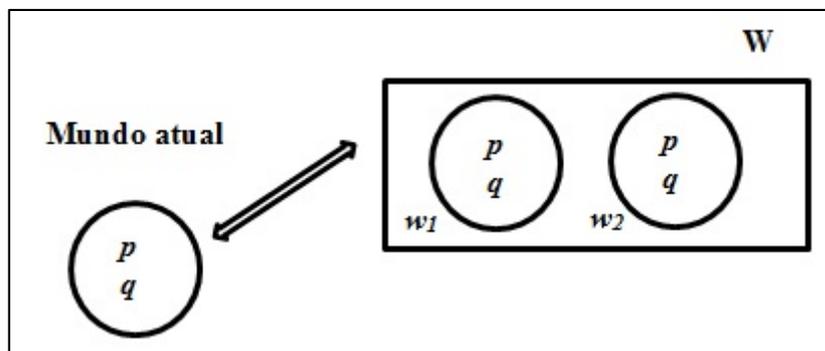


Diagrama 2

Note que não apenas a proposição (i) é tornada verdadeira, mas também a proposição (ii). Traduzindo (ii) para o idioma modal, temos: *em todos os mundos próximos em que p é verdadeira, q também é*. A estrutura no Diagrama 2 fornece, portanto, um modelo para a atribuição de confiabilidade do detector.

Podemos agora transpor o caso do detector para a formação de crença. O bipe do detector passa a ser a crença do sujeito. Assim, diremos que um mecanismo cognitivo, ou processo de formação de crenças, é confiável pelo menos se não produz uma crença de que α quando α não é o caso. E também se pelo menos produz a crença de que α quando α é o caso. Ora, essa última exigência é nada mais do que a condição de aderência – $\alpha \square \rightarrow B\alpha$ – uma das subcondições da sensibilidade.

O proponente da segurança poderia nos chamar atenção para o fato de que essa mesma estrutura torna verdadeira a condicional da segurança – $B\alpha \square \rightarrow \alpha$. Nos mundos possíveis mais próximos ao atual em que o sujeito acredita que α , α é o caso. No Diagrama 2, substitua q por Bp , e fica claro que a condicional da segurança é satisfeita. Isso quer dizer então que a segurança também é uma condição necessária para a confiabilidade? É exatamente isso o que sugere o proponente da segurança ao dizer que o detector é confiável porque *difícilmente teria apitado sem que fosse o caso de haver ali um metal*.

Retornemos ao exemplo do detector. Suponha que o vendedor vá fazer uma demonstração do detector a fim de me garantir que ele é confiável. Ele passa o detector por várias superfícies e mostra que sempre que o detector apita, há ali um metal. Quando

pergunto, “o que me garante que um bipe não seja um alarme falso?”. O vendedor responde: “são raras as situações em que esse detector dará um alarme falso; talvez você tenha de usá-lo em Marte para ver um alarme falso!”. Ele complementa dizendo que o bipe é uma indicação confiável da presença do metal. Ora, mas se o bipe é uma indicação confiável da presença do metal, então a presença do metal tem de ser anterior ao bipe; pois é a presença do metal que é responsável pelo bipe. De modo que a maneira mais natural de entender a bipe do detector – levando-se em conta que o metal causa o bipe – é através da condição de aderência.

Reconheço que a intuição aqui é bastante escorregadia. A ideia de considerar o bipe como um indicador confiável de metal parece ser captada pela condicional *Se o detector emitisse um bipe, então haveria um metal ali*. Considere então a seguinte condicional, *Se houvesse fumaça, haveria fogo*. Essa condicional parece captar a ideia que de a fumaça é uma indicação confiável de fogo. Mas uma vez que sabemos que pode haver fumaça sem fogo, a fumaça só será um bom indicador se o ambiente monitorado estiver livre de máquinas de fumaça ou qualquer outra coisa que produza fumaça suficiente para parecer fumaça de incêndio. Em outras palavras, a fumaça é um bom indicador de fogo quando, no ambiente adequado, (i) se não houvesse fogo, não haveria fumaça, e (ii) se houvesse fogo, haveria fumaça. Não queremos que a fumaça seja um alarme falso. O mesmo pode ser dito para o detector. Na medida em que um detector é confiável ele dificilmente emite um alarme falso. Mas dizer que o detector não produz alarmes falsos simplesmente porque ele sempre apita na presença de metal não é suficiente. Ao avaliarmos um detector durante uma compra, por exemplo, gostaríamos que o vendedor nos mostrasse o que acontece quando o detector passar por uma superfície em que não há metal por perto. Queremos nos assegurar de que se não há metal, não há bipe.

Depois então de nos assegurarmos que o detector é bom porque apita na presença de metal, queremos nos assegurar que ele não daria alarme falso. Ou seja, para que ele seja considerado realmente bom, queremos ver se no caso em que não há metal, mas outra coisa parecida, ele não apita. Vamos considerar agora um segundo conjunto de mundos próximos, $W' = \{w_3, w_4\}$ em que $\neg p$ é o caso – *i.e.*, a proposição ⟨Há um metal por perto⟩ é falsa. Nesses mundos também é o caso que $\neg q$ – *i.e.*, a proposição ⟨O detector apita⟩ é falsa. Assim, temos o seguinte diagrama:

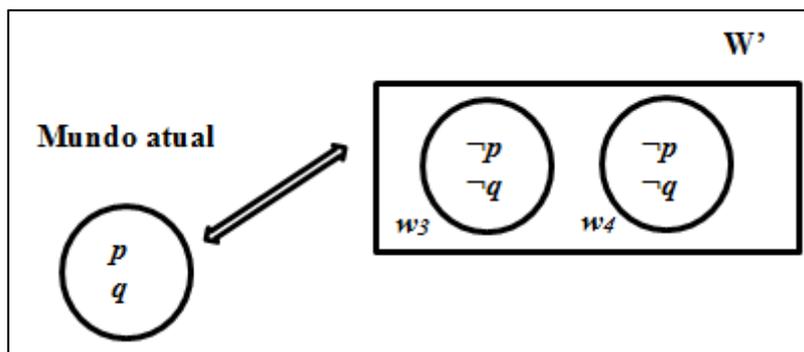


Diagrama 3

Note que essa estrutura também torna verdadeira a proposição de que não há qualquer mundo próximo em que o detector apite e não haja metal por perto. E se substituirmos $\neg q$ por $\neg Bp$ – *i.e.*, pela proposição ⟨S não acredita que há metal por perto⟩ – temos a condicional da variância satisfeita – $\neg p \Box \rightarrow \neg Bp$.

Espero que as considerações acima tenham deixado claro que a nossa noção intuitiva – pré-teórica – de confiabilidade é bem captada pela sensitividade. Fazendo a união de ambas as estruturas apresentadas nos diagramas 2 e 3, temos um modelo para a confiabilidade:

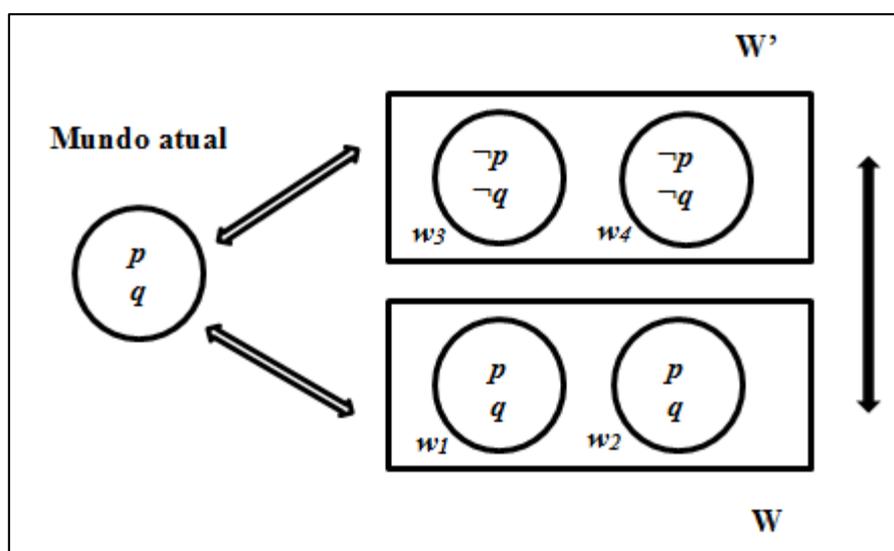


Diagrama 4

O Diagrama 4 representa o teste feito pelo vendedor para mostrar ao seu cliente que o detector é confiável. Primeiro ele mostra que na presença de metal, o detector apita. Depois, mostra

que, num ambiente em que há algo que poderia ser confundido com metal, mas que não há metal, o detector não apita – *i.e.*, não produz alarme falso.

O argumento, contudo, ainda está incompleto. Falta explicar a “ilusão cognitiva” do proponente da segurança.

Em primeiro lugar, não penso que a acusação de ilusão cognitiva por parte de Sosa seja inteiramente justa. Ele diz que “alguns de nós podem ser levados a aceitar a exigência de sensibilidade por ela ser facilmente confundida com a exigência correta, a de segurança, sucumbindo assim a uma ilusão cognitiva” (p. 148). Essa acusação só faria sentido se pelo menos parte dos proponentes da sensibilidade visasse defender a sensibilidade contra a segurança. Mas não foi isso o que ocorreu. A bem da verdade, a condição de segurança passou a fazer frente à sensibilidade apenas depois da publicação do artigo de Sosa. De modo que a acusação seria mais adequada se levasse em consideração as nossas intuições modais acerca da confiabilidade, como apresentado acima. Algo como isto. “Ao considerarmos intuitivamente a confiabilidade, somos muito naturalmente levados a pensar em situações contrafactuais de dois tipos: situações em que o objeto a ser rastreado ou detectado está presente e situações em que o objeto não está presente. Isso nos leva a pensar que as condicionais da aderência e variação são pelo menos necessárias (e talvez conjuntamente suficientes) para dar corpo às nossas intuições. Mas como a segurança é a condição correta, fomos enganados pela intuição.”

Porém, essa nova acusação fica com o ônus de explicar por que não precisamos levar em conta, ao caracterizar teoricamente a confiabilidade, as situações contrafactuais – *i.e.*, os mundos possíveis mais próximos – em que o objeto a ser rastreado não está presente. É claro que a resposta do proponente da segurança envolve insistir que a sensibilidade leva em conta mundos possíveis demasiado distantes, como por exemplo, mundos possíveis em que o sujeito é um cérebro numa cuba recebendo estímulos eletroquímicos para acreditar que está rodeado de objetos externos – e demais cenários céticos extravagantes. Levar em conta esses mundos invariavelmente nos obriga (i) a rejeitar o princípio de fechamento e (ii) ter de considerar como insensitivas crenças como a de que não sou um cérebro numa cuba, e crenças indutivas como a de sei que o lixo que joguei pela calha atingiu a caçamba. Pensar em termos da condição de segurança, contudo, ainda não afasta a plausibilidade de termos de considerar a condição de variação para caracterizar a confiabilidade. Penso que o exemplo do detector

tenha deixado suficientemente claro a necessidade da variação. Talvez aceitar (i) e (ii) seja o preço que tenhamos de pagar. Porém, como já defendi no Capítulo III, a sensibilidade não exclui o conhecimento indutivo. E quanto ao caso do fechamento e ceticismo, argumentarei no Capítulo V que não é óbvio que considerar como insensitivas crenças em negações de hipóteses céticas, ao mesmo tempo em que consideramos ter conhecimento de proposições sobre o mundo externo, nos obrigue a rejeitar completamente o fechamento.

Mas já agora, considerando que a aderência capta bastante bem a nossa exigência de que em situações próximas nas quais o objeto esteja presente o mecanismo detecte tal objeto, a condição de segurança é que passa a ser a ilusão. Como ilustrado no Diagrama 2, os mundos- p mais próximos são também mundos- Bp . Isso está de acordo com a segurança. Mas isso só acontece quando pressupomos que os mundos- Bp mais próximos são também mundos- p . Isto é, quando pressupomos que os mundos mais próximos relevantes para a avaliação da contrafactual $Bp \square \rightarrow p$ são apenas aqueles mundos em que é verdade que $Bp \wedge p$. Mas é claro que o proponente da segurança não quer dizer a trivialidade de que S sabe que α somente se nos mundos mais próximos em que S acredita verdadeiramente que α então α é o caso. Ele quer dizer que S não teria facilmente formado uma crença falsa de que α . A aderência, pelo contrário, não implica essa trivialidade. Ela diz que nas circunstâncias em que α é o caso, S acredita que α (lembrando, é claro, do método de formação de crença e do ambiente).

O principal resultado deste capítulo foi: a segurança não é melhor que a sensibilidade. Se for considerada como uma condição rival ao conhecimento, é satisfeita por *default*. E na medida em que não for trivial, é melhor vista como a subcondição de aderência. Além do mais, ainda que interpretada como a aderência, não seria suficiente para explicar a confiabilidade dos nossos mecanismos cognitivos. Os proponentes da segurança pensam que sua abordagem explica o conhecimento indutivo e oferece uma resposta ao argumento cético que não implica a rejeição do fechamento. Já vimos que EFAS dá conta do conhecimento indutivo. Mas a segurança não pode dar conta seja do que for, uma vez que não é uma condição robusta o suficiente. No capítulo seguinte veremos que mesmo a acusação de rejeição do fechamento epistêmico não é tão séria quanto parece à primeira vista.

V

CETICISMO E FECHAMENTO

Uma abordagem do conhecimento deveria iluminar os argumentos céticos e mostrar onde reside sua força. Caso essa abordagem nos leve à rejeição desses argumentos, seria melhor que não fosse de modo tão fácil e superficial. Pensar que o cético negligencia algo óbvio, atribuir-lhe um simples erro, ou uma confusão, ou uma falácia é recusar-se a reconhecer o poder de sua posição e a força que ela pode exercer sobre nós. Com isso, perdemos a oportunidade de colher seus insights e obter auto-conhecimento na compreensão de por que seus argumentos nos seduzem. Ademais, não podemos afastar o fantasma do ceticismo sem antes ouvir aquilo que ele tem a nos revelar.

Robert Nozick, *Philosophical Explanations*

Prévia

Se EFAS for plausível, então não pode haver conhecimento sem o funcionamento apropriado das nossas faculdades cognitivas em ambientes cognitivos adequados ou favoráveis. E um ambiente é favorável quando a condição de sensibilidade é cumprida: mantendo-se aproximadamente o mesmo ambiente e o mesmo método de formação de crença, se não fosse o caso que α o sujeito não acreditaria que α , e se fosse o caso que α o sujeito acreditaria que α . Além disso, embora a segurança pareça ser superior à sensibilidade, vimos que a intuição de segurança é melhor entendida coma a sub-condição de aderência. Resta-nos examinar se a resposta nozickiana ao ceticismo (a resposta padrão baseada na sensibilidade) é de fato implausível como defende grande parte dos epistemólogos contemporâneos. Essa suposta implausibilidade se deve ao putativo fato de uma teoria da sensibilidade implicar a invalidade do princípio de fechamento, uma vez que prevê que sabemos proposições sobre o mundo externo muito embora não saibamos as negações das hipóteses céticas radicais – sendo essas últimas implicadas pelas primeiras.

Neste capítulo argumentarei que embora a sensibilidade implique a invalidade de uma versão mais ampla do fechamento, daí não segue que não possamos usar uma versão mais restrita.

Como veremos, há alguma motivação para o abandono de um princípio de fechamento mais geral que é independente (mas bastante próxima) de uma resposta ao argumento cético; e a sensibilidade é capaz de explicar esses casos. Isso sugere que o abandono do fechamento não é uma manobra desesperada para evitar o ceticismo. Contudo, na medida em que a sensibilidade explica porque o fechamento falha para um conjunto determinado de proposições e não para outras, nada nos impede de formular um princípio mais restrito que atenda certas exigências. Chamarei tal postura de *nozickianismo moderado*. É nozickiana porque mantém que o princípio de fechamento epistêmico mais geral é violado. E moderada porque permite que um princípio restrito (ainda com alcance considerável) esteja em funcionamento em várias de nossas práticas epistêmicas.

1. O argumento cético

Consideremos o argumento cético radical padrão:

- 1- Se sei que tenho mãos, então sei que não sou um cérebro numa cuba.
- 2- Não sei que não sou um cérebro numa cuba.
- 3- Logo, não sei que tenho mãos.

Ambas as premissas são plausíveis. Façamos “BIV” abreviar a proposição (Sou um cérebro numa cuba). A razão para pensar que a premissa 1 é verdadeira pode ser vista se pensarmos em sua negação: *sei que tenho mãos & não sei que não-BIV*. Uma vez que a proposição (Tenho mãos) implica (Não sou um cérebro numa cuba), pelo menos na medida em que reconheço essa implicação, eu não deveria sustentar a negação de 1. Proferi-la parece conversacionalmente inadequado; é sustentar aquilo que DeRose (1995) chamou de uma “conjunção abominável”.

Quanto a 2, a sua plausibilidade decorre da apreciação do seguinte cenário. Suponha que um cientista maligno tenha me raptado, retirado meu cérebro do meu crânio e o colocado numa cuba cheia de nutrientes. A cuba está ligada a um supercomputador que envia estímulos ao meu cérebro de modo que formo crenças perceptuais exatamente como se fossem produto do mundo externo. Em tal situação, embora minhas crenças sobre o mundo externo sejam massivamente falsas, eu acredito que são verdadeiras. Não há qualquer marca em meus

estados mentais que me permitam dizer se essas crenças são produto de engodo. Esse cenário é uma possibilidade. O que me garante que tal possibilidade não seja atual? Afinal, todas as minhas crenças sobre o mundo externo poderiam ser produto de uma simulação de computador. E nesse caso, eu agiria exatamente como se fossem crenças verdadeiras. Assim, na medida em que aquilo que me parece uma mão ser compatível tanto com tal estado ser produto de um objeto externo (a minha suposta mão) quanto de uma simulação computacional, não tenho razão para pensar que não sou um cérebro numa cuba.

O argumento cético é um típico paradoxo filosófico: parte de premissas plausíveis e termina, via raciocínio dedutivo, com uma conclusão inaceitável. Mas se ambas as premissas são plausíveis e o raciocínio é correto, o que fazer? Onde está o erro? É de se esperar que a negação de qualquer dessas premissas venha acompanhada de uma teoria filosófica que seja pelo menos tão plausível quanto eram inicialmente plausíveis as premissas.

2. Externismo e ceticismo

2.1. A exigência cética

O externismo epistemológico é a teoria segundo a qual os principais componentes do conhecimento são externos à mente do sujeito. De acordo com o externismo, se um sujeito sabe uma proposição qualquer, então uma condição ou conjunto de condições substanciais ao conhecimento, C , externos à mente do sujeito, foram cumpridos. Por exemplo, se S sabe que α , então a sua crença de que α foi produzida por um mecanismo cognitivo funcionando apropriadamente num ambiente adequado. E o mais importante, o sujeito não precisa saber que C foi cumprida, ou mesmo formar qualquer crença sobre a confiabilidade de seus mecanismos cognitivos. Por exemplo, antes de me sentar para escrever este texto, eu estava a preparar uma xícara de café. Ao pegar os ingredientes no armário, eu sabia que o pó de café, o coador, etc. estavam lá. A minha crença de que tais coisas estavam lá é o que me moveu até o armário para pegá-las. Porém, no momento em que eu me lembrava de que elas estavam lá – e, por conseguinte, sabia – eu não formei qualquer crença sobre a confiabilidade da minha memória. E se fosse necessário que eu soubesse que minha memória estava a funcionar apropriadamente naquele momento para saber que estava ali o pó de café, invariavelmente um regresso infinito se daria. Pois, para saber α eu teria de saber que α foi produzida de maneira confiável; e para saber que α foi produzida de maneira confiável, eu teria de saber que a minha crença de que α foi produzida de maneira confiável foi também produzida de maneira

confiável, e assim *ad infinitum*. Ao recusar a exigência de saber que C foi cumprida, o externismo evita tal situação indesejável.

E de maneira ainda mais interessante, evita o desafio cético radical. Considere a premissa central do argumento cético:

(1) Se S sabe que tem mãos, então sabe que não é um BIV.

Aparentemente tal premissa não diz qualquer coisa sobre a crença de S ser ou não formada de maneira confiável ou ter sido formada por um mecanismo cognitivo funcionando apropriadamente, etc. etc. Mas é isso, em última instância que está em jogo. Afinal, se sou um cérebro numa cuba sendo enganado sobre o mundo à minha volta, num sentido relevante alguma condição externa substancial deixou de ser cumprida.

Para ver que isso é assim, vamos considerar primeiro a clássica hipótese cartesiana do gênio maligno. Suponhamos que eu seja uma *res cogitans* tendo minhas aparentes crenças perceptuais sendo formadas pelo poder do gênio maligno. Suponhamos também que todas essas crenças sejam falsas. Em primeiro lugar, se considerarmos somente a confiabilidade, o mecanismo cognitivo que as está produzindo será infiel, uma vez que tende a produzir mais crenças falsas do que verdadeiras. Contudo, se o gênio maligno me transportasse para um mundo normal – um mundo em que sou tanto uma *res cogitans* quanto uma *res extensa* – meus mecanismo perceptuais se tornariam confiáveis, uma vez que tenderiam agora a produzir mais crenças verdadeiras do que falsas. Isso sugere que o ambiente é essencial para a confiabilidade; isto é, o mecanismo tem de estar funcionando apropriadamente num ambiente adequado. Assim, a pergunta “Você sabe que não está sendo enganado por um gênio maligno” pode ser posta da seguinte maneira:

(PC) Você sabe se (a) as suas crenças perceptuais estão sendo formadas por um mecanismo que tende a produzir mais crenças verdadeiras do que falsas, e se (b) o ambiente no qual essas crenças estão sendo produzidas é adequado para o funcionamento apropriado desse mecanismo?

Classicamente o cenário do gênio maligno tem sido usado para evidenciar que não somos capazes de oferecer uma resposta positiva tanto a PCa quanto a PCb. Vejamos. De que maneira poderíamos responder a PCa? Certamente uma resposta *a priori* não está disponível,

de modo que só podemos respondê-la *a posteriori*, com base em crenças perceptuais. Mas isso é circular, uma vez que estaríamos usando a percepção para justificar a percepção. E quanto a PCb, como respondê-la? Novamente só *a posteriori*. Eu teria de oferecer evidência sobre o meio, o que por si já pressupõe que tenho conhecimento sobre o meio. Novamente uma resposta circular. Assumindo que argumentos circulares são em geral epistemicamente viciosos, não é possível responder a PC.

O que foi dito acima se aplica ao cenário do cérebro numa cuba. Meu cérebro, ao ser mantido na cuba, mantém seu processamento de informações normalmente. De modo que os inputs dado pelo cientista, na medida em que são do mesmo tipo que a informação resultante do processamento da informação externa na retina, produzem o mesmo tipo de output que produziriam caso fossem dados inputs do mundo externo. Assim, nesse cenário cético, é verdade que meus mecanismos perceptuais estão funcionando apropriadamente. Contudo, não é o caso que estejam funcionando num ambiente adequado. O meu mecanismo visual, por exemplo, foi desenhado evolutivamente para captar determinados comprimentos de ondas luminosas, a determinada distância, etc. etc. Assim, num cenário em que sou um cérebro numa cuba, o ambiente não é favorável para a formação de crenças verdadeiras.

Portanto, em última instância, quando o cético nos pergunta se sabemos que não estamos num cenário cético, está querendo uma resposta a PC.

Ora, se considerarmos o externismo, então ainda que não tenhamos resposta a dar a PC, daí não se segue que não saibamos que não temos mãos. Saber as negações das hipóteses céticas não é uma exigência, portanto, que o externismo impõe ao conhecimento (cf. SOSA 1994; HILL 1996). O ceticismo, por conseguinte, não é algo com que o epistemólogo externista deva se preocupar. É claro que isso não é tudo. Os externistas ainda nos devem uma explicação de pelos menos por que o desafio cético é tão atraente e poderoso.

2.2. Externismo e fechamento

Note que, na medida em que o externista não precisa se preocupar com o ceticismo, é porque considera a Premissa 1 falsa. Mas isso implica a conjunção abominável. Além do mais, implica a rejeição do seguinte princípio de fechamento epistêmico:

(PF) Se S sabe que α , e sabe que $\alpha \rightarrow \beta$, então S sabe que β .

PF nos diz basicamente que o conhecimento é fechado sob implicação conhecida. A noção de fechamento provém da teoria de conjuntos: um conjunto S é fechado sob uma operação qualquer sse o resultado da aplicação dessa operação aos elementos de S é também um elemento de S . Por exemplo, a soma de quaisquer dois números naturais tem como resultado também um número natural. Do mesmo modo, PF nos diz que dado o conjunto de proposições Γ conhecidas por um sujeito, esse sujeito saberá todas aquelas proposições β que ele derivar ou deduzir de Γ .¹⁰³

É importante não esquecer de que o sujeito precisa derivar ou deduzir β a partir de Γ . Por exemplo, suponha que um observador de pássaros amador tenha lido num livro de ornitologia que apenas os tuiuiús possuem papo pelado e vermelho. Numa visita ao pantanal, ele vê uma ave de papo pelado e vermelho. Mas suponha que nosso observador, distraído, não tenha conectado as proposições (Apenas tuiuiús têm papo pelado e vermelho) e (Aquele pássaro possui papo pelado e vermelho). Ele se maravilhou com o pássaro e pensou “com uma imponência dessas, só pode ser o tuiuiú”. Assim, embora a sua crença de que o pássaro é um tuiuiú seja verdadeira, ela não pode ser conhecimento, uma vez que foi formada pelo sentimento de maravilhamento e não por qualquer mecanismo cognitivo. Isso a despeito do nosso observador saber que apenas tuiuiús têm papo pelado e vermelho e que o pássaro diante dele tem papo pelado e vermelho. Para que a sua crença de que aquele pássaro é um tuiuiú contar como conhecimento, nesse caso, ela tinha de ter sido inferida a partir das suas outras duas crenças que são conhecimento. E, certamente, essa inferência teria de ser competente, uma vez que um raciocínio confuso ou mal feito não é confiável. Assim, PF seria melhor formulado do seguinte modo:

(PF*) Se S sabe que α , sabe que $\alpha \rightarrow \beta$, deduz β de maneira competente a partir de α e $\alpha \rightarrow \beta$, e com isso passa acreditar que β , então S sabe que β .^{104,105} (cf. WILLIAMSON 2000, p. 117; HAWTHORNE 2004, pp. 32-35)

¹⁰³ O fechamento epistêmico também pode ser formulado em termos de outras propriedades epistêmicas como a justificação, garantia e racionalidade. Por exemplo, façamos “ $J\alpha$ ” representar “A crença de S de que α está justificada”; assim: (PFJ) [$J\alpha \wedge J(\alpha \rightarrow \beta)$] $\rightarrow J\beta$.

¹⁰⁴ Essa provavelmente não é a versão definitiva do princípio, uma vez que há algumas outras tentativas de contraexemplos a PF* (e.g. HAWTHORNE 2004 e KVANVIG 2006). Para modificações, veja DAVID & WARFIELD (2008). Não precisamos seguir por esse caminho; PF* é suficiente para a nossa discussão.

¹⁰⁵ Hawthorne (2004, pp. 33-34) distingue duas versões do fechamento: o fechamento de *premissa única* (FPU) e o fechamento de *múltiplas premissas* (FMP):

Contudo, levando-se em conta PF*, a simples caracterização do externismo não violará o fechamento. Note que tudo o que o externismo diz é que não precisamos saber que não-BIV (ou que as condições de confiabilidade e ambiente favorável foram cumpridas) para saber várias proposições sobre o mundo externo. Assim, o externismo em geral diz apenas que há mundos possíveis próximos em que S sabe que α muito embora não saiba que as condições externas tenham sido cumpridas. Mas a condição pode não ter sido cumprida simplesmente porque S não formou quaisquer crenças sobre a confiabilidade de seus processos cognitivos e seu funcionamento no ambiente. E se S não forma uma crença correspondente, obviamente não pode saber que não-BIV ou que as condições externas não foram cumpridas. O externismo implica na falsidade de PF, mas não de PF*. PF* diz que no caso em que S acredita em não-BIV com base numa inferência competente a partir daquilo que ele já sabe, S sabe que não-BIV.

Um externista que adote PF* não está dizendo que saber que não-BIV seja uma condição necessária para saber uma proposição qualquer sobre o mundo externo. Antes, está apenas dizendo que é suficiente para saber que não-BIV que acreditemos que não-BIV com base naquilo que já sabemos. Esse tipo de externismo tem uma vantagem dupla: não precisa responder ao ceticismo, embora seja capaz de fazê-lo. Contudo, não é esse o caso com a sensibilidade.

2.3. Sensitividade e fechamento

A chave da resposta nozickiana ao ceticismo está na insensitividade da nossa crença em hipóteses céticas. Consideremos a hipótese BIV. De acordo com a sensibilidade, se sei que não-BIV, então nos mundos mais próximos em que sou um cérebro numa cuba, não acredito que não sou um cérebro numa cuba.¹⁰⁶ No mundo mais próximo em que sou um cérebro numa cuba, acredito que tenho mãos, uma vez que estou sendo enganado pela simulação computacional. E a partir dessa crença e da crença de que se tenho mãos então não sou um cérebro numa cuba, infiro que não sou um cérebro numa cuba. Assim, no mundo mais

(FPU) Se S sabe que α , e deduz β de maneira competente a partir de α , e com isso passa acreditar que β , então S sabe que β .

(FMP) Se S sabe que $\alpha_1 \dots \alpha_n$, e deduz β de maneira competente a partir de $\alpha_1 \dots \alpha_n$, e com isso passa acreditar que β , então S sabe que β .

Note que PF* é uma instância de FMP. Seja como for, podemos, para os nossos fins, ignorar tal distinção.

¹⁰⁶ Isto é, se \neg BIV, então $\neg(\neg$ BIV) $\square \rightarrow \neg$ B(\neg BIV), sendo essa contrafactual equivalente a BIV $\square \rightarrow \neg$ B(\neg BIV).

próximo em que BIV, acredito que \neg BIV.¹⁰⁷ Portanto, a minha crença de que \neg BIV não conta como conhecimento.

Por outro lado, a crença de que tenho mãos é sensitiva, uma vez que nos mundos mais próximos em que não tenho mãos, mundos em que nasci maneta ou as perdi num acidente, não acredito que tenho mãos. E igualmente, naqueles mundos mais próximos (incluindo o atual!) em que tenho mãos, acredito que tenho. Portanto, a minha crença de que tenho mãos é conhecimento.

Juntando os resultados acima (e fazendo “ p ” representar a proposição (Tenho mãos)):

(\neg PF*) Sei que p , sei que $(p \rightarrow \neg$ BIV), deduzo de maneira competente que \neg BIV a partir de p e $(p \rightarrow \neg$ BIV), passo a acreditar em \neg BIV com base nessa dedução, **porém** não sei que \neg BIV.

Esse não é um resultado inesperado ou mesmo desagradável a Nozick:

É claro que qualquer abordagem que inclua como condição necessária para o conhecimento a condição subjuntiva $[\neg\alpha \Box \rightarrow \neg B\alpha]$, terá a consequência de que o conhecimento não é fechado sob implicação lógica conhecida. (NOZICK 1981, p. 207)¹⁰⁸

Na verdade, esse é exatamente o resultado visado pelos proponentes da sensitividade a fim de evitar o resultado cético:

É plausível (alguns diriam até que é inteiramente óbvio) que saibamos coisas sobre os nossos arredores materiais [...] E não é de todo claro que tenhamos tal conhecimento se, ao mesmo tempo, mantemos o fechamento. Alguém precisa ceder. Esse raciocínio não impressionará ao cético, é claro; mas caso se possa argumentar que a rejeição do fechamento [...] é a única maneira de se evitar o ceticismo, isso deveria convencer aqueles filósofos que acham o ceticismo tão “bizarro” ou “abominável” quanto a

¹⁰⁷ Levando-se em consideração a condição de aderência, $\alpha \Box \rightarrow B\alpha$, a hipótese BIV passa no teste: nos mundos próximos em que \neg BIV, acredito que \neg BIV.

¹⁰⁸ Nozick (p. 207) argumenta linhas antes nessa passagem (que omiti propositalmente) que a falha do fechamento do conhecimento de implicações conhecidas se deve ao fato de a condição de variação não ser fechada sob implicação conhecida. Mas esse argumento pode ser uma potencial falácia da composição: do fato de uma condição necessária ao conhecimento não ter uma propriedade P, não se segue que o conhecimento não tenha P. Ted Warfield (2004) e Tim Kraft (2014) defendem que Nozick de fato comete a falácia da composição, ao passo que Anthony Brueckner (2004), Peter Murphy (2006) e M. Yan (2013) defendem que o argumento de Nozick é um entinema, e procuram, cada um ao seu modo, explicitar a premissa implícita. Não tenho espaço aqui para avaliar essa discussão. E nem precisamos tomar partido, uma vez que, como o próprio Warfield aponta, a teoria de Nozick ainda assim implica a falsidade do fechamento.

rejeição do fechamento. [...] A única maneira de preservar o conhecimento de verdades comezinhas, aquelas verdades que todos dizem conhecer, é [...] abandonar o fechamento. (DRETSKE 2005 [2013], p. 32)

2.4. Intuição contrária

A despeito de a falha do fechamento parecer dar apoio à conjunção abominável, muitos epistemólogos têm sido convencidos pela força intuitiva do fechamento. Na introdução de um artigo contra o fechamento, Steven Hales nós dá uma ideia:

Muito se tem dito acerca do princípio de que o conhecimento é fechado sob implicação conhecida [...] E por uma boa razão – é um princípio bastante plausível. É devido a isso que Richard Feldman escreve, “alguma versão do princípio de fechamento (...) é certamente verdadeira. Com efeito, a ideia desse princípio ser falso espanta a mim e a outros filósofos, pois é uma das poucas ideias plausíveis a emergir na filosofia nos últimos tempos” (Feldman 1994, 1). Se o princípio é verdadeiro, então ele constitui uma descoberta valorosa para a epistemologia. Princípios desse tipo têm sido usados pelos fundacionistas para mostrar como o conhecimento derivado se baseia em fundações imediatamente cognoscíveis; Gettier dependeu de um princípio como esse para atacar o conhecimento como crença verdadeira justificada; e diz-se que os céticos dependem dele para construir seu argumento contra o conhecimento empírico. Além do mais, tal princípio nos permitiria deduzir aquilo que alguém sabe de outras coisas que já sabe. Isto é, deduzir no sentido forte – o fechamento não apenas possibilita a inferência de crença ou conhecimento provável, mas a dedução da existência ou presença deles com toda força que a lógica pode oferecer. Essa é uma das grandes notas promissórias que a lógica doxástica deixou para a epistemologia. (HALES 1995, p. 185)

E mesmo aqueles que reconhecem os eventuais problemas que surgem com o fechamento hesitam:

[N]ão estou absolutamente convencido da validade do princípio de fechamento dedutivo epistêmico, [...] embora não seja capaz de pensar em uma objeção; mas [nesse caso], os problemas aparentes aos quais ele conduz [o ceticismo] são ou apenas aparentes ou são melhores resolvidos abandonando-se outros princípios menos óbvios. (STINE 1976 [1999], p. 147)

Além do mais, os defensores (e demais simpatizantes) chamam atenção para uma função central do fechamento, a expansão (via dedução) do conhecimento:

[O] fechamento parece intuitivamente plausível. Se sei que todos os homens são mortais e que Sócrates é um homem, então parece que sei, ou pelo menos estou em posição de saber, que Sócrates é mortal porque sei, ou pelo menos estou em posição de saber as consequências lógicas daquilo que sei. Todo estudante de geometria euclidiana emprega implicitamente o fechamento quando diz saber um determinado teorema porque o derivou de outros teoremas. Em geral, parece que podemos sempre expandir o nosso conhecimento através da dedução. (KLEIN 2004, p. 165)

2.5. Segurança e fechamento

Mencionei anteriormente que alguns externistas gostariam de manter o fechamento. Esse é o caso dos neo-moorianos (SOSA 1999; PRITCHARD 2002, 2005, 2007, 2008). Os neo-moorianos concordam com a condicional do argumento cético, porém pensam poder saber que \neg BIV com base no conhecimento de proposições sobre o mundo externo, *e.g.*, a proposição de que tenho mãos. O princípio externista empregado, contudo, será o de segurança: a crença de S de que α é segura se, e somente se, nos mundos mais próximos em que S acredita que α é também o caso que α .

Consideremos a proposição ⟨Tenho mãos⟩. De acordo com a segurança, se sei que tenho mãos, não há qualquer mundo próximo em que formo a minha crença de que tenho mãos do mesmo modo como formei no mundo atual e essa crença é falsa. E em todos esses mundos próximos em que a proposição ⟨Tenho mãos⟩ é verdadeira, a proposição ⟨Sou um BIV⟩ é falsa. Considerando que os mundos próximos são aqueles mundos em que ⟨Tenho mãos⟩ é verdadeira, não há qualquer mundo próximo em que a proposição ⟨Sou um BIV⟩ seja verdadeira.

Levemos em conta agora que S forma a crença de que \neg BIV através do argumento mooriano: deduziu a crença de que \neg BIV de maneira competente a partir das proposições ⟨Tenho mãos⟩ e ⟨Se tenho mãos, então \neg BIV⟩. Como os mundos relevantes para a avaliação da crença de S são os mundos em que S tem mãos, nesses mesmos mundos a proposição de que S é um BIV não pode ser verdadeira. Logo, em todos os mundos próximos em que S acredita que \neg BIV com base no argumento mooriano é caso que \neg BIV. Por conseguinte, S sabe que \neg BIV.

Supondo que a segurança seja uma condição *prima facie* plausível, há tentativas de se mostrar que ela viola o fechamento. Por exemplo, Peter Murphy (2005) ofereceu a seguinte

variante dos celeiros de Kripke. Na roça dos celeiros falsos há 999 celeiros vermelhos reais e 1 celeiro falso azul. S olha para um celeiro azul e forma a crença correspondente. S também sabe que se há um celeiro azul em sua frente, há um celeiro em sua frente. S então deduz de maneira competente que há um celeiro em sua frente. A sua crença de que há um celeiro azul em sua frente é segura, uma vez que nos mundos mais próximos em que S acredita que há um celeiro azul em sua frente, é verdade que há um celeiro azul em sua frente. Porém, a sua crença de que há um celeiro em sua frente é insegura. Há um mundo próximo, o mundo em que S olha para o celeiro vermelho – que é um celeiro falso – acredita que há um celeiro em sua frente, muito embora seja falso que haja um celeiro em sua frente. Contudo, o exemplo não leva em consideração o modo como S formou a sua crença de que há um celeiro no mundo atual (PRITCHARD 2005; SOSA 2002; WILLIAMSON 2000; BERNECKER 2012). Naqueles mundos mais próximos em que S acredita que há um celeiro em sua frente inferindo-a das proposições \langle Há um celeiro azul em minha frente \rangle e \langle Se há um celeiro azul em minha frente, então há um celeiro em minha frente \rangle , não é o caso que a sua crença de que há um celeiro em sua frente seja falsa.

Contudo, como argumentei no capítulo anterior, se a segurança tem algum apelo intuitivo é porque ela é melhor vista como (e na pior das hipóteses, foi confundida com) a condição de aderência: $\alpha \Box \rightarrow B\alpha$. É fácil ver que a aderência não viola o fechamento. Se α implica β , não haverá qualquer mundo próximo em que α seja verdadeira e β falsa. E se nesses mundos S forma a sua crença do mesmo modo que no mundo atual, não poderá ter uma crença falsa. Portanto, ainda que a preservação do fechamento seja uma vantagem, não será a segurança que a irá obter.

2.6. Semi-sensitividade, fechamento e \neg BIV

A crença de que \neg BIV é supostamente insensitiva. Isso se deve à própria hipótese BIV ser posta de tal modo que nenhum sujeito esteja em posição de afirmá-la ou negá-la. Naqueles mundos em que sou um BIV, acredito que \neg BIV. E mesmo que eu fosse um BIV, não haveria mundos próximos em que acredito que sou um BIV. Mas nos concentremos, por agora, em dois pontos: a noção de um ambiente favorável e a subcondição de aderência. Talvez seja possível manter o fechamento e a sensibilidade.

Podemos apresentar o argumento a favor da impossibilidade de se saber $\neg BIV$ (“argumento $\neg K(\neg BIV)$ ” por brevidade) do seguinte modo:

- 1) Se $K\alpha$, então $B\alpha$ é sensitiva.
- 2) $B(\neg BIV)$ é insensitiva.
- 3) Logo, $\neg K(\neg BIV)$.

Em que condições uma crença é sensitiva? Nos casos em que tanto a variação quanto a aderência são cumpridas:

(SEN) $B\alpha$ é sensitiva sse (i) $\neg\alpha \Box \rightarrow \neg B\alpha$ e (ii) $\alpha \Box \rightarrow B\alpha$

Levemos em conta agora o ambiente adequado para se formar a crença de que $\neg BIV$. Na hipótese original, S forma a crença de que $\neg BIV$ com base no raciocínio mooriano. O seu cérebro está funcionando apropriadamente, porém dentro da cuba. Poderíamos dizer que tal ambiente não é adequado, o que me parece razoável. Contudo, no que diz respeito à capacidade de raciocínio, apenas o que é interno à mente de S é que conta, o ambiente externo à partida não conta. É claro então que se levarmos em conta primeiro a hipótese BIV, nunca teremos um ambiente favorável para a formação de qualquer crença empírica sobre o mundo externo.

Proponho que invertamos as coisas. Vamos considerar primeiro a hipótese de que tenho mãos e o ambiente favorável é à formação de tal crença. Sendo assim, cenários BIV estão descartados. Em tais ambientes, o resultado do meu raciocínio mooriano não será uma crença falsa. Assim, não há mundos BIV próximos em que eu acredite que $\neg BIV$. Note que a condição de aderência é satisfeita – $\neg BIV \Box \rightarrow B(\neg BIV)$: naqueles mundos em que $\neg BIV$, o ambiente é favorável à formação da minha crença de que tenho mãos e ao raciocínio, acredito que $\neg BIV$. E essa restrição devida ao ambiente faz com que a variação seja vacuamente satisfeita: não há qualquer mundo próximo que torne falsa a condicional $BIV \Box \rightarrow \neg B(\neg BIV)$.

Chamemos a uma crença “semi-sensitiva” se, e somente se, for (i) aderente e (ii) não-variante. Por sua vez, uma crença é não-variante se e somente se a condicional da variação for vacuamente satisfeita. Uma crença ser não-variante não é a mesma coisa que ser invariante. Uma crença é invariante se e somente se a condicional da variação for falsa. E diremos, por

fim, que uma crença é insensitiva se, e somente se, ou (i) não for aderente ou (ii) for invariante.

Ora, de acordo com as definições acima, a crença de que $\neg BIV$ é semi-sensitiva, e não insensitiva. Portanto, a premissa 2 do argumento $\neg K(\neg BIV)$ é falsa.

Considere agora a seguinte reformulação da primeira premissa do argumento $\neg K(\neg BIV)$:

1*) Se $K\alpha$, então ou $B\alpha$ é sensitiva ou $B\alpha$ é semi-sensitiva.

O quão convencido estou da verdade de (1*)? Talvez não muito. Contudo, vale notar que nossas crenças em verdades necessárias *a posteriori* são semi-sensitivas.¹⁰⁹ Ora, como a premissa 2 de $\neg K(\neg BIV)$ é falsa porque a crença de que $\neg BIV$ é semi-sensitiva, temos agora o seguinte argumento:

1*) Se $K\alpha$, então ou $B\alpha$ é sensitiva ou $B\alpha$ é semi-sensitiva.

2*) $B(\neg BIV)$ é semi-sensitiva.

3*) Logo, $\neg K(\neg BIV)$.

que é claramente um argumento inválido.

Talvez uma razão para aceitarmos a semi-sensitividade seja algo como a intuição dos proponentes da segurança. Nesse caso, poderíamos dizer que os proponentes da segurança tinham em mente algo como a semi-sensitividade, muito embora a tenham tentado formular como $B\alpha \square \rightarrow \alpha$. Mas, uma vez que consideremos tal apenas como um equívoco,¹¹⁰ talvez eles se dispusessem a chamar uma crença de segura quando for semi-sensitiva.¹¹¹ Nesse caso, diríamos que S sabe que α se, e somente se, a sua crença de que α é sensitiva ou segura.

¹⁰⁹ No que diz respeito ao conhecimento de verdades necessárias, parte da motivação de Nozick para a introdução da condição de aderência foi exatamente para dar conta das verdades necessárias, uma vez que a condição de variação é vacuamente satisfeita. Contudo, ele não pensou na mesma estratégia para o conhecimento de $\neg BIV$.

¹¹⁰ Não estou certo se um proponente da segurança aceitaria tal diagnóstico. Contudo, especulo que o fato (por um motivo que me escapa) de ao longo dos anos os epistemólogos terem focado quase toda atenção à condição de variação e a chamado de sensitividade, e o fato de a aderência parecer apenas um remendo para o conhecimento de verdades necessárias, tenham sido responsáveis pelo esquecimento da aderência.

¹¹¹ Vale lembrar que a segurança teria também de ser formulada respeitando as condições de restrição do método de formação e do ambiente estipulados pela teoria da função apropriada (ou qualquer outro tipo próximo de confiabilismo). Cf. Capítulo IV.

Ainda que tal posição pareça implausível, e, não insistirei no contrário, na seção 3.2 abaixo ofereci um argumento a favor de sua implausibilidade.

3. Contraexemplos ao fechamento e insensitividade

No caso do argumento acima não parecer convincente, podemos tentar outra via: restringir o fechamento a um conjunto de proposições. Nesse caso, diríamos que o conhecimento é fechado sob a implicação conhecida, *com exceção* de proposições tais e tais. Contudo, essa manobra soa bastante *ad hoc*, uma vez que parece ser uma tentativa desesperada de compatibilizar sensibilidade e fechamento. Por isso seria razoável esperar que algo como

(PF_{ad hoc}) Se S sabe que α , sabe que $\alpha \rightarrow \beta$, deduz β de maneira competente a partir de α e $\alpha \rightarrow \beta$, e com isso passa acreditar que β , então S sabe que β , *a menos* que β seja \neg BIV.

fosse evitado.

No que se segue, apresentarei algumas violações ao fechamento que podem ser apreciadas o tanto quanto possível independentemente da sensibilidade. Essas violações ocorrem, não obstante, porque as crenças nas implicações envolvidas são insensitivas. E na medida em que essas proposições puderem ser agrupadas uniformemente com proposições que descrevem a negação de um cenário cético, poderemos completar a cláusula “a menos que β seja...” de maneira não *ad hoc*.

3.1. Contraexemplos e dificuldades

Além do puzzle cético, há outros casos que ajudam a motivar o abandono do fechamento. A minha intenção não é discuti-los em pormenor – o que seria matéria para outra tese – mas apenas tentar afastar a impressão de que o abandono do fechamento é uma manobra desesperada para se evitar o ceticismo, como sugere a passagem de Dretske (2005) citada acima. Apresentarei dois contraexemplos e duas dificuldades para o fechamento relacionadas aos contraexemplos.

Começemos pelos contraexemplos. O primeiro é clássico caso da ZEBRA (DRETSKE 1970). Considere as proposições seguintes.

(Z) O animal na jaula é uma zebra.

($\neg M$) O animal na jaula não é uma mula pintada para parecer uma zebra.

Lindinha vai com seu pai ao zoológico e olha para a jaula em que há um equino, com cerca de 1,3 metros de altura, listrado em preto em branco (preto com listras brancas!) e acredita que é uma zebra. Ela tem boa evidência para acreditar que é uma zebra. Lindinha, portanto, sabe que (Z). Seu pai, porém, maliciosamente, pergunta-lhe se ela sabe que aquele animal não é uma mula muito bem pintada para parecer uma zebra. Talvez a zebra do zoológico tenha morrido e para que a jaula não ficasse vazia, os administradores do zoológico decidiram fazer esse truque. Afinal, fosse uma zebra ou uma mula disfarçada de zebra, Lindinha teria a mesma evidência. Ela, portanto, não sabe que ($\neg M$). Sendo assim, ainda que Lindinha se atentasse para o fato de que (Z) implica ($\neg M$), e inferisse de maneira competente que ($\neg M$), a partir de seu conhecimento de que (Z) e (Z) \rightarrow ($\neg M$), parece plausível ainda negar que Lindinha saiba que ($\neg M$). Isso quer dizer, então, que a antecedente de PF* é satisfeita mas não a sua consequente. PF* é, portanto, inválido.

Muitos pensam que ZEBRA não é convincente. Mas há casos que me parecem mais convincentes. Considere (D), a proposição <Tenho 50 reais no bolso>. Sei que (D), uma vez que acabei de sair da padaria com um troco de 50 reais. Considere agora ($\neg F$), a proposição <A cédula em meu bolso é um papel especial preparado para parecer uma nota de 50 reais>. Não sou capaz de distinguir uma nota real de uma nota falsificada, de modo que se eu tivesse uma nota falsa no bolso, teria do mesmo modo acreditado que tinha 50 reais no meu bolso. E isso a despeito de eu saber que (D) implica ($\neg F$).

Chamarei “proposição zebrada” qualquer proposição que o sujeito use para negar que está num cenário *quase* cético.¹¹² Um cenário que descreve uma alternativa enganadora à proposição alvo de conhecimento – uma proposição que diz que as coisas se *parecem* ao sujeito como se fosse aquilo que a proposição alvo descreve e que a evidência perceptual disponível ao sujeito não é suficiente para discriminar entre as duas. São exemplos de proposições zebradas: <Não é o caso que aquele indivíduo é um Nexus 6>, <Não é o caso que aquele seja um celeiro de fachada preparado para parecer com um celeiro real>, <Não é o caso

¹¹²Um cenário é *quase cético* quando a hipótese alternativa (*quase cética*) puder, num contexto apropriado, ser excluída pelo sujeito. Por exemplo, Lindinha poderia ter excluído a alternativa de o animal ser uma mula disfarçada se fosse uma zoóloga experiente. No caso da hipótese cética, porém, parece não haver um contexto apropriado em que o sujeito possa afastá-la. A distinção, portanto, entre negações de hipótese céticas e proposições zebradas é uma questão de grau.

que o relógio de bolso do Sr. X seja à bateria disfarçado para parecer de corda), (Não é o caso que aquele rapaz que chamei pelo nome de Tweedledee seja seu irmão gêmeo, Tweedledum, se fazendo passar pelo primeiro), etc.

É claro que, assim como no caso do argumento mooriano, um opositor poderia apelar à intuição contrária. Lindinha sabe que ($\neg M$). Peter Klein (1995), por exemplo, acusa Dretske de não distinguir entre dois princípios – o *fechamento* e a *transmissão de evidência*. Segundo ele, Dretske está correto em apontar que a evidência que S tem para saber que α não é adequada para S saber que β (quanto α implica β) – isto é, a evidência não é transmitida das premissas para a conclusão. Contudo, o fechamento (corretamente interpretado) diz que as premissas do argumento, uma vez que contam como conhecimento,¹¹³ é que funcionam como evidência a favor da conclusão.¹¹⁴

Mas não penso que essa manobra vá funcionar. Digamos que uma evidência *e* confirma uma hipótese *h* (ou crença, ou proposição, etc.) somente se *e* aumentar a probabilidade de *h* – i.e., $P(h | e) > P(e)$.¹¹⁵ Por exemplo, entro no elevador e formo a crença macabra de que o elevador vai cair. Inicialmente, podemos considerar que essa crença não dispõe de qualquer base razoável. Mas suponha que durante a subida, o elevador comece a emperrar, que se inicie uma pane elétrica e, por fim, eu escute barulho de cabos se rompendo. Agora tenho alguma evidência para acreditar que o elevador irá despencar. Ou seja, essas informações aumentaram a probabilidade da minha crença ser verdadeira. Suponha, contudo, que eu me lembre que o zelador do prédio é um antigo amigo meu e que adora me pregar peças (inclusive as mais sem graças). Ele seria muito bem capaz de aprontar comigo naquele elevador. Sendo assim, essa nova informação faz com que a probabilidade da minha crença ser verdadeira diminua – e, talvez, a ponto de eu deixar de acreditar que o elevador despencará. Dito de outro modo, essa nova informação *anula* a garantia que eu tinha para a minha crença.¹¹⁶ Empreguemos agora essa mesma linha de raciocínio à crença do mooriano de que $\neg BIV$. Suponha que inicialmente ele considere a força o argumento cético e, por isso,

¹¹³Na verdade, Klein discute o fechamento da justificação.

¹¹⁴ Não confundir essa ideia com a tese de que conhecimento é evidência ($E = K$), defendida por Timothy Williamson (2000).

¹¹⁵ Para fins argumentativos, usarei “probabilidade” e “grau de crença” como sinônimas. Com isso, contudo, não estou supondo que a probabilidade de uma hipótese seja a medida de um grau de crença puramente subjetivo – como fazem os bayesianos subjetivistas. Pode muito bem ser o caso de que o grau de crença de um sujeito represente adequadamente a probabilidade objetiva de uma hipótese – como supõem os bayesianos objetivistas. Seja como for, não precisamos tomar partido aqui.

¹¹⁶ Estou supondo, obviamente, que evidência infirmadora funciona como um anulador.

seu grau de crença em \neg BIV é relativamente baixo. Esperamos então que o argumento mooriano eleve (a um grau razoável) a probabilidade da sua crença de que \neg BIV. O mooriano terá de olhar para as suas mãos e constatar que sabe isso. Ora, mas uma vez que ele reconheceu que não consegue afastar a hipótese cética inicialmente, ele não pode atribuir um grau de crença razoavelmente elevado à proposição de que tem mãos. Responder ao cético é como lembrar do meu amigo aldrabão. Se, nesse último caso, reconhecer que ele pode estar pregando-me uma peça anula minha crença na queda do elevador, reconhecer também que posso ser vítima de um gênio maligno anula a minha crença de que tenho mãos! É claro que há uma saída: dizer que não preciso afastar a hipótese cética a fim de que eu atribua um grau de crença razoável à hipótese de que tenho mãos. Em outras palavras: posso saber que tenho mãos e não saber que \neg BIV. Mas isso implica na rejeição do fechamento, e a princípio o mooriano não é simpático a tal resultado.¹¹⁷

Passemos ao segundo contraexemplo, PROPOSIÇÕES LOTÉRICAS (VOGEL 1990).¹¹⁸ Uma proposição α é lotérica pelo menos se (i) α é altamente provável, e (ii) ainda assim S não é capaz de saber α . Por exemplo, ao fazer um jogo na Mega Sena, sei que as chances de eu acertar os seis números é ínfima (algo como uma em 50 milhões!). Assim, a proposição \langle Não ganharei na Mega Sena \rangle é altamente provável. Não obstante, não sei que perderei; afinal, se soubesse, eu não teria feito o jogo! Considere agora a seguinte proposição que sei, dada a minha condição de desempregado: \langle Não irei a Acapulco nas próximas férias \rangle . Mas \langle Não irei a Acapulco nas próximas férias \rangle implica \langle Não ganharei na Mega Sena \rangle , supondo obviamente que farei o jogo. Ora, dado o fechamento, eu teria de dizer que não sei que não irei a Acapulco; o que é implausível. Esquematicamente:

(\neg A) Não irei a Acapulco nas próximas férias.

(\neg L) Não ganharei na Mega Sena.

¹¹⁷ Jonathan Vogel (1990) pensa que ZEBRA não constitui um contraexemplo ao fechamento. Seu argumento é basicamente o de que o *conhecimento de fundo* do sujeito o dá “justificação para negar que o animal seja uma mula, de modo que [o sujeito] sabe que não é” (p. 14). Mas há pelo menos duas dificuldades. (i) Ainda que o conhecimento de fundo do sujeito lhe dê a justificação, essa justificação é proposicional, e o conhecimento exige justificação doxástica (se exigir alguma justificação). Assim, teríamos de formular um princípio de fechamento que levasse em conta o conhecimento de fundo. (ii) Supondo que o conhecimento de fundo forneça tudo o que evidencialmente necessário e suficiente para saber que o animal não é uma mula pintada, esse conhecimento não seria produto de uma inferência com base nas premissas como exige PF*.

¹¹⁸ Hawthorne (2004) oferece o tratamento mais completo até agora às proposições lotéricas. Ele defende uma posição – o *invariantismo moderado sensitivo* – que supostamente explica a força das proposições lotéricas, mantém o fechamento, não sucumbi ao ceticismo, e não é contextualista.

Sei que ($\neg A$), sei que ($\neg A$) implica ($\neg L$), mas não sei que ($\neg L$). A menos que rejeitemos o fechamento, teríamos de considerar que não sei que ($\neg A$).

Outro caso. Com todas as reservas feitas e passagens compradas, (como um bom mineiro!) sei que (G): nas férias estarei em Guarapari. E (G) implica que ($\neg C$): não terei um ataque cardíaco fulminante no meu último dia de trabalho. Contudo, não sei que ($\neg C$); de acordo com o dito popular, “não dá para saber quando encontraremos a morte”. Porém, não parece razoável eu dizer que não sei que estarei em Guarapari nas férias só porque não sei quando vou morrer.

Passemos às dificuldades. A primeira delas envolve o CONHECIMENTO FÁCIL (COHEN 2002). Considere:

(V) Aquela mesa é vermelha.

($\neg B$) Aquela mesa não é branca iluminada para parecer vermelha.

Lindinha vai com o pai até uma loja de decoração, aponta para uma mesa a certa distância e diz que vai querer “aquela mesa vermelha”. Ela viu o objeto à boa distância, à iluminação adequada, etc. e lhe pareceu uma mesa vermelha. Lindinha tem evidência adequada para acreditar que (V). Porém, essa mesma evidência – a mesa lhe parecer (ser apresentada como) vermelha – não é boa o bastante para ela acreditar que a mesa não é branca iluminada para parecer vermelha. Afinal, àquela distância, se houvesse uma fonte de luz vermelha iluminando uma mesa branca, a mesa lhe pareceria vermelha. Portanto, Lindinha não sabe que ($\neg B$). Isso a despeito de (V) implicar ($\neg B$). E mesmo que Lindinha recorra a um argumento mooriano, dizendo saber que (V), saber que (V) implica ($\neg B$), e que por isso sabe que ($\neg B$), ainda assim a evidência disponível parece insuficiente. Do contrário, seria muito *fácil* obter conhecimento de proposições como ($\neg M$), ($\neg V$) e ($\neg BIV$).¹¹⁹

Além do conhecimento fácil, o fechamento torna também fácil excluir toda e qualquer evidência contra determinada crença, considerando-a como EVIDÊNCIA ENGANADORA (HARMAN 1973; COLLINS 2006), isso constitui a segunda dificuldade. Considere o seguinte caso. Antes de sair de casa conferi todas as portas e janelas para garantir que estavam trancadas. Assim, sei que a minha casa está trancada. E da proposição de que a minha casa

¹¹⁹ Para mencionar novamente o caso da nota falsa, se oferecer um argumento mooriano para dizermos que sabemos que a nota não é falsa fosse o bastante, então não precisaríamos de dispositivos para checar a autenticidade das notas.

está trancada, segue-se que ela não foi arrombada. Suponha, porém, que vejo um famoso larápio da vizinhança passar com uma edição comemorativa raríssima em vinil das gravações completas de Robert Johnson; provavelmente eu era um dos únicos no Brasil a possuir tal cópia. Tenho agora alguma evidência de que entraram em minha casa, algo que só poderiam ter feito arrombando-a. Ora, mas se sei que minha casa está trancada e sei que se a minha casa está trancada então não foi arrombada, então sei que a minha casa não foi arrombada (e meu disco está a salvo!). Essa evidência tem, portanto, de ser enganadora. Uma vez mais, suponha que adquiro nova evidência de que minha casa foi arrombada – a minha vizinha me liga dizendo que ouviu um baralho de gavetas sendo reviradas dentro da minha casa. Ora, mas se sei que a minha casa está trancada, então sei... e, portanto, essa evidência é também enganadora. Esse procedimento pode ser repetido inúmeras vezes. Mas isso produz um resultado inaceitável: frente a qualquer conjunto de evidência contra a minha crença, posso ignorá-lo. Podemos generalizar o argumento:

- | | |
|--|----------------------|
| 1. Sei que α | [Premissa] |
| 2. Sei que α implica β | [Premissa] |
| 3. Tenho evidência a favor de $\neg\beta$ | [Premissa] |
| 4. Sei que β | [1,2 PF*] |
| 5. Se sei que β , então qualquer evidência a favor de $\neg\beta$ é enganadora | [Premissa] |
| 6. \therefore Não tenho evidência a favor de $\neg\beta$ | [4,5 – 3, I-] |

CONHECIMENTO FÁCIL e EVIDÊNCIA ENGANADORA podem ser gerados tanto com proposições zebradas quanto com proposições lotéricas. Sendo assim, no caso de insistirmos no fechamento, teríamos como resultado o conhecimento fácil de que não ganharei na loteria (no caso de eu ter feito uma aposta, é claro!); ou o conhecimento fácil de que não terei um enfarto nos anos vindouros, etc. E teríamos também de considerar que um diagnóstico médico de problemas cardíacos é uma evidência enganadora contra a minha crença de que não terei um ataque cardíaco.¹²⁰

É interessante que houvesse uma explicação uniforme de por que os casos acima violam o fechamento. Poderíamos começar com a sugestão de que a única classe de

¹²⁰ Há ainda outros puzzles e paradoxos que envolvem o fechamento, *e.g.*, o paradoxo de McKinsey (McKINSEY 1991), o paradoxo da loteria (KYBURG 1961), e o paradoxo do prefácio (MAKINSON 1965). Não os apresentarei aqui, contudo; isso alongaria desnecessariamente esta seção.

proposições para as quais haverá falha de fechamento são as proposições lotéricas, e que todos os outros casos apresentados envolvem proposições lotéricas. Mas como já notado por Vogel (1990) negações de hipóteses céticas não são proposições lotéricas. Ele sustenta que uma proposição α é lotérica sse (i) α for altamente provável, (ii) $\neg\alpha$ não for anormal, (iii) as alternativas a α forem tão prováveis quanto α , e (iv) for arbitrário preferir α a qualquer uma das suas alternativas.¹²¹ Por exemplo, no caso da proposição ⟨Não ganherei na Mega Sena⟩, (i) ela é altamente provável, (ii) caso eu ganhasse o prêmio, isso não seria anormal, (iii) para qualquer outro apostador da Mega Sena as suas chances de perder são tão altas quanto a minha, e (iv) por isso seria arbitrário que eu acreditasse que somente eu perderia o jogo. Porém, no que diz respeito à proposição ⟨Não sou um cérebro numa cuba⟩ não é possível dizer se para qualquer outra pessoa a sua chance de ser um cérebro numa cuba é aproximadamente a mesma da minha. No caso da loteria, tenho a garantia de que (ainda que todos os apostadores tenham a mesma ínfima chance de ganhar) alguém vai ganhar. Porém, não sei de qualquer pessoa que passou a ser um cérebro numa cuba. Aparentemente, dizer “Fulano foi o último a ser raptado e ter virado um cérebro numa cuba” não tem o mesmo peso que dizer “Sicrano foi o último ganhador da Mega Sena”. Por conseguinte, não há como atribuir um valor à probabilidade de alguém ser (ou não) um cérebro numa cuba. Além disso, uma vez que não sabemos qual a frequência com que as pessoas passam a ser incubas, não podemos determinar a sua normalidade ou anormalidade.

Vogel (p. 17) também argumenta que ZEBRA não é um caso em que esteja envolvida uma proposição lotérica. Para dizer o mínimo, encontrar uma mula disfarçada de zebra num zoológico seria completamente anormal. Não há a mínima razão para pensar que, embora improvável, alguém vá encontrar uma mula disfarçada num zoológico.

No que segue veremos de que modo a sensibilidade é capaz de fornecer um tratamento uniforme aos casos de falha do fechamento.

3.2. Um diagnóstico: insensibilidade

Já vimos que a crença em \neg BIV é insensitiva. Há pelo menos um mundo próximo em que sou um cérebro numa cuba, raciocino a partir do argumento mooriano, e acredito na proposição

¹²¹ Na verdade, Vogel (p. 16; 18) apresenta apenas três condições: (I) anormalidade, (II) confiança da evidência estatística, e (III) não-arbitrariedade. Essas condições são as condições (ii)-(iv) acima. Introduzi a condição (i) para tornar as coisas mais explícitas (que fica também óbvia na discussão de Vogel).

⟨Sou um cérebro numa cuba⟩. O mesmo se dará com os casos apresentados na seção anterior. Assim, na medida em que todos aqueles casos são casos genuínos de falha do fechamento, o são em virtude das proposições zebradas e das proposições lotéricas serem insensitivas. Vamos a eles.

Comecemos pelas proposições lotéricas, mais especificamente pela proposição lotérica original, ⟨Meu jogo na loteria perderá⟩. Dretske (1971) começa por motivar a sua teoria das razões conclusivas – que é a proto-versão da sensibilidade – notando que a condicional contrafactual

Se eu fosse ganhar na loteria, então a chance de eu ganhar não seria $1/m$ (em que m é o número total de apostas)

é falsa. Dado que a loteria é justa, e dado o grande número de apostas, cada um dos bilhetes tem a mesma chance de ganhar. Suponha que 10.000 bilhetes tenham sido vendidos. O meu bilhete tem a chance de $1/10.000 = 0,0001$ de ser sorteado, assim como cada um dos outros 9.999 restantes. Muito embora seja ínfima a probabilidade do meu bilhete ser sorteado, o mesmo vale para os outros bilhetes. E como o sorteio é aleatório, nenhum dos bilhetes (nem mesmo o sorteado) tem uma probabilidade diferente de 0,0001. Assim, ainda que eu ganhasse, a chance de eu vencer seria $1/m$. Portanto, não sei que meu bilhete vai perder. Posto no jargão dos mundos possíveis: há pelo menos um mundo próximo no qual ganho na loteria e as minhas chances não são diferentes de $1/m$.

Para tornar as coisas ainda mais claras: se sei que o meu bilhete vai perder, então *os mundos mais próximos em que meu bilhete ganhar (os mundos em que ele não perde) são mundos em que não acredito que ele vai perder (acredito que vai ganhar) com base na informação de que as chances do meu bilhete ser sorteado são de $1/m$* . Mas é claro que há pelo menos um mundo em que ganho e ainda acredito que não ganharia. É fácil ver isso. Cada possibilidade de um bilhete ser sorteado conta como um mundo possível. Temos então 10.000 mundos possíveis. Em todos esses mundos acredito que meu bilhete perderá. Mas há um mundo em que o meu bilhete é sorteado; e lá eu ainda acredito que ele perderá. Portanto, não sei que meu bilhete vai perder. Minha crença é *insensitiva*.

Vejamos outro exemplo de proposição lotérica, o caso do carro roubado (VOGEL 1990, pp. 15-16). Assim como no caso anterior, não posso saber a proposição ⟨Meu carro não

foi roubado quando estacionado na área X), em que X é uma área qualquer com alto índice de roubo de carro. Suponha que as estatísticas de roubo de carro na área X sejam 2 carros roubados por dia (em qualquer hora), não tendo os ladrões preferência em modelo, marca e estado de conservação. Suponha ainda que, por dia, mil carros sejam estacionados ali. As chances de o meu carro ser roubado são pequenas, 0,002. Seria razoável que eu acreditasse que naquele momento meu carro não foi roubado. Mas os outros 999 carros têm aproximadamente os mesmos 0.002 de chance de serem roubados. Ainda que meu carro não tenha sido roubado, se tivesse sido, eu continuaria acreditando que não foi. Ou seja: há pelo menos um mundo em que meu carro foi roubado e que acredito que não foi roubado, pois reconheço as ínfimas chances de 2 em 1.000 dele ser roubado. Novamente, minha crença é insensitiva.¹²²

Passemos às proposições zebradas. Nos mundos possíveis mais próximos em que não há uma zebra, mas uma mula disfarça na jaula, e nesses mundos Lindinha forma a sua crença apenas com base na aparência do animal, ela acredita falsamente que há uma zebra na jaula. Ou seja, a crença de Lindinha de que o animal não é uma mula disfarçada é insensitiva. Ou considere o caso em que formo a crença de que Tweedledee está do outro lado da rua. Nos mundos mais próximos em que é Tweedledum que está lá, continuo acreditando que é Tweedledee. Assim, minha crença é novamente insensitiva.

Resumindo: a minha crença de que \neg BIV, as minhas crenças em proposições lotéricas e em proposições zebradas são todas insensitivas. Além do mais, proposições lotéricas e zebradas podem ser usadas para construir cenários céticos. A diferença entre elas parecer ser apenas de grau. Podemos, portanto, agrupar essas proposições numa classe minimamente uniforme – chamemos-lhe ZLB (de “zebra”, “lotérica” e “BIV”). Podemos agora completar a cláusula “a menos que β seja...” com ZLB.

Ora, mas uma vez que \neg BIV poderia ser vista como semi-sensitiva, parece que as proposições lotéricas e zebradas também deveriam ser vistas assim. Ou seja, uma vez que consideramos o ambiente adequado para a formação, por exemplo, da crença de que aquele animal é uma zebra, não haverá qualquer mundo próximo em que acredito que o animal é uma

¹²² Obviamente que poderíamos fazer restrições para tornar a crença sensitiva. Por exemplo, suponha que meu sistema anti-furtos envie uma mensagem para o meu celular caso meu carro seja arrombado e haja uma tentativa de ligação. Além do mais a sua localização é constantemente monitorada via GPS. Nesse caso, a minha crença de que meu carro não foi roubado seria sensitiva dado o modo como forma a crença.

zebra e seja falso que o animal é uma zebra (porque é uma mula disfarçada). E como só os mundos em que o animal é uma zebra são os relevantes, a condição de aderência é satisfeita e a de variação vacuamente satisfeita.

Essa manobra, contudo, não funcionará para as proposições lotéricas. Há mundos em que faço o jogo, acredito que vou perder, mas ganho. A minha crença de que não vou ganhar na loteria é invariante. E há mundos em que faço o jogo, acerto o resultado, mas continuo acreditando que perderei; a minha não é aderente. Seja como for, crenças em proposições lotéricas são insensitivas.

E já agora, podemos levantar a objeção do conhecimento fácil. Considerar a semi-sensitividade faz com que saibamos muito facilmente que \neg BIV e que não é o caso que o animal é uma mula disfarçada. Uma motivação externista para aceitar o conhecimento fácil talvez fosse dizer que uma vez que não precisamos saber de antemão que não estamos num cenário cético para saber proposições sobre o mundo externo, saber que não estamos num cenário cético não pode ser tão difícil. O problema com essa resposta, penso, reside na própria dialética do debate entre externistas e internistas. Uma das mais reconhecidas queixas para com as teorias internistas do conhecimento é que sempre acabam por dar no ceticismo. Exigir, por exemplo, que o sujeito tenha de ter acesso em primeira pessoa a todos os fatores que tornam uma crença conhecimento é demasiado forte. E como mencionei no início deste capítulo, o externismo abre mão de tal exigência, o que faz com que ele não precise refutar o cético antes de afirmar o que quer que seja sobre o mundo fora de sua mente. Isso sugere que o externista reconhece desde o início a dificuldade de refutar hipóteses céticas.

É claro que há a possibilidade de se perceber o quão fácil é saber que \neg BIV depois que se adota o externismo e se chega à conclusão de que ela é semi-sensitiva. Contudo, vale lembrar que se S refuta uma hipótese cética, isso implica que ele sabe que seu mecanismo cognitivismo está funcionando apropriadamente num ambiente adequado. A esse ponto o externista não conseguirá evitar as acusações de *facilitação* [*bootstrapping*] (VOGEL 2000; COHEN 2002; SOSA 2009). Suponha que o medidor de gasolina do meu Fusca seja confiável. Assim, sempre que eu formo a crença de que o tanque está cheio através da leitura do medidor, o tanque está cheio. Depois de muitas vezes tendo acreditado corretamente que o tanque está cheio, infiro a partir disso que o medidor é confiável. Contudo, a minha crença de que o medidor é confiável foi inferida a partir de crenças que pressupõem a confiabilidade do

medidor. Em última instância, a minha crença de que o medidor é confiável depende do medidor ser confiável, o que é circular. O meu conhecimento de que o medidor é confiável não pode ser tão fácil assim. Do mesmo modo, quando digo, por exemplo, que minha visão é confiável (e está funcionando no ambiente adequado), não há outro modo de fazê-lo a não ser dependendo de outras crenças visuais.¹²³ Portanto, não parece ser uma boa ideia insistir que crenças ZLB sejam semi-sensitivas.

4. Fechamento restrito

Penso ter ficado claro haver uma boa razão para considerar que ZLB forma um conjunto não-arbitrário: são proposições insensitivas implicadas por proposições empíricas sobre o mundo externo que sabemos. E penso também haver boas razões para vivermos com a ideia de que tais proposições não podem ser conhecidas por qualquer argumento mooriano – deixo aberta a possibilidade de haver outra maneira de se sabê-las.¹²⁴ Vamos considerar então uma formulação restrita do fechamento (BAUMANN 2012, pp. 20-21):

(PF**) Se S sabe que α , sabe que $\alpha \rightarrow \beta$, deduz β de maneira competente a partir de α e $\alpha \rightarrow \beta$, e com isso passa acreditar que β , então S sabe que β , *a menos* que β seja antecedentemente desconhecida por S, tomada por garantida e pressuposta pela crença e pelo conhecimento de S de que α .

Consideremos o adendo “a menos que β seja antecedentemente desconhecida por S, tomada por garantida e pressuposta pela crença e pelo conhecimento de S de que α ”. Isso é o que de fato acontece com as proposições ZLB. Essas proposições são todas (i) antecedentemente desconhecidas por S antes dele oferecer o argumento mooriano, e (ii) são pressupostas pelo conhecimento que S tem de pelo menos uma das premissas do argumento mooriano.

Vejamos. Ao acreditar que há uma zebra na jaula e inferir que não é uma mula disfarçada, Lindinha não sabia que não havia ali uma mula, afinal se soubesse não teria feito a inferência. Contudo, ao ir ao zoológico, as pessoas geralmente pressupõe que não há animais

¹²³ De todos os problemas epistemológicos recalcitrantes, o problema da circularidade parece ser o mais intratável. Intuitivamente, algo de muito errado parece estar envolvido em crenças que são justificadas circularmente. Alguns externistas dedicaram algum esforço à circularidade (van CLEVE 1984; ALSTON 1986; BERGMANN 2006), muito embora a maior parte dos filósofos pareça evitar tal questão.

¹²⁴ No caso das proposições zebradas, parece haver de fato outros meios de sabê-las sem se recorrer ao argumento mooriano. Por exemplo, a análise do material genético do animal nos dirá se é uma mula ou não. Ou posso saber que a pessoa que vejo não é o Twedleedum porque seu pai me disse que ele está viajando para Pindamonhangaba.

disfarçados por lá. Do mesmo modo, quando acredito não vou a Acapulco, certamente assumo que não terei dinheiro para ir, ainda que seja um dinheiro ganho na loteria. Quando digo que meu carro está estacionado na rua X, na vaga y, estou claramente supondo que ele não foi roubado. No discurso cotidiano costumamos usar expressões bastante elucidativas, como por exemplo: “Vou pegar o carro e ir embora; supondo, é claro, que não tenha sido roubado!”. Daí o meu raciocínio para concluir que ele não foi roubado ser inefetivo.

O diagnóstico de Crispin Wright (2000; 2003) para a suposta falha do argumento mooriano é feito nessa linha. Muito embora ele certamente se recusaria a considerar que algo como PF* ou PF** sejam princípios de fechamento. Ele diria que são princípios de transmissão de garantia. Isso porque a transmissão de garantia diz respeito à aquisição de garantia de uma crença a partir de outras crenças garantidas, ao passo que o fechamento diz respeito apenas ao conhecimento de proposições implicadas por outras proposições já conhecidas pelo sujeito. Considere as seguintes proposições:

- (K) Éder Jofre nocauteou Katsutoshi Aoki.
- (C) Éder Jofre venceu a luta.
- (L) Uma luta de boxe estava acontecendo.

Suponha que Cassius Clay liga o rádio e escuta que (K), Éder Jofre nocauteia seu adversário. Cassius então infere que (C), e assim sabe que (C) com base em (K). E note que (C) implica (K). Suponha agora que alguém pergunte a Cassius como ele sabe que estava tendo uma luta de boxe, e ele responde que uma vez que ele ouviu que Éder Jofre nocauteou o adversário, e sabe que se Éder Jofre venceu, então ele sabe que estava tendo uma luta. Mas sendo seu interlocutor um tanto sagaz, replica “Ora, mas se você sabe que Éder Jofre venceu porque nocauteou, certamente já sabia ou pressupunha de alguma forma que estava acontecendo uma luta de boxe. Seria estranho que você acreditasse que Éder Jofre venceu a luta se uma luta não estivesse acontecendo.” Isso sugere que o argumento de Cassius Clay é ruim para fornecer garantia à sua crença de que uma luta estava acontecendo, isto é, há falha de transmissão de garantia das premissas para a conclusão. Por outro lado, diria Wright, o fechamento continua intacto, uma vez que diz que Se Cassius Clay sabe que (C), então ele já sabia (L). Em suma, princípios de transmissão são mais amplos que o fechamento; se há transmissão há fechamento, mas não o contrário.

Embora interessante, não seguirei a linha de raciocínio de Wright. E a razão é a seguinte: a resposta nozickiana ao ceticismo é a de que S sabe que tem mãos muito embora não saiba que \neg BIV. Não deveríamos esperar que houvesse outra maneira de se saber que \neg BIV a não ser através de inferência. Wright, por outro lado, pensa que temos o *direito epistêmico* [entitlement], adquirido não-inferencialmente e muito menos perceptualmente, de acreditar em proposições como \neg BIV, que o universo existe há mais do que cinco minutos, que há outras mentes, etc. Esse direito epistêmico surge do fato de que quando digo algo como “Sei que tenho mãos” inevitavelmente estou comprometido com a proposição de que não sou um BIV.¹²⁵ Vou simplesmente pressupor aqui que PF** é um princípio de fechamento.¹²⁶

Baumann (pp. 22-23), porém, parece não ver muita vantagem em um princípio como PF**, uma vez que ele encontra no próprio Nozick uma versão restrita do fechamento pressuposta nas condições para o conhecimento inferencial:

S sabe que β via inferência a partir de α sse:

- (i) S sabe que α ;
- (ii) S infere β a partir de α de maneira competente, e
- (iii) com isso vem a acreditar que β ,
- (iv) $\neg\beta \square \rightarrow \neg B\alpha$ (ou não inferiria β a partir de α), e
- (v) $\beta \square \rightarrow B\alpha$ (e inferiria β a partir de α caso estivesse a inferir ou β ou $\neg\beta$ a partir de α).

As condições (iv) e (v) são condições de rastreamento para as premissas: se S sabe β via inferência a partir de α , então a crença de S na conclusão rastreia a verdade das premissas. E partir disso podemos formular o que ele chama de “Nozick-Fechamento”:

(NF) Se S sabe que α , sabe que $\alpha \rightarrow \beta$, deduz β de maneira competente a partir de α e $\alpha \rightarrow \beta$, e com isso passa acreditar que β , $\neg\beta \square \rightarrow \neg B\alpha$ (ou não inferiria β a partir

¹²⁵ A estratégia de Wright é conhecida com “estratégia fulcral” devido à influência das observações de Wittgenstein sobre as proposições fulcrais [*hinge propositions*] em *On Certainty* (1969), proposições que têm de ser pressupostas para que possamos dizer saber várias outras. Para uma discussão da epistemologia das proposições fulcrais, veja WRIGHT (2004), PRITCHARD (2014; 2015), COLIVA & MOYAL-SHARROCK (2016) e SALVATORE (2017).

¹²⁶ Peter Baumann (2012, p. 21, n.16) e Fred Dretske (2005), por exemplo, ignoram se há qualquer diferença substancial entre fechamento e transmissão. Pritchard (2014, p. 190-191), por outro lado, formula o que seria um princípio de transmissão como uma versão de fechamento.

de α), $\beta \square \rightarrow B\alpha$ (e inferiria β a partir de α caso estivesse a inferir ou β ou $\neg\beta$ a partir de α), então S sabe que β .

NF produz os mesmos resultados que PF**. A crença de Lindinha de que não há uma mula disfarçada não rastreia a verdade da sua crença de que o animal é uma zebra. A minha crença de que não ganharei na loteria não rastreia a minha crença de que vou a Acapulco. Assim como a minha crença de que \neg BIV não rastreia a verdade da minha crença de que tenho mãos. Claramente, contudo, NF e PF** não são equivalentes. Um nozickiano radical poderia querer manter que não há qualquer tipo de pressuposição ou algo próximo de uma proposição fulcral quando o sujeito sabe que tem mãos. Por outro lado, um defensor da sensibilidade que pretenda afastar a acusação de aceitar uma conjunção abominável teria PF** como um importante aliado. Esboçarei essa estratégia na seção seguinte. Antes de avançar, porém, quero tornar claro que a principal razão pela qual alguns epistemólogos se viam relutantes em abandonar o fechamento – veja as citações da seção 2.4. – continua intacta seja com PF** ou NF: a saber a ampliação do conhecimento via inferência.

Se temos boas razões para pensar que o conjunto ZLB é uniforme e que não saber tais proposições não ameaça o nosso conhecimento de proposições empíricas, um princípio de fechamento do conhecimento que nos permita saber uma ampla gama de proposições que se seguem daquilo que já sabemos (com exceção de ZLB) é nada implausível e muito menos *ad hoc*. Não precisamos de um princípio mais geral do que o necessário.

5. Crença razoável em \neg BIV

Nesta seção ofereço um esboço de uma estratégia que responda a acusação de sustentar uma conjunção abominável. Concordo que há algo de desconcertante em alguém dizer “Sei que tenho mãos mas não sei que não sou um BIV”. Contudo, me parece perfeitamente sensato alguém que afirme “Sei que tenho mãos, não sei que não sou um BIV, mas tenho uma crença razoável de que não sou um BIV”. Chamarei essa nova conjunção de *conjunção razoável*. Alguém que profere a conjunção razoável poderá dizer, em defesa própria, algo como se segue: “O conhecimento de proposições empíricas exige sensibilidade, e a minha crença de que \neg BIV é insensitiva; porém, quando digo que sei uma proposição, estou a pressupor que não estou sendo enganado, que minhas faculdades cognitivas estão funcionando

apropriadamente, etc. Se não fosse razoável pressupô-las, eu teria de sustentar uma conjunção ainda mais abominável: sei que tenho mãos e não é razoável acreditar que não sou um BIV!”.

Mas por que é razoável? Se reconheço que a minha crença de que tenho mãos pressupõe que não sou um BIV, é bastante razoável que eu acredite que não sou um BIV. E essa razoabilidade é concedida mesmo a crenças falsas. Por exemplo, quando crianças dizem que ganharam o que pediram de natal para o Papai Noel é porque pressupõem que o Papai Noel existe. E quando elas dizem que Papai Noel existe porque do contrário elas não teriam ganhado o presente, costumamos achar razoável tal inferência. E achamos razoável porque elas levaram em conta considerações explicativas. Vou defender que algo parecido se passa com a minha crença de que \neg BIV. Acredito que \neg BIV porque a hipótese BIV não explica tão bem a minha crença de que tenho mãos quanto a hipótese de que meu cérebro está funcionando apropriadamente num ambiente adequado. Na verdade, vou defender que dado as minhas crenças sobre o mundo externo, tal evidência confirma mais a hipótese de que *não* sou um cérebro numa cuba do que a hipótese de que sou um cérebro numa cuba. Isso porque, embora ambas tenham praticamente o mesmo poder preditivo, a hipótese de que não sou um cérebro numa cuba é mais simples e, por conseguinte, tem uma probabilidade prévia maior.

Façamos “CFA” representar a hipótese de que *meu cérebro está num corpo funcionando apropriadamente no mundo externo*, o argumento pode ser resumido da seguinte forma:

- 1- Se, dado E, uma hipótese H_1 é mais provável que uma rival H_2 , então é razoável acreditar que $\neg H_2$.
- 2- Dado E, CFA é mais provável que BIV.
- 3- Logo, é razoável acreditar que \neg BIV.

A chave para esse argumento é a teoria bayesiana da confirmação.¹²⁷ De acordo com o bayesianismo, a evidência E confirma uma hipótese H sse $P(H|E) > P(H)$. Isto é, a probabilidade posterior da hipótese – a probabilidade da hipótese depois da aquisição da evidência – tem de ser maior do que a sua probabilidade prévia – a probabilidade antes da

¹²⁷ Estou me baseando aqui na mesma estrutura que Richard Swinburne (1979) constrói seu argumento probabilístico a favor da existência de Deus. Um argumento parecido encontra-se em HUEMER (2015). Há argumentos bastante próximos a esses, porém ao invés de se basearem em considerações bayesianas, baseiam-se em considerações puramente explicativas. São versões abduativas de argumentos anti-céticos. (cf. VOGEL 1990, 2005).

consideração da evidência. Se duas hipóteses, H_1 e H_2 , concorrem para explicar E , preferiremos aquela que for mais provável dado E . Isto é, se $P(H_1 | E) > P(H_2 | E)$, então é razoável aceitar que H_1 em detrimento de H_2 .

O teorema de Bayes nos fornece uma maneira de calcular as probabilidades posteriores:

$$P(H | E) = \frac{P(E | H)}{P(E)} \cdot P(H)$$

Os dois fatores da direita são cruciais, o *poder explicativo* de H , que é representado pelo quociente $P(E | H) / P(H)$ – o quociente da *likelihood* de H e a probabilidade prévia da E – e a *probabilidade prévia* de H .

Queremos saber se E confirma mais CFA do que BIV, isto é, se $P(CFA | E) > P(BIV | E)$. Contudo, dado que tanto CFA quanto BIV têm as mesmas conseqüências observacionais, elas têm o mesmo poder explicativo. De modo que não importa o valor que atribuamos ao poder explicativo de ambas na equação. O que terá de fazer a diferença – se é que há alguma – será a probabilidade prévia de cada uma das hipóteses. Obviamente, então, que teremos de deixar de lado algo como o princípio de indiferença: dadas duas hipóteses concorrentes, suas prévias são sempre 0,5. Aceitar o princípio de indiferença aqui implica no ceticismo. Mas então como decidir qual das hipóteses tem uma prévia maior?

A *probabilidade intrínseca* (cf. SWINBURNE 1979). Há basicamente dois critérios para a probabilidade intrínseca: (i) adequação ao conhecimento de fundo, K , e (ii) simplicidade. (i) Quanto mais H se adequa a K , maior $P(H | K)$. Porém, se H for bastante ampla, K tende a se tornar mera evidência tautológica. Seja como for, CFA e BIV parecem ser igualmente amplas. (ii) Uma hipótese mais simples é sempre mais provável que uma mais complexa. Não é pacífico entre os filósofos o que é exatamente a simplicidade,¹²⁸ mas podemos considerar pelo menos os seguintes critérios. (a) *simplicidade ontológica*. Uma teoria será mais simples do que outra se postular menos tipos de entidades. (b) *simplicidade sintática*. Uma teoria será mais simples sintaticamente que outra se seu aparato formal for mais facilmente manipulável. Vamos, contudo, considerar apenas a simplicidade ontológica.

¹²⁸ Para uma defesa da simplicidade como critério para verdade veja SWINBURNE (1997).

Dado que CFA e BIV são igualmente amplas, a nossa questão da probabilidade intrínseca se resume à simplicidade. Qual dessas hipóteses é a mais simples?

Começemos por BIV. Três coisas são pelo menos postuladas: meu cérebro, a cuba e o supercomputador. Mas tem de haver mais: há toda tecnologia que permite ao supercomputador mandar inputs ao meu cérebro. E há também leis da natureza. E certamente há alguém que não é um cérebro numa cuba que foi o responsável por raptar o meu cérebro. E claro, há a intenção de alguém em mandar esses inputs ao meu cérebro.

Passemos então à CFA. Assim como BIV ela postula o mundo externo, regido pelas mesmas leis da natureza, e com pelo menos um cérebro dentro de um crânio. Porém, não postula uma tecnologia avançada para encubar cérebros e muito menos uma intenção por parte de alguém em incubar cérebros. Ao invés, os inputs que meu cérebro recebe provêm diretamente do mundo externo e são recebidos através dos meus órgãos. Além do mais, a arquitetura do meu aparato cognitivo é produto de milhões de anos de evolução.

Ambas as hipóteses concordam com a existência do mundo externo, mas divergem quanto à origem das minhas crenças e experiências: uma diz que são, em última instância, produto da intenção de alguém em me enganar, e a outra diz que as minhas crenças e experiências são em sua maioria produto de inputs do mundo externo. O que é mais simples então: (BIV) = mundo externo + **intenção enganadora** ou (CFA) = mundo externo + **funcionamento apropriado do cérebro num ambiente apropriado**?

Parece claro que CFA é mais simples. Postular a intenção de alguém em me enganar quanto ao mundo externo é mais complexo do que postular que meu cérebro está dentro do meu crânio recebendo inputs dos objetos externos via meus órgãos perceptivos. A razão é simples: a tecnologia que permitiu me incubarem tem de ter sido baseada no estudo do funcionamento do cérebro. Assim, antes de haver a tecnologia da cuba, já havia cérebros que recebiam input do mundo externo.

É claro que se a tecnologia do incubamento já estiver disponível, nunca saberei, uma vez que os cientistas nunca me deixarão saber tal. A esperança é que aconteça algum bug! Note, porém, que teríamos de pressupor que o mundo tinha um passado tal como acreditamos agora, e que uma série de fatos complexos ocorreu envolvendo o desenvolvimento da tecnologia do incubamento e da intenção de uma ou mais mentes em capturar cérebros e

incubá-los. Não é impossível, mas certamente muito menos econômico para explicar as nossas crenças e experiências.

Portanto, CFA é mais simples que BIV. Logo, $P(\text{CFA} | E) > P(\text{BIV} | E)$. Ou seja, CFA confirma mais E do que BIV. E, por conseguinte, é razoável acreditar que $\neg\text{BIV}$.

Talvez ainda houvesse uma maneira de considerar BIV tão plausível quanto CFA. Dizer que sempre fui um BIV. Nunca tive corpo e, por isso, nunca fui raptado. Porém, ainda assim teríamos de postular a existência do meu cérebro, da cuba e do enganador. Teríamos de dizer que o enganador domina a tecnologia, etc. etc. E isso não parece ajudar. A hipótese ainda é mais complexa. Uma tentativa mais radical seria apelar para o idealismo. Minhas experiências perceptuais são o meu estar consciente de um padrão qualquer de dados dos sentidos – padrões fornecidos por um programa de computador ou pela mente de Deus. Mas nesse caso deixamos de discutir sobre o conhecimento e passamos a discutir sobre a natureza do mundo que conheço. O idealista não flerta com o ceticismo. Ele o evita dizendo que o conhecimento do mundo objetivo é o conhecimento de um padrão de dados dos sentidos. Por outro lado, se insistirmos nessa hipótese idealista, fica difícil escapar da acusação de Putnam (1981) de que a hipótese BIV é autorrefutante. Supondo que o epistemólogo externista mantenha o mesmo espírito na semântica, e sustente que o conteúdo dos nossos estados mentais é parcialmente determinado pelo mundo externo, considerar a hipótese de que sempre fomos BIV se torna inconsistente. Pois suponha que sempre fui um BIV. Assim, o significado dos termos que uso fazem referência ao padrão de bits da simulação de computador, não ao mundo externo. Assim, “cérebro” e “cuba” se referem a padrões de bits, e não há objetos externos. Ora, mas na hipótese BIV “cérebro” e “cuba” se referem a objetos no mundo externo. Assim, se sou um BIV a minha hipótese BIV não quer dizer o mesmo que a hipótese original. Mas se estou considerando a hipótese original, não posso estar me referindo ao mundo dos bits. Só me resta estar me referindo ao mundo externo. E, portanto, o mundo externo tem de existir.

Note que se o argumento probabilístico acima foi bem sucedido, não se segue que sei que $\neg\text{BIV}$. Ainda estou assumindo que a crença de que $\neg\text{BIV}$ é insensitiva. Mas disso não se segue que não seja razoável. Se “ser razoável” quer dizer pelo menos ter alguma evidência para preferir uma hipótese à outra, então o argumento acima é capaz de estabelecer que acreditar que $\neg\text{BIV}$ é razoável.

Por fim, ainda que longe de ser um princípio tão forte quanto PF** e ainda que não resolva nossos problemas quanto ao conhecimento de ZLB, o seguinte princípio parece aceitável:

(PCR) Se a crença de S de que β , obtida via inferência a partir de α , é insensitiva, mas S tem alguma evidência para preferir β a $\neg\beta$, então a crença de S de que β é razoável (embora não seja conhecimento)

Desse modo, parece-me que podemos anular a acusação de sustentar a conjunção abominável ao proferirmos a conjunção razoável: *sei que tenho mãos, não sei que $\neg BIV$, porém a minha crença de que $\neg BIV$ é razoável.*

6. Sensitividade e a força da hipótese cética

Resta ainda que a nossa abordagem explique a força do argumento cético. A explicação é simples, como já vimos; a crença em $\neg BIV$ é insensitiva. Nesse ponto a abordagem nozickiana é concessiva. Mas não concessiva a ponto de tornar o conhecimento impossível. A sensitividade (ou a EFAS) usa o mesmo mecanismo para dar conta do nosso conhecimento, da falta de conhecimento e da impossibilidade de conhecimento.

Vimos no início deste capítulo que cenários céticos como BIV e do gênio maligno servem para enfatizar a dificuldade de sabermos que (i) nosso aparato cognitivo está funcionando apropriadamente, e (ii) que está funcionando num ambiente favorável. José Luis Bermúdez (2008) faz notar que, na primeira Meditação, Descartes atinge a conclusão cética mesmo antes de apresentar a hipótese do gênio maligno. O argumento da ilusão chama a atenção para (i) – na medida em que coloca em xeque a confiabilidade da percepção –, e o argumento do sonho para (ii) – na medida em que a percepção pode ser perfeitamente confiável num cenário de sonho e ainda assim nos enganar:

Descartes atinge agora o último grau do ceticismo. É surpreendente, porém, que ele tenha atingido essa conclusão dois parágrafos antes de fazer menção à hipótese do gênio maligno, que é geralmente considerada a razão de Descartes para a suspensão do assentimento de todas as suas crenças prévias. Já antes de introduzir a hipótese do gênio maligno, Descartes enfatiza que a conclusão cética fora estabelecida com satisfação. Ele se refere ao “fato de todas as suas opiniões terem sido postas em dúvida, como visto” e diz que “Não basta ter feito tais considerações, é preciso ainda que cuide de lembrar-me delas”. A hipótese do gênio maligno é parte do “esforço de

lembrá-las”. É um dispositivo que ele introduz a fim de interromper a sua divagação e não um argumento em favor da conclusão de que todas as suas crenças são instáveis. (BERMÚDEZ 2008, pp. 60-61)

Ou seja, a hipótese do gênio maligno tem valor heurístico. Serve para enfatizar a possibilidade de as condições externas ao conhecimento não estarem sendo cumpridas.¹²⁹

A EFAS, ao estabelecer condições para o conhecimento, deixa claro como o conhecimento é possível. Se uma crença verdadeira é conhecimento, então foi produzida por um mecanismo cognitivo que está funcionando apropriadamente num ambiente adequado e favorável, ou seja, num ambiente que permite o rastreamento. E uma vez que EFAS é uma teoria externista, não há a exigência que o sujeito esteja ciente de que tais condições externas tenham sido cumpridas.

É claro que muitas vezes é necessário nos assegurarmos previamente do funcionamento das nossas capacidades cognitivas e do ambiente. Por exemplo, se tomo uma dosagem alta de um remédio que me faz não ser capaz de distinguir verde de vermelho, não é uma boa ideia que eu pegue o trânsito sob o efeito desse remédio. Em muitos casos não saberei se tenho de parar no sinal ou avançar. Em casos como esse, dizemos que a proposição <Tomei uma dosagem x do remédio R> funciona como anulador para as minhas crenças de que o sinal está verde e de que o sinal está vermelho. Porém, se não tomo tal remédio ou algo parecido, a mera possibilidade de eu tê-lo ingerido sem saber não funciona como um anulador para a minha crença de que o sinal está vermelho. Seria absurdo exigir que tenho de ser capaz de eliminar essa possibilidade a fim de saber que o sinal está vermelho. Do mesmo modo, não faz sentido que eu tenha de ser capaz de eliminar hipóteses céticas radicais para saber que tenho mãos. Isso tudo a despeito de ser completamente razoável que eu acredite não ter ingerido a droga, e que eu acredite não ser um BIV.

Poder-se-ia objetar, contudo, que explicar como o conhecimento é possível não é o mesmo que mostrar que o conhecimento é efetivo, ou seja, do fato de mostrarmos em que condições um sujeito tem conhecimento não se segue que ele efetivamente o tenha. Uma

¹²⁹ Sosa (2017) chama atenção para o mesmo ponto em sua interpretação da “epistemologia da virtude cartesiana”. Ele aponta várias passagens das *Meditações* que indicam que Descartes “não apenas buscava a verdade mas também a *aptidão*. E a aptidão requer competência que seja boa o bastante, que seja confiável o bastante. É preciso assegurar a si próprio que tal aptidão foi alcançada, pois ela é necessária para a sua confiança de que evitaste o erro e alcançaste a certeza [...] Em todos os aspectos estruturais importantes a epistemologia da virtude cartesiana tem a mesma opinião [...] do *perspectivismo de virtudes*” (pp. 18-19).

coisa, continua a objeção, é imaginar um sujeito em condições favoráveis e dizer “S sabe que α ”, outra coisa é ser esse sujeito, estar em condições favoráveis e dizer “S sabe que α ”. Esse último caso, um caso de auto-atribuição de conhecimento, envolve conhecimento de segunda ordem: sei que sei que α . E só é possível mostrar que o conhecimento é efetivo se antes eu mostrar que tenho conhecimento de primeira ordem. Mas a minha crença de que sei que α é insensitiva, uma vez que nos mundos próximos em que não sei que α (por conta de não ser o caso que α) ainda assim acredito que sei que α . Portanto, a perspectiva externista não responde de fato ao desafio cético, que é mostrar que efetivamente temos conhecimento. Vale citar aqui a famosa queixa de Barry Stroud (2000, p.119):

[M]esmo que seja verdade que podemos saber algo sem saber que o sabemos, o filósofo teórico do conhecimento não pode se limitar a insistir nesse ponto e esperar que a aceitação de uma abordagem “externista” do conhecimento seja plenamente satisfatória. Se pudesse, estaria a colocar-se na posição de alguém que diz: “Não sei se compreendo o conhecimento humano ou não. Se aquilo em que acredito a seu respeito for verdade e se minhas crenças a seu respeito forem produzidas da maneira que minha teoria considera estar certa, então sei como é que o conhecimento humano se produz, e, nesse sentido, compreendo-o. Mas se as minhas crenças não forem verdadeiras, ou não forem alcançadas dessa maneira, não o compreendo”. Pergunto-me qual das hipóteses estará correta. Pergunto-me se compreendo o conhecimento humano ou não. Ora, essa não é uma posição satisfatória a alcançar no estudo do conhecimento humano – ou seja do que for.

Quanto a essa objeção, confesso que me parece incontornável. E isso em parte explica a força do desafio cético. Mas permita-me uma distinção entre *desafio cético* e *argumento cético*. Responder ao desafio cético é mostrar que sabemos a negação de hipóteses céticas. Ao passo que responder ao argumento cético é mostrar que uma de suas premissas é implausível. E o externismo de fato não responde ao desafio cético. Pois mostra que uma de suas premissas é implausível: não preciso saber que \neg BIV para saber que tenho mãos.

É claro que essa resposta só tem força se já pressupusermos que temos conhecimento. E dessa perspectiva, só é razoável abandonar essa suposição se nos convenceremos de que o conhecimento é impossível. E o argumento cético não tem essa força contra o externista. O externismo, e em particular EFAS, mostra como o conhecimento é possível. É claro que podemos estar completamente errados e que não tenhamos qualquer conhecimento sobre o mundo externo. E se for esse o caso, também não saberemos. E é aqui que a exigência de

Stroud me parece desmesurada: na medida em que entender o conhecimento humano pressupuser que tenhamos de saber que sabemos, a sua queixa contra o externismo invariavelmente nos empurrará para o ceticismo. Reconheço que do ponto de vista externista torna-se difícil saber que sabemos; mas nada impede que tenhamos crenças razoáveis sobre o conhecimento humano. Parece-me satisfatório que o epistemólogo seja aquele que tem crenças razoáveis sobre o conhecimento humano. (Ainda que nunca seja capaz de *saber* se são razoáveis ou não!).

Prefiro pensar que somos como aventureiros seguindo várias trilhas que prometem dar em Eldorado. Ainda que não haja qualquer Eldorado, vale a pena ir à procura. Por outro lado, aqueles que concebem a refutação do ceticismo como a tarefa primordial da teoria do conhecimento dificilmente concordarão com tal resposta. Talvez eles pensem que não valha a pena teorizar sobre algo incerto. E aqui atingimos um ponto no qual não me parece haver debate proveitoso.¹³⁰ Quanto a mim, prefiro a incerteza da riqueza de Eldorado à miséria certa da acomodação.

¹³⁰ Isso não quer dizer que nada mais possa ser dito. Por exemplo, Sosa (2017) ao interpretar a estratégia cartesiana como uma estratégia da epistemologia das virtudes, não nega a legitimidade de se começar do zero, muito embora faça ressalvas importantes. A mais importante delas é que Descartes não estaria suspendendo o conhecimento de primeira ordem, mas a penas o conhecimento de segunda. Isso quer dizer que, num sentido relevante, essa epistemologia cartesiana das virtudes também parte do pressuposto de que temos conhecimento.

CODA

Podemos resumir os resultados da EFAS – *Epistemologia da Função Apropriada Sensitiva* – para o conhecimento empírico como se segue:

A crença verdadeira de S de que α é conhecimento sse:

[Conhecimento não inferencial]

- (i) S formou a crença de que α através de um mecanismo cognitivo, C, funcionando apropriadamente; e
- (ii) C estava funcionando num ambiente tal que:
 - a. se não fosse o caso que α , S não acreditaria que α , e
 - b. se fosse o caso que α , S acreditaria que α .

Ou

[Conhecimento inferencial]

- (iii) S inferiu a crença de que α de maneira competente a partir de um conjunto de crenças $\{\beta_1, \dots, \beta_n\}$;
- (iv) S sabe cada uma das β_i (inferencialmente ou não-inferencialmente);
- (v) $P(\alpha \mid [\beta_1 \wedge \dots \wedge \beta_n]) > P(\alpha)$;
- (vi) se não fosse o caso que α , S não acreditaria que $(\beta_1 \wedge \dots \wedge \beta_n)$ é verdadeira ou não formaria a crença de que $(\beta_1 \wedge \dots \wedge \beta_n)$, e
- (vii) se fosse o caso que α , S acreditaria que $(\beta_1 \wedge \dots \wedge \beta_n)$ é verdadeira.

Note que as condições (v)-(vii) permitem o conhecimento indutivo (como argumentado no Capítulo III); e estão de acordo com a versão restrita do fechamento (Capítulo V):

(PF**) Se S sabe que α , sabe que $\alpha \rightarrow \beta$, deduz β de maneira competente a partir de α e $\alpha \rightarrow \beta$, e com isso passa acreditar que β , então S sabe que β , *a menos* que β seja antecedentemente desconhecida por S, tomada por garantida e pressuposta pela crença e pelo conhecimento de S de que α .

Assim, EFAS dá os vereditos corretos sobre os casos de conhecimento empírico, sem com isso se preocupar em ter de refutar o desafio cético.

BIBLIOGRAFIA

AYER, A. *The Problem of Knowledge*. Macmillan, 1956.

ALSTON, W. "Epistemic Circularity". *Philosophy and Phenomenological Research* 47 (1): 1-30, 1986.

BAUMANN, P. "Reliabilism – Modal, Probabilistic or Contextualist". *Grazer Philosophische Studien*, 79, 77-89, 2009.

BECKER, K. *Epistemology Modalized*. Routledge, 2007.

BEEBEE, H. & SABBARTON-LEARY, N. "On the Abuse of the Necessary A Posteriori". In BEEBEE & SABBARTON-LEARY (eds.) *The Semantics and Metaphysics of Natural Kinds*. Routledge. pp. 159-79, 2010.

BENACERRAF, P. "Mathematical Truth". *Journal of Philosophy* 70 (19): 661-679, 1973.

BERGMANN, M. *Justification Without Awareness: A Defense of Epistemic Externalism*. Oxford University Press, 2006.

BERMÚDEZ, J. L. "Cartesian Skepticism: Arguments and Antecedents". In GRECO (ed.), 2008.

BERNECKER, S. "Sensitivity, Safety and Closure". *Acta Analytica* 27 (4): 367-381, 2012.

BLACK, T. & BECKER, K. (orgs.) *The Sensitivity Principle in Epistemology*. Cambridge University Press, 2012.

BLACK, T. & MURPHY, P. "In Defense of Sensitivity". *Synthese* 154 (1):53-71, 2007.

BLOUW, P., BUCKWATER, W. & TURRI, J. "Gettier Cases: A Taxonomy". In BORGES et. al. (2017), 242-252.

BOGARDUS, T. & MARXEN, C. "Yes, Safety is in Danger". 2014 - *Philosophia* 42 (2): 321-334, 2014.

- BOGARDUS, T. "Knowledge under Threat". *Philosophy and Phenomenological Research* 88 (2): 289-313, 2012.
- BOGHOSSIAN, P. *Fear of Knowledge*. Oxford University Press, 2006.
- BORGES, R., ALMEIDA, C. & KLEIN, P. (eds.) *Explaining Knowledge: New Essays on the Gettier Problem*. Oxford University Press, 2017.
- BOTHAM, T. M. "Plantinga and favorable mini-environments". *Synthese* 135: 431-441, 2003.
- BOYCE, K. & MOON, A. "In Defense of Proper Functionalism: Cognitive Science Takes on Swampman". *Synthese* 193 (9): 2987–3001, 2016.
- BROCANO-BERROCAL, F. "Is Safety in Danger?". *Philosophia* 42 (1): 1-19, 2014.
- BROCANO-BERROCAL, F. "Luck". *Internet Encyclopedia of Philosophy*, 2016.
- BRUECKNER, A. "Strategies for Refuting Closure for Knowledge". *Analysis* 64 (4): 333–335, 2004.
- BRUECKNER, A. *Essays on Skepticism*. Oxford University Press, 2010.
- CHISHOLM, R. *Theory of Knowledge*. Prentice Hall, 1966.
- COHEN, S. "Basic Knowledge and the Problem of Easy Knowledge". *Philosophy and Phenomenological Research* 65 (2): 309-329, 2002.
- COLIVA, A. & MOYAL-SHARROCK, D. "Introduction: Hinge Epistemology". *International Journal for the Study of Skepticism* 6 (2-3): 73-78, 2016.
- COLLINS, J. "Epistemic Closure Principles". *Internet Encyclopedia of Philosophy*, 2006.
- COMESAÑA, J. "Unsafe Knowledge". *Synthese* 146 (3): 395-404, 2005.
- CRAIG, E. *Knowledge and the State of Nature: An Essay in Conceptual Synthesis*. Oxford University Press, 1990.
- DAVID, M. & WARFIELD, T. "Knowledge-Closure and Skepticism". In SMITH, Q. (ed.), *Epistemology: New Essays*. Oxford University Press, 2008.

- DeROSE, K. & WARFIELD, T. (eds.) *Skepticism: A Contemporary Reader*. Oxford University Press, 1999.
- DeRose, K. "Insensitivity is Back, Baby!". *Philosophical Perspectives* 24 (1):161-187, 2010.
- DeRose, K. "Solving the Skeptical Problem". *Philosophical Review* 104 (1):1-52, 1995.
- DeRose, K. "Sosa, Safety, Sensitivity, and Skeptical Hypothesis". In GRECO (ed.), 2004.
- DRETSKE, F. "Conclusive Reasons". *Australasian Journal of Philosophy*, Vol. 49, n. 1, 1971.
- DRETSKE, F. "Epistemic Operators". *The Journal of Philosophy*, Vol. 67, No. 24, pp. 1007-1023, 1970.
- DRETSKE, F. "The Case against Closure". In STEUP, TURRI & SOSA (2014).
- DUTANT, J. "The Legend of Justified True Belief Analysis". *Philosophical Perspectives*, 29 (1): 95-145, 2015.
- GELLMAN, J. "A new Gettier-type refutation of Nozick's analysis of knowledge". *Principia* 8 (2), pp. 279-283, 2004.
- GETTIER, E. "Is Justified True Belief Knowledge?". *Analysis* vol. 23, n. 6, 1963.
- GOLDBERG, S. "Reliabilism". In CULLISON, A. (ed.) *The Bloomsbury Companion to Epistemology*, Bloomsbury, 2015.
- GOLDBERG, S. "Sensitivity from Others". In BLACK & BECKER (2012).
- GOLDMAN, A. "A Causal Theory of Knowing". *Journal of Philosophy*, 64 (12): 357-372, 1967.
- GOLDMAN, A. "Discrimination and Perceptual Knowledge". *The Journal of Philosophy* Vol. 73, No. 20, pp. 771-791, 1976.
- GOLDMAN, A. "Review of Philosophical Explanations". *The Philosophical Review*, Vol. 92 (1): 81-88, 1983.

GOLDMAN, A. "Social Epistemology". In *Stanford Encyclopedia of Philosophy*, Fall 2006 ed. Disponível em:

<https://stanford.library.sydney.edu.au/archives/fall2008/entries/epistemology-social/>

GOLDMAN, A. "What is Justified Belief". In George Pappas (ed.) *Justification and Knowledge*. D. Reidel. pp. 1-25, 1979.

GOLDMAN, A. *Epistemology and Cognition*. Harvard University Press, 1986.

GRANDY, R. E. & WARNER, R. "Paul Grice". *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/grice/> .

GRECO, J. & REIBSAMEN, J. "Reliabilist Virtue Epistemology". In SNOW (2018).

GRECO, J. (ed.) *Sosa and his Critics*. Willey-Blackwell, 2004.

GRECO, J. "Better Safe than Sensitive". In BLACK & BECKER (eds.), 2012.

GRECO, J. "Knowledge as Credit for True Belief". In Michael DePaul & Linda Zagzebski (eds.), *Intellectual Virtue: Perspectives From Ethics and Epistemology*. Clarendon Press. pp. 111-134, 2003.

GRECO, J. "The Nature of Ability and the Purpose of Knowledge". *Philosophical Issues* 17.1: 57–69, 2007.

GRECO, J. "Worry about Pritchard's Safety". *Synthese*, Vol. 158, No. 3, pp. 299-302, 2007.

GRECO, J. *The Oxford Handbook of Skepticism*. Oxford University Press, 2008.

GRENNE, R. & BALMERT, N. A. "Two notions of warrant and Plantinga's solution to the Gettier problem". *Analysis*, 57, 2, pp. 132-139, 1997.

GRICE, P. *Studies in the way of words*. Harvad University Press, 1989.

HALES, S. "Epistemic Closure Principles". *The Southern Journal of Philosophy*, XXXIII, 1995.

- HARMAN, G. *Thought*. Princeton University Press, 1973.
- HAWTHORNE, J. "The Case for Closure". In STEUP, TURRI & SOSA (2014).
- HAZE, T. "Two New Counterexamples to the Truth-Tracking Theories of Knowledge". *Logos and Episteme* 6 (3): 309-311, 2015.
- HAZLETT, A. (ed.) *The Gettier Problem at 50. Philosophical Studies*, edição especial, 172 (1), 2015.
- HAZLETT, A. "The maturation of the Gettier problem". *Philosophical Studies*, 172: 1-6, 2015.
- HETHERINGTON, S. (ed.) *The Gettier Problem*. Cambridge University Press, 2018.
- HETHERINGTON, S. "A Fallibilist and Wholly Internalist Solution to the Gettier Problem". *Journal of Philosophical Research* 26: 307-324, 2001.
- HETHERINGTON, S. "Actually Knowing". *Philosophical Quarterly* 48 (193): 453-469, 1998.
- HETHERINGTON, S. "Knowing Failably". *Journal of Philosophy* 96 (11): 565-587, 1999.
- HETHERINGTON, S. *How to Know: A Praticalist Conception of Knowledge*. Willey-Blackwell, 2011.
- HETHERINGTON, S. *Knowledge and the Gettier Problem*. Cambridge University Press, 2016.
- HILL, C. "Process reliabilism and cartesian skepticism". *Philosophy and Phenomenological Research* 56 (3): 567-581, 1996.
- HOWTHORNE, J. "The Epistemology of Possible Worlds: A Guide Tour". *Philosophical Studies*, 84: 183-202, 1996.
- HUEMER, M. "Serious theories and skeptical theories: Why you are probably not a brain in a vat". *Philosophical Studies* 173, (4): 1031-1052, 2016.
- IMMERMAN, D. "Williamson, Closure, and KK". *Synthese*, 2018.

- KELP, C. "Knowledge and Safety". *Journal of Philosophical Research* 34: 21-31, 2009.
- KIM, B. & McGRATH, M (eds.) *Pragmatic Encroachment in Epistemology*. Routledge, 2019.
- KLEIN, P. "A Proposed Definition of Propositional Knowledge". *Journal of Philosophy*, 68 (16): 471-482, 1971.
- KLEIN, P. "Closure Matters: Academic Skepticism and Easy Knowledge". *Philosophical Issues* 14 (1):165–184, 2004.
- KLEIN, P. "Immune Belief Systems". *Philosophical Topics*, 14 (1): 259-280, 1986.
- KLEIN, P. "Knowledge, Causality, and Defeasibility". *Journal of Philosophy*, 73 (20): 792-812, 1976.
- KLEIN, P. "Misleading 'Misleading Defeaters'". *Journal of Philosophy*, 76 (7): 382-386, 1979.
- KLEIN, P. "Skepticism and Closure: Why the Evil Genius Argument Fails". *Philosophical Topics* 23 (1): 213-236, 1995.
- KRAFT, T. "Transmission arguments against closure are still fallacious". *Synthese*, 191: 2617–2632, 2014.
- KRAFT, T. "Transmission Arguments Against Knowledge Closure Are Still Falacious". *Synthese* 191 (12): 2617-2632, 2014.
- KRIPKE, S. *Philosophical Troubles: Collected Papers Vol. 1*. Oxford University Press, 2011.
- KVANVIG, J. (ed.) *Warrant and Contemporary Epistemology: Essays in Honor of Plantinga's Theory of Knowledge*. Rowman and Littlefield, 1996.
- KVANVIG, J. "Closure Principles". *Philosophy Compass* 1 (3) :256–267, 2006.
- KYBURG, H. *Probability and the Logic of Rational Belief*. Wesleyan University Press, 1961.
- LACKEY, J. "We don't deserve credit for everything we know". *Synthese*, 158 (3): 345-361, 2007.

LACKEY, J. "Why We Don't Deserve Credit for Everything We Know". *Synthese* 158.3: 345–361, 2007.

LANDESMAN, C. *Skepticism: The Central Issues*. Willey-Blackwell, 2002.

LAUDAN, L. "A Confutation of Convergent Realism". *Philosophy of Science* 48 (1): 19-49, 1981.

LE MORVAN, P. "Knowledge before Gettier". *British Journal for the History of Philosophy*, 25 (6): 1216-1238.

LEHRER, K. & PAXSON, T. "Knowledge: Undefeated Justified True Belief". *Journal of Philosophy* 66 (8): 225-237, 1969.

LEHRER, K. "The Fourth Condition on Knowledge: A Defense". *Review of Metaphysics*, 24 (1): 122-128, 1970.

LEITE, A. "'Skepticism, Sensitivity, and Closure, or Why the Closure Principle is Irrelevant to External World Skepticism". *Croatian Journal of Philosophy*, vol. IV, n. 12, 2004.

LEITE, A. "Epistemological Externalism and the Project of Traditional Epistemology". *Philosophy and Phenomenological Research*, vol. LXX, n. 3, 2005.

LEVINSON, S. *Pragmatics*. Cambridge University Press, 1983.

LEWIS, D. *Counterfactuals*. Blackwell, 1973.

LUPER-FOY, S. "The Epistemic Predicament: Knowledge, Nozickian Tracking and Scepticism". *Australasian Journal of Philosophy* 62 (1): 26 – 49, 1984.

LUPER-FOY, S. *The Possibility of Knowledge: Nozick and his Critics*. Rowman and Littlefield, 1987.

LYCAN, W. "Evidence one does not possess". *Australasian Journal of Philosophy*, 55 (2):114-126, 1977.

LYCAN, W. "On the Gettier problem problem". In Stephen Hetherington (ed.), *Epistemology Futures*. Oxford University Press. pp. 148-168, 2006.

- MAKINSON, D. "The Preface Paradox". *Analysis* 25 (6): 205, 1965.
- McKINSEY, M. "Anti-individualism and Privileged Access". *Analysis* 51 (1): 9-16, 1991.
- MEDAWAR, P. *The Limits of Science*. Oxford University Press, 1984.
- MURPHY, P. "A Strategy for Assessing Closure". *Erkenntnis* 65 (3): 365 – 383, 2006.
- NAGEL, T. *The Last Word*. Oxford University Press, 1997.
- NETA, R. & ROHRBAUGH, G. "Luminosity and the Safety of Knowledge". *Pacific Philosophical Quarterly* 85 (4): 396–406, 2004.
- NOZICK, R. *Philosophical Explanations*. Harvard University Press, 1981.
- PAPPAS, S. & SWAIN, M. "Some conclusive reasons against 'conclusive reasons'". *Australasian Journal of Philosophy* 51 (1), 72-76, 1973.
- PLANTINGA, A. "Warrant and accidentally true belief". *Analysis* 57, 2. pp. 140-45, 1997.
- PLANTINGA, A. "Warrant and Accidentally True Belief". *Analysis* 57 (2): 140–145, 1997.
- PLANTINGA, A. *Warrant and Christian Belief*. Oxford University Press, 2000.
- PLANTINGA, A. *Warrant and Proper Function*. Oxford University Press, 1993b.
- PLANTINGA, A. *Warrant: The Current Debate*. Oxford University Press, 1993a.
- PLANTINGA, A. *Where the Conflict Really Lies: Science, Religion, and Naturalism*. Oxford University Press, 2011.
- POLLOCK, J. "Defeasible Reasoning". In ADLER, J. & RIPS, L. *Reasoning: Studies of Human Inference and its Foundations*. Oxford University Press, 2008.
- POPPER, K. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, 1962.
- POPPER, K. *Objective Knowledge: An Evolutionary Approach*. Oxford University Press, 1975.

PRITCHARD, D. & WHITTINGTON, L. J. (eds.) *The Philosophy of Luck*. Willey-Blackwell, 2015.

PRITCHARD, D. "Anti-Luck Epistemology and the Gettier Problem". *Philosophical Studies*, 172, 93-111, 2015.

PRITCHARD, D. "Anti-luck Epistemology". *Synthese*, Vol. 158, No. 3, pp. 277-297, 2007.

PRITCHARD, D. "Anti-luck virtue epistemology". *The Journal of Philosophy* 109, 247-279, 2012a.

PRITCHARD, D. "Entitlement and the Groundlessness of Our Believing". In DOOD, D. & ZARDINI, E. (eds.) *Contemporary Perspectives on Scepticism and Perceptual Justification*. Oxford University Press, 2014.

PRITCHARD, D. "Epistemic Risk", *Journal of Philosophy*, 113, 550-71, 2016.

PRITCHARD, D. "In defense of modest anti-luck epistemology". In BLACK, T. & BECKER, K. (orgs.), 2012b.

PRITCHARD, D. "Sensitivity, Safety, and Anti-Luck Epistemology". In GRECO (ed.), 2008.

PRITCHARD, D. "The Modal Account of Luck". *Metaphilosophy* 45 , 594-619, 2014. Reimpresso em PRITCHARD & WHITTINGTON (eds.), 2015, pp. 145-166.

PRITCHARD, D. "Virtue Epistemology and Epistemic Luck, Revisited". *Synthese* 154 (1): 53-71, 2008.

PRITCHARD, D. *Epistemic Angst: Radical Skepticism and the Groundlessness of our Believing*. Princeton University Press, 2015.

PRITCHARD, D. *Epistemic Luck*. Oxford University Press, 2005.

PSILLOS, S. *Knowing the Structure of Nature*. Palgrave, 2009.

ROTH, M. D. & ROSS, G. *Doubting: Contemporary Perspectives on Skepticism*. Kluwer Academic Publishers, 1990.

ROUSH, S. *Tracking Truth: Knowledge, Evidence and Science*. Oxford University Press.

RYSIEW, P. “Epistemic Contextualism”. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/contextualism-epistemology/> .

SAGAN, C. *The Demon-Haunted World: Science as a Candle in the Dark*. Ballantine Books, 1996.

SALVATORE, N. “Hinge Epistemology: An Anti-Skeptical Skepticism?”. *Sképsis*, 15, pp. 62-68, 2017.

SANFORD, D. “Knowledge and Relevant Alternatives: Comments on Dretske”. *Philosophical Studies*, Vol. 40, No. 3, Current Issues in Epistemology, pp. 379-388, 1981.

SELLARS, W. *Science, Perception and Reality*. Humanities Press, 1963.

SHOPE, R. *The Analysis of Knowing: A Decade of Research*. Princeton University Press, 1983.

SNOW, N. E. *The Oxford Handbook of Virtue*. Oxford University Press, 2018.

SOSA, E. “How to Defeat Opposition to Moore”. *Philosophical Perspectives*, 13: 141-153, 1999.

SOSA, E. “Philosophical Scepticism and Epistemic Circularity”. *Aristotelian Society Supplementary Volume* 68 (1): 263 – 307, 1994.

SOSA, E. “Philosophical Skepticism and Epistemic Circularity”. *Proceedings of the Aristotelian Society*, Vol. 68, pp. 263-307, 1994.

SOSA, E. “Tracking, Competence, and Knowledge”. In MOSER, P. (ed.), *The Oxford Handbook of Epistemology*, 2002.

SOSA, E. *A Virtue Epistemology: Apt Belief and Reflective Knowledge, vol.1*. Oxford University Press, 2007.

SOSA, E. *Epistemology*. Princeton University Press, 2017.

SOSA, E. *Knowledge in Perspective*. Cambridge University Press, 1991.

SOSA, E. *Reflective Knowledge: Apt Belief and Reflective Knowledge, vol.2*. Oxford University Press, 2009.

STALNAKER, R. "A theory of conditionals". RESCHER, N. (org) *Studies in logical theory*, pp. 98-112. Blackwell, 1968.

STARMANS, C. & FRIEDMAN, O. "The Folk Conception of Knowledge". *Cognitio*, 124 (3): 272-283, 2012.

STARR, W. "Counterfactuals". *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/fall2019/entries/counterfactuals/> .

STINE, G. "Skepticism, Relevant Alternatives, and Deductive Closure". *Philosophical Studies* 29 (4): 249-261. Reimpresso em DeROSE & WARFIELD (eds.) 1999.

STROUD, B. "Scepticism, 'Externalism', and the Goal of Epistemology". *Proceedings of the Aristotelian Society*, Vol. 68, pp. 263-307, 1994.

STROUD, B. *Understanding Human Knowledge*. Oxford University Press, 2000.

SWAIN, M. "Epistemic Defeasibility". *Philosophical Topics*, 14 (1): 259-280, 1974.

SWINBURNE, R. *Simplicity as Evidence for Truth*. Marquette University Press, 1997.

SWINBURNE, R. *The Existence of God*. Oxford University Press, 1979.

TOLLY, J. "Swampman: A Dilemma for Proper Functionalism". *Synthese*, 2018. <https://doi.org/10.1007/s11229-018-1684-0>

TURRI, J. "Manifest Failure: The Gettier Problem Solved". *Philosopher's Imprint*, vol. 11, n. 8, 2011.

VAIDYA, A. "The Epistemology of Modality", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/modality-epistemology/> .

Van CLEVE, J. "Reliability, Justification, and the Problem of Induction". *Midwest Studies in Philosophy* 9 (1): 555-567, 1984.

- VOGEL, J. "Cartesian Skepticism and Inference to the Best Explanation". *Journal of Philosophy* 87 (11): 658-666, 1990.
- VOGEL, J. "Reliabilism Leveled". *Journal of Philosophy* 97 (11): 602-623, 2000.
- VOGEL, J. "Subjunctivists". *Philosophical Studies* 134 (1): 73 – 88, 2007.
- VOGEL, J. "The enduring trouble with tracking". In BLACK & BECKER (eds.), 2012.
- VOGEL, J. "Tracking, Closure, and Inductive Knowledge". In LUPER-FOY (ed.), 1987.
- WARFIELD, T. "When Epistemic Closure Fails and Does Not Fail: A Lesson From the History of Epistemology". *Analysis* 64 (1): 35–41, 2004.
- WILLIAMSON, T. *Knowledge and its Limits*. Oxford University Press, 2000.
- WITTGENSTEIN, L. *On Certainty* (ed. Anscombe and von Wright). Harper Torchbooks, 1969.
- WRIGHT, C. & DAVIES, M. "On Epistemic Entitlement". *Aristotelian Society Supplementary Volume* 78: 167-245, 2004.
- WRIGHT, C. "Cogency and question-begging: Some reflections on McKinsey's paradox and Putnam's proof". *Philosophical Issues* 10 (1): 140-63, 2000.
- WRIGHT, C. "Some reflections on the acquisition of warrant by inference". In NUCCETELLI, S. (ed.), *New Essays on Semantic Externalism and Self-Knowledge*. MIT Press. pp. 57-78, 2003.
- WRIGHT, C. "Warrant for nothing (and foundations for free)?" *Aristotelian Society Supplementary Volume* 78 (1): 167–212, 2004.
- YAMADA, T. "The Epistemic Closure Principle and the Assessment Sensitivity of Knowledge Attributions". In REBUSCHI, M. et.al *Interdisciplinary Works in Logic, Epistemology, Psychology and Linguistics: Dialogue, Rationality, and Formalism*. Springer, 181-199, 2014.
- YAN, M. "When Does Epistemic Closure Fail?". *Analysis* 73 (2):260-264, 2013.
- YOURGRAU, P. "Knowledge and Relevant Alternatives". *Synthese*, 175-179, 1983.

ZAGZEBSKY, L. "The Inescapability of Gettier Problems". *Philosophical Quarterly*, vol. 44, 65-73, 1994.

ZAGZEBSKY, L. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. Cambridge University Press, 1996.

ZALABARDO, J. *Scepticism and Reliable Belief*. Oxford University Press, 2012.