



**PROCEEDINGS of the  
Brazilian Research Group on  
EPISTEMOLOGY**

---

**2016**

 editora fi

**Proceedings of the Brazilian Research Group on Epistemology** é uma publicação anual de pesquisa em epistemologia analítica contemporânea (ambas em suas vertentes tradicionais e formais), organizada e editada por membros da academia brasileira. Seu principal objetivo é divulgar o trabalho de epistemólogos latino-americanos para a comunidade internacional, numa tentativa de criar um novo canal de discussão com pesquisadores de outros continentes.

**Editores chefes:**

Luis Rosa e André Neiva

**Comitê editorial:**

Cláudio de Almeida  
Felipe de Matos Müller  
Kátia Martins Etcheverry  
Carlos Augusto Sartori  
Tiegue Vieira Rodrigues  
Alexandre Meyer Luz  
Tito Alencar Flores  
Rogel Esteves de Oliveira

 **editora fi**  
[www.editorafi.org](http://www.editorafi.org)



**PROCEEDINGS**  
**of the Brazilian**  
**Research Group on**  
**ÉPISTEMOLOGY**

The *Proceedings of the Brazilian Research Group on Epistemology* is an annual publication of research in contemporary analytic epistemology (both traditional and formal approaches), organized and edited by members of the Brazilian academia. It aims mainly at publishing the work of Latin-American epistemologists to the international community, in an attempt to create a new channel of debate with researchers from other parts of the globe.

## **Editors-in-chief:**

---

Luis Rosa e André Neiva

## **Editorial board:**

---

Cláudio de Almeida  
Felipe de Matos Müller  
Kátia Martins Etcheverry  
Carlos Augusto Sartori  
Tiegue Vieira Rodrigues  
Alexandre Meyer Luz  
Tito Alencar Flores  
Rogel Esteves de Oliveira



**PROCEEDINGS  
of the Brazilian  
Research Group on  
EPISTEMOLOGY**

---

**2016**

**$\Psi$**  editora fi

**Diagramação e capa:** Lucas Fontella Margoni

**Revisão técnica:** Stéphane Dias



Todos os livros publicados pela  
Editora Fi está sob os direitos da  
Creative Commons 4.0

[https://creativecommons.org/licenses/by/4.0/deed.pt\\_BR](https://creativecommons.org/licenses/by/4.0/deed.pt_BR)

---

#### Dados Internacionais de Catalogação na Publicação (CIP)

---

ROSA, Luis; NEIVA, André. (Orgs.)

Proceedings of the Brazilian Research Group on Epistemology  
[recurso eletrônico] / Luis Rosa; André Neiva (Orgs.) -- Porto  
Alegre, RS: Editora Fi, 2016.

206 p.

ISBN - 978-85-5696-001-6

Disponível em: <http://www.editorafi.org>

1. Epistemologia. 2. Lógica. 3. Teoria do Conhecimento.
4. Testemunho. 5. Revista. I. Título. II. Série.

CDD-121

---

Índices para catálogo sistemático:

1. Epistemologia 121

# Sumário

O PROBLEMA EPISTÊMICO DA CONTINGÊNCIA

---

**Felipe M. S. Miguel . 9**

KLEIN ON KNOWLEDGE FROM FALSEHOOD

---

**J. R. Fett . 31**

ALGUMAS RAZÕES EM DEFESA DA POSIÇÃO NÃO-  
CONFORMISTA NO DESACORDO ENTRE PARES

---

**José Leonardo Ruivo . 45**

SOBRE UMA OBJEÇÃO À CONDIÇÃO DE  
SEGURANÇA

---

**Lucas Roisenberg Rodrigues . 66**

RELAÇÃO DE SUPORTE

---

**Luiz Paulo Da Cas Cichoski . 83**

CONFIANÇA E TESTEMUNHO COMO  
INTERPESSOAIS: UMA TENTATIVA DE  
ANALOGIA ENTRE ÉTICA E EPISTEMOLOGIA DO  
TESTEMUNHO

---

**Patricia Ketzer . 111**

DEONTOLOGISMO EPISTÊMICO, CETICISMO E  
EPISTEMOLOGIA MORAL, AUTOCONHECIMENTO  
E FALIBILISMO INTROSPECTIVO:  
UMA TENTATIVA DE APROXIMAÇÃO

---

**Ricardo Rangel Guimarães . 139**

ANTI-ANTI-REDUCTIONIST CONSIDERATIONS  
ABOUT THE JUSTIFICATION OF TESTIMONIAL  
BELIEFS

---

**Luis Rosa & André Neiva . 161**

A AGÊNCIA EPISTÊMICA REFLEXIVA DE  
ERNEST SOSA

---

**Thiago Rafael Santin . 171**



# O PROBLEMA EPISTÊMICO DA CONTINGÊNCIA

---

*Felipe M. S. Miguel*<sup>1</sup>

## **Introdução**

O problema epistêmico da contingência, também conhecido como "problema da irrelevância dos fatores causais," ou "problema da variabilidade histórica,"<sup>2</sup> é um problema que passou a receber atenção de epistemólogos contemporâneos somente recentemente<sup>3</sup> e que vem sendo

---

<sup>1</sup> Doutorando pela PUCRS/CAPES.

<sup>2</sup> Bogardus (2013: 371).

<sup>3</sup> O problema, contudo, ao menos em certas formulações, pode ser encontrado em, por exemplo, Filo de Alexandria, que, segundo, Nathan Ballantyne, “confessa que ele não se surpreenderá se a ‘multidão volúvel e heterogênea’ crer ‘no que quer que seja que tenha sido transmitido a ela,’ seja pelos pais, mestres ou cultura. Filo parece preocupado, porém, com o fato de que filósofos, ‘que professam perseguir o que há de claro e verdadeiro nas coisas, estão divididos em brigadas e pelotões e registram dogmas que são discordantes.’ Fatores sociais parecem inculcar diferentes dogmas em diferentes grupos — e filósofos não estão isentos. Isso leva Filo a adotar uma resposta pirrônica típica, resultando no cancelamento dos opostos e suspensão de juízo.” Filósofos renascentistas e modernos, como Montaigne, Descartes, Locke, entre outros, também se preocuparam com o que Ballantyne chama de “problema da variabilidade histórica”, com “cada um considerando a influência da cultura em seu pensamento.” John Stuart Mill, por exemplo, em *On Liberty*, “captura vividamente um pensamento sobre variabilidade: ‘E o mundo, para cada indivíduo, significa a parte dele com que ele entra em contato; seu partido, sua seita, sua igreja, sua sociedade de classes. Nem é a sua fé nessa autoridade coletiva estremecida de qualquer modo por ele se conscientizar de que outras eras, países, seitas, igrejas, classes, e partidos pensaram, e pensam neste mesmo momento, exatamente o contrário. Nunca o perturba que mero acidente tenha decidido qual desses mundos numerosos é objeto de sua confiança, e que as mesmas

tratado com crescente interesse devido ao seu potencial de gerador de ceticismo em relação a crenças afetadas por fatores contingentes ou causalmente irrelevantes, como o lugar e o momento em que a crença é formada<sup>4</sup>. O problema é normalmente apresentado através de um contrafactual do tipo "se você tivesse nascido em lugar tal ao invés do lugar no qual você nasceu, você não acreditaria no que você acredita," ou "se você tivesse estudado em outra universidade, você não acreditaria na teoria na qual você acredita." Um exemplo apresentado por Adam Elga (2008), originalmente formulado por G. E. Cohen, ilustra bem o problema:

#### Analítico/Sintético

Um estudante joga uma moeda para o alto para determinar se ele deveria estudar em Harvard ou Oxford. Ele vai para Oxford e acaba endossando a distinção entre analítico e sintético. Mas ele descobre que, se tivesse ido para Harvard, ele teria rejeitado a distinção analítico/sintético, embora ele teria sido exposto às mesmas evidências relevantes (Elga, 2008: 1).

---

causas que fazem dele um frequentador de igreja em Londres, teriam feito dele um budista ou confucianista em Pequim. A ideia de Mill parece ser de que algumas de nossas convicções têm causas culturais ou históricas, e que essas causas são uma questão de 'mero acidente'. Ao refletir na maneira pela qual certas crenças são influenciadas por tais causas, a nossa 'fé' em nossas ideias deveria ser 'estremecida' ou 'perturbada.'" (Ballantyne, 2013: 241).

<sup>4</sup> Miriam Schoenfield resume a importância do problema do seguinte modo: "Eu acho que é extremamente importante que nós descubramos como deveríamos responder quando ficamos sabendo [da influência dos fatores irrelevantes]. Pois se for o caso de que, ao ficarmos sabendo que uma crença foi causada por uma influência irrelevante, a coisa racional a se fazer é abrir mão dessa crença, as consequências seriam bastante drásticas. Isso se deve ao fato de que muitas de nossas crenças que foram causadas por fatores irrelevantes serem crenças religiosas, morais, políticas e filosóficas." (Schoenfield, 2012: 2).

---

A questão, então, é se a conscientização de que a crença depende de um fator causal irrelevante deveria levar a pessoa a suspender juízo ou reduzir o seu grau de crença na proposição em questão. A posição de Elga, como veremos, é de que a redução ou não do grau de crença dependerá de estarmos violando ou não os nossos próprios padrões de raciocínio. Contanto que estejamos sendo fieis aos nossos padrões, não seria necessário reduzir o grau da crença. Há, contudo, outras posições sendo defendidas em relação ao problema da contingência, em particular as posições de Joshua Schechter, de Roger White, e de Miriam Schoenfield, Tomás Bogardus e Nathan Ballantyne. Na seção seguinte, procurarei resumir a posição de cada um desses autores.

### **Adam Elga**

O artigo não-publicado de Adam Elga, *Lucky to Be Rational*, é o ponto de partida da discussão sobre o problema epistêmico da contingência. Nesse artigo, Elga apresenta três casos de crenças que são formadas sob a influência de fatores arbitrários ou irrelevantes, nas quais os agentes percebem que um fator irrelevante desempenhou papel importante na formação da crença. Além do caso apresentado acima, Elga utiliza os seguintes casos:

#### Conservadorismo/Liberalismo

Fred percebe que se os seus pais tivessem morado numa vizinhança mais conservadora, ele teria – com base na mesma evidência – crido em ideias políticas bastante diferentes daquelas que ele crê. Além disso, o método através do qual os seus pais optaram entre vizinhanças liberais e conservadoras foi o de arremço de moeda para cima (Elga, 2008: 1).

### Teísmo/Ateísmo

Jill crê em Deus, ao contrário de seu irmão Jack. Eles foram criados juntos e expostos às mesmas evidências. Os seus pais confessaram a Jack que “quando vocês eram crianças, nós demos a ambos tratamentos que influenciaram o desenvolvimento de seus cérebros. Jill recebeu o ‘tratamento teísta’, que fez com que argumentos para a existência de Deus parecessem mais persuasivos para ela. Você recebeu o ‘tratamento ateu’, com o efeito oposto. Nós jogamos uma moeda para cima para determinar quem receberia qual tratamento” (Elga, 2008: 1).

A questão, em cada caso, então, é se a conscientização de que a crença foi formada por influência de fatores irrelevantes deveria levar o agente a reduzir o seu grau de confiança na proposição em questão. No primeiro caso, ele percebe que a sua crença de que a distinção entre analítico e sintético é verdadeira foi influenciada pelo fato dele ter estudado em Oxford ao invés de Harvard. No segundo exemplo, o agente se conscientiza de que as suas crenças políticas foram influenciadas pelo fato dele ter vivido em determinada vizinhança. No terceiro caso, o agente é informado de que o fato dele ser um teísta se deve a ele ter recebido um tratamento cerebral teísta.

Para Elga, ao se conscientizar que a sua crença foi influenciada por fatores contingentes, o agente tem duas opções: concluir que a sua crença se deve, de fato, às influências circunstanciais e, assim, reduzir o seu grau de crença na proposição em questão, ou concluir que os fatores irrelevantes, por sorte, o conduziram ao caminho da verdade e, portanto, continuar a manter a crença na proposição. A resposta adequada, segundo Elga (2008: 2), “depende de se o fator irrelevante colocou o agente em risco de não ser fiel aos seus próprios padrões de raciocínio”.

---

Em defesa de sua resposta, Elga se concentra no caso Analítico/Sintético, procurando mostrar como o fornecimento dos detalhes do caso pode ajudar na formulação da resposta mais adequada ao problema da contingência. Elga propõe dois cenários: um em que os alunos de Harvard e Oxford possuem os mesmos padrões epistêmicos, e um em que os alunos possuem padrões epistêmicos diferentes. No cenário de Mesmos Padrões, os alunos de ambas as universidades são educados de modo a desenvolverem os mesmos padrões epistêmicos, mas os alunos violam esses padrões ao adotarem as opiniões dos especialistas locais. No de Padrões Diferentes, os alunos de cada universidade adotam padrões epistêmicos fundamentais diferentes – os de Harvard adotam os padrões de Harvard, e os de Oxford adotam os padrões daquela universidade, de tal modo que alunos de ambas adotam as opiniões dos especialistas locais porque eles compartilham dessas opiniões.

Para Elga, no caso dos Mesmos Padrões, os fatores irrelevantes fazem com que o agente viole os seus próprios padrões de raciocínio, através da aceitação apressada da opinião dos especialistas de sua universidade. Nesse caso, a atitude correta é reduzir o grau de convicção na crença. No caso dos Padrões Diferentes, por outro lado, os fatores irrelevantes atuam através da influência sobre quais padrões de raciocínio o agente virá a adotar, e não o levando a violar os seus padrões de raciocínio. Nesse caso, a conscientização da influência dos fatores irrelevantes não deveria levar o agente a reduzir o seu grau de confiança na crença em questão.

Elga acredita ter chegado a uma solução satisfatória, mas não desprovida de desconforto. Satisfatória porque o que o problema da contingência mostraria é que a descoberta de que fatores irrelevantes influenciaram as nossas crenças tornaria mais evidente a existência de padrões de raciocínio coerentes incompatíveis com os

nostros próprios padrões. O diagnóstico oferecido por Elga para o desconforto, por outro lado, é que os casos de influência de fatores irrelevantes tornariam explícita a possibilidade de “erro global” e, portanto, de ameaça cética.

## Joshua Schechter

Em seu breve artigo não-publicado *Luck, Rationality, and Explanation: A Reply to Elga's Lucky to Be Rational*,<sup>5</sup> Schechter rejeita a conclusão de Elga de que a resposta adequada ao problema da contingência depende dos nossos próprios padrões de raciocínio, isto é, de que a resposta adequada seria o agente reduzir o seu grau de crença somente se os fatores irrelevantes o levassem a desprezitar os seus padrões de raciocínio. Para Schechter (2008: 2), teria passado despercebido a Elga “uma conexão importante entre justificação e explicação” que, se levada em conta de forma adequada, nos levaria a concluir que, frente à conscientização de que fatores contingentes afetam as nossas crenças, a atitude correta é a redução do nosso grau de crença, mesmo se estivermos respeitando os nossos padrões de raciocínio.

De acordo com Elga, a conscientização de que determinadas crenças nossas foram influenciadas por fatores irrelevantes ou contingentes traria à nossa atenção a existência de diversos padrões de raciocínio coerentes que são incompatíveis com o nosso próprio padrão de raciocínio. Isso, contudo, não deveria produzir maiores consequências. Na perspectiva de Elga, o problema da contingência somente tornaria saliente determinada preocupação cética imemorialmente presente no exercício

---

<sup>5</sup> Tanto o artigo de Elga como o de Schechter foram apresentados inicialmente na *Bellingham Summer Philosophy Conference*, de 2008, com o artigo de Schechter tendo sido apresentado como comentário ao de Elga.

---

filosófico, não devendo produzir pressão adicional para se reduzir o nosso grau de confiança na crença.

Schechter, por outro lado, apresenta um diagnóstico um pouco diferente. Para ele, o desconforto produzido pela conscientização de que determinada crença foi formada por influência de fatores irrelevantes se deve à conscientização de que não possuímos explicação para a confiabilidade do padrão de raciocínio relevante.

Portanto, para Schechter, o problema da contingência não é um problema de ameaça cética ou meramente do desconforto da descoberta de que há diversos padrões de raciocínio coerentes, mas um problema relacionado à desconfortável situação de nos vermos sem a possibilidade de fornecer explicação para a confiabilidade do nosso padrão de raciocínio. E “em muitos casos, quando descobrimos um fator causal irrelevante de uma crença, faz-se necessária redução no grau de confiança na crença, ainda que,” contrário ao que é defendido por Elga, “estejamos agindo de acordo com os nossos padrões de raciocínio” (Schechter, 2008: 11).

### **Roger White**

White começa a sua abordagem do problema da contingência em seu artigo *You just Believe that Because...* fazendo uma distinção entre dois sentidos nos quais seria possível indagar do porquê de determinada crença: o da indagação das razões que justificam a crença, e o da indagação da explicação causal num sentido mais amplo. É a relevância deste último tipo de indagação que é o foco da investigação que interessa a White em seu artigo. O propósito de White é tentar entender por que haveria um problema epistêmico com crenças adquiridas por fatores contingentes, tal como o ambiente no qual a pessoa adquiriu a crença, e indagar se esse problema poderia levar

à conclusão de que teríamos de duvidar da maioria das crenças formadas por esse método.

Uma distinção importante que White procura fazer é entre o problema da contingência, a reavaliação de crenças e o problema do desacordo. A reavaliação de crenças ocorre quando o agente repensa as bases de determinada crença que até então ele sustentava sem questionamento e é facilmente distinguível dos demais problemas. O problema da contingência e do desacordo, por outro lado, estão claramente conectados, segundo White. A princípio, esses dois problemas se distinguiriam na medida em que o problema do desacordo diria respeito ao efeito que a existência de pessoas com opiniões diferentes sobre determinada proposição tem sobre a racionalidade da crença nessa proposição, enquanto o problema da contingência diria respeito “à significância epistemológica (no que diz respeito a informação) que causas de crenças têm à parte de outros fatores tais como desacordo” (White, 2010: 4). White, porém, sugere que um exame mais detalhado revelaria grande dificuldade em se distinguir claramente o problema da contingência do problema do desacordo. Enquanto o problema da contingência envolveria a existência de correlação (os alunos de Oxford acreditam em  $p$ , enquanto os alunos de Harvard acreditam em  $\textit{não-p}$ ), o problema do desacordo não envolveria correlação (as opiniões dos alunos estariam aleatoriamente distribuídas entre  $p$  e  $\textit{não-p}$ ). Mas qual seria a diferença relevante entre ambos? Aparentemente, na situação de correlação seria possível esboçar uma explicação para a diferença entre crenças, mas nada além disso, e isso não seria suficiente para apaziguar as preocupações trazidas por casos como Analítico/Sintético. Nesse ponto, White conclui provisoriamente que a etiologia da crença em Analítico/Sintético não adicionaria nada além do problema do desacordo.



---

Tendo deixado de lado momentaneamente a questão da inter-relação entre os problemas do desacordo e da contingência, White parte para a exploração de algumas das possíveis explicações para o que exatamente tornaria a influência de fatores causais irrelevantes problemática. Entre elas: de que seria um problema de insensibilidade (isto é, um problema com o não rastreamento da verdade), de relevância explicativa (isto é, para que a minha crença em  $p$  seja justificada seria preciso que a explicação última da minha crença que  $p$  fizesse referência a se  $p$ ), um problema de meios independentes (isto é, um problema da falta de um meio independente de se avaliar o percurso que levou o agente à verdade), e um problema relacionado a sorte. White conclui que nenhuma dessas hipóteses explicaria adequadamente o desconforto produzido pelo problema da contingência e retorna à conclusão que ele havia alcançado provisoriamente de que o problema da contingência não pareceria adicionar nada além do que é colocado pelo problema do desacordo:

Mas agora nós simplesmente retornamos à questão do desacordo. O que está direcionando a preocupação específica aqui não tem nada a ver com fatos que [o agente do caso Analítico/Sintético] descobriu sobre a etiologia de suas próprias crenças. Trata-se apenas do fato de que ele tem evidência de que há filósofos muito inteligentes e bem-informados que diferem em suas opiniões (White, 2010: 30).

### **Miriam Schoenfield**

Em *Permission to Believe: Why Permissivism Is True and What it Tells us about Irrelevant Influences on Belief*, Miriam Schoenfield defende uma posição permissivista para o problema do desacordo entre pares e procura mostrar como essa posição ajuda a dissipar algumas das

preocupações acarretadas pelo problema da contingência. Schoenfield defende que há casos em que a resposta permissivista se aplica ao caso do desacordo de pares e outras em que isso não ocorre, e que essa mesma distinção se aplicaria ao problema dos fatores causais irrelevantes.

Permissivismo é a posição de “que, às vezes, há mais de uma resposta racional para um determinado conjunto de evidências” (Schoenfield, 2014: 2). No que diz respeito ao caso específico do desacordo entre pares, o permissivista defende, essencialmente, que numa situação em que dois agentes são pares epistêmicos, isto é, compartilham o mesmo conjunto de evidências e são igualmente bons avaliadores desse conjunto de evidências, e discordam em relação a se  $p$  é verdadeira, seria permissível que eles mantivessem essa posição de desacordo em relação a  $p$  racionalmente. Mas, de acordo com Schoenfield, essas condições de permissividade não se aplicariam necessariamente a todos os casos de desacordo entre pares. Nesses casos, isto é, nos casos não-permissivos, ao contrário dos casos em que as condições se aplicam, seria necessário reduzir a confiança na crença.

Schoenfield defende que essa posição permissivista pode ser transposta para o caso dos fatores irrelevantes de tal modo que, em casos em que é permissível, dada a evidência, acreditar tanto em  $p$  quanto em  $\text{não-}p$ , a conscientização de que a crença foi causada por um fator irrelevante não deveria produzir um anulador para a racionalidade dessa crença. Em outras circunstâncias, no entanto, em que temos razões para crer que os fatores irrelevantes nos levaram a raciocinar de maneira irracional, essa posição permissivista não se aplica e, portanto, a resposta correta é reduzir o nosso grau de confiança na crença.

O argumento de Schoenfield é, em suma, o seguinte:

P1. O permissivismo é verdadeiro.

P2. Se o permissivismo é verdadeiro, a ideia de que se deveria reduzir significativamente a confiança em casos permissivos de influências irrelevantes é sem motivação.

P3. Se a ideia de que se deve reduzir significativamente a confiança em casos permissivos de influências irrelevantes é sem motivação, não se deve reduzir a confiança em tais casos.

C1. Não se deve reduzir significativamente a confiança com base em influências irrelevantes em casos permissivos.

Tomemos uma pessoa que toma consciência de que ela somente tem as crenças que ela tem (políticas, econômicas, morais, religiosas ou filosóficas) e somente considera convincentes os argumentos a favor dessas crenças devido à influência de sua comunidade. Essa pessoa passa a perceber que, caso ela tivesse sido socializada em outra comunidade e exposta ao mesmo conjunto de evidências, ela acreditaria na negação da proposição que ela crê. Nesse caso, defende Schoenfield, ambas as posições seriam racionais e essa pessoa não precisaria reduzir significativamente o seu grau de crença. Como resume Schoenfield,

Se você começa com E e você sabe que, com base em E, é permissível crer em p e permissível crer em não-p, ficar sabendo que a sua crença em p foi causada por um fator irrelevante não te dá uma razão para reduzir a sua confiança em p (Schoenfield, 2014: 13).

## **Tomás Bogardus**

Bogardus está interessado especificamente no problema da contingência para crenças religiosas. O ponto de partida de sua investigação é o debate sobre o efeito que a existência de pluralismo religioso tem para a posição exclusivista do ponto de vista religioso. O argumento do pluralista seria essencialmente o seguinte:

(1) Se você tivesse nascido e crescido em outro lugar, em outro momento, você teria crenças religiosas diferentes.

Portanto,

(2) As suas crenças religiosas não contam como conhecimento genuíno.

Esse argumento, no entanto, é inválido. Isso fica claro se substituirmos “crenças religiosas” pelo nome da cidade na qual você nasceu. Essa substituição levaria à conclusão de que você não teria conhecimento genuíno do nome da cidade onde você nasceu.

Para Bogardus, o problema da contingência para crenças religiosas é um problema de método de formação de crença. Bogardus vê o argumento pluralista contra o exclusivismo religioso como sendo baseado na ideia de que o método de formação da crença religiosa (socialização) não seria confiável, enquanto o método através do qual o pluralista chegaria à sua crença (reflexão racional) seria confiável. Desse modo, Bogardus nota que o argumento do pluralista teria uma força menor do que talvez fosse pretendido originalmente: ele estaria afirmando que somente o exclusivista não-reflexivo não teria conhecimento. Caso contrário, isto é, caso o pluralista

---

buscasse incluir no seu argumento exclusivistas reflexivos, o argumento do pluralista se auto-refutaria.

Com isso em mente, Bogardus investiga três possíveis diagnósticos sobre qual seria exatamente o problema envolvido no problema da contingência. A primeira possibilidade é que o problema da contingência seria um problema de segurança na aquisição de conhecimento. Nessa interpretação, o que o pluralista estaria dizendo é que o método de aquisição de crenças religiosas seria “modalmente frágil”: se o exclusivista religioso tivesse nascido num lugar ou momento em que outra religião fosse predominante, e tivesse formado as suas crenças através do mesmo método que ele de fato utilizou, ele facilmente poderia ter acreditado falsamente (de acordo com a sua própria visão) nessa outra religião. Com isso, a crença religiosa do exclusivista não teria sido formada de forma segura e, portanto, ela não constituiria conhecimento.

Bogardus assinala que ambas as inferências desse “argumento da segurança” são inválidas. Em primeiro lugar, do fato de que a crença do exclusivista poderia facilmente ter sido falsa não se pode inferir que o método era inseguro quando ele foi efetivamente utilizado. Em segundo lugar, segurança não seria necessária para conhecimento<sup>6</sup>; portanto, não seria uma inferência válida concluir que por não ter sido formada de forma segura uma crença não constituiria conhecimento. Desse modo, Bogardus (2013: 391-392) conclui que “a contingência de nossas crenças religiosas não mostra que elas foram formadas de modo inseguro; e mesmo se tivessem, segurança não seria necessária para conhecimento”.

---

<sup>6</sup> Bogardus oferece nesse artigo um resumo de sua defesa da não necessidade de segurança oferecida em outro artigo: Bogardus (2014), ‘Knowledge under Threat’.

Tendo descartado a interpretação do problema da contingência para crenças religiosas como um ataque à segurança da crença religiosa, Bogardus examina a possibilidade do problema da contingência ser um problema de acidentalidade de formação de crenças. Seguindo a análise de Peter Unger de conhecimento em termos de acidentalidade, Bogardus explora a possibilidade da crítica do pluralista ser no sentido da crença do exclusivista ser “mero acidente” ou resultado de um “acidente de nascimento”. O argumento seria, portanto, de que, porque a crença do exclusivista emerge de uma acidentalidade, essa crença não configuraria conhecimento. O problema com essa interpretação, segundo Bogardus, é que, assim como no caso da análise de segurança acima, do contrafactual

(3) Se o exclusivista religioso tivesse nascido num lugar ou momento em que outra religião fosse predominante, e tivesse formado as suas crenças através do mesmo método que ele de fato utilizou, ele facilmente poderia ter acreditado falsamente nessa outra religião.  
não é possível concluir que

(4) É acidental que o exclusivista esteja certo (se estiver).

Para demonstrar isso, Bogardus oferece o seguinte contraexemplo: se você fosse amigo de Copérnico em 1515 e ele, por medo de represália, compartilhasse os dados por ele coletados que justificariam a crença no heliocentrismo somente com você, e através disso você viesse a saber que o heliocentrismo é verdadeiro, teria sido somente por acidente (você estar vivo naquele momento e naquelas circunstâncias) que você teria adquirido esse

---

conhecimento. Nesse caso, o tipo de acidentalidade não é o mesmo descrito em (4).

Além disso, a inferência de (4) a

(5) As crenças religiosas do exclusivista não configuram conhecimento

seria válida somente se conhecimento excluísse o tipo de acidentalidade expresso em (4). Mas, de acordo com a argumentação apresentada por Bogardus (2013: 385)<sup>7</sup>, “um sujeito pode saber que p ainda que seja acidental que ele esteja certo a respeito de ser o caso que p”. Sendo assim, ambas as inferências seriam inválidas e a formulação do problema da contingência em termos de acidentalidade falharia. Como resume Bogardus (2013: 392), “a contingência de nossas crenças religiosas não mostra que elas são acidentalmente verdadeiras, caso sejam verdadeiras; e mesmo se forem, não-acidentalidade não seria necessária para conhecimento”.

Por fim, Bogardus examina uma terceira opção, a de que o problema da contingência seria essencialmente um problema de desacordo de pares ou, como ele prefere colocar, um problema de simetria. E é essa terceira opção, Bogardus sustenta, que oferece o diagnóstico mais adequado para o problema em discussão. Sob essa perspectiva, o pluralista estaria argumentando que haveria uma simetria entre os seguidores das diferentes religiões. Ou como coloca Bogardus,

Se eu aceito o cristianismo como um resultado de educação quando criança ou socialização, mas venho a saber que muitos outros aceitam muitas outras diferentes religiões na mesma base [...], não

---

<sup>7</sup> Devido a limitações de espaço, limito-me aqui a direcionar o leitor às páginas 385-387 do texto de Bogardus.

seria inaceitavelmente arbitrário ou petição de princípio ainda assim manter as minhas crenças da minha educação infantil e socialização? O fato de que essa educação infantil e socialização são minhas não pode apropriadamente romper a simetria epistêmica, e então eu deveria abandonar as minhas crenças religiosas (Bogardus, 2013: 389).

Embora o argumento da simetria esteja em posição favorável em relação aos diagnósticos anteriores, ele também, defende Bogardus, não é capaz de fornecer um anulador para a crença religiosa.

O argumento sugerido por Bogardus é o seguinte,

(6) Se você tivesse nascido num lugar ou momento em que outra religião fosse predominante, e tivesse formado as suas crenças através do mesmo método que você de fato utilizou, você facilmente poderia ter acreditado falsamente nessa outra religião.

(7) Muitas pessoas que de fato nasceram e cresceram em outro lugar, em outra época, e formaram crenças religiosas usando o mesmo método que você de fato utilizou formaram crenças religiosas diferentes.

(8) Não está claro se você está numa posição epistêmica melhor que o seu eu contrafactual ou que as pessoas que discordam de você.

Portanto,

(9) As suas crenças religiosas não constituem conhecimento.

O veredito de Bogardus (2013: 392) em relação a esse argumento da simetria é que ele “ou se auto-refuta ou se aplica somente a uma fração da população que mantém



---

crenças religiosas de maneira que não é reflexiva e que não está ciente [...] das premissas devido a negligência ou irracionalidade”. Mas por que ele se auto-refutaria? Ele se auto-refutaria porque esse mesmo argumento se aplicaria ao ateísmo. As suas premissas se aplicariam a todos nós. Aplicar-se-iam até mesmo à afirmação da solidez do argumento. Desse modo, o escopo do argumento teria de ser reduzido a crentes religiosos que não são reflexivos. Mas mesmo assim, muitos crentes religiosos que não são reflexivos não seriam atingidos pelo argumento (“...ou se aplica somente a uma fração da população...”), pois “se um crente que não é reflexivo não é culpavelmente consciente das premissas desse argumento, suas crenças religiosas podem muito bem ser racionais a despeito da verdade das premissas” (Bogardus, 2013: 391)<sup>8</sup>.

### **Nathan Ballantyne**

Em *The Problem of Historical Variability*, Nathan Ballantyne procura investigar se há realmente algo de problemático em termos de racionalidade com crenças influenciadas por nossos históricos pessoais [*backgrounds*] devido ao caráter de algum modo acidental ou arbitrário dessas crenças. Ballantyne (2013: 239) prefere chamar esse problema de problema da variabilidade histórica ou, simplesmente, *variabilidade* – “a ideia de que as nossas convicções em disputa variam de acordo com diferentes históricos pessoais”. No artigo, Ballantyne procura tornar mais claros os pressupostos envolvidos no problema da variabilidade e formula dois argumentos que recomendam

---

<sup>8</sup> A defesa dessas conclusões aqui certamente parece demasiadamente apressada, mas, devido a limitações de espaço, limito-me aqui a direcionar o leitor às páginas 390-1 do texto de Bogardus para uma apresentação mais adequada do argumento.

dúvida ou ceticismo em relação a crenças formadas desse modo.

O primeiro argumento, que Ballantyne chama de *Argumento da Simetria*, é obtido através da eliminação gradual das diferenças epistêmicas relevantes entre você e o seu eu contrafactual, o que garante certa simetria:

*P1* – se você tem razão para crer que a sua crença que *p* é tal que se o seu histórico pessoal tivesse sido diferente em certos aspectos, então você não teria aceitado *p*, ainda que você tivesse usado a mesma evidência para *p* e as habilidades cognitivas relevantes a crer apropriadamente em *p* que você na verdade usou, *então* a sua crença que *p* é irracional.

*P2* – você tem razão para crer que a sua crença que *p* é tal que se o seu histórico pessoal tivesse sido diferente em certos aspectos, então você não teria aceitado *p*, ainda que você tivesse usado a mesma evidência para *p* e as habilidades cognitivas relevantes a crer apropriadamente em *p* que você na verdade usou.

*C* – a sua crença que *p* é irracional.

Mas por que pensar que *P2* é plausível? Zé decide entre se matricular em filosofia na UFRGS ou na PUCRS através de um arremesso de moeda. Independentemente do lado em que a moeda caiu, dentro de alguns anos Zé e o seu eu contrafactual terão as mesmas evidências relevantes e habilidades cognitivas no que diz respeito a certas teses. Alguns anos mais tarde, Zé reflete sobre aquele arremesso de moeda e conclui que se a moeda tivesse caído com o outro lado para cima ele não aceitaria hoje a teoria *p*, embora ele tivesse as mesmas evidências relevantes e habilidades cognitivas no que diz respeito a *p*<sup>9</sup>. De acordo

---

<sup>9</sup> Portanto, uma diferença importante entre o problema do desacordo entre pares e o da variabilidade seria que, enquanto o primeiro exige desacordo apenas *atual* com outra pessoa, o segundo funciona mesmo com desacordo *possível* com o seu eu contrafactual.

---

com Ballantyne, a situação de Zé é análoga à situação de muitos filósofos com relação a suas crenças filosóficas. Sendo assim, esses filósofos deveriam aceitar P2. Mas para Ballantyne, a aceitação de que nós teríamos eus contrafatuais discordantes com as mesmas evidências relevantes e habilidades cognitivas é algo com enorme potencial cético.

Ballantyne, no entanto, acredita que por melhor que seja o argumento da simetria, ele falharia em capturar algo importante sobre variabilidade:

Nós percebemos que, se a vida tivesse seguido uma direção diferente, então nós teríamos evidências diferentes no que diz respeito a uma vasta gama de situações e proposições. Isso se deve ao fato de que, na maioria dos casos, nós e os nossos eus contrafatuais estaríamos em situações evidencialmente assimétricas (Ballantyne, 2013: 248).

Ballantyne procura, então, formular um segundo argumento que leve em conta o que ele chama de arbitrariedade evidencial. As proposições  $p$  e não- $p$  são evidencialmente arbitrárias se, no que diz respeito a evidência  $E$ ,  $E$  não apoia uma dessas proposições ao invés da outra. Racionalidade repele arbitrariedade. Se em situação como essa alguém crê em uma das proposições, essa pessoa estaria realizando algo epistemicamente impróprio.

Imaginemos agora que o arremesso de moeda de Zé acabou levando-o a cursar filosofia na UFRGS, onde ele obteve evidência  $E$ . Mas que, se ele tivesse ido para a PUCRS, ele obteria evidência  $E^*$ . Ao contrário do primeiro caso envolvendo simetria evidencial, temos agora um caso de assimetria. E Zé parece estar bem fundamentado em crer em  $p$  dadas as evidências obtidas na UFRGS, e o seu eu contrafactual em não- $p$  dadas as evidências obtidas na

PUCRS. Segundo Ballantyne, nesse caso de assimetria evidencial o que produz perplexidade é o fato de que Zé poderia ter facilmente ter obtido outro conjunto de evidências, afinal, ele cursou filosofia na UFRGS, obtendo evidências E ao invés de E\*, por causa de um mero arremesso de moeda.

A partir dessas e outras<sup>10</sup> constatações, Ballantyne formula o segundo argumento, que ele chama de *Argumento da Arbitrariedade*:

P1 – se você tem razão para crer que a sua crença que a proposição p é tal que (i) se o evento e1 ocorresse, você creia em p e (ii) se o evento e2 ocorresse, você não aceitaria p, e (iii) você não tem razão para pensar que e1 torna mais provável que e2 que você agora crê verdadeiramente e não falsamente se p, e (iv) você não tem razão para pensar que tanto e1 como e2 tornam altamente provável que você creará verdadeiramente e não falsamente no que diz respeito a p, *então* a sua crença que p é irracional.

P2 – você tem razão para crer que a sua crença que a proposição p é tal que (i) se o evento e1 ocorresse, você creia em p e (ii) se o evento e2 ocorresse, você não aceitaria p, e (iii) você não tem razão para pensar que e1 torna mais provável que e2 que você agora crê verdadeiramente e não falsamente se p, e (iv) você não tem razão para pensar que tanto e1 como e2 tornam altamente provável que você creará verdadeiramente e não falsamente no que diz respeito a p.

C – a sua crença que p é irracional.

O que diferenciaria os dois argumentos é que somente no argumento da arbitrariedade: (1) é possível que você e o seu eu contrafactual tenham evidências relevantes e habilidades cognitivas dissimilares; (2) é perguntado se a

---

<sup>10</sup> Que serão omitidas aqui devido a limitações de espaço.

---

sua história te coloca numa situação vantajosa para conhecer p em comparação com uma história alternativa. Ballantyne conclui que os dois argumentos apresentados são bastante plausíveis, mas que, ainda que eles falhassem, a variabilidade continuaria sendo um problema a ser enfrentado por epistemólogos. A mais importante lição a ser tirada disso tudo, defende Ballantyne (2013: 257), é que “não deveríamos ter muita confiança a respeito de assuntos difíceis e em disputa”.

## **Conclusão**

Neste ensaio, procurei apresentar o problema da contingência e as respostas que têm sido oferecidas até o momento. Em particular, procurei investigar no que exatamente consistiria o problema e qual seria a resposta adequada do agente que se vê numa situação de descoberta de que determinada crença sua foi formada sob a influência de fatores causais irrelevantes ou contingentes. Vimos que diversas respostas para o problema têm sido propostas até o momento, e que, mais do que isso, a própria formulação do problema e a compreensão do que está em jogo nos exemplos que expressam o problema tem angariado pouco acordo. O problema da contingência constitui-se num campo de estudos ainda pouco explorado e que merece a atenção crescente dos pesquisadores no campo de epistemologia, especialmente no que diz respeito ao que distinguiria exatamente esse problema do problema mais amplo do desacordo racional.

## **Referências**

Ballantyne, Nathan (2013). ‘The Problem of Historical Variability’. In: Machuca, Diego (Ed.). *Disagreement and Skepticism*, Routledge Studies in Contemporary Philosophy, New York: Routledge, 239-259.

Bogardus, Tomás (2013). ‘The Problem of Contingency for Religious Belief’, *Faith and Philosophy* 30(4): 371-392.

Bogardus, Tomás (2014). ‘Knowledge under Threat’, *Philosophy and Phenomenological Research* 88(2): 289–313.

Elga, Adam (2008). ‘Lucky to Be Rational’, Paper Presented at the Bellingham Summer Philosophy Conference. Publicado na internet, disponível através da seguinte URL:  
<<http://www.princeton.edu/~adame/papers/bellingham-lucky.pdf>>.

Schechter, Joshua (2008). ‘Luck, Rationality, and Explanation: a Reply to Elga’s “Lucky to Be Rational.”’ Publicado na internet, disponível através da seguinte URL:  
<<http://www.brown.edu/Departments/Philosophy/onlinepapers/schechter/LuckRationalityExplanation.pdf>>.

Schoenfield, Miriam (2014). ‘Permission to Believe: Why Permissivism is True and What it Tells Us about Irrelevant Influences on Belief’, *Noûs* 48(2): 193-218. Citações no texto são da versão disponível no website da autora:  
<http://www.miriamschoenfield.com/research/permission-to-believe-82812.pdf>

White, Roger (2010). ‘You just Believe that Because...’, *Philosophical Perspectives* 24(1): 573-615.

# KLEIN ON KNOWLEDGE FROM FALSEHOOD

---

*J. R. Fett*<sup>1</sup>

## 1. Introduction

How do inferential knowledge and falsehoods are epistemically related to each other? Inferential knowledge is the kind of propositional knowledge that one gets via reasoning. Reasoning, in turn, is a cognitive process in which a belief is produced by virtue of and based on other(s) belief(s)<sup>2</sup>. As regards the role that falsehoods play in the production of inferential knowledge, there is an entrenched opinion that is shared by the philosophical community. Traditionally, epistemologists hold that false beliefs do not play any positive role in the production of inferential knowledge. Rather, their presence in a piece of reasoning is understood as very harmful or even totally devastating when it comes to the acquisition of knowledge via inference.

The beginning of what we call contemporary epistemology gave rise to a literature that clearly supports that philosophical tradition concerning knowledge and false beliefs. It suffices to see that in every Gettier case in which there is a falsehood playing a relevant role in the reasoning that the protagonist instantiates, it is exactly the false belief [that constitutes] the knowledge-preventing element<sup>3</sup>. The intuition behind this verdict is that falsehoods are *bad indicators of truth*, and the only way they can give one the

---

<sup>1</sup> E-mail: jrfett01@gmail.com

<sup>2</sup> See Davis, 1999: 428. See also Audi, 2011, chapter 8.

<sup>3</sup> See Gettier, 1963.

truth is when luck intervenes. But since Plato, knowledge is thought of as being incompatible with luck.<sup>4</sup> Thus false beliefs can never give one knowledge, even though sometimes, as in Gettier cases, they can give one the truth.

The intuition to which we referred above, that falsehoods can never give one knowledge – though they can give one the truth when luck intervenes – is the core of the first approach aimed at solving the Gettier problem, the No-False-Lemmas account<sup>5</sup>. Mark Clark, its proponent, upon facing the counterexamples Gettier had put forward, found that what explains why the protagonists lack knowledge when they are Gettierized is that their reasoning depended essentially on a false premise, which *only by chance* led them to the truth. If the relevant falsehood were removed from the reasoning, none of them would get around to believing a true proposition, as they luckily did. The fragility exhibited in their truly believing is what makes us refrain from calling them *knowers*.

The conclusion that one cannot get knowledge from false beliefs was supported by so compelling an argument that, though it had not served well as an anti-Gettier weapon – due to its limited scope – its core intuition has been incorporated by every other approach to the Gettier problem that has been advanced since then. So that in the post-Gettier literature, it is a common place that there cannot be knowledge from falsehood. This central intuition can be properly expressed as follows.

**Orthodoxy:** Inferential knowledge of a conclusion requires true relevant premises<sup>6</sup>.

---

<sup>4</sup> Plato expresses this view in his *Menon*.

<sup>5</sup> See Clark, 1963.

<sup>6</sup> Besides the post-Gettier literature on the analysis of knowledge, Orthodoxy is backed by an Aristotelian thesis according to which only



---

Relatively recently, however, some very intriguing cases flying in the face of the Orthodoxy were put forward. Consider one of them, offered by Risto Hilpinen.

**Temperature:** A mother suspects that her child has fever, and when she measures the temperature and looks at the thermometer, she takes it to read 40.0 degrees Celsius. Let us call the child 'c' and the thermometer again 'M'. If the thermometer is fairly accurate and the mother has reasonably good eyesight, we can say under these circumstances that she knows that the child has temperature, i.e., that  $T(c) > 37.0$ . (To say that the child has temperature is just another way of saying that the temperature of the child is more than 37 degrees Celsius.) But the mother need not have perfect eyesight and the thermometer need not be completely accurate (few ordinary thermometers are): the actual thermometer reading might be  $M(c) = 39.8$ , and the actual temperature of the child might be 39.2 degrees Celsius. [...] This example suggests that a person can know things not only on the basis of (valid) inference from what he or she knows, but in some cases even on the basis of inference from what is not known (or even true), provided that the latter (evidential) propositions are sufficiently close to the truth (Hilpinen, 1988: 163-164).

Hilpinen invites us to think about the hypothetical situation in which the thermometer reading is 39.8° but the actual temperature of the child is 39.2°. He suggests that in this situation the mother *can* come to know that her child has temperature based essentially (and solely) on a false belief for which she gains evidence via the thermometer

reading. The mother infers from the false premise that her child has 39.8° the true and justified belief that her child has fever. It seems that she *does* know her conclusion. It seems clear that she gets knowledge from falsehood.

We saw above a piece of reasoning with a falsehood as its premise, and it seems that one can get knowledge through that reasoning – that is, that one can get knowledge from falsehood, *contra* Orthodoxy. This datum demands that something in the epistemological theorizing concerning the relation between falsehoods and inferential knowledge needs to be reconsidered in light of that possibility.

Ted Warfield and Peter Klein contributed to highlighting this threat against the Orthodox View on inferential knowledge. It turns out that what since Aristotle had been taken as “set in stone”, as it were, has been challenged by a variety of counterexamples. Let us take a look at some of them.

**Handout:** Counting with some care the number of people present at my talk, I reason: ‘There are 53 people at my talk; therefore my 100 handout copies are sufficient’. My premise is false. There are 52 people in attendance — I double counted one person who changed seats during the count. And yet I know my conclusion (Warfield, 2005: 407-408).

**The Appointment Case:** On the basis of my apparent memory, I believe that my secretary told me on Friday that I have an appointment on Monday with a student. From that belief, I infer that I do have an appointment on Monday. Suppose, further, that I do have an appointment on Monday, and that my secretary told me so. But she told me that on Thursday, not on Friday. I know that I have such an appointment even though I

---

inferred my belief from the false proposition that my secretary told me on Friday that I have an appointment on Monday (Klein, 2008: 36).

It is highly intuitive for a number of people who have considered these cases that they exemplify the unknown (until then) possibility of getting knowledge from falsehood. In the first case, I gain knowledge by counting the people in the audience and reasoning based on that counting that my copies will suffice, though I double counted one of the participants. In the second case, I gain knowledge that I have an appointment on Monday even though I mistake the day in which I was informed of that.

At the end of the day, maybe we will be disabused of that entrenched idea we called Orthodoxy by the pull of the putative cases of knowledge from falsehood. But now let us clean up the whole scenario a bit. After we do that, we will be in a better position to assess the power of these counterexamples.

## 2. Overdetermination

Epistemologists are familiar with one kind of situation in which one can get inferential knowledge from a reasoning that involves false beliefs. There are cases in which the falsehood is present but it is not a knowledge-preventing element. And what is more, in these cases, despite there being a falsehood working in the reasoning, its presence is *inessential* to the acquisition of knowledge. These are the so-called epistemic overdetermination cases. To illustrate the point, consider this example.

**Havit/Nogot Overdetermination Case:** Suppose that I am doxastically justified in believing that Havit owns a Ford (which is true) and also justified in believing that Nogot owns a Ford (which is false). On the basis of those two beliefs, I infer, and

thereby come to know, that someone in the class owns a Ford (Klein, 2008: 41).

Epistemic overdetermination is what happens when the subject has two or more independent sources available, each of them being sufficient to cause or justify her belief. In the example above, I have two independent sources available for coming to believe that someone in the class owns a Ford, each of them, in turn, being sufficient to doxastically justify my target-belief.

A well-known metaphor that can help us understand the epistemic overdetermination phenomenon suggests imagining a water flow that results from a faucet whose valves of cold and hot water are both open. Imagine further that opening only one of the two valves suffices to initiate a water flow. Once one valve is open, the other is inessential as far as causing and sustaining a water flow are concerned.

Following Keith Lehrer (1979: 219), we will call the dispensable falsehoods in overdetermination cases ‘harmless falsehoods’, since they do not spoil the subject’s good position to acquire knowledge. Cases of harmless falsehoods are cases of knowledge *despite* falsehood. We are interested in cases of knowledge coming *essentially from* falsehood. So, now we are in a better position to assess the status of any purported case we have come across along the way.

Although many have found putative cases of knowledge arising from false beliefs and have shown how those cases go against the entrenched Orthodox View in the epistemology of reasoning, no one has put forward a theory explaining what place that datum will occupy in the epistemological theorizing up until 2008. Peter Klein was the pioneer here<sup>7</sup>. Our next section will explore in details

---

<sup>7</sup> See Klein, 2008.

his theory of Useful Falsehoods, aimed at dealing with our new intuition on the relation between inferential knowledge and false beliefs.

### 3. Useful Falsehoods

Useful falsehoods are the third kind of falsehood that we have identified until now. As we have seen, there are *knowledge-preventing* falsehoods, like those figuring in Gettier cases. As we have also seen, some falsehoods are *harmless*, since they do not spoil the good epistemic position in which the subject stands. Peter Klein holds that there are also *knowledge-producing* falsehoods, which he baptizes as “useful falsehoods”. According to Klein, the following analysis is supposed to account for the capacity of those false beliefs to give us knowledge.

The belief that *uf* is a useful falsehood to S (for acquiring knowledge that *b*) by producing a doxastically justified belief that *b* iff:

1. *uf* is false.
2. The belief that *uf* is doxastically justified for S.
3. The belief that *uf* is essential in the causal production of the belief that *b*.
4. *uf* propositionally justifies *b*.
5. *uf* entails a true proposition, *t*.
6. *t* propositionally justifies *b*.
7. Whatever doxastically justifies the belief that *uf* for S also propositionally justifies *t* for S (Klein, 2008: 48).

In a case of useful falsehood, if the protagonist gains knowledge of the target-belief, he does so only by virtue of the relevant falsehood. The falsehood in each of those cases is *indispensable* to the potential acquisition of

knowledge – that is, you cannot rule out the evidential path in which the falsehood figures. And more importantly, useful falsehoods are able to lead us to knowledge because they satisfy those conditions, 1-7, which are stated in the analysis above. In what follows, we will explain the rationale for including them.

First off, analytically, useful falsehoods must be false beliefs; otherwise, our talk would be nonsense. This explains why condition 1 is required.

It is also the case that in order to produce a doxastically justified belief for S, useful falsehoods must themselves be doxastically justified beliefs for S. After all, it is typically assumed by the majority of epistemologists that a belief is inferentially doxastically justified only if the belief that justifies it is itself a doxastically justified belief. Were this not the case, how would the latter transmit epistemic status to the former? In other words, how would the latter be entitled by the former if the former itself were not entitled? That common understanding of the way in which epistemic entitlement works and how justification is transmitted via inference explains why condition 2 is required.

The reason why condition 3 is required is that it distinguishes useful falsehoods from harmless falsehoods. As we have already seen, useful falsehoods are essential for furnishing the target-belief to S, since it is the falsehood which causes and justifies the target-belief, whereas harmless falsehoods can be removed from the reasoning without impairing destructively the positive epistemic status of the target-belief.

Condition 4 is accounted for by the logical relation between propositional and doxastic justification. Since useful falsehoods produce doxastically justified beliefs for S, and since doxastic justification implies propositional justification, useful falsehoods propositionally justify those beliefs they produce.

Finally, we arrived at condition 5. The reason why it is included in the analysis of useful falsehoods harks back to a remark made by Risto Hilpinen while he was considering the Temperature Case, which we considered earlier. Upon imagining some counterfactual scenarios in which the protagonist clearly gains knowledge from falsehood, Hilpinen wrote:

[...] This example suggests that a person can know things not only on the basis of (valid) inference from what he or she knows, but in some cases even on the basis of inference from what is not known (or even true), provided that the latter (evidential) propositions are sufficiently close to the truth (Hilpinen, 1988: 164).

According to Klein, it is exactly this “closeness to truth”, to which Hilpinen referred, what makes useful falsehoods able to produce knowledge for the subject. But what does it mean for a belief being “sufficiently close to truth”? The meaning of this relation is supposed to be explained by conditions 5-7 taken together.

To explain the metaphor suggesting closeness to truth, incarnated through conditions 5-7, Klein uses the notion of ‘inferential path’. An inferential path it is the actual or hypothetical way taken in an inference process – in other words, the steps of a process of reasoning, from its premises to its conclusion. The point is that Klein thinks that every case of useful falsehood will have the following inferential paths (Klein, 2008: 37):

- |    |                 |      |               |     |
|----|-----------------|------|---------------|-----|
| A. | $x \rightarrow$ | $uf$ | $\rightarrow$ | $h$ |
| B. | $x \rightarrow$ | $t$  | $\rightarrow$ | $h$ |

In every case of useful falsehood, path A is the path which is actually taken, where the evidence,  $x$ , doxastically justifies the useful false belief,  $uf$ , which in turn, doxastically

justifies the conclusion, *b*. In every case of useful falsehood there is also an inferential path like B, which is only hypothetical, but which is not taken by S. In path B, evidence, *x*, propositionally justifies a true proposition, *t*, which, in turn, propositionally justifies a conclusion, *b*. The point is that the conclusion *b* that is produced in path A has positive epistemic status due to the mere existence of path B.

As Claudio de Almeida (2004, section 1) has pointed out, a useful falsehood's being sufficiently close to truth is something like "the availability of an epistemically relevant truth in a useful falsehood case". According to Klein's proposal, this epistemically relevant truth must be entailed by the useful falsehood, as condition 5 requires, in order for it, the falsehood, to be able to produce knowledge. Thus, S knows *b* only because there is a true proposition *t* which is entailed by *uf* and gives *uf* the power of epistemizing *b*<sup>8</sup>.

Condition 6 on useful falsehoods helps us understand in what way the truth *t* makes its contribution in this extraordinary scenario in order for S to acquire knowledge that *b*. We should not forget that *t* does not contribute in the causation of *b*. Rather, *t* contributes just by virtue of its mere existence and by virtue of propositionally justifying *b*. According to Klein's account, though *t* does not contribute in the causation of *b*, *t* contributes by guaranteeing the existence of an epistemically "healthy" path which, in turn, epistemizes *b*.

Finally, to fully explain the presence of condition 7 among the requirements, notice that the truth *t* has to be at least propositionally justified in order for it to contribute with the doxastic justification of *b*. It suffices to remember that a justifier must be itself justified so as to be able to

---

<sup>8</sup> For an objection to this aspect of Klein's account of useful falsehoods, see (De Almeida, forthcoming).



offer epistemic entitlement to some belief. At this point, a requirement of an inferential path that justifies  $t$  is needed. And to guarantee that  $t$  has this epistemic status, the most plausible thing to do is to require, as condition 7 does, that whatever propositionally justifies  $uf$  must propositionally justify  $t$  as well.

But how is that possible? It is possible because the falsehood  $uf$  will be always equivalent to the conjunction ( $uf \text{ } \mathcal{E} \text{ } t$ ), which guarantees that  $t$  will have positive epistemic status whenever  $uf$  has. [Let us?] See how that is so in the cases we have considered, The Appointment Case and the Handout:

“My secretary told me on Friday that I have an appointment on Monday”  $\equiv$  [(“My secretary told me on Friday that I have an appointment on Monday”) & (“My secretary told me that I have an appointment on Monday”)]<sup>9</sup>.

“There are 53 people at my talk; therefore my 100 handout copies are sufficient.”  $\equiv$  [(“There are 53 people at my talk; therefore my 100 handout copies are sufficient”) & (“There are fewer than 100 people at my talk; therefore my 100 handout copies are sufficient”)].

The falsehood  $uf$  will be always equivalent to the conjunction ( $uf \text{ } \mathcal{E} \text{ } t$ ). Besides that, whenever ( $uf \text{ } \mathcal{E} \text{ } t$ ) is justified for S, then both  $uf$  e  $t$  will be individually justified for S. Now we can see that whenever S has good justification for believing  $uf$ , S will have good justification for believing ( $uf \text{ } \mathcal{E} \text{ } t$ ), and so will have good justification for believing  $t$ . Thus, whenever  $uf$  has positive epistemic status,

---

<sup>9</sup> Cf. Klein, 2008: 43.

$t$  will have positive epistemic status. This explains the presence of condition 7<sup>10</sup>.

#### 4. Putting the theory to work

In the remainder of this paper, we will look again at the putative cases of knowledge from falsehood presented earlier in order to assess whether or not the theory of useful falsehoods is able to accommodate our intuitions regarding those very cases.

First, reconsider the Handout case. There, the useful falsehood is “there are 53 people at my talk”. This proposition though false is able to give me knowledge that my copies are sufficient only because there is a truth “in the vicinity”, so to speak, that is related to that useful falsehood as required by conditions 1-7 of the account proposed by Klein. The falsehood “there are 53 people at my talk” entails the truth “there are fewer than 100 people at my talk”, which in turn epistemizes my conclusion.

Now, remember The Appointment Case. There, the useful falsehood is “my secretary told me on Friday that I have an appointment on Monday with a student”. This proposition though false is able to give me knowledge that I have an appointment on Monday with a student only because there is a truth “in the vicinity”, so to speak again, that is related to the useful falsehood as required by conditions 1-7 of the account proposed by Klein. The falsehood in this case, “my secretary told me on Friday that I have an appointment on Monday with a student”, entails the truth “my secretary told me that I have an appointment on Monday”, which in turn epistemizes my conclusion.

---

<sup>10</sup> See the whole argument in Klein, 2008: 53-54.

---

## 5. Concluding Remarks

It is time to sum up. As stated in the abstract, our aim was to present a surprising, new datum in the epistemology of reasoning, namely there being putative cases of knowledge from falsehood. Flying in the face of the Orthodox View, those cases show subjects gaining knowledge based essentially on false premises. As was mentioned, no theory of knowledge we are familiar with was able to accommodate those cases. Peter Klein was the first to come up with a theory of knowledge from falsehood, which we tested against the cases and found *prima facie* plausible and satisfactory.

As for the impact that the problem of knowledge from falsehood has on the epistemological theorizing on inferential knowledge, we conclude with Klein (2008: 56) himself when he says that this account is “only the first step in understanding the important and initially surprising fact that a false belief can play an essential role in producing knowledge”<sup>11</sup>. We gained reasons to hold that some falsehoods *are not* misleading indicators of truth. Some falsehoods can serve as a good pathway to acquire knowledge.<sup>12</sup>

## References

- Audi, Robert (2011). *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, 3rd Edition, London: Routledge.
- Coffman, E. J (2008). ‘Warrant without Truth?’, *Synthese* 162(2): 173-194.

---

<sup>11</sup> See Coffman, 2008; Fitelson, 2010; and De Almeida (forthcoming) for a more up-to-date view on the state of play over the problem of knowledge from falsehood.

<sup>12</sup> I would like to express my gratitude to the Brazilian funding agency CNPq for the scholarship that funded my work on this project.

- Clark, Michael (1963). 'Knowledge and Grounds: A Comment on Mr. Gettier's Paper', *Analysis* 24(2): 46-48.
- Davis, Wayne (1999). 'Inferential Knowledge'. In: Audi, Robert (Ed.). *The Cambridge Dictionary of Philosophy*, Second Edition, Cambridge: Cambridge University Press: 428.
- De Almeida, Claudio (2004). 'On Useful Falsehoods'. Published on the internet, available through the following URL: <<http://certaindoubts.com/on-useful-falsehoods/>>.
- De Almeida, Claudio (forthcoming). 'Knowledge, Benign Falsehoods, and the Gettier Problem'. In: Borges, Rodrigo; De Almeida, Claudio; Klein, Peter (Eds.). *Explaining Knowledge: New Essays on the Gettier Problem*, Oxford: Oxford University Press.
- Fitelson, Branden (2010). 'Strengthening the Case for Knowledge from Falsehood', *Analysis* 70(4): 666-669.
- Gettier, Edmund (1963). 'Is Justified True Belief Knowledge?', *Analysis* 23(6): 121-123.
- Hilpinen, Risto (1988). 'Knowledge and Conditionals', *Philosophical Perspectives* 2, Epistemology: 157-182.
- Klein, Peter (2008). 'Useful False Beliefs'. In: Smith, Quentin (Ed.). *Epistemology: New Essays*. Oxford: Oxford University Press, 25-61.
- Lehrer, Keith (1979). *Knowledge*, Oxford: Clarendon Press.
- Warfield, Ted (2005). 'Knowledge from Falsehood', *Philosophical Perspectives* 19, Epistemology: 405-416.

# ALGUMAS RAZÕES EM DEFESA DA POSIÇÃO NÃO-CONFORMISTA NO DESACORDO ENTRE PARES

---

*José Leonardo Ruivo*  
(*Doutorando/PUCRS*)<sup>1</sup>

## **Introdução**

Ao longo de nossas vidas tomamos contato com várias informações provenientes das mais diversas fontes. É de se esperar que, dada a diversidade e qualidade das fontes e informações, muitas vezes as pessoas entrem em *desacordo* acerca de determinado assunto. Podemos pensar inúmeras áreas e temas em que encontramos pessoas discordando.

Desse fato alguém poderia apresentar uma *explicação de natureza psicológica* sobre a ocorrência do desacordo. Mas explicações desse tipo seguem um padrão *descritivo*. Por exemplo, um psicólogo poderia dizer que as pessoas entram em desacordo pelo prazer de discutir, independente dos argumentos envolvidos. Sua teoria precisaria explicar uma tendência natural das pessoas discutirem sem considerar os argumentos envolvidos, assim como mostrar que essa tendência natural se expressa em casos de desacordo.

Explicações de natureza psicológica como essa, embora busquem *descrever* o que ocorre nos casos de desacordo, não nos diz qual atitude estamos *intitulados* a ter

---

<sup>1</sup> Email: [jleonardo.ruivo@gmail.com](mailto:jleonardo.ruivo@gmail.com).

nesses casos. Dito de outro modo, ela parece carecer de um aspecto normativo acerca do desacordo. Para deixar isso mais claro, imaginemos a seguinte situação: você está a conversar com um colega sobre os benefícios e malefícios do cigarro. Em certo momento você diz que o cigarro não faz mal à saúde, e seu colega discorda. Mas como essa discordância afetaria as opiniões em questão? Ambos estão intitulados a suspender as suas opiniões sobre o assunto? Ou, a despeito das conclusões contraditórias, ambas opiniões são igualmente válidas?

As últimas questões possuem um peso normativo, fato que foi notado por muitos epistemólogos. Nesse sentido eles buscaram padrões normativos nas situações que envolvem disputas de opinião. Por isso o fenômeno do desacordo tornou-se o centro de uma série de questões epistemológicas. Mas quais são essas questões?

## 1. Questões Epistêmicas

A fim de introduzir o debate, apelamos, na seção anterior, para exemplos que indicam *intuitivamente* o que está em jogo quando falamos no desacordo. Dado nosso interesse analítico na questão, cabe identificar de modo mais preciso os termos do problema: ou seja, cabe especificar o que é exatamente o interesse normativo e epistêmico sobre o desacordo.

Nosso interesse sobre o desacordo é de natureza epistêmica porque visa avaliar a *atitude doxástica* de um sujeito frente às razões que ele dispõe. Dito de outro modo, queremos saber se um sujeito *está intitulado/justificado*<sup>2</sup> a crer

---

<sup>2</sup> É importante notar que “estar justificado” é um termo ambíguo. Podemos estar nos referindo as crenças que um indivíduo possui atualmente ou, então, as crenças que um indivíduo está intitulado a ter, independente de possuí-las atualmente ou não. Tal distinção, introduzida por Roderick Firth (1978) é importante para salientar a natureza da avaliação do sujeito epistêmico. Aqui estamos interessados

---

no conteúdo proposicional P em um determinado momento. Para avaliar a intitulação do sujeito precisamos examinar as razões envolvidas, ou seja, as evidências que um sujeito dispõe, naquele momento, para saber se ele está ou não está intitulado a crer que P. Em suma, tal avaliação considera que a atitude doxástica precisa ser *proporcional* as evidências que o sujeito dispõe. Desse modo ou: (1) a evidência disponível *dá suporte* para que um sujeito S creia em uma proposição P; ou (2) a evidência disponível *dá suporte* para que um sujeito S creia em uma proposição não-P; ou, (3) a evidência disponível *é neutra* em relação da crença de um sujeito S com respeito a uma proposição P<sup>34</sup>.

Por exemplo, se S vê uma parede branca à sua frente, e seu sistema cognitivo está funcionando normalmente, isso é evidência para ele crer que vê uma parede branca à sua frente, o que ilustra a possibilidade (1). Esse mesmo caso ilustra a possibilidade (2), ou seja, S está intitulado a formar a crença de que ele não vê uma parede amarela à sua frente. Contudo, se ele vier a descobrir que está em um cenário enganador, digamos, que existe uma lâmpada na sala que mascara a cor real da parede, então ele está em uma situação do tipo (3), ou seja, não está intitulado a formar qualquer crença sobre a cor da parede.

Antes de avançarmos, cabem duas observações. A primeira é que nosso modelo pretende exaurir as

---

na segunda relação, de intitulação, conhecida como justificação proposicional, e não na primeira, chamada de justificação doxástica.

<sup>3</sup> Aqui a relação de suporte visa explicar, de modo intuitivo, o que significa estar justificado. Notamos, contudo, que a discussão é muito mais profunda, especialmente no que diz respeito se aquilo que justifica uma crença é um processo confiável ou uma representação mental de uma proposição tomada como verdadeira. Para uma leitura introdutória sobre o tema veja Feldman (2003), capítulos 4 e 5; e Steup (2005).

<sup>4</sup> Alguns epistemólogos trabalham com graus de crença ao invés do modelo por nós adotado. Acreditamos que tudo que será dito aqui aplica-se a esses casos, feita as devidas modificações.

possibilidades relevantes para a questão. Ou seja, ele indica explicitamente as possibilidades em que um sujeito *está justificado a crer*. Desse modo, seguindo nosso exemplo, se S está em um cenário enganador e vem a formar a crença de que a parede é amarela, então ele não está intitulado a fazer isso e, portanto, não está justificado a crer, ou mesmo, ele é irracional em formar essa crença.

A segunda observação diz respeito à natureza do modelo. Poderíamos avaliar os casos de desacordo sob outra perspectiva, por exemplo, visando normatividade prática. Isso seria o caso se nosso interesse fossem critérios para avaliar o papel de fatores práticos no peso das razões em disputa (como, p.ex. o quão importante é para um sujeito estar correto). Não é o caso aqui que estamos menosprezando a importância desse tipo de análise, mas simplesmente escolhendo um foco em detrimento de outro. Apesar disso, cabe notar que ao restringir a avaliação epistêmica às possibilidades (1-3), estamos dizendo que é possível explicá-las levando em conta somente o suporte evidencial - ou seja, sem apelar para *razões práticas*<sup>5</sup>.

Até então estivemos às voltas em explicar o que envolve uma avaliação epistêmica. Na próxima seção iremos apresentar uma “fenomenologia do desacordo”, a fim de delimitar os momentos e as características relevantes do mesmo para a investigação epistêmica.

---

<sup>5</sup> Uma conclusão que se segue é que, se um sujeito crê em algo levando em conta *somente* razões práticas, diremos que ele não está justificado (epistemicamente) a crer com base nas evidências que ele dispõe. O exemplo clássico dessa situação é a *Aposta de Pascal* que parte da seguinte conjectura: qual é a melhor situação, eu crer que Deus existe e ele não existir de fato ou eu não crer que ele existe e ele existir? Na segunda, estarei condenado ao inferno enquanto que, na primeira, evito pecar. Logo, é melhor eu crer que Deus existe. Ou seja, nesse caso alguém crê na existência de Deus *somente* por razões práticas, e não por razões epistêmicas.



## 2. Desacordo

O que é um desacordo? Intuitivamente, um desacordo envolve a ponderação de opiniões *aparentemente* contraditórias sobre um determinado tema. Por que aparentemente? Porque, muitas vezes, *o que gera* essa situação diz respeito a falta de uma ou mais destas características: *clareza* sobre o que está sendo dito; *informação* mais ampla sobre o tema; e, *capacidade/habilidade* em considerar o problema. Voltaremos a isso adiante.

Além de questões que dizem respeito àquilo que gera um desacordo, podemos pensar que um desacordo não precisa, necessariamente, de duas ou mais pessoas para ocorrer. Há casos onde o sujeito, através de uma *reflexão*, descobre que possui crenças inconsistentes sobre um determinado tema. Ou então, há casos onde o sujeito discorda de *algo que ele lê*, escrito por alguém que já faleceu<sup>6</sup>.

Nos casos em que envolve duas ou mais pessoas, podemos pensar em: situações aonde você *é uma das partes envolvidas* ou situações em que você *é um observador das partes envolvidas*. A diferença significativa desse segundo tipo de caso é que você não precisa ter qualquer opinião sobre o assunto em questão - poderia estar assistindo um desacordo entre especialistas da biogenética simplesmente para saciar sua curiosidade sobre o que se discute atualmente acerca do tema<sup>7</sup>.

---

<sup>6</sup> Note-se que podemos combinar essas situações imaginando um caso onde o sujeito discorda consigo mesmo, não em relação a crenças inconsistentes atuais, mas em relação a crenças passadas e atuais (por exemplo, uma pessoa lendo um texto que escreveu em um passado distante).

<sup>7</sup> Esse é o chamado problema novato versus especialistas. Ele não será nosso foco de interesse aqui. Sobre isso ver Goldman, 2011; Etcheverry, 2012.

A última consideração acerca dos tipos possíveis de desacordo diz respeito aos *momentos* que estão presentes nessas situações. Como vimos, há casos de desacordo que envolvem somente uma pessoa. Mas podemos imaginar casos cuja questão pode ser compartilhada. Nesse sentido, queremos sugerir dois graus de compartilhamento: um primeiro grau relativo à *consciência da existência* do desacordo e, um segundo grau, relativo à *consciência das razões* pelas quais existe o desacordo.

Ainda que breve, a tipologia apresentada é significativa quando pensamos em uma “fenomenologia do desacordo”. Embora tenhamos buscado nessa seção instrumentalizar minimamente o leitor, tal tarefa foi pensada com a finalidade de compreender e avaliar criticamente a natureza do tipo de desacordo que é foco da literatura epistemológica, a saber, o desacordo entre pares.

O que há de particular no desacordo entre pares? Em primeiro lugar, trata-se de um desacordo que *não é gerado* por falta de informação, clareza ou capacidade dos envolvidos. Isso quer dizer que, do ponto de vista dos sujeitos envolvidos, ambos creem que estão igualmente bem informados, que estão sendo suficientemente claros e que ambos possuem as mesmas capacidades cognitivas. Ou, dito de outro modo, as partes dissidentes são *pares epistêmicos* e, por isso, estão em *simetria*, no sentido de não terem, aparentemente, uma vantagem epistêmica sobre o outro<sup>8</sup>. Aparentemente porque pode resultar que, após o desacordo inicial, virem a descobrir que eles não eram pares epistêmicos porque uma das três condições não foi

---

<sup>8</sup> Não precisamos nos demorar muito sobre a definição de par epistêmico, embora caiba um esclarecimento. Tal noção é altamente contextual, no sentido de que ela não pode ser muito idealizada, sob risco de não instanciar nenhum caso de desacordo atual. Por outro lado, se a noção for muito frouxa, ela não permite uma simetria suficiente para que o desacordo seja epistemicamente intrigante (gerado por falta de clareza, informação ou habilidade).

---

satisfeita. Além disso, se as partes dissidentes estão em simetria relativo a clareza e informação que possuem, é porque ambos têm *consciência das razões* do outro, e discordam.

Imagine a seguinte situação: (4) você, em isolamento, formou a *crença P com base em uma evidência E*, minuciosamente analisada. (5) Contudo mais tarde veio a conhecer um par epistêmico seu que formou, também em isolamento, a *crença de que não-P, também com base na evidência E*, após minuciosa análise<sup>9</sup>. Dado (4) e (5), chegamos ao problema epistemológico do desacordo entre pares: (Q1) qual é a atitude racional que você está intitulado a ter com respeito a P?

### 3. A Perspectiva do Peso Equivalente

Ao refletir sobre (Q1), algumas saídas vem à mente. Talvez a primeira seja de que os pares precisam buscar mais informações antes de tomarem partido. Mas cabe notar que esse tipo de saída, por si só, não diz respeito a racionalidade epistêmica dos pares. Investigar é um tipo de atividade prática e, nesse caso, não estamos buscando algum tipo de normatividade prática (como vimos na seção 1). O que se busca é uma resposta epistêmica com respeito a situação de desacordo que os pares se encontram. Assim, se as partes decidirem por investigar mais, vamos supor aqui, para fins de argumentação, que tomaram tal decisão com a base epistêmica de que ambos não estavam intitulados a formar a crença de que P com base na evidência disponível, ou seja, nossa possibilidade (3).

---

<sup>9</sup> A condição de análise minuciosa aqui visa qualificar o tipo de embasamento em questão: a relação entre E e as proposições em questão é clara, bem informada e realizada competentemente pelos envolvidos.

Isso já aponta como iremos encarar o desafio e buscar uma resposta ao problema em questão. Como vimos nas seções anteriores, se o campo da justificação epistêmica relaciona-se com as possibilidades (1-3), então responde-se a (Q1) com um princípio que explique qual atitude doxástica o sujeito está intitulado a ter em um caso de desacordo entre pares. Afinal, o sujeito é racional se mantiver sua atitude doxástica ou não?

Mas tal tarefa não é fácil e aqui a literatura se divide<sup>10</sup>. O princípio que estruturou toda a discussão sobre desacordo entre pares e ainda conta com muitos adeptos é a *Perspectiva do Peso Equivalente* que diz:

*Perspectiva do Peso Equivalente* (EW): Em casos de desacordo entre pares, deve-se dar peso igual para as opiniões do par e própria<sup>11</sup>.

Vejam como um defensor de (EW) responde a (Q1).

*Desacordo* (EW): Em primeiro lugar é bastante plausível supor que em um caso como descrito em (4-5), *ex hypothesi*, nenhum dos partidos possa apelar para um critério objetivo que quebre a simetria. Para que isso fique mais claro, iremos trabalhar com a perspectiva da primeira pessoa. Se eu creio em P com base na evidência E, eu tenho uma razão para crer que P, a saber, a *evidência de primeira ordem* E. Todavia, no momento em que eu tenho a *consciência do desacordo*, e que esse alguém é meu par, (EW) diz que as razões para crer devem ser pesadas

---

<sup>10</sup> Seguiremos os princípios tal como apresentados por Thomas Kelly (2011). No original, Equal Weight View (EW), No Independent Weight View (NIW), Symmetrical No Independent Weight View (SNIW), e, Asymmetrical No Independent Weight View (ANIW).

<sup>11</sup> Richard Feldman (2006, 2007) marcou o início desse debate defendendo tal posição.

---

de modo igual - uma vez que se tratam de pares. Para que essas razões possam ser pesadas de modo igual, cada uma das partes precisa tomar consciência *das razões do outro*. No instante que essas razões são divulgadas, cada uma das partes passa a ter uma *nova evidência*: você passa a crer que seu par crê em não-P com base em E. Essa *evidência de ordem superior* passa, agora, a ser uma razão que você tem para crer que não-P, a saber: que seu par epistêmico crê em não-P com base em E. Uma vez que o critério de racionalidade implica em descartar inconsistências<sup>12</sup>, se eu tenho uma razão para crer que P e uma razão para crer que não-P, a atitude racional é a suspensão da minha crença sobre P. Através do critério de simetria, reflexivamente, a atitude racional do seu par será também a suspensão da crença sobre P.

A última afirmação mostra porque (EW) tem como consequência um ceticismo localizado: se a atitude racional dos pares em desacordo é a suspensão da crença, então é impossível o desacordo racional entre pares. Todos os casos de desacordo são somente desacordos aparentes, gerados devido a um erro de uma das partes. Note que essa interpretação forte está comprometida com a ideia de que: se há um erro de uma das partes, então; ou um dos dissidentes foi considerado erroneamente como sendo um par, ou os dois dissidentes foram erroneamente considerados como sendo competentes.

Na literatura (EW) recebeu o rótulo de uma posição *conformista* sobre o desacordo racional. No que segue de nosso trabalho, iremos apresentar os argumentos de Thomas Kelly em favor da posição *não conformista* sobre o desacordo entre pares. Antes de analisá-la diretamente cabe

---

<sup>12</sup> Vamos assumir isso a despeito dos problemas envolvidos com o paradoxo do prefácio e o paradoxo da loteria. Sobre a relação entre crenças racionais e inconsistentes ver Rodrigues (2012).

olharmos mais de perto algumas perspectivas não conformistas.

#### 4. Perspectivas Não Conformistas

Sobre os princípios que iremos elencar aqui, cabe notar que eles servirão de baliza para a discussão. Nesta seção apresentaremos uma série de perspectivas não conformistas e, na seção seguinte, iremos avaliá-las comparativamente.

Se, como vimos, a (EW) diz respeito a um princípio bastante forte, iremos iniciar examinando um outro, igualmente forte mas que vai na direção oposta, a *Perspectiva do peso não independente*:

*Perspectiva do peso não independente* (NIW): Em pelo menos alguns casos de desacordo entre pares pode ser perfeitamente racional não conferir peso algum a opinião do partido oposto.

Como um defensor da (NIW) responde a (Q1)? Retomemos o seguinte momento do desacordo entre pares aonde ambos *conhecem as razões do desacordo*. Do ponto de vista da primeira pessoa: você crê em P com base em E e crê que seu par crê em não-P com base em E. Nesse caso, se (NIW) é verdadeira, logo *é possível* ser racional para você manter a crença em P a despeito das razões que seu par apresenta. Mas isso não pode ser feito arbitrariamente, sob risco de dogmatismo. Assim, o defensor de (NIW) precisa apresentar um critério claro para que seu princípio seja válido.

Talvez a saída mais promissora para o defensor da (NIW) seja que ele introduza uma assimetria legítima, como, por exemplo, que você tenha avaliado corretamente o suporte evidencial e seu par não. Com tal *especificação dos*

---

*casos, pode-se fortalecer (NIW) em direção a Perspectiva do peso assimétrico e não independente:*

*Perspectiva do peso assimétrico e não independente (ANIW):* Em pelo menos alguns casos de desacordo entre pares pode ser perfeitamente racional não conferir peso algum a opinião de um par enquanto a própria opinião é a resposta racional à evidência original.

Desse modo, o resultado de (ANIW) contrasta com o resultado de (EW) porque, nesse ponto, um dos pares deve se converter a posição do outro, sob risco de irracionalidade.

Cabe uma observação acerca da relação entre simetria e pares epistêmicos. Havíamos dito anteriormente que (EW) parece conectar paridade e simetria de tal modo que o conformista terá de defender que é o agnosticismo que mantém a paridade  $e$ , conseqüentemente, a impossibilidade do desacordo racional. Por que o agnosticismo mantém a paridade? Porque, se existem casos de desacordo que persistem, eles servem para nos mostrar que no futuro, quando descobrirmos se algum dos dois estava de fato correto ou ambos estavam errados em manter a sua posição, então descobriremos qual dos dois, se algum, era realmente competente epistemicamente. Nesse sentido, suspender o juízo mantém a paridade, porém, com a consequência cética.

Todavia os não conformistas parecem lidar com uma noção de par que os leva a uma petição de princípio pois: como alguém pode ser um par epistêmico, igualmente competente e, ainda assim, “um ser mais par do que o outro”? Uma das razões em favor dos não conformistas é o fato de que eles lidam com uma noção de par que é contextualmente determinada, ou seja, considerando as dificuldades de uma noção muito idealizada ou muito frouxa (cf. n.7). Mas se a noção de par dos não

conformistas tem a motivação virtuosa de lidar com atribuições de competências (e, por conseguinte, de pares) que são falíveis, a teoria ainda precisa explicar como *ao longo* do desacordo entre pares *permanece a paridade* ainda que *a simetria seja quebrada*. Tal explicação será dada na próxima seção.

Antes iremos indicar uma radicalização da (NIW), a *Perspectiva do peso simétrico e não independente*. O princípio diz:

*Perspectiva do peso simétrico e não independente* (SNIW): Em alguns casos de desacordo entre pares ambos partidos da disputa podem ser perfeitamente racionais mesmo se nenhum dá peso algum para a opinião do outro partido.

Ou seja, se (NIW) autoriza que *algum* dos pares é racional ao não conferir peso a opinião do outro; (SNIW) radicaliza apontando que *ambos* são racionais a despeito de, por exemplo, você crer que P com base em E e seu par crer que não-P com base em E<sup>13</sup>. Se (EW) tem como consequência um ceticismo localizado, (SNIW) tem como consequência um relativismo localizado porque avalia como igualmente racionais crenças contraditórias dos pares, a despeito de compartilharem a mesma evidência.

Com isso em mente, adentramos na próxima seção para explicar: como, para um não conformista, permanece a paridade ainda que a simetria seja quebrada; e, qual

---

<sup>13</sup> A título de nota cabe adiantar que não iremos aprofundar a discussão sobre a (SNIW) porque para avaliarmos tal princípio precisaríamos examinar outro, a *Tese da Unicidade*, que diz: (UT) um corpo de evidência justifica, no máximo, uma proposição de um conjunto de proposições incompatíveis e justifica, no máximo, uma atitude doxástica referente a qualquer proposição particular (o princípio foi formulado por Feldman, 2007: 205). Dado nosso propósito introdutório acreditamos que podemos passar ao largo de tal discussão sem afetar nossos objetivos. (sobre SNIW e UT ver Kelly, 2011: 187-190)



---

princípio responde corretamente a (Q1) especificando adequadamente (de modo que não seja vago) quais casos instanciam, quais não instanciam e por que.

## 5. Evidência Total

Vimos que o não conformista tem o ônus de explicar como permanece a paridade ainda que a simetria possa ser quebrada. Kelly (2005) apresenta uma explicação plausível para isso. Partindo da ideia de atribuição de paridade epistêmica que seja, em alguma medida, determinada pelo contexto, é fácil ver que somente isso não é suficiente, pois considerar alguém como par epistêmico está diretamente relacionado a tomar alguém como *competente* em determinado assunto. Embora não entraremos aqui em uma longa explicação de um modelo de atribuição de competência, podemos supor que tal modelo, se pretende respeitar o critério contextual, funciona como uma espécie de generalização indutiva.

Como isso garante, para um não conformista, que a simetria seja quebrada sem retirar a atribuição de paridade epistêmica? A resposta não conformista: porque quando eu quebro a simetria em um caso de desacordo entre pares eu não solapo a ideia de que meu par é um excelente avaliador sobre determinada questão em geral, mas, simplesmente, que *naquela ocasião* ele falhou. Ainda que não ofereçamos um modelo que diga quantas falhas (ou quantos acertos) são necessários para retrain (ou conceder) a atribuição de competência, ele é suficiente para nossos propósitos porque explica como é possível que, em um caso de desacordo entre pares, quando a simetria é quebrada, a paridade pode permanecer. Mas o não conformista ainda não explicou *como* a simetria pode ser quebrada.

Para isso precisamos avaliar os princípios não conformistas que apresentamos anteriormente. Vamos tomar (ANIW) como sendo a versão mais promissora de

(NIW). Vamos avaliar tal princípio com base no nosso caso de desacordo entre pares:

*Desacordo (ANIW):* É bastante plausível supor que em um caso como descrito em (4-5), *ex hypothesi*, nenhum dos partidos possa apelar para um critério objetivo que quebre a simetria. Agora, se eu creio em P com base na evidência E, eu tenho uma razão para crer que P, a saber, a evidência de primeira ordem E. (Cabe lembrar que, nesse caso, a crença foi formada de modo *bem informado* e *competentemente*). Todavia, no momento em que eu tenho a *consciência do desacordo*, e que esse alguém é meu par, (ANIW) diz que em alguns casos as razões para crer devem ser pesadas de modo diferente considerando *qual é a resposta racional à evidência E*. Para que essas razões possam ser pesadas, cada uma das partes precisa tomar consciência *das razões do outro*. No instante que essas razões são divulgadas, cada uma das partes passa a ter uma *nova evidência*: você passa a crer que seu par crê que não-P com base em E. Essa *evidência de ordem superior* passa, agora, a ser uma razão que você tem para crer que não-P, a saber: que seu par epistêmico crê que não-P com base em E. Contudo, considerando que esse é um caso licenciado por (ANIW) é perfeitamente racional eu descartar a evidência de ordem superior uma vez que a minha evidência de primeira ordem responde racionalmente a evidência E.

Tal exemplo mostra por que (ANIW) é falsa (ou, pelo menos, altamente problemática). Dado que, *ex hypothesi*, nenhum dos partidos pode apelar para um critério objetivo que quebre a simetria, então como pode ser que um dos pares se coloque do ponto de vista de lugar nenhum, como se *para ele* estivesse disponível um critério objetivo de confirmação? E, mesmo que seja verdade que

---

em um futuro (próximo ou distante) eu e meu par tenhamos acesso a esse critério objetivo de informação, *no momento do desacordo* não é racional julgar que foi o par que avaliou erroneamente a evidência. Ou seja, ao fim e ao cabo (ANIW) descarta, de modo arbitrário, a evidência de ordem superior que o par oferece.

Talvez se pudesse argumentar que uma prova que objetivamente garante o suporte de P com base na evidência E seria uma prova forte o suficiente para que não fosse facilmente descartada. Porém, sabemos que a fenomenologia do raciocínio não costuma acompanhar a relação de suporte: tanto quem raciocina mal como quem raciocina impecavelmente podem sentir-se igualmente certos, a despeito do primeiro ter uma certeza psicológica e o segundo, uma certeza epistêmica. E o problema aqui exige que essa falta de transparência tenha algum lugar no tratamento epistêmico.

Essa, talvez, seja a importância mais significativa da evidência de segunda ordem: quando dois pares estão em consenso acerca de determinada relação de suporte, isso é uma evidência que funciona em favor da primeira ordem. Por que isso? Considere um caso em que as pessoas estão em um consenso acerca de uma determinada relação de suporte. É de se esperar que, entre aqueles que creem que P com base em E, muitos não o façam de modo *independente*, mas porque consideram o consenso como uma boa evidência para crer que P. O que há de desconfortável nesse caso de consenso cego é que é possível que esse tipo de caso seja muito difícil de distinguir dos casos em que o sujeito raciocinou de modo independente.

Se (ANIW) é falsa isso deixa o defensor de (EW) em situação melhor<sup>14</sup>? Como vimos no do *Desacordo (EW)* o

---

<sup>14</sup> Considerando, para nossos propósitos que (ANIW) é a versão mais promissora de (NIW). A outra saída do defensor de (NIW), como vimos na seção anterior, seria apelar para (SNIW).

impacto de considerar a evidência de ordem superior tende a pressionar a evidência de primeira ordem em direção a suspensão da crença. É bastante plausível supor que tal situação possa ser generalizada: proporcional ao número de pares que opina sobre a sua crença de primeira ordem, maior peso a ordem superior terá. Mas, se (EW) tem uma conclusão correta nesses casos, sua explicação é problemática. Considere o seguinte caso:

*Acordo contra (EW):* É bastante plausível supor que em um caso como descrito em (4-5), *ex hypothesi*, nenhum dos partidos possa apelar para um critério objetivo que quebre a simetria. Agora, se eu creio em P com base na evidência E, eu tenho uma razão para crer que P, a saber, a evidência de primeira ordem E. A crença foi formada de modo *bem informado e competentemente*. Todavia, a evidência E dá suporte para não-P. Meu par, que formou a crença de modo bem informado e competentemente, chegou ao resultado de que a evidência E dá suporte para P: ele tem uma razão para crer que P, a saber, sua evidência de primeira ordem. No instante que essas razões são divulgadas, cada uma das partes passa a ter uma *nova evidência*: você passa a crer que seu par crê que P. Essa *evidência de ordem superior* passa, agora, a ser *uma nova razão* que você tem para crer que P, a saber: que seu par epistêmico crê em P com base em E. Através do critério de simetria, reflexivamente, a atitude racional do seu par será considerar a evidência de ordem superior, uma vez que a minha evidência de primeira ordem responde racionalmente a evidência E.

No caso *Acordo contra (EW)* os pares avaliam erroneamente a evidência de primeira ordem mas, a despeito do seu suporte evidencial, a segunda ordem lhes dará maior confiança. O defensor de (EW) poderia contra-argumentar que o problema desse caso não é (EW), mas o

---

fato de que há um erro *das duas partes*. Mas, dado os termos do problema, *ex hypothesi*, não há um ponto de vista de independente que garantisse um critério objetivo de confirmação. Em suma, o defensor de (EW) precisa dizer como sua explicação diz respeito a uma tendência da ordem superior triunfar sobre a evidência de primeira ordem, *levando em consideração a primeira ordem*, e sem o apelo inverso a suposição problemática de (ANIW), que provas que não garantem objetivamente o suporte de P com base na evidência E seriam provas suficientemente fortes para que fossem facilmente descartadas.

Aonde isso nos leva? É por conta da dialética envolvida na argumentação entre (ANIW) e (EW) que Kelly apresenta a *Teoria da Evidencia Total*:

*Teoria da Evidencia Total*: (TE) “o que é racional para crer *depende* tanto da *evidência de primeira ordem* quanto da de *ordem superior* que é disponibilizada pelo fato do par epistêmico crer naquilo que ele crê.” (Kelly, 2011: 201).

O que (TE) introduz no é que, nos casos de desacordo, no momento em que um dos pares considera P com base na evidência E, sua evidência total é *somente* sua evidência E. Contudo, após tomarem consciência das razões do desacordo, sua evidência total E\* é: (6) a evidência de primeira ordem E; (7) o fato de que você crê que P é verdadeira; (8) o fato de que seu par crê que H é falsa<sup>15</sup>.

Para Kelly tal modelo deixa claro o erro de (EW) e (ANIW). Para (ANIW), nos casos de desacordo entre pares, a atitude racional sobre P dada a evidência E\* é idêntica a atitude racional sobre P dada a evidência E, a despeito do efeito de E\*. Isso está errado porque E\* tem efeito sobre o que é racional para crer em casos de

---

<sup>15</sup> Cf. (Kelly, 2011: 202) e (Kelly, 2005: 190)

desacordo. Esse efeito, explicitado por (7) e (8), tem um impacto de forçar a racionalidade em direção ao agnosticismo.

Mas generalizar esse impacto como o defensor de (EW) faz não é o modo correto de compreender as coisas, a despeito do seu veredito. O que ocorre quando a força da evidência de ordem superior solapa a força da evidência de primeira ordem diz respeito a uma força psicológica que precisa, também, ser levada em consideração.

Dito de outro modo, (EW) parece nos forçar a uma compreensão de que crer nos outros é pior do que formar a crença sozinho, de modo independente. Contudo, essa é uma perspectiva distorcida do fenômeno porque não leva em conta a raiz do fenômeno:: crer no que os outros creem *porque* os outros creem *pode* não contar como evidência adicional<sup>16</sup>.

Mas se o defensor da (TE) consegue descrever o problema, como ele responde a (Q1), ou seja, quando a opinião do par deve ou não deve contar nos casos de desacordo? A fim de esclarecer seu ponto, Kelly apresenta o seguinte exemplo:

*Caso do restaurante:* Você vai a um restaurante com vários amigos. Você e um desses seus amigos tem o hábito de fazer o cálculo da divisão da conta para todos. Ao final da refeição, como de hábito, vocês fazem esse cálculo independentemente um do outro. Você crê que a conta individual é R\$ 43 por pessoa, e isso é uma resposta perfeitamente plausível para a questão: o quanto cada um deve

---

<sup>16</sup> Um exemplo disso é o “caso do guru” (Goldman, 2011). Trata-se de um líder de um grupo de adeptos que tudo que o líder crê, os adeptos creem. Mas os adeptos creem acriticamente, eles são como que clones intelectuais. Esse caso Goldman utiliza para atacar a ideia de que o número de especialistas importa para a atribuição de confiabilidade/autoridade epistêmica.

---

pagar nessa refeição. Contudo, seu amigo apresenta a resposta de que a conta individual é R\$450, uma quantidade que passa significativamente o valor total da conta.

O defensor da (TE) responde ao *Caso do restaurante* do seguinte modo: dada a totalidade das evidências que estão disponíveis para você que são relevantes para saber o valor da conta, é irracional dar qualquer crédito para a proposição de que a divisão total da conta é R\$450, a despeito do fato de que é isso que o seu par crê. Nesse caso, diz Kelly, são as considerações não psicológicas que afundam as considerações psicológicas na insignificância epistêmica<sup>17</sup>.

### **Considerações Finais**

Buscamos apresentar a posição não conformista de Thomas Kelly acerca do desacordo racional entre pares. Por conta do que foi dito, temos fortes razões para crer que a abordagem do autor é relevante porque evidencia uma série de problemas ao conformismo, a posição tradicionalmente aceita no debate sobre desacordo.

Ao formular a teoria da evidência total Kelly mostra que a evidência de primeira ordem deve desempenhar um papel balanceado com a evidência de ordem superior. Isso aponta para uma saliência do sujeito em questão, o que parece minar a possibilidade da existência de uma norma geral para a revisão de crença das formas que consideramos nas seções 2 e 3. Ao fazer com que nosso foco esteja voltado para a evidência total, Kelly aponta também a importância de que o epistemólogo considere, em algum grau, uma fenomenologia do desacordo, ou seja, quais são os momentos envolvidos no desacordo entre pares.

---

<sup>17</sup> Cf. Kelly (2011: 207-208)

Por último, qual é a importância epistêmica do desacordo? Esperamos ter mostrado para o leitor que ele é significativo, a despeito da relação de simetria. No caso de um desacordo que não é entre pares, ele é importante para que os envolvidos tenham acesso a novas evidências. Já no caso em que envolvem pares epistêmicos, ele é importante para que a força probativa dos argumentos possa ser testada e aprimorada.<sup>18</sup>

## Referências

- Etcheverry, Kátia (2012). ‘O Novato e a Disputa entre Especialistas: um Problema para a Epistemologia Social’. In: Müller, Felipe; Rodrigues, Tiegue (Orgs.). *Epistemologia Social: Dimensão Social do Conhecimento*, Porto Alegre: EdiPUCRS, 69-84.
- Feldman, Richard (2003). *Epistemology*, Upper Saddle River, New Jersey: Prentice Hall.
- Feldman, Richard (2006). ‘Epistemological Puzzles about Disagreement’. In: Hetherington, Stephen (Ed.). *Epistemology Futures*, Oxford: Oxford University Press, 216-236.
- Feldman, Richard (2007). ‘Reasonable Religious Disagreements’. In: Antony, Louise M. (Ed.). *Philosophers without Gods: Meditations on Atheism and the Secular Life*, New York: Oxford University Press, 194-214.
- Firth, Roderick (1978). ‘Are Epistemic Concepts Reducible to Ethical Concepts?’. In: Goldman, Alvin; Kim,

---

<sup>18</sup> Gostaria de agradecer a Cláudio de Almeida, Felipe Muller, Rogel Oliveira, Thiago Santin, Luiz Cichoski, Luis Rosa e Gregory Gaboardi pelas sugestões e comentários.



- 
- Jaegwon (Eds.). *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt*, Springer Science & Business Media, Dordrecht: Reidel, 215-229.
- Goldman, Alvin (2011). 'Experts: Which Ones Should You Trust?'. In: Goldman, Alvin; Whitcomb, Dennis. (Eds.). *Social Epistemology: Essential Readings*, New York: Oxford University Press, 109-133.
- Kelly, Thomas (2005). 'The Epistemic Significance of Disagreement'. In: Gendler, Tamar S.; Hawthorne, John (Eds.). *Oxford Studies in Epistemology, vol. 1*, New York: Oxford University Press, 167-196.
- Kelly, Thomas (2011). 'Peer Disagreement and Higher Order Evidence'. In: Goldman, Alvin; Whitcomb, Dennis. (Eds.). *Social Epistemology: Essential Readings*, New York: Oxford University Press, 183-217.
- Lackey, Jennifer (2010). 'A Justificationist View about Disagreement's Epistemic Significance'. In: Haddock, Adrian; Millar, Alan; Pritchard, Duncan (Eds.). *Social Epistemology*, Oxford: Oxford University Press, 298-325.
- Rodrigues, Lucas R. (2012). *Inconsistência e Racionalidade: uma Introdução ao Paradoxo do Prefácio*. 75 f. Dissertação (Mestrado em Filosofia) – Faculdade de Filosofia e Ciências Humanas, PUCRS, Porto Alegre.
- Steup, Matthias (2005). 'Epistemology' In: Zalta, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy*, <<http://plato.stanford.edu/entries/epistemology/>>.

# SOBRE UMA OBJEÇÃO À CONDIÇÃO DE SEGURANÇA<sup>1</sup>

---

*Lucas Roisenberg Rodrigues*<sup>2</sup>

## **Introdução**

Desde que Edmund Gettier (1963) propôs contraexemplos devastadores à análise do conhecimento como crença verdadeira justificada, os filósofos têm debatido insistentemente acerca de qual deve ser a análise correta do conceito de conhecimento. A imensa plausibilidade daqueles contraexemplos, como já foi notado logo na primeira onda de reações ao artigo de Gettier, estava assentada na tese de que o conhecimento é incompatível com a crença acidentalmente verdadeira. Conforme estes exemplos também deixam claro, a justificação não é suficiente para excluir tal tipo de coincidência ou acidentalidade.

Alguns exemplos devem esclarecer o que está em jogo com o problema levantado por Gettier. Digamos que João consulte um relógio. Até onde sabe, o relógio em questão nunca apresentou qualquer defeito, e recebe manutenção constante. Não é a primeira vez que João o utilizou, e podemos até supor que o relógio apresenta um excelente histórico de acertos. Assim, não há qualquer

---

1 O argumento central deste artigo me ocorreu, pela primeira vez, nas aulas do professor Cláudio Almeida em 2014, em uma apresentação do colega Felipe Miguel no programa de pós-graduação em filosofia da PUCRS, e me foi lembrada alguns meses após pelo professor Cláudio em uma de suas disciplinas. Agradeço tanto ao professor Cláudio Almeida, como ao colega Felipe Miguel, por terem ajudado a desenvolver as ideias e argumentos contidos neste texto.

<sup>2</sup> E-mail: [lucasroisenberg@gmail.com](mailto:lucasroisenberg@gmail.com)

razão para João duvidar do resultado indicado pelo relógio.

Por hipótese, o mostrador do relógio indica que são três horas da tarde. Com base na indicação, João forma a crença correspondente de que são três horas da tarde. Ele tem razões ou evidências para crer, e estas parecem ser tão boas quanto evidências ordinárias podem ser. De um ponto de vista exclusivamente interno, não há nenhuma falha intelectual: qualquer sujeito racional, dispondo das mesmas evidências e informações que ele possui, formaria a mesma crença.

Suponhamos agora que são três horas de tarde, e a crença de João, portanto, é verdadeira. Todavia, sem que ele saiba, o relógio parou de funcionar *exatamente* vinte e quatro horas antes de ser consultado. Apesar de João proceder racionalmente, e a sua crença ser justificada, não diríamos que ele *sabe*. João formou uma crença correta por coincidência ou sorte.

Consideremos agora outro contraexemplo, o famoso exemplo dos falsos celeiros apresentado por Goldman (1976). Henry está dirigindo por uma estrada na zona rural. Enquanto dirige, observa, ao longo do caminho, certas estruturas que parecem idênticas a celeiros. Ele olha para um destes objetos, e forma a crença verdadeira de que existe um celeiro no local observado por ele. Contudo, sem que ele saiba, existem muitas fachadas de celeiros nessa região rural, que os moradores construíram para enganar os viajantes. Henry havia observado o único celeiro genuíno nessa estrada repleta de fachadas. Ele simplesmente teve *sorte*. No entanto, se tivesse observado uma das muitas fachadas ali existentes, o que não seria nenhuma surpresa caso acontecesse, então ele a teria tomado por um celeiro genuíno.

Nos dois contraexemplos que examinamos, a sorte desempenha um papel fundamental na aquisição da crença verdadeira, e, ao mesmo tempo, parece ser a intervenção da sorte que impede que a crença assim formada seja um caso

de conhecimento, apesar da justificação a ela anexada. É a constatação de que a justificação não basta por si só para excluir a acidentalidade da crença verdadeira que nos coloca a seguinte questão, que é o núcleo do problema de Gettier: quais propriedades devem substituir ou complementar a justificação em uma correta análise do conhecimento, a fim de excluir a acidentalidade?

## 2 A condição de segurança

Uma explicação bastante promissora do problema de Gettier foi proposta por Nozick (1981). Segundo o autor, uma crença candidata a ser um caso de conhecimento deve satisfazer o seguinte princípio:

**Sensibilidade:** se fosse falso que P, então S não acreditaria que P com base no método M.

No presente contexto, *método* refere-se à maneira ou processo que o sujeito cognitivo empregou para formar a crença de que P. É geralmente aceito que o princípio de sensibilidade deve ser relativizado ao método ou modo de formação de crenças empregado, a fim evitar contraexemplos que uma versão simplificada do princípio de sensibilidade está submetida, tais como este apresentado por Nozick:

Uma avó nota que seu neto está bem quando ele vem visitá-la; mas se ele estivesse morto ou doente, outros diriam a ela que ele está bem para aliviar o seu sofrimento. Todavia, isto não quer dizer que ela não saiba que seu neto está bem (ou, ao menos, que ele está caminhando) quando ela o vê. Claramente, nós devemos reelaborar nossas condições para levar em conta explicitamente os modos e métodos de formar crença. (Nozick, 1981: 179)

O princípio da sensibilidade parece dar conta de

---

maneira bastante satisfatória dos contraexemplos do tipo Gettier. João não sabe que são três horas da tarde, pois ele creia que são três horas da tarde *qualquer que fosse o horário*. É da natureza de um relógio quebrado que ele indique o mesmo horário em todos os momentos do dia. Se fosse falso que são três horas da tarde, e João utilizasse o relógio para formar a sua crença a respeito do horário, ele ainda assim acreditaria que são três horas da tarde.

De maneira semelhante, o princípio da sensibilidade também responde adequadamente ao contraexemplo dos celeiros. É falso que se não existisse um celeiro genuíno no local que Henry observou, e ele formasse a sua crença mantendo o mesmo método ou modo de formação de crenças, Henry não acreditaria que existe um celeiro genuíno.

O princípio de sensibilidade é um condicional subjuntivo ou contrafactual<sup>3</sup>. A semântica dos contrafactuais é um tópico bastante controverso, porém a análise mais bem desenvolvida e aceita deve-se aos trabalhos de Stalnaker (1968) e Lewis (1973)<sup>4</sup>. Consideremos um condicional contrafactual, tal como o condicional ‘ $P > Q$ ’<sup>5</sup>. O condicional ‘ $P > Q$ ’ é verdadeiro no mundo @ quando Q é verdadeiro no P-mundo mais próximo ao mundo @. A relação de proximidade é definida

---

3 Embora nem todos os condicionais subjuntivos sejam estritamente contrafactuais, e nem todos os contrafactuais estejam apresentados no modo gramatical subjuntivo, seguirei aqui a prática corrente de usar *contrafactual* e *subjuntivo* intercambiavelmente.

4 Deve-se notar aqui que Nozick oferece uma interpretação alternativa da semântica dos condicionais contrafactuais. Entretanto, a alteração de Nozick não será relevante para a análise das objeções de que tratamos neste artigo, e podemos, sem maiores problemas, aceitar a semântica proposta por Lewis e Stalnaker.

5 Utilizarei ‘@’ para referir-me ao mundo atual, ‘>’ para referir-me ao condicional contrafactual e ‘P-mundo’ para referir-me ao (s) mundo (s) possível (eis) em que P.

em termos de similaridade: quanto mais similar é um mundo, mais próximo ele é.

Para considerar se um condicional subjuntivo é verdadeiro, devemos não apenas levar em conta o que é verdadeiro ou é o caso no mundo atual, mas também devemos observar o que é verdadeiro em circunstâncias ou mundos possíveis alternativos<sup>6</sup>. Por “mundos relevantes” entendo o conjunto dos mundos possíveis em que uma determinada condição deve ser verdadeira a fim de que a crença do sujeito cognitivo satisfaça um determinado princípio ou condição modal.

No que diz respeito ao princípio de sensibilidade, o conjunto de mundos relevantes é o conjunto de dos mundos possíveis mais próximos ao mundo atual em que a proposição-alvo é falsa, e a condição a ser satisfeita é a não-crença na verdade da proposição-alvo. Conforme vemos, a teoria de Nozick pertence à "família" das epistemologias modais, isto é, teorias epistemológicas que tentam analisar e/ou explicar o conceito de conhecimento por meio de algum princípio que conecta a crença do sujeito a respeito da proposição-alvo e a verdade do proposição-alvo em um conjunto de mundos possíveis relevantes.

Apesar das muitas virtudes da condição de sensibilidade, ela foi alvo de poderosas objeções<sup>7</sup>. Em razão dos muitos problemas encontrados no princípio de

---

6 Não pretendo discutir o tópico da metafísica dos mundos possíveis, isto é, de sua natureza ontológica. É suficiente, para nossos propósitos, considerar que mundos possíveis correspondem às circunstâncias ou cenários em que o mundo @ seria diferente e, portanto, não seria o mesmo @.

7 Entre os problemas mais graves enfrentados destacam-se os problemas do conhecimento indutivo, apresentado do conhecimento de segunda-ordem, os contraexemplos do celeiro vermelho e a rejeição do princípio de fecho do conhecimento pela implicação conhecida. Uma excelente apresentação dessas objeções à teoria do rastreamento, entre outras, pode ser encontrada em Vogel (2012)

sensibilidade, foi sugerida a sua substituição por uma condição semelhante, porém que, ao menos segundo seus defensores, não padece de defeitos a que está submetida a primeira. Refiro-me aqui à condição de segurança que foi defendida, entre outros, por Pritchard (2005), Sosa (1999a) e Williamson (2000). O núcleo do princípio de segurança é a tese de que se S sabe que P, então P não seria falso em circunstâncias próximas ou semelhantes às circunstâncias em que ele atualmente forma a crença de que P.

A condição de segurança foi definida de diversas maneiras, nem todas compatíveis entre si, mas a seguinte proposta de Sosa servirá para nossos propósitos:

Dizemos que uma crença de S que P *segura* sse: S não acreditaria que P a menos que fosse o caso que P. (Alternativamente, a crença de S de que P é 'segura' sse: como matéria de fato, embora talvez não como matéria de estrita necessidade, S não acreditaria que P a menos que fosse o caso que P (Sosa, 1999a: 378).

A caracterização acima pode ser tomada como equivalente ao seguinte condicional subjuntivo:

**Segurança:** se S acreditasse que P, então seria verdadeiro que P.

A condição de segurança, assim como o princípio de sensibilidade, também pretende resolver o problema de Gettier. Retomando o exemplo anterior, João acredita verdadeiramente que são três horas da tarde porque consultou um relógio que parou de funcionar precisamente doze horas antes. A crença de João não é segura, pois se ele formasse a crença de que P em circunstâncias muito próximas àquelas em que ele formou a sua crença, então ele poderia formar uma crença falsa. A possibilidade de erro envolvida aqui não deve ser entendida em termos lógicos.

Ela está circunscrita aos mundos próximos ou muito similares ao mundo atual.

É a possibilidade de erro em mundos próximos ao mundo atual que caracteriza os casos Gettier, e que é responsável pela acidentalidade da crença do sujeito cognitivo. Tomando como base a condição de segurança, poderíamos dizer que uma crença é verdadeira por sorte ou acidente, quando muito facilmente, ou que, em circunstâncias muito semelhantes, que quase foram atualizadas, a crença teria sido falsa. Assim como o princípio de sensibilidade anteriormente examinado, a condição de segurança oferece uma explicação plausível da intervenção da sorte epistêmica na produção dos exemplos de Gettier.

Apesar de facilmente confundíveis, as condições da segurança e da sensibilidade são distintas. A diferença fundamental entre o princípio de segurança e o princípio de sensibilidade reside no conjunto de mundos relevantes delimitado por cada um deles. Os mundos relevantes para a satisfação da condição de segurança são aqueles em que o sujeito S acredita que P, e a condição modal é a verdade da proposição-alvo. Conforme já exposto, os mundos relevantes para a satisfação da sensibilidade são os mundos em que a proposição-alvo é falsa, e a condição modal é a não-crença na verdade da proposição-alvo. A sensibilidade diz respeito ao que aconteceria se a proposição-alvo fosse falsa. A segurança diz respeito ao que aconteceria se o sujeito acreditasse na proposição-alvo.

Além de distintas, as duas condições são também independentes. Segurança não se segue de sensibilidade, pois de  $P > Q$  não se segue  $\neg Q > \neg P$ .

No que diz respeito aos condicionais indicativos, é válido o seguinte raciocínio:

1. Se eu estou doente, então eu não vim à aula.
2. Logo, se eu vim à aula, então eu não estou doente.



Embora o condicional indicativo seja logicamente equivalente à sua correspondente contrapositiva, o mesmo não é correto a respeito dos condicionais subjuntivos. Por exemplo, é inválido o seguinte raciocínio:

1. Se a PUCRS estivesse funcionando, então eu estaria na PUCRS.
2. Logo, se eu não estivesse na PUCRS, então a PUCRS não estaria funcionando.

É inválido o raciocínio acima, pois é falso que se eu não estivesse na PUCRS, então ela não estaria funcionando, uma vez que a PUCRS funcionaria mesmo que eu viesse a estar em outro lugar.

Assim como acontece com a sensibilidade, o princípio da segurança em sua forma simples deve ser complementado com referências ao modo de formação de crenças do sujeito. A condição de segurança é satisfeita quando não facilmente aconteceria de o sujeito S acreditar que P com base no modo ou método M e ainda assim P ser falsa.

Há muito a se dizer em favor da condição de segurança, conforme já assinalaram alguns de seus defensores, tais como Sosa (1999a) e Pritchard (2005). Eles alegam que o princípio explica de maneira simples e elegante por que o sujeito não possui conhecimento nos chamados ‘casos Gettier’, mostrando como e sob que condições o conhecimento é incompatível com a sorte epistêmica.

Também alegam os autores que princípio também sucede onde propostas muitos semelhantes, tais como a teoria das razões conclusivas, de Dretske (1971), ou a teoria do rastreamento, de Nozick (1981), parecem produzir resultados extremamente contraintuitivos. É afirmado que,

ao contrário daquelas teorias, o princípio de segurança não apenas preserva o princípio de fechamento do conhecimento sob implicação conhecida, como também produz um robusto resultado anticético, compatível com o conhecimento tanto de proposições ordinárias do senso comum, como com o conhecimento de que as hipóteses céticas, tais como as hipóteses do gênio maligno, são falsas<sup>8</sup>. Por fim, defende-se que a teoria da segurança tem se mostrado resistente a contraexemplos, tais como o problema do conhecimento indutivo ou do conhecimento de segunda-ordem, que são objeções bastante sérias à teoria do rastreamento.

### 3 A objeção de Comesaña

Apesar das suas alegadas vantagens sobre a sensibilidade, o princípio de segurança não está imune a sérias objeções. Uma delas foi apresentada em Comesaña (2005). O contraexemplo é o seguinte:

FESTA DE HALLOWEEN: haverá uma festa de Halloween na casa de Andy, e eu fui convidado. A casa de Andy é muito difícil de encontrar, então ele contrata Judy para permanecer em uma bifurcação e direcionar as pessoas para a casa (o trabalho de Judy é dizer às pessoas que a festa é na casa situada na rua à esquerda). Sem que eu saiba, Andy não deseja que Michael vá à festa, e diz para Judy que, se ela vir Michael, então ela deve dizer a ele a mesma coisa que diz aos outros (que a casa é na rua à esquerda), mas deve imediatamente telefonar para Andy a fim que a festa seja transferida para a casa

---

8 Não interessa muito aqui os detalhes da formulação de um princípio de fecho do conhecimento pela implicação conhecida. A ideia central é a tese de que se S deduz validamente que Q a partir de P, e S sabe que P, então S sabe que Q.

---

de Adam, que é na rua à direita. Eu seriamente considero fingir ser Michael, mas, no último momento, eu desisto. Quando eu alcanço a bifurcação, eu pergunto Judy onde será a festa, e ela me diz que será na rua à esquerda. (Comesaña, 2005: 397).

Parece que, no exemplo tal como descrito, S sabe que a festa irá ocorrer na rua à esquerda<sup>9</sup>. Contudo, a crença de S não é segura, de acordo com as caracterizações prévias da condição de segurança., pois S facilmente poderia ter acreditado falsamente que a festa irá ocorrer na rua à esquerda, uma vez que existe um mundo W muito próximo ao mundo @, tal que em W o sujeito S acredita falsamente que a festa será na rua à esquerda com base no mesmo método que ele utiliza no mundo atual. O mundo W é justamente o mundo em que S decide ir à festa fingindo ser Michael. Como essa é uma situação de conhecimento inseguro, o princípio de segurança não pode ser uma condição necessária para o conhecimento.

#### 4 A objeção de Bogardus

Em um artigo recente, Bogardus (2012), apresentou algumas objeções ao contraexemplo de Comesaña, e também a outros, semelhantes, apresentados por Neta e Rohrbaugh (2004) e Kelp (2009)<sup>10</sup>. De acordo com Bogardus, esses autores cometiam uma falácia, qual seja,

---

9 Comesaña apresentou o caso da festa de Halloween na primeira pessoa do singular. Para tornar a exposição mais fácil, utilizarei a terceira pessoa, e referir-me-ei ao sujeito cognitivo envolvido simplesmente como 'S'.

10 Não trato aqui, em detalhes, dos contraexemplos de Kelp, Neta e Rohrbaugh. Todavia, as objeções que Bogardus propôs ao contraexemplo de Comesaña aplicam-se, *mutatis mutandis*, aos contraexemplos de Kelp, Rohrbaugh e Neta.

“cada um move falaciosamente do fato de que S estava em risco epistêmico logo antes de formar a crença para a conclusão de que a crença de S foi formada inseguramente” (Bogardus, 2012: 289).

Vejam os como a crítica de Bogardus aplica-se ao contraexemplo apresentado por Comesaña. De acordo com Bogardus (2012),

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

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

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

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

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

Do fato de que *S* não finge ser Michael no mundo @ no momento *T* em que *S* forma a crença de que *P* não se segue que não devemos considerar mundos possíveis alternativos em que *S* decide fingir ser como Michael antes do momento *T*. Evidentemente, não podemos considerar que toda verdade em @ deve ser mantida nos mundos relevantes para a satisfação da condição de segurança, pois, se esse fosse o caso, conhecimento seria idêntico a crença verdadeira. Dito de outra forma, do fato de que *S* não está vestido como Michael no mundo @ em *T* quando *S* forma a crença de que *P* não segue que o conjunto dos mundos possíveis relevantes para a crença de que *P* não inclui mundos possíveis em que *S* decide vestir-se como Michael *antes* de *T*. Sem uma razão adicional que legitime essa inferência, o que *não* está presente no artigo de Bogardus, a sua objeção perde muito da sua força inicial.

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

forma a crença em @, e, contudo, é falso que P em W<sup>11</sup>.

Assim, no mundo @:

1. Em T<sub>1</sub>, S delibera se irá vestir-se como Michael.
2. Em T<sub>2</sub>, S decide *não* se vestir como Michael.
3. Em T<sub>3</sub>, S acredita que P com base em um método M.

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

1. Em T<sub>1</sub>, S delibera se irá vestir-se como Michael.
2. Em T<sub>2</sub>, S decide vestir-se como Michael.
3. Em T<sub>3</sub>, S acredita que P com base em um método M.

De acordo com o exemplo de Comesaña, a crença de S de que P formada no tempo T<sub>3</sub> no mundo @ não é segura, isto é, não satisfaz a condição de segurança tal como vimos anteriormente, pois existe um mundo W muito próximo ao mundo @ tal que, em W, o sujeito S acredita falsamente que P da mesma maneira que ele acredita em @.

Segundo Bogardus (2012), o sujeito S corre o risco de formar uma crença falsa no tempo T<sub>1</sub> no mundo @; porém, no tempo T<sub>3</sub>, quando o sujeito S forma a crença de que P, ele não corre mais o risco de formar uma crença falsa, uma vez que ele *já decidiu* não fingir ser o Michael em T<sub>2</sub>. De acordo com ele, é irrelevante a existência do mundo

---

11 Para facilitar a exposição, adotei a seguinte terminologia. ‘P-mundos-relevantes’ refere-se aos mundos relevantes para o conhecimento de que P, isto é, aos mundos em que P é verdadeiro se S sabe que P. ‘Q’ refere-se à proposição ‘o sujeito S decidiu vestir-se como Michael’, e ‘P’ é a proposição ‘a festa será na rua à esquerda’. ‘P-mundo’ é o mundo em que P é verdadeiro, ‘Q-mundo’ é o mundo em que Q é verdadeiro, e assim por diante.

W em que S acredita falsamente que P. O fato responsável pela irrelevância do mundo W é a existência de uma determinada condição, digamos a condição C, tal que C é satisfeita em @ e não é satisfeita em W. No contraexemplo da festa de Halloween, a condição C consiste em não estar disfarçado como Michael.

Não penso que esta é, em última análise, uma objeção convincente. O problema, naturalmente, é que Bogardus parece considerar que, uma vez que a condição C não foi atualizada em um momento  $T_2$  antes do momento  $T_3$  em que S forma a crença de que P, a existência de um mundo W em que  $\neg C$  torna-se irrelevante para a avaliação acerca da segurança da crença de que P formada em  $T_3$  no mundo @. Conforme vimos, esta inferência é problemática, mas, de qualquer modo, não foi defendida no artigo de Bogardus. Tais considerações são suficientes para rejeitar *prima facie* o argumento de Bogardus.

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

A descrição e a individuação do método de formação de crenças compõem um dos problemas mais delicados que enfrentam todas as teorias que empregam o conceito de método de formação de crenças, o conhecido problema da generalidade. Dúvidas à parte sobre um critério *geral* da individuação de métodos, as nossas intuições no caso particular não favorecem a inserção de C como parte do método utilizado por S. Tanto em W como em @, o sujeito

S parece formar exatamente a mesma crença da mesma maneira, qual seja, baseando-se no testemunho de que P dado por Judy. Não parece razoável que fatos como a aparência ou a vestimenta de S possam ser relevantes para determinar *que maneira ou método de formação de crenças foi utilizado no caso*. Novamente, sem uma razão adicional que justifique a inserção de C como parte de M, é difícil ver como a objeção de Bogardus poderia manter-se.

## 5 Conclusão

Ao longo deste artigo, expusemos a condição de segurança, a objeção de Comesaña e a réplica de Bogardus. Existem duas lacunas na objeção de Bogardus, que colocam sob suspeição as suas objeções ao contraexemplo de Comesaña. Em primeiro lugar, ele assume tacitamente que se o sujeito S não finge ser Michael no mundo atual @, quando ele forma a crença de que P no momento T, então o mundo possível em que finge ser Michael antes de formar a crença de que P é irrelevante para avaliar a segurança da crença de que P formada em momentos anteriores a T. Trata-se de uma afirmação duvidosa e que, no entanto, é fundamental para a objeção.

Em segundo lugar, a melhor alternativa plausível em defesa da objeção de Bogardus, a inserção do fato de que C entre as características do método M, não parece promissora. A vestimenta de S não está incluída na caracterização do método M, a menos que queiramos simplesmente definir o método M de maneira *ad hoc*, de forma a tornar a crença sempre verdadeira. Apesar de o próprio Bogardus apresentar em seu artigo outras objeções contra o princípio de segurança, espero ter demonstrado que as objeções originais continuam mantendo sua força e eficácia como algumas mais poderosas críticas à condição de segurança.



---

**Referências**

- Bogardus, Tomas (2012). 'Knowledge under Threat', *Philosophy and Phenomenological Research* 88(2): 289–313.
- Comesaña, Juan (2005). 'Unsafe Knowledge', *Synthese* 146(3): 395–404.
- Dretske, Fred (1971). 'Conclusive Reasons', *Australasian Journal of Philosophy* 49(1): 1–22.
- Gettier, Edmund (1963). 'Is Justified True Belief Knowledge?', *Analysis* 23(6): 121–123.
- Goldman, Alvin (1976). 'Discrimination and Perceptual Knowledge', *The Journal of Philosophy* 73(20): 771–791.
- Lewis, David (1973). *Counterfactuals*, Malden: Blackwell Publishing.
- Kelp, Christoph (2009). 'Knowledge and Safety', *Journal of Philosophical Research* 34: 21–31.
- Neta, Ram; Rohrbaugh, Guy (2004). 'Luminosity and the Safety of Knowledge', *Pacific Philosophical Quarterly* 85(4): 396–406.
- Nozick, Robert (1981). *Philosophical Explanations*, Cambridge MA: Harvard University Press.
- Pritchard, Duncan (2005). *Epistemic Luck*, New York: Oxford University Press.
- Sosa, Ernest (1999a). 'How Must Knowledge Be Modally Related to What Is Known?', *Philosophical Topics* 26(1): 373–384.

Sosa, Ernest (1999b). 'How to Defeat Opposition to Moore', *Noûs, Supplement: Philosophical Perspectives* 33: 141–153.

Stalnaker, Robert C. (1968). 'A Theory of Conditionals'. In: *American Philosophical Quarterly Monograph Series*: (2) 98-112, Oxford: Blackwell.

Vogel, Jonathan (2012). 'The enduring trouble with tracking'. In: Becker, K., Black, T (eds.). *The sensitivity principle in epistemology*, 122-151, New York: Oxford University Press

Williamson, Timothy (2000). *Knowledge and its Limits*, New York: Oxford University Press.

# RELAÇÃO DE SUPORTE

---

*Luiz Paulo Da Cas Cichoski*<sup>1</sup>

## Introdução

A relação de suporte se dá entre *uma crença* e *as razões pelas quais* um sujeito mantém essa crença. A crença repousa sobre essas razões, isto é, essas razões servem de base para a crença. Epistemólogos estão interessados justamente em precisar que relação é essa que se dá entre crenças e razões. Dado que muitas relações podem se dar entre crenças e razões, a questão central aqui é determinar aquela relação de suporte que “quando obtida entre uma crença [C] e uma razão R pela qual a crença é mantida, essa obtenção é aquilo que torna o caso de que R é a razão (ou ao menos *uma* razão) pela qual [C] é mantida”<sup>2</sup> (Neta, 2011: 109).

A relação de suporte figura como peça-chave de qualquer teoria da justificação. Particularmente, essa relação parece marcar a distinção entre crenças justificadas e crenças justificáveis<sup>3</sup>:

---

<sup>1</sup> Doutorando em Filosofia – Pontifícia Universidade Católica do Rio Grande do Sul; E-mail: luizpcichoski@gmail.com.

<sup>2</sup> “(...) when it obtains between a belief B and a reason R for which the belief is held, its obtaining is what makes it the case that R is the reason (or at least *a* reason) for which B is held.”

<sup>3</sup> Ian Evans (2013) questiona se essa é realmente uma interpretação do que seria a relação de suporte. Para Evans, existe uma diferença entre (a) uma relação de suporte e (b) a relação entre crenças justificadas e as razões em virtude das quais as crenças estão justificadas. Evans chama (b) de “suporte apropriado” diferente de (a) que tende a captar uma relação mais fraca entre crenças e razões, não exigindo que essas razões sejam boas razões para a manutenção da crença. Dado o objetivo introdutório do presente trabalho, seguiremos a interpretação mais usual do termo que não faz essa distinção mais fina e parece tomar a “relação de suporte” aos moldes de (b), como ficará evidente no decorrer desta seção.

Uma crença é meramente justificável para uma pessoa S quando S possui razões suficientes para justificar a crença mas não fez qualquer conexão apropriada entre as razões e a crença e, conseqüentemente, permanece injustificada em manter a crença. A conexão apropriada se dá quando a crença está sustentada nas razões. Uma crença está justificada para S quando S possui razões suficientes para justificar a crença e fez uma conexão apropriada entre sua crença e suas razões (i.e., a crença está sustentada nas razões)<sup>4</sup> (Korcz, 1997, p. 171).

Neta (2011) ainda chama atenção para duas outras importantes distinções: (1) razões pelas quais (*for which*) o sujeito crê são diferentes da razão por que (*why*) o sujeito crê: a razão por que alguém crê que o suspeito é culpado pode ser (a crença no<sup>5</sup>) o fato de que o suspeito faz com

---

<sup>4</sup> “A belief is merely justifiable for a person S when S possesses reasons sufficient to justify the belief, but has not made any appropriate connection between the reasons and the belief, and consequently remains unjustified in holding the belief. The appropriate connection would be the belief’s being based on the reason. A belief is justified for S when S possesses reasons sufficient to justify the belief and has made an appropriate connection between her reasons and her belief (i.e., the belief is based on the reasons).”

<sup>5</sup> Existe uma disputa sobre a natureza das razões. O que são razões? Parece que, em geral, epistemólogos adotam um compromisso com a visão *Psicológica*, onde razões são objetos psicológicos, em especial crenças. Assim, as razões que um indivíduo tem para crer são eventos e estados mentais do sujeito (veja Turri (2009) para uma discussão do ponto e uma defesa dessa visão). Em especial nas discussões sobre razões para agir (e não para crer), muitos autores (Dancy (2000); Alvarez (2010)) defendem que razões são fatos concretos do mundo (tal como o fato de a maçaneta ter impressões digitais do suspeito). Existe ainda um outro grupo (Scanlon (1999); Dickinson (2012)) sustentando que o conteúdo das nossas crenças é que são as razões pelas quais agimos ou acreditamos, ou seja, razões são proposições.

---

que o sujeito se lembre de uma pessoa da qual não gosta; entretanto, essa não é uma razão pela qual alguém crê que o suspeito é culpado (excetuando-se os casos em que o sujeito possui outras razões que sustentam a crença de que pessoas que lembram outras pessoas das quais ele não gosta são especialmente suscetíveis de serem culpadas de crimes). (2) razões pelas quais (*for which*) o sujeito crê são diferentes de razões para (*for*) o sujeito crer. Alguém pode crer que o suspeito é culpado em função do (a crença no) fato de que as impressões digitais do suspeito estão por toda a arma do crime. Entretanto, no caso de a arma do crime ser a maçaneta da porta da frente da residência do suspeito, essa parece não ser uma razão para crer que o suspeito é culpado. Da mesma forma, (a crença no) o fato de que o DNA do suspeito estava sobre toda a vítima é uma razão para alguém acreditar que o suspeito é culpado. Entretanto, se esse sujeito tem uma desconfiança quanto a evidências sobre DNA, essa não é uma razão pela qual ele manterá a crença de que o suspeito é culpado.

Neta (2011: 110) estabelece a relação entre essas distinções<sup>6</sup>:

(...) uma razão pela qual alguém acredita em alguma coisa é sempre uma razão por que alguém crê, mas nem todas as razões por que alguém

---

Para uma discussão abrangente sobre essas diferentes propostas veja Alvarez (2010) e Dickinson (2012). No presente trabalho, nos manteremos neutros sobre esse aspecto.

<sup>6</sup> Atente para o fato de que Neta está preocupado com um tipo especial de relação de suporte, aquele chamado de “suporte apropriado” na nota 2. A distinção relevante é a exigência de que a razão seja o “tipo de coisa certa para ser uma razão” que suporte a crença, enquanto visões mais largas da relação de suporte, tal como defendida por Evans, segundo a qual aceita-se que crenças sejam sustentadas por experiências, percepções, desejos, memórias, preconceito, etc. A diferença fundamental está no fato de que a visão mais larga não exige que a relação de suporte se dê apenas entre crenças e *boas razões*.

acredita em alguma coisa são razões pelas quais alguém crê. Uma razão pela qual alguém crê em alguma coisa é, algumas vezes, mas não sempre, uma razão para alguém crer. Uma razão para alguém crer em alguma coisa é, algumas vezes, mas não sempre, uma razão pela qual alguém crê. Juntos, esses pontos colocam restrições substanciais para qualquer teoria adequada de relação de suporte: a crença [C] de uma criatura pode estar em uma relação de suporte com uma razão R *somente se* R é uma razão por que a criatura mantém [C] e *somente se* R é o tipo de coisa certa para ser uma razão para a criatura manter [C]. Nós podemos colocar esse mesmo ponto dizendo que [C] é sustentada sobre R somente se R faz as duas coisas a seguir: (i) R racionalmente suporta a manutenção de [C] e (ii) R explica por que [C] é mantida.<sup>7</sup>

Tendo esclarecido alguns pontos sobre o que deveria ser explicado por uma teoria da relação de suporte, passaremos a analisar três propostas: (i) Teoria Causal de Relação de Suporte; (ii) Teoria Contrafactual de Relação de Suporte; e (iii) Teoria Doxástica de Relação de Suporte.

---

<sup>7</sup> “(...) a reason for which someone believes something is always a reason why she believes it, but not all reasons why someone believes something are reasons for which she believes it. A reason for which someone believes something is sometimes, but not always, a reason for her to believe it. And a reason for someone to believe something is sometimes, but not always, a reason for which she believes it. Together, these points place substantial constraints on any adequate theory of the basing relation: a creature’s belief B can stand in the basing relation to a reason R *only if* R is the reason why the creature holds B, and *only if* R is the right sort of thing to be a reason for a creature to hold B. We might put this same point by saying that B is based on R only if R does both of the following two things: (i) R rationally supports holding B and, (ii) R explains why B is held.”

---

## Teoria Causal de Relação de Suporte

Vimos que, se uma razão  $R$  é uma razão pela qual alguém mantém uma crença  $C$  (ou seja, a relação entre  $R$  e  $C$  é do tipo que queremos explicar: é uma relação de suporte), então se segue que a razão  $R$  é uma razão porque alguém mantém a crença  $C$  (ou seja, ser uma razão pela qual alguém mantém uma crença é condição suficiente para que essa razão também seja uma razão por que alguém mantém a crença). Será que essa relação nos aponta para a necessidade de que a razão  $R$  tenha causado a crença  $C$  desse sujeito?

Muitos epistemólogos parecem responder afirmativamente a essa pergunta. Eles acreditam que a relação de suporte é uma espécie de relação causal; logo, quando existe uma relação de suporte entre  $R$  e  $C$ , deve existir uma relação causal entre  $R$  e  $C$ <sup>8</sup>. Segundo Korcz (2010), Paul Moser (1989: 157) parece defender uma Teoria Causal de Relação de Suporte: “A crença ou assentimento de que  $P$  por  $S$  é sustentada em sua razão proposicional justificadora  $Q =_{df}$  a crença ou assentimento de que  $P$  por  $S$  é causalmente sustentada de maneira não desviante por sua crença ou assentimento de que  $Q$  e sua associação de  $P$  e  $Q$ .”<sup>9</sup>

---

<sup>8</sup> Não devemos confundir uma defesa de Teoria Causal de Relação de Suporte com uma Teoria Causal de Justificação ou Conhecimento. Keith Korcz (2010) argumenta que alguém pode tomar a Teoria Causal de Relação de Suporte apenas para explicar as razões utilizadas por um sujeito para formar ou avaliar uma crença, sendo essa abordagem independente das razões que, de fato, contribuem para a justificação da crença. Além disso, justificação pode exigir mais do que suporte de boas razões, como o cumprimento de deveres epistêmicos, coerência do sistema cognitivo etc.

<sup>9</sup> “ $S$ 's believing or assenting to  $P$  is based on his justifying propositional reason  $Q =_{df}$   $S$ 's believing or assenting to  $P$  is causally sustained in a nondeviant manner by his believing or assenting to  $Q$ , and by his associating  $P$  and  $Q$ .”

Korcz enfatiza que a proposta de relação de suporte de Moser é uma modalidade bastante forte por tratar de razões justificadoras. Outro aspecto para o qual Korcz chama a atenção é a presença de duas condições: (1) uma condição de sustentação causal e (2) uma condição de associação. A ideia da segunda é que o sujeito deveria estar ciente (*aware*) do suporte que a razão provê para a crença. Moser (1989: 141–142) define essa condição de associação da seguinte forma:

S satisfaz uma relação de associação entre E e P =<sub>df</sub>  
(i) S tem uma consciência *de re* do suporte de E para P e (ii) como resultado não desviante dessa consciência, S está em um estado disposicional em que, se ele focasse sua atenção somente em sua evidência para P (enquanto tudo o mais permanecesse constante), ele focaria sua atenção em E<sup>10</sup>.

Essa “consciência *de re*” é uma consciência direta do suporte de E para P. Ela não é proposicional e se assemelha ao tipo de consciência que temos quando estamos com sede ou com frio antes de pensarmos “Estou com sede” ou “Estou com frio”. A característica não-proposicional se deve ao fato de que não precisamos pensar em palavras para ter esse tipo de consciência, e o caráter direto provém do fato de que não precisamos de qualquer sensação (imagem mental ou *sense datum*) que faça o intermédio entre estar com sede e estar ciente de que se está com sede.

---

<sup>10</sup> “S occurrently satisfies an association relation between E and P =<sub>df</sub>  
(i) S has a *de re* awareness of E’s supporting P, and (ii) as a nondeviant result of this awareness, S is in a dispositional state whereby if he were to focus his attention only on his evidence for P (while all else remained the same), he would focus his attention on E.”



Korcz (1997) acredita que essa condição de associação possui ingredientes internalistas muito rigorosos. Mesmo que ela possa ser cumprida por essa “consciência *de re*” (que parece menos exigente), não parece que algo desse tipo ocorra no cotidiano das pessoas. Korcz dá como exemplo a crença de que o sol possui manchas. Pode ser que alguém esteja ciente de que acredita que o sol possui manchas porque leu isso em um livro de astronomia. Entretanto, ao analisar o caso com mais cuidado, o sujeito logo perceberá que essa crença está sendo sustentada por uma longa lista de razões. Além do livro de astronomia, existe um conjunto de evidências que atestam que o livro é confiável, que o sol existe, que ele havia visto uma palestra sobre isso, que havia assistido a um documentário sobre o fluxo dos campos magnéticos do sol, etc.

Outro problema para as teorias causais de relação de suporte está na existência de cadeias causais desviantes. Esse problema se origina na Filosofia da Ação, com Donald Davidson. Uma cadeia desviante está presente quando é possível para um agente (a) ter a razão R para realizar a ação A, (b) realizar A e (c) R causar a realização de A sem que R seja uma razão pela qual o agente realizou A. Alguns exemplos esclarecem como essa situação é possível de acontecer (Wilson, Shpall, 2012):

(1) Garçom brincalhão – um garçom quer dar um susto em seu chefe. Ele percebe que há uma pilha de copos de vidro por perto e pensa que derrubar a pilha provocaria um grande susto no chefe. Entretanto, ao pensar nas consequências de seus atos (na repreensão do chefe furioso), o garçom fica desestabilizado e cambaleia involuntariamente em direção à pilha de copos, esbarrando nela e fazendo grande estrago.

(2) Criança malcriada – entediado com a apresentação da ópera, um menino decide atrapalhar o maestro do espetáculo forçando uma tosse. Entretanto, o

garoto fica nervoso (peso na consciência ou medo do castigo que o espera) e começa a tossir involuntariamente.

(3) Assassino sortudo – Betty quer matar Jughead. Além disso, ela tem a intenção de matar Jughead de uma tal forma que pareça um acidente inesperado. Betty atira e não acerta Jughead, entretanto, a bala ricocheteia e acerta um galho de árvore muito próximo de Jughead. No interior do galho havia uma colmeia de marimbondos. Os marimbondos atacam Jughead até a morte.

Esses exemplos buscam mostrar que a ação realizada pelos sujeitos é involuntária, ainda que a intenção tenha participado da cadeia causal que culmina na ação. Os casos (1) e (2) apresentam um “desvio causal primário”, tendo em vista que a causa desviante afeta o comportamento físico/motor do indivíduo. O caso (3) apresenta um “desvio causal secundário”, considerando-se que a causa desviante ocorre entre o comportamento do agente e o resultado da intenção.

Uma adaptação do problema da causa desviante para o caso da relação de suporte é o exemplo de Alvin Plantinga (1993: 69):

De repente, vendo Sylvia, eu formo a crença de que eu a estou vendo; como resultado disso, eu me atrapalho e derrubo minha xícara de chá, queimando minha perna. Então, eu formo a crença de que minha perna dói; mas, ainda que minha primeira crença seja uma causa (em parte) da última, não é o caso que eu aceite a última sobre a base evidencial da primeira<sup>11</sup>.

---

<sup>11</sup> “Suddenly seeing Sylvia, I form the belief that I see her; as a result, I become rattled and drop my cup of tea, scalding my leg. I then form the belief that my leg hurts; but though the former belief is a (part) cause of the latter, it is not the case that I accept the latter on the evidential basis of the former.”

---

No exemplo de Plantinga, a causa é desviante no sentido de que a primeira crença faz parte da cadeia causal que origina a segunda crença; entretanto, o mero fato de uma crença possuir um papel causal não parece suficiente para que se dê uma relação de suporte entre ela e a crença que emerge ao longo dessa cadeia causal: “Nem toda cadeia causal de razões para crenças estabelece uma relação de suporte, e aquelas que não conseguem fazê-lo são cadeias causais desviantes”<sup>12</sup> (Korcz, 1997, p. 173).

Neta argumenta que os ataques provenientes das cadeias causais desviantes não são suficientes para mostrar a falsidade das teorias causais de relação de suporte. O ataque mais forte é feito a partir do contraexemplo apresentado pelo caso do advogado cigano, elaborado por Keith Lehrer (1971: 311-312):

O exemplo envolve um advogado que está defendendo um homem acusado de ter cometido oito assassinatos hediondos. Os assassinatos têm características similares; em cada caso, a vítima era um estudante de filosofia de Oxford que foi asfixiado até a morte com uma cópia das Investigações Filosóficas. Existe evidência conclusiva de que o cliente do advogado é o culpado dos sete primeiros assassinatos. Todos, incluindo o advogado, estão convencidos de que o homem em questão cometeu todos os oito crimes; apesar disso, o homem se diz inocente de todos.

Entretanto, o advogado é um cigano com fé absoluta nas cartas. Em uma noite, ele consulta as cartas sobre seu caso, e as cartas lhe dizem que seu cliente é inocente do oitavo assassinato. Ele verifica novamente, e as cartas lhe dão a mesma resposta. Ele começa a se convencer de que o cliente é inocente de um dos oito assassinatos. Como

---

<sup>12</sup> “Not every causal chain from reason to belief establishes a basing relation, and those that fail to do so are deviant causal chains.”

resultado, ele estuda as evidências com uma perspectiva diferente, bem como com grande cuidado e descobre uma linha de raciocínio muito complicada, porém completamente válida, partindo da evidência e chegando à conclusão de que o seu cliente não cometeu o oitavo assassinato. (Ele não tinha condições de obter uma oitava cópia das Investigações Filosóficas). Esse raciocínio confere ao advogado conhecimento. (...)

Entretanto, os fatores emocionais são tão poderosos e o raciocínio tão complicado que ninguém mais está convencido. Todos têm um desejo irresistível de crer que o assassino de todas as oito vítimas foi encontrado. De fato, e esse é um ponto crucial, se não fosse por sua fé inabalável nas cartas, até mesmo o advogado seria influenciado por esses fatores emocionais e duvidaria de que seu cliente é inocente do oitavo assassinato. É apenas por causa de sua fé nas cartas que o raciocínio lhe dá conhecimento<sup>13</sup>.

---

<sup>13</sup> “The example involves a lawyer who is defending a man accused of committing eight hideous murders. The murders are similar in character, in each case the victim is an Oxford philosophy student who has been choked to death with a copy of *Philosophical Investigations*. There is conclusive evidence that the lawyer’s client is guilty of the first seven murders. Everyone, including the lawyer, is convinced that the man in question has committed all eight crimes, though the man himself says he is innocent of all. However, the lawyer is a gypsy with absolute faith in the cards. One evening he consults the cards about his case, and the cards tell him that his client is innocent of the eighth murder. He checks again, and the cards give the same answer. He becomes convinced that his client is innocent of one of the eight murders. As a result he studies the evidence with a different perspective as well as greater care, and he finds a very complicated though completely valid line of reasoning from the evidence to the conclusion that his client did not commit the eighth murder. (He could not have obtained an eighth copy of *Philosophical Investigations*.) This reasoning gives the lawyer knowledge. . . . However, the emotional factors are so powerful and the reasoning so complicated that no one else is convinced. They all have an overwhelming desire to believe that

Essa tentativa de contraexemplo foi muito criticada e divide a intuição dos epistemólogos envolvidos na discussão. Desta forma, é muito relevante especificar o que Lehrer estava tentando fazer ao apresentar esse caso. O objetivo de Lehrer era mostrar que a relação de suporte não precisa envolver causalção<sup>14</sup>. O caso tenta exemplificar uma

---

the murderer of all eight victims has been found. Indeed, and this is the crucial point, if it were not for his unshakable faith in the cards, the lawyer himself would be swayed by those emotional factors and would doubt that his client was innocent of the eighth murder. It is only because of his faith in the cards that the reasoning gives him knowledge.”

<sup>14</sup> Lehrer (1990) possui outro exemplo semelhante em que busca extrair as mesmas conclusões:

**Sr. Raco:** Suponha um homem, Sr. Raco. Ele é preconceituoso e acredita, com base no seu preconceito, que os membros de uma raça estão predispostos a desenvolverem determinada doença. Suponha ainda que essa crença é uma convicção inabalável tão forte que nenhum conjunto contrário de evidências poderia diminuir sua convicção, nem mesmo evidências favoráveis poderiam fortalecer a posição de Sr. Raco. Imagine que Sr. Raco se torna um médico e começa a estudar essa doença em particular. Suponha que Sr. Raco consegue fazer um levantamento de toda a literatura sobre essa enfermidade e descubra que as evidências, bastante conclusivas, confirmam sua convicção. As evidências científicas mostram que somente membros da raça em questão são suscetíveis a essa doença. Imagine que, apesar de ter se tornado um especialista perfeitamente capaz de entender as normas das evidências científicas, infelizmente, quanto a essa questão, Sr. Raco permanece tão preconceituoso quanto antes. Ainda assim, ele entende e avalia as evidências tão bem quanto qualquer outro especialista da área e, portanto, possui (boas) razões que justificam sua crença. Ele descobriu que sua convicção foi confirmada pelas evidências científicas. Ele sabe que somente os membros daquela raça são suscetíveis à doença em questão. Ainda assim, as razões que justificam sua crença não a explicam causalmente. A crença é resultado do preconceito, e não das razões, mas ela é confirmada por razões que provêm justificção para a crença. O preconceito dá ao Sr. Raco convicção, mas as razões lhe dão justificção.

situação desse tipo ao mostrar que a causa da crença do advogado de que seu cliente é inocente do oitavo assassinato é sua fé nas cartas (e o fato de que as cartas lhe disseram que seu cliente era inocente do oitavo assassinato). Entretanto, para Lehrer, a crença do advogado de que seu cliente é inocente do oitavo assassinato é suportada pela complicada linha de raciocínio.

As maiores críticas ao caso se concentram nesta última consideração. Para muitos epistemólogos, o advogado não está justificado em crer que seu cliente é inocente do oitavo assassinato, por causa da complicada linha de raciocínio; logo, o advogado não poderia ter conhecimento de que seu cliente é inocente do oitavo assassinato. Aqui fica clara a grande confusão que as teorias

---

Esses exemplos buscam minar qualquer proposta de teoria causal do conhecimento. Particularmente, Lehrer está interessado em criticar propostas confiabilistas (Goldman, 1967). Para ele, essas teorias incorrem em uma falácia causal. O problema central está na distinção entre teorias que versam sobre explicações de porquê alguém crê em algo e teorias que explicam como alguém sabe algo. Para Lehrer, teorias causais confundem esses dois tipos de explicação. Buscam na explicação da (geração da) crença o *status* epistêmico apropriado para que essa crença se configure como um caso de conhecimento. Lehrer condena qualquer teoria causal por não dar a devida atenção a essa distinção. Para ele, a crença pode ser explicada por qualquer fator alheio à epistemologia (*razões* de ordem prática, política, erótica); entretanto, a explicação de como alguém sabe algo deve depender de evidências. Uma evidência explica como alguém sabe algo se essa evidência justifica (sustenta) a crença desse sujeito.

O que os exemplos do advogado cigano e do Sr. Raco buscam demonstrar são casos em que é possível que uma crença justificada seja causalmente independente da evidência que a justifica. Esses casos mostram como a evidência não explica por que a pessoa mantém a crença, ainda que a justificação para a crença seja sustentada pela evidência. Contra as teorias causais, Lehrer argumenta, a partir desses casos, que a evidência que justifica a crença de uma pessoa pode ser adquirida depois que a crença tenha sido formada (voltaremos a essa questão na nota 17).

de relação de suporte enfrentam. Será que uma relação de suporte garante justificação? Parece que essa reação ao exemplo de Lehrer é exagerada, dado que a relação de suporte parece uma condição necessária de justificação, e não uma condição suficiente<sup>15</sup>.

## Teoria Contrafactual de Relação de Suporte

Concordando com Lehrer e seu caso do advogado cigano, mas buscando manter as intuições em favor de uma Teoria Causal de Relação de Suporte, Marshall Swain desenvolve uma Teoria Contrafactual de Relação de Suporte. Korcz (1997) resume a teoria da seguinte forma: uma crença está suportada sobre uma razão se a razão causou a crença ou a teria causado. A relação causal contrafactual apropriada é aquela em que existe um mundo possível  $W$  no qual a crença ocorre, a razão ocorre, nada além da razão causa a crença e a situação epistêmica do sujeito continua a mesma. Assim, se a razão causou a crença em  $W$ , e  $W$  é muito próximo do mundo atual (mais próximo que outros mundos onde a razão não causa a crença), então a crença está suportada sobre essa razão no mundo atual.

Swain (1979) define o suporte da crença de um sujeito  $S$  de que  $h$  sobre um conjunto de razões causais  $R$  em um tempo  $t$  da seguinte forma:

1.  $S$  acredita que  $h$  em  $t$ ; e
2. Para cada membro  $r_j$  de  $R$ , existe um tempo  $t_n$  (que pode ser idêntico ou anterior a  $t$ ) tal que  $S$  tem (ou tinha)  $r_j$  em  $t_n$ ; e, ou (b, i) existe uma cadeia causal do fato de  $S$  estar no estado de razão  $r_j$  em  $t_n$  para  $S$  crer que  $h$  em  $t$ , ou

---

<sup>15</sup> Reveja notas 2, 5 e 7.

(b, ii)  $S$  estar no estado de razão  $r_j$  em  $t_n$  é um sobredeterminante genuíno da crença de  $S$  de que  $h$  em  $t$ , ou

(b, iii)  $S$  estar no estado de razão  $r_j$  em  $t_n$  é um pseudo-sobredeterminante da crença de  $S$  de que  $h$  em  $t$ <sup>16</sup>.

Obviamente as cláusulas (b, i), (b, ii) e (b, iii) precisam de esclarecimento. Especificamente, precisamos esclarecer o que significa “cadeia causal”, “sobredeterminante genuíno” e “pseudo-sobredeterminante”. Swain (1979) define esses termos da seguinte forma:

(Cadeia causal) Em que  $c, d_1, d_2, \dots, d_n, e$  é uma sequência de eventos ocorrentes distintos (mas não necessariamente uma sequência temporal em que  $c$  e  $e$  podem ser os únicos membros), essa sequência é uma cadeia causal se e somente se:

- (a) se  $c$  não tivesse ocorrido,  $d_1$  não teria ocorrido,
- (b) para cada  $n$ , se  $d_{n-1}$  não tivesse ocorrido, então  $d_n$  não teria ocorrido, e
- (c) se  $d_n$  não tivesse ocorrido, então  $e$  não teria ocorrido.

---

<sup>16</sup> “1. S believes that h at t; and

2. for every member  $r_j$  of R, there is some time  $t_n$  (which may be identical with or earlier than t) such that

S has (or had)  $r_j$  at  $t_n$ ; and either

(b, i) there is a causal chain from S’s being in reason state  $r_j$  at  $t_n$  to S’s believing that h at t, or

(b, ii) S’s being in reason state  $r_j$  at  $t_n$  is a genuine overdeterminant of S’s believing that h at t, or

(b, iii) S’s being in reason state  $r_j$  at  $t_n$  is a pseudo-overdeterminant of S’s believing that h at t.”



(Sobredeterminante genuíno) Em que  $c$  e  $e$  são eventos ocorrentes,  $c$  é um sobredeterminante genuíno de  $e$  se e somente se: existe algum conjunto de eventos ocorrentes  $D = [d_1, d_2, \dots, d_n]$  (eventualmente, tendo apenas um membro) tal que:

(a') se  $c$  não tivesse ocorrido e se qualquer membro  $d_i$  de  $D$  tivesse ocorrido, mas nenhum outro membro de  $D$  tivesse ocorrido, e se  $e$  tivesse ocorrido de qualquer forma, então teria havido uma cadeia causal de eventos realmente ocorrentes e distintos de  $d_i$  até  $e$ ; e

(b') se nenhum membro de  $D$  tenha ocorrido e se  $c$  e  $e$  tenham ocorrido de qualquer forma, então teria havido uma cadeia causal de eventos realmente ocorrentes e distintos de  $c$  até  $e$ .

(Pseudo-sobredeterminante) Em que  $c$  e  $e$  são eventos ocorrentes,  $c$  é um pseudo-sobredeterminante de  $e$  se e somente se:

(a'') não existe cadeia causal de  $c$  até  $e$ ; e

(b'') existe um conjunto de eventos ocorrentes  $D = [d_1, d_2, \dots, d_n]$  (eventualmente, tendo apenas um membro) tal que existe uma cadeia causal de cada  $d_i$  em  $D$  até  $e$ , e se nenhum membro de  $D$  tenha ocorrido, mas  $c$  e  $e$  tenham ocorrido de qualquer forma, então teria havido uma cadeia causal de  $c$  até  $e$ <sup>17</sup>.

---

<sup>17</sup> "(Causal chain) Where  $c, d_1, d_2, \dots, d_n, e$  is a sequence of distinct occurrent events (but not necessarily a temporal sequence, and where  $c$  and  $e$  may be the only members), this sequence is a causal chain if and only if:

(a) if  $c$  had not occurred,  $d_1$  would not have occurred

(b) for each  $n$ , if  $d_{n-1}$  had not occurred, then  $d_n$  would not have occurred, and

Neta (2011) sintetiza a proposta de Swain da seguinte forma: para que a crença de um sujeito S de que h seja suportada sobre um conjunto de razões causais R em um tempo t, a crença de S deve ser dependente de R contrafactualmente em uma destas três maneiras: (i) a crença ter sido causada por R, (ii) R ser uma causa sobredeterminante da crença (R causa a crença e continuaria causando em situações contrafactuais onde R e a crença ocorrem), ou (iii) ter sido causada por R, caso a atual cadeia causal não tivesse acontecido.

A ideia fica mais clara quando reavaliarmos o caso do advogado cigano de Lehrer: a complicada linha de raciocínio não causou a crença do advogado de que seu cliente era inocente do oitavo assassinato; entretanto, a complicada linha de raciocínio pseudo-sobredetermina essa crença. Se o advogado não tivesse lido as cartas, mas tivesse descoberto a complicada linha de raciocínio e tivesse

---

(c) if  $d_n$  had not occurred, then e would not have occurred.

(Genuine overdeterminant) Where c and e are occurrent events, c is a genuine overdeterminant of e if and only if: There is some set of occurrent events  $D = [d_1, d_2, \dots, d_n]$  (possibly having only one member) such that:

(a') if c had not occurred and if any member  $d_i$  of D had occurred, but no other members of D had occurred, and if e had occurred anyway, then there would have been a causal chain of distinct actually occurrent events from  $d_i$  to e; and

(b') if no member of D had occurred, and if c and e had occurred anyway, then there would have been a causal chain of distinct actually occurrent events from c to e.

(Pseudo-overdeterminant) Where c and e are occurrent events, c is a pseudo-overdeterminant of e if and only if:

(a'') there is no causal chain from c to e; and

(b'') there is some set of occurrent events  $D = [d_1, d_2, \dots, d_n]$  (possibly having only one member) such that there is a causal chain from each  $d_i$  in D to e, and if no member of D had occurred, but c and had occurred anyway, then there would have been a causal chain from c to e."

acreditado na inocência de seu cliente, então haveria uma cadeia causal a partir da descoberta do advogado da complicada linha de raciocínio até a crença do advogado de que seu cliente era inocente do oitavo assassinato. Assim, a teoria de Swain deixa clara a tentativa de Lehrer, ao mostrar que a crença do advogado de que seu cliente era inocente do oitavo assassinato está suportada sobre a complicada linha de raciocínio descoberta pelo advogado, ainda que essa complicada linha de raciocínio não tenha causado a crença de que o seu cliente era inocente do oitavo assassinato<sup>18</sup>.

---

<sup>18</sup> O próprio Lehrer (1990) discorda dessa conclusão. Os casos por ele construídos têm por objetivo opor qualquer tipo de teoria causal de suporte, contrafactual ou não. Lehrer concede que, nos casos ordinários, a proposta de Swain está correta, isto é, ainda que a razão não tenha originado a crença, essa razão está fazendo o suporte causal. Entretanto, Lehrer não está convencido de que, nos casos que ele apresenta, essa relação está clara. Parece insuficiente a conclusão de que a (boa) razão parcialmente explique ou sustente a crença, pois, nos casos de Lehrer, é possível que a (boa) razão não adicione nada à força crença. Lehrer acredita que preconceito ou uma fé inabalável (nas cartas) podem ser tudo que explique e sustente a crença.

Lehrer também é contrário a essa interpretação do caso do advogado cigano. Para ele, a fé inabalável nas cartas é uma condição necessária para que o sujeito tenha conhecimento nesse caso. No caso de o advogado não ter lido as cartas e estar diante (somente) da complicada linha de raciocínio, Lehrer permanece cético quanto ao desfecho da vida mental desse indivíduo. Para ele, nada garante que o sujeito passe, então, a crer (que o seu cliente é inocente do oitavo assassinato) com base na (boa) razão (da complicada linha de raciocínio).

Lehrer defende sua posição a partir da adição de um novo elemento à história: ao perder suas crenças baseadas na fé inabalável que tinha nas cartas, o sujeito fica louco e deixa de ser influenciado pela razão (o que não parece fazer jus à proposta de Swain; esse é um mundo mais distante do que aquele onde o indivíduo não leu as cartas e mesmo assim está diante da complicada linha de raciocínio). Aceitando o adendo de Lehrer ao caso, de fato, mesmo que o advogado venha a crer que seu cliente é inocente do oitavo assassinato ao estar diante da complicada linha de raciocínio, não é possível que ele esteja

Em virtude de ser uma reformulação de teorias causais, a Teoria Contrafactual de Relação de Suporte proposta por Swain não escapa do problema das cadeias causais desviantes. Entretanto, outro contraexemplo formulado por Joseph Tolliver ataca as inovações propostas por Swain, em especial a ideia de que uma razão que satisfaz a condição de ser uma pseudo-sobredeterminante de uma crença cumpre uma condição suficiente para que essa razão estabeleça uma relação de suporte com a crença. Korcz (2010) apresenta o contraexemplo de Tolliver da seguinte maneira:

(...) suponha um estudante de física que aprendeu que a partir do período de um pêndulo (i.e., o tempo que ele leva para completar um balanço) alguém pode calcular seu comprimento e vice-versa. O estudante mede um pêndulo particular e descobre que o pêndulo tem um comprimento  $C$  e calcula que o pêndulo deve ter um período  $P$ . O estudante também tem duas crenças gerais sobre pêndulos, a saber (1) que se  $x$  é um pêndulo de período  $P$ , então  $x$  é um pêndulo de comprimento  $C$  e (2) que se  $x$  é um pêndulo de comprimento  $C$ , então  $x$  é um pêndulo de período  $P$ . Podemos supor que está claro nesse caso que a crença do estudante em relação ao período é suportada (ao menos em parte) sobre sua crença em relação ao comprimento, mas sua crença em relação ao comprimento não está suportada sobre sua crença em relação ao período. Mas, Tolliver afirma, a crença do estudante em relação ao período pseudo-sobredetermina sua crença em relação ao comprimento do pêndulo e, portanto, é contada, na teoria de Swain, como o suporte de sua crença em relação ao comprimento do pêndulo. Isso é assim

---

sustentando sua crença nessa razão, dado que ele ficou louco, e a loucura é responsável por explicar suas crenças.

---

porque, segundo Tolliver, se a causa real da crença do estudante em relação ao comprimento não tivesse ocorrido e se o estudante ainda tivesse ambas as crenças em relação ao período e ao comprimento, então sua crença em relação ao período teria causado sua crença em relação ao comprimento<sup>19</sup>.

## **Teoria Doxástica de Relação de Suporte**

Alguns filósofos tentaram formular teorias doxásticas da relação de suporte. Essas teorias defendem que a relação de suporte requer do sujeito a manutenção de crenças sobre a relação epistêmica que existe entre uma crença de seu sistema e as razões sobre as quais ela está suportada. Mais precisamente, essas teorias defendem que manter uma meta-crença de que a razão que suporta a crença é uma boa razão para suportá-la é uma condição suficiente para que a crença esteja sustentada nessa razão.

---

<sup>19</sup> “(...) suppose a physics student has learned that from the period of a pendulum (i.e., the time it takes to complete a swing) one can calculate its length and vice versa. The student measures a particular pendulum and discovers that it has a length  $L$ , and calculates that it must have period  $P$ . The student also has two general beliefs about pendulums, namely (1) that if  $x$  is a pendulum of period  $P$ , then  $x$  is a pendulum of length  $L$ , and (2) that if  $x$  is a pendulum of length  $L$ , then  $x$  is a pendulum of period  $P$ . We may suppose that it is clear in this case that the student's belief about the period is based (at least in part) on her belief about its length, but her belief about its length is not based on her belief about the period. But, Tolliver claims, the student's belief about the period pseudo-overdetermines her belief about the length of the pendulum, and hence gets counted, on Swain's theory, as the basis of her belief about the length of the pendulum. This is so because, according to Tolliver, if the actual cause of the student's belief about the length had not occurred, and if the student still had both her belief about the period and her belief about the length, then her belief about the period would have caused her belief about the length.”

Uma dessas teorias foi formulada por Joseph Tolliver (1982). Tolliver buscou acomodar seu contraexemplo do pêndulo (que atacava a Teoria Contrafactual de Swain) e o contraexemplo do advogado cigano (que atacava a Teoria Causal). Neta (2011: 114) apresenta a proposta de Tolliver da seguinte forma:

S suporta sua crença que [c] sobre a razão r no tempo t se e somente se:

(i) S acredita que [c] em t e S acredita em r em t,

(ii) S acredita que a verdade de r é evidência para a verdade de [c], e

(iii) caso S estimasse que a probabilidade de [c] é igual a h ( $0 < h < 1$ ) em t, se S viesse a acreditar que r pela primeira vez em t, então a estimativa de S sobre a probabilidade da proposição “a probabilidade de [c] é maior ou igual a h” seria maior em t do que era antes de t<sup>20</sup>.

Korcz (2010) observa que a condição (iii) tem como objetivo dar conta do contraexemplo do pêndulo. Essa condição pretende especificar a direção da relação evidencial de suporte. Dado que o estudante chegou primeiro a uma crença sobre o comprimento do pêndulo e somente depois fez o cálculo para inferir o período, a estimativa da probabilidade de que o pêndulo tem um período P aumentou. Entretanto, a crença do estudante sobre o período do pêndulo não aumentará a estimativa da

---

<sup>20</sup> “S bases her belief that b on reason r at time t if and only if:

(i) S believes that b at t and S believes that r at t,

(ii) S believes that the truth of r is evidence for the truth of b, and

(iii) where S’s estimate of the likelihood of b equals h ( $0 < h < 1$ ) at t, if S were to come to believe that r for the first time at t, then S’s estimate of the likelihood of the proposition “the likelihood of b is greater than or equal to h” would be greater at t than it was prior to t.”

---

probabilidade de que o pêndulo tem comprimento  $C$ , porque a crença do estudante sobre o comprimento é necessária para calcular o período.

Apesar de desviar sua própria crítica, a condição (iii) abre espaço para outros contraexemplos. Neta (2011) aponta o caso de Jones que sempre acreditou que um analgésico é efetivo porque, sempre que tomava o analgésico, ele sentia que o analgésico funcionava. Entretanto, Jones leu um artigo científico que dizia que o analgésico, de fato, era efetivo, entretanto, também dizia que os consumidores do analgésico eram péssimos julgadores da efetividade dos analgésicos que tomavam. Diante da leitura do artigo, Jones continua crendo que o analgésico é efetivo, mas o suporte sobre o qual essa crença estava mantida mudou: antes, era o julgamento que Jones fazia do histórico de efetividade do analgésico quando o consumia, e, agora, é o estudo científico divulgado na revista que Jones leu.

Até aqui nada surpreendente e nada que deponha contra a condição (iii). Entretanto, suponha que, antes de ler o artigo, Jones atribuía uma estimativa de probabilidade sobre a efetividade do analgésico de 0,95, e que agora Jones atualizou sua estimativa de probabilidade da efetividade do analgésico para 0,85. Naturalmente, esperamos que Jones tenha uma estimativa de probabilidade da proposição “a probabilidade de que o analgésico seja efetivo é maior ou igual à 0,95” menor do que aquela que ele tinha antes de ler o artigo científico. Agora sim temos um problema para a proposta de Tolliver. Dada a atualização da crença de Jones, ele não satisfaz a cláusula (iii), ainda que seja muito questionável que Jones tenha deixado de ter uma razão que suporte sua crença sobre a efetividade do analgésico. Dessa forma, a condição (iii) parece muito forte para uma teoria de relação de suporte.

Alguém poderia pensar que uma teoria doxástica ainda teria condições de se sustentar ao abandonar a

condição (iii) da proposta de Tolliver. No entanto, a maior parte das críticas que uma teoria doxástica recebe diz respeito à condição (ii). Korcz (2010) apresenta três possíveis objeções que exploram a implausibilidade da condição (ii):

(a) Posse de conceitos epistêmicos – a condição (ii) parece muito forte ao exigir que formemos meta-crenças sobre relações epistêmicas entre crenças e razões. Crianças pequenas ou pessoas menos instruídas podem carecer dos conceitos epistêmicos exigidos para a formação das meta-crenças e, ainda assim, serem capazes de ter razões para acreditar. Os proponentes de teorias doxásticas poderiam responder a essa objeção (1) apelando para a possibilidade de que formamos meta-crenças sem o emprego de conceitos epistêmicos (de maneira semelhante à estratégia da “consciência *de re*”, de Moser) ou (2) dizendo que esse tipo de meta-crença seria mantida em condições apropriadas mas que não são necessárias em todos os casos.

(b) Ezekiel e a meta-crença que não estabelece uma relação de suporte –

Por exemplo, suponha que Ezekiel pertence a um culto religioso e acredita servilmente em qualquer coisa que o líder do culto, Exidor, lhe diga. Um dia, Exidor diz a Ezekiel que a crença de Ezekiel em Deus é uma boa razão para que Ezekiel acredite em tudo o mais naquilo que ele crê, e Ezekiel servilmente vem a acreditar nisso. Entretanto, seria muito contra-intuitivo aceitar que, agora, qualquer coisa que Ezekiel crê esteja, de fato, sustentada sobre sua crença em Deus<sup>21</sup> (Korcz, 2010).

---

<sup>21</sup> “For example, suppose Ezekiel belongs to a religious cult and slavishly believes whatever the cult leader, Exidor, tells him. One day



(c) Razões subconscientes – parece bastante plausível que tenhamos algumas razões para crer das quais não estamos cientes. Talvez essas razões estejam no nosso subconsciente, e dificilmente teremos acesso a elas. Se essas razões estão longe do nosso acesso, parece bastante difícil que consigamos formar meta-crenças sobre a relação que essas razões têm com nossas outras crenças.

## Conclusão

No presente trabalho, buscamos apresentar uma visão geral do que é a relação de suporte. Acredita-se que o esclarecimento de como razões sustentam crenças pode fornecer elementos cruciais para a construção de teorias da justificação. Guiados por esse objetivo, epistemólogos têm defendido versões bastante restritas de relação de suporte, nas quais é necessário que as crenças estejam relacionadas com boas razões. Apresentamos três propostas que fazem jus a essa aspiração.

A teoria causal parece bastante plausível à primeira vista; estaríamos dispostos a assentir que as razões que geram minhas crenças são aquelas que a sustentam. Mas parece ter problemas em especificar exatamente como se dá a relação causal entre crenças e razões. Um obstáculo importante para essas teorias parece ser a prática cotidiana de revisão de crenças. Dado que, em geral, utilizamos crenças como razões para sustentar outras crenças, aos revisarmos nosso sistema de crenças frequentemente trocamos as razões (crenças) que sustentam algumas

---

Exidor tells Ezekiel that his belief in God is a good reason to believe everything else Ezekiel believes, and Ezekiel slavishly comes to believe this. Nonetheless, it would seem highly counter-intuitive to accept that, now, everything else Ezekiel believes is in fact based on his belief in God.”

(outras) crenças. Como vimos com Lehrer, a falha das teorias causais está em querer destacar propriedades epistêmicas do processo de criação das crenças, dada a constatação cotidiana de troca de razões para sustentar uma crença que já possuíamos a muito tempo.

A teoria contrafactual tenta corrigir esse problema das teorias causais mais simples, ao mostrar que, no caso em que os eventos ocorrentes atuais não houvessem ocorrido, teríamos uma cadeia causal conectando as crenças e as razões em análise. O contraexemplo de troca de razões é facilmente explicado pela teoria contrafactual. Se a história da vida mental do sujeito mostra que  $R_1$  foi responsável por formar a crença  $C$  em  $t_1$ , mas, em  $t_2$ ,  $R_1$  vem a ser substituída por  $R_2$  (caso de troca de razões por revisão do sistema de crenças, por exemplo), o defensor da teoria contrafactual argumentaria que, caso esse indivíduo jamais tivesse se deparado com  $R_1$ , mas passasse a ter a razão  $R_2$ , formaria a crença  $C$  com base em  $R_2$ . Isso mostra que uma cadeia causal genuína poderia se dar entre  $R_2$  e  $C$ , logo é possível que  $C$  e  $R_2$  estejam em relação de suporte.

Entretanto, as teorias contrafactuais parecem não dar conta de outros contraexemplos. Ela parece refém de exemplos que mostram os problemas de modelos nomológico-dedutivos de explicações científicas (Hempel, 1965). Um exemplo famoso é a explicação da sombra do poste (Woodward, 2014). Alguém pode derivar o tamanho da sombra projetada por um poste a partir da altura do poste e do ângulo do sol sobre o horizonte, a partir das leis da propagação da luz. Nesse caso, o ângulo do sol e a altura do poste explicam o tamanho da sombra. Mas, utilizando a mesma fórmula, poderíamos derivar a altura do poste a partir do ângulo do sol e do tamanho da sombra. Apesar de a conta estar correta e respeitando a lei científica que estamos utilizando, esse caminho não parece fornecer uma explicação. Parece muito irrazoável dizer que o tamanho da sombra explica a altura do poste.

---

Exemplos como esse sugerem que explicações possuem direções assimétricas. Em Filosofia da Ciência, isso é um problema para a teoria nomológico-dedutiva, que parece insensível a essa assimetria. No caso das teorias contrafactuais de relação de suporte, isso também parece ocorrer, como demonstrado no exemplo de Tolliver. Como a teoria contrafactual pretende preservar as intuições de uma teoria causal, ela parece refém do mesmo tipo de problema. A teoria também busca oferecer uma explicação; se os casos mostram uma insensibilidade da teoria em identificar relações assimétricas de explicação causal, a teoria não cumpre adequadamente com seu objetivo.

A última teoria de relação de suporte apresentada tenta ancorar a identificação da relação de suporte em outra crença do indivíduo, uma meta-crença capaz de detectar essa relação. As estratégias que apelam para níveis superiores geralmente carregam consigo outros problemas. Uma estratégia desse tipo parece gerar demandas muito altas para que sujeitos ordinários. Além disso, nossa vida mental não parece tão transparente para nossa própria consciência. Por mais que a introspecção possua uma série de privilégios, está longe de ser unânime um acordo em torno da sua infalibilidade e capacidade de alcance (parece que algumas crenças ou aspectos da nossa vida mental não são identificados pela introspecção).

Assim, a discussão em torno da relação de suporte mostra-se um tópico com poucas certezas e definições, um campo aberto para novas propostas ou ajustes daquelas aqui apresentadas.

## **Agradecimentos**

Gostaria de agradecer aos colegas da disciplina de Epistemologia Analítica III (2014/2), ministrada pelo professor Cláudio Gonçalves De Almeida na Pontifícia Universidade Católica do Rio Grande do Sul, onde uma

versão preliminar desse texto foi lida. Obviamente, nenhum deles é responsável pelos eventuais erros do presente trabalho. Agradeço também aos organizadores desse volume pela iniciativa em divulgar os trabalhos dos novos epistemólogos brasileiros, bem como pelo convite em fazer parte deste projeto.

## Referências

- Alvarez, Maria (2010). *Kinds of Reasons: an Essay in the Philosophy of Action*, Oxford: Oxford University Press.
- Dancy, Jonathan (2000). *Practical Reality*, New York: Oxford University Press.
- Dickinson, Jeremy A. (2012). *An Essay on the Ontology of Reasons*, Tese (Filosofia). Colégio de Artes e Ciências, Universidade de Siracusa, Siracusa.
- Evans, Ian (2013). 'The Problem of the Basing Relation', *Synthese* 190(14): 2943–2957.
- Goldman, Alvin (1967). 'A Causal Theory of Knowing', *The Journal of Philosophy* 64(12): 357–372.
- Hempel, Carl (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Korcz, Keith A. (1997). 'Recent Work On The Basing Relation', *The American Philosophical Quarterly* 34(2): 171-191. Publicado na internet, disponível através da seguinte URL:  
<<http://www.ucs.louisiana.edu/~kak7409/RWBR.html>>.

- 
- Korcz, Keith A. (2010). 'The Epistemic Basing Relation'. In: Zalta, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL: <<http://plato.stanford.edu/archives/spr2010/entries/basing-epistemic/>>.
- Lehrer, Keith (1971). 'How Reasons Give us Knowledge, or the Case of the Gypsy Lawyer', *The Journal of Philosophy* 68(10): 311–313.
- Lehrer, Keith (1990). *Theory of Knowledge*, Boulder, CO: Westview.
- Moser, Paul (1989). *Knowledge and Evidence*, Cambridge, UK: Cambridge University Press.
- Neta, Ram (2011). 'The Basing Relation'. In: Bernecker, Sven; Pritchard, Duncan (Eds.). *The Routledge Companion to Epistemology*, London and New York: Routledge, 109-118.
- Plantinga, Alvin (1993). *Warrant: The Current Debate*, Oxford: Oxford University Press.
- Scanlon, Thomas (1999). *What We Owe to Each Other*, Cambridge, MA: Belknap Press.
- Swain, Marshall (1979). 'Justification and the Basis of Belief'. In: Pappas, George (Ed.). *Justification and Knowledge*, Dordrecht: D. Reidel, 25-50.
- Tolliver, Joseph (1982). 'Basing Beliefs on Reasons', *Grazer Philosophische Studien* 15: 149–161.
- Turri, John (2009). 'The Ontology of Epistemic Reasons', *Noûs* 43(3): 490-512.

Wilson, George; Shpall, Samuel (2012). 'Action'. In: Zalta, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL: <<http://plato.stanford.edu/entries/action/>>.

Woodward, James (2014). 'Scientific Explanation'. In: Zalta, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL: <<http://plato.stanford.edu/entries/scientific-explanation/>>.

# CONFIANÇA E TESTEMUNHO COMO INTERPESSOAIS: UMA TENTATIVA DE ANALOGIA ENTRE ÉTICA E EPISTEMOLOGIA DO TESTEMUNHO

---

*Patricia Ketzer*<sup>1</sup>

A procura por uma definição de confiança nos fez voltar às origens do termo na busca de compreender seu uso em Epistemologia. Muitos autores apontavam para definições oriundas de outras áreas, como Sociologia ou Filosofia Política, mas uma análise breve do uso do conceito nos debates sobre o testemunho apontou para Filosofia Moral. A utilização de terminologia oriunda da Ética na Epistemologia remonta à História da Filosofia. Inicia-se com John Locke (1975), e, posteriormente, dá origem a toda uma tradição, conhecida como ‘deontologismo epistêmico’, que busca definir conhecimento fazendo uso de conceitos éticos. Existem três opções possíveis quando utilizamos terminologia da Ética em Epistemologia. Pode-se utilizar os conceitos no mesmo sentido que eles possuem em suas áreas de origem, reduzindo os conceitos epistêmicos a conceitos éticos. Entretanto, Roderick Firth (1978) apresenta uma defesa contundente de que conceitos epistêmicos não são redutíveis a conceitos éticos. Firth defende que é possível

---

<sup>1</sup> Email: [patriciaketzer@gmail.com](mailto:patriciaketzer@gmail.com).

utilizá-los de forma analógica. A analogia seria a segunda opção; transpõem-se esses conceitos da Ética, aplicando-os na discussão epistemológica. Outra alternativa seria um uso exclusivamente epistêmico, que desconsidera as características morais dos conceitos.

Neste trabalho, avaliaremos a possibilidade de um uso analógico. Para tal, iremos apresentar a concepção formulada por Benjamin McMyler, em seu livro *Testimony, Trust and Authority* (2011). McMyler propõe o que ele nomeia de ‘modelo de segunda pessoa’, mas sua proposta está sujeita às mesmas objeções que se pode apontar a qualquer concepção interpessoal, pois considera o conceito moral de confiança nos debates sobre testemunho. Deste modo, o filósofo será representativo das visões interpessoais no decorrer deste trabalho; através de sua proposta, avaliaremos a viabilidade de um uso analógico do conceito de confiança em Epistemologia. Apresentaremos, brevemente, o conceito de confiança em Ética. Em seguida, apresentaremos a proposta interpessoal de McMyler. E, por fim, destacaremos as principais objeções a essa proposta e ao uso do conceito de confiança.

O conceito de confiança em Filosofia Moral pode ser distinguido de dois modos: fiar-se e confiar (Cf. Baier, 1986). Trata-se de conceitos distintos, com características próprias. O conceito de fiar-se remete não apenas a pessoas, mas também a coisas. Nos fiamos no despertador, ou no semáforo, mas também nos fiamos no transeunte desconhecido para nos informar onde fica a estação de metrô mais próxima. Essa relação não envolve vínculos profundos e quando desapontada não gera atitudes reativas. Quando nos fiamos em alguém não nos ressentimos se essa pessoa não age conforme o esperado ou não nos informa corretamente sobre a próxima estação de metrô, apenas nos sentimos desapontados, do mesmo modo que não nos sentimos traídos quando o despertador não funciona e nos atrasamos para o trabalho.



---

Confiar é um tipo de relacionamento diferente, que gera atitudes reativas quando desapontado, tais como raiva e ressentimento. Quando confiamos em um amigo e ele nos desaponta, nós nos sentimos traídos. Ainda, confiança envolve boa vontade, comprometimento e, segundo alguns autores (Cf. McLeod, 2011), otimismo para com a pessoa confiada. Quem confia tem que estar aberto aos riscos que essa confiança pressupõe; a tentativa de eliminar o risco através de reflexão racional seguidamente enfraquece a atitude de confiança, reduzindo-a a fiabilidade.

Confiar implica em se colocar em uma situação de vulnerabilidade à traição. Frequentemente, confiança é incompatível com reflexão racional. Por exemplo, aquela confiança que depositamos nas pessoas com o objetivo de que ajam de modo fidedigno. Confio que meu amigo não retornará a fazer uso de substâncias ilícitas. Confio, ainda que saiba que ele já recaiu muitas vezes e que existe grande probabilidade de recair novamente. Confiança pode nos fazer resistentes a evidências.

O conceito de confiança em Filosofia Moral é definido a partir destas características: uma relação interpessoal de fiabilidade na boa vontade do outro, que envolve aceitação da vulnerabilidade. Confiar é assumir a responsabilidade de identificar a fiabilidade e a boa vontade do outro, e ter o direito de delegar algo a esse outro. Por outro lado, o confiado assume a responsabilidade de corresponder ao que se espera dele com boa vontade.

Confiança moral implica em aceitação dos riscos e em resistência a evidências. A tentativa de eliminar riscos através de reflexão racional enfraquece o conceito, pois é característica da confiança que se esteja vulnerável à traição e aberto aos riscos de confiar em alguém. A busca por razões para confiar pode anular a confiança, reduzindo-a a fiabilidade. Além disso, ao confiar, é comum que desconsideremos evidências contrárias em relação à pessoa confiada. E em Epistemologia, assumimos que é errado

crer com base em evidências insuficientes. O conceito de confiança moral não é dirigido para a verdade, pois muitas vezes confiamos por questões que são moralmente significativas, mas incompatíveis com fatores epistêmicos. Não podemos reduzir o conceito moral de confiança nos debates em Epistemologia, pois não maximiza os fins epistêmicos de adquirir crenças verdadeiras e evitar crenças falsas.

Ainda, se aceitamos, como alguns autores sugerem (Cf. McLeod, 2011), que confiar exige otimismo em relação à pessoa confiada, teremos a violação da integridade intelectual do indivíduo, que irá crer de forma desproporcional às evidências, esperando demais do sujeito em quem confia, tornando-se uma pessoa crédula. Otimismo e pessimismo não são atitudes aceitas em Epistemologia, porque pressupõem esperar demais ou de menos, em vez de levar em conta as evidências disponíveis. Deste modo, descartamos a possibilidade de uma redução do conceito moral na Epistemologia do Testemunho.

As visões interpessoais de testemunho utilizam o conceito de confiança de modo análogo ao uso em Filosofia Moral. Trata-se da transposição do conceito de confiança moral aplicado à Epistemologia do Testemunho, em uma tentativa de considerar os aspectos morais do conceito, juntamente com fatores epistêmicos. Analisaremos a proposta de Benjamin McMyler como representante das concepções interpessoais.

Concepções interpessoais de testemunho consideram a intencionalidade dos agentes, de modo que testemunho é concebido no âmbito de práticas religiosas, sociais ou políticas e de contextos educacionais. Esta definição de testemunho tem sido vista como implicando competência e boa vontade epistêmica. McLeod (2011) afirma que podemos compreender confiança a partir de várias perspectivas, mas elas são coerentes apenas se compartilham características da confiança interpessoal.

## A Visão Interpessoal de Benjamin McMyler: O Modelo de Segunda Pessoa

McMyler (2011) defende que aprender a partir do testemunho é fundamentalmente uma capacidade epistêmica social, um empreendimento cooperativo entre falante e ouvinte. Somente a crença testemunhal admite conceder razões para crer. E, se é possível se conceder razões, significa que não somos os únicos responsáveis pela justificação de nossas crenças. Se perguntado sobre como sabe, você tem o direito de transferir a resposta à outra pessoa, o que ele nomeia *direito epistêmico de conceder*. McMyler (2011) define esse direito como uma obrigação de os sujeitos envolvidos na relação testemunhal responderem a desafios sobre suas crenças, ou abandoná-las. A peculiaridade da justificação testemunhal é justamente que o ouvinte possa recorrer ao falante original quando questionado, lançando para ele o desafio de responder à questão.

Entretanto, McMyler (2011) destaca que nem todo desafio epistêmico pode ser concedido ao falante original. “Um ouvinte tem o direito de conceder somente aqueles desafios que envolvem a produção de evidência contra o que tem afirmado saber” (McMyler, 2011: 63). Em uma relação testemunhal, em que um falante afirma (direta ou indiretamente) que  $p$  para um ouvinte, o falante assume o ônus epistêmico<sup>2</sup> de  $p$ , mas o ouvinte assume o ônus epistêmico<sup>3</sup> em relação à fidedignidade do falante. Assumir o ônus epistêmico em relação à fidedignidade do falante,

---

<sup>2</sup> O que McMyler parece compreender por ônus epistêmico (*epistemic burden*) é a obrigação de responder às objeções que possam ser levantadas para as proposições proferidas.

<sup>3</sup> No caso do ouvinte, ele deve responder às objeções sobre a fidedignidade do testificador.

entretanto, não quer dizer que o ouvinte esteja indo às suas próprias conclusões sobre se  $p$  é verdadeiro, exatamente por isso ele pode conceder ao falante quando questionado sobre  $p$ . O ônus epistêmico é compartilhado entre falante e ouvinte, de modo que o ouvinte pode conceder ao falante, mas se o falante não der conta do questionamento levantado, ambos perdem justificação.

O que distingue justificação testemunhal de outros tipos de justificação é o direito epistêmico de conceder, de modo que o grande problema em Epistemologia do Testemunho seria explicar esse direito. Quando temos o direito de conceder ao falante? Por que temos e por que não temos esse direito?

Vamos pensar no seguinte caso: eu escutei Marcelo dizer a Ana, na saída para o almoço, que tinha o *Compêndio de Epistemologia* em sua mochila. No almoço, André comenta comigo que precisa fazer uma consulta ao livro, e eu lhe digo que Marcelo está com ele. Ele me pergunta como eu sei, e apresenta um desafio<sup>4</sup> para minha crença: “Eu mencionei para Marcelo que precisava do livro e ele não me ofereceu!”, afirma André. Eu tenho o direito epistêmico de conceder a Marcelo o questionamento de André, mesmo que Marcelo não tenha dito para mim que estava com o livro, porque ele afirmou publicamente para Ana, e eu o ouvi.

Mas existem alguns casos nos quais ouvir alguém dizer que  $p$  não basta para que eu possa conceder a essa pessoa se questionado sobre  $p$ . Um dos casos que McMyler (2011) destaca é a espionagem. Eu ouvi Ana afirmar que  $p$  porque implantei uma escuta em seu telefone, de modo que isso me impede de conceder a ela qualquer autoridade. As palavras de Ana podem até servir de evidência para minha crença de que  $p$ , mas eu não posso conceder a ela, pois estou chegando a  $p$  por minhas próprias conclusões. Nesse

---

<sup>4</sup> McMyler utiliza essa terminologia; no inglês, *challenge*.

---

caso, não há a relação intencional necessária para justificação testemunhal<sup>5</sup>.

A defesa de McMyler (2011: 66) é a de que justificação testemunhal exige endereçamento a um ouvinte, seja através do proferimento público de uma sentença, seja através de proferimento direto a determinado ouvinte. Mas, em casos no quais não há intenção de informar, como no caso da escuta no telefone, não há justificação testemunhal. Pode-se admitir que, nesses casos, a declaração do falante, mesmo sem intenção, pode servir como evidência, mas, como McMyler defende que o que caracteriza a relação testemunhal enquanto tal é a possibilidade de conceder ao falante, e esses casos não se aplicam, então não são casos de justificação testemunhal.

Endereçamento é uma condição necessária para justificação testemunhal, não uma condição suficiente. Se o problema do testemunho é explicar o privilégio de segunda pessoa, que permite conceder desafios de volta ao falante, a resposta é relativamente simples: em uma relação testemunhal, o falante assume a responsabilidade epistêmica (*epistemic responsibility*)<sup>6</sup> de enfrentar os desafios referentes à sua afirmação, e o ouvinte reconhece a suposta responsabilidade epistêmica do falante ao se colocar disposto a conceder autoridade.

Essa relação requer a cooperação entre falante e ouvinte; se essa cooperação não existe, seja porque o falante não assume a responsabilidade de enfrentar o desafio, seja porque o ouvinte não lhe concede autoridade

---

<sup>5</sup> O falante não possui a intenção de informar que p.

<sup>6</sup> Apesar de McMyler não fornecer uma definição clara do que está entendendo por *responsabilidade epistêmica*, no contexto das relações interpessoais de testemunho parece tratar-se de nada além de arcar com o ônus epistêmico. O filósofo defende que o testemunho é *sui generis* epistemológico, e, a partir disso, trata todos os conceitos pressupostos na relação interpessoal de testemunho (responsabilidade, autoridade, confiança, direito) como epistemológicos.

para fazê-lo, então o ouvinte não terá justificação testemunhal. Para McMyler (2011), essa atividade cooperativa entre falante e ouvinte é uma condição necessária para justificação testemunhal, mas é também o que lhe dá sua característica mais peculiar, que a torna diferente de outras formas de justificação.

Para esclarecer, McMyler (2011) traça uma distinção entre *contar*<sup>7</sup> algo a alguém e *meramente* dizer algo a alguém. Quando eu conto que p e marco essa declaração por expressões como “Eu acredito que p” ou “Eu penso que p”, estou assumindo a responsabilidade de responder aos desafios pelo meu ouvinte. Mas pode haver casos nos quais eu não esteja segura sobre p e não queira assumir a responsabilidade; nesses casos, posso ainda afirmar p, mas espero que meu ouvinte chegue a p por suas próprias conclusões. Ele não só não pode conceder a mim, como tampouco irá fazê-lo, se, ao investigar sobre p, chegar à conclusão de que p é verdadeiro. Assim, não se trata de justificação testemunhal. Se não existe uma atividade cooperativa, que inclui a intenção de informar, a pressuposição da responsabilidade do falante por parte do ouvinte, o falante assumindo essa responsabilidade, e o direito do ouvinte de conceder, não há justificação testemunhal.

Quando o ouvinte utiliza o ato de fala de outra pessoa apenas para contar como evidência para sua crença, mas chega às conclusões por inferência própria não se trata de justificação testemunhal. A concepção elaborada por McMyler considera que o ouvinte tem a obrigação de buscar suas próprias conclusões sobre a fidedignidade (*trustworthiness*) do falante, mas também que ele não está na posição de chegar a suas próprias conclusões sobre o conteúdo do testemunho do falante.

---

<sup>7</sup> Em inglês *telling*.

A concepção de testemunho formulada pelo filósofo envolve confiança no falante, responsabilidade do ouvinte sobre em quem confiar, responsabilidade do falante sobre o que afirma e direito do ouvinte de conceder ao falante. É uma contribuição interpessoal, que envolve responsabilidade e autoridade, e concebe o testemunho como uma capacidade epistêmica social.

No livro *Testimony, Trust and Authority* (2011), McMyler apresenta a abordagem de segunda pessoa, iniciando por uma distinção muito relevante. Ele afirma que casos de conhecimento testemunhal envolvem confiar em um falante, mas que “confiar (*trusting*) em outra pessoa envolve mais do que meramente fiar-se (*relying*) na pessoa” (McMyler, 2011: 91). Fiabilidade (*reliability*) aparece em uma perspectiva mais individualista, e é considerada o suficiente para formar uma razão epistêmica. Trata-se de uma espécie de dependência epistêmica entre os agentes e suas fontes de conhecimento. Fiabilidade (*reliability*) é o que confere *status* de evidência a algo, e fiar-se é a atitude do agente para aceitar algo como evidência. Já confiar (*trust*) implica boa vontade do falante em prol das necessidades do ouvinte, e está em uma perspectiva interpessoal.

Geralmente, quando falamos dos nossos próprios sentidos, estamos falando de fiabilidade (*reliability*), a mesma terminologia que utilizamos para coisas materiais. Fiabilidade expressa uma relação mais básica com o mundo; nos fiamos que as regras morais básicas serão cumpridas e que o mundo irá funcionar da forma adequada. Mas, caso isso não ocorra, não nos sentimos traídos, apenas decepcionados. Quando o ônibus não passa no horário, por exemplo, me sinto decepcionada, mas isso não faz com que eu deixe de acreditar no funcionamento das instituições. Com meus sentidos se passa o mesmo, me sinto decepcionada quando me engano, mas não me sinto traída. Entretanto, quando confio em alguém e essa pessoa

me engana deliberadamente, me sinto traída, pois é a violação de uma relação mais profunda.

McMyler (2011: 91) afirma que podemos nos fiar (*rely*) em alguém sem confiar (*trust*) nessa pessoa. Quando pedimos informações a completos estranhos na rua, estamos nos fiando neles. Não confiamos, não se estabelece nenhuma relação de direitos e deveres, mas, ainda assim, acreditamos quando nos indicam a direção e seguimos suas indicações. Para McMyler (2011), confiar nos outros para adquirir justificação testemunhal envolve mais do que meramente fiar-se de que as pessoas estejam falando coisas verdadeiras. Confiar envolve, pelo menos, fiar-se em que o outro diga coisas verdadeiras por sua boa vontade em relação a nós. Assim, confiar envolve também pressupor a boa vontade do outro em me auxiliar, mais do que simplesmente crer que o que o outro diz é verdadeiro.

A perspectiva de segunda pessoa considera que, em muitos casos, o ouvinte aceita as palavras do falante e as toma para si, assumindo o testemunho do falante com base em autoridade, e, nesses casos, parece que há confiança. São casos nos quais o ouvinte não busca uma conclusão própria para sua crença, e por isso são casos que merecem um tratamento diferenciado. Os adeptos desse modelo reconhecem que existe diferença entre casos em que o ouvinte apenas fia-se nas palavras do falante para encaminhá-lo à verdade e casos em que o ouvinte confia no falante e não busca suas próprias conclusões.

Nesta perspectiva, testemunho é considerado razão epistêmica genuína, no sentido amplo de algo que conta como razão para verdade ou probabilidade de uma proposição, mas não no sentido de colocar o sujeito em uma posição de chegar às suas próprias conclusões sobre a proposição. No modelo de segunda pessoa, a expressão de um falante de que  $p$  é razão para o ouvinte crer que  $p$ , mas não de modo a chegar às suas próprias conclusões sobre  $p$ .



---

O objetivo dessa proposta é exatamente fornecer uma explicação de casos de justificação nos quais o sujeito não está chegando às conclusões por si próprio, mas está confiando em um falante como guia para adquirir crenças verdadeiras. Sendo que o conceito de confiança tradicionalmente implica boa vontade, na abordagem de segunda pessoa o conceito é compreendido como fiabilidade na boa vontade do outro.

Quando confiamos no testemunho dos outros para nos guiarem para verdade, a boa vontade é compreendida em termos de responsabilidade epistêmica. Quando o ouvinte confia no falante, ele delega a este a responsabilidade de responder aos desafios sobre sua crença. Para McMyler (2011), testemunhar implica em assumir a responsabilidade sobre suas afirmações, enquanto que confiar no testemunho implica fiar-se na responsabilidade do falante.

O ouvinte tem o direito epistêmico de conceder os desafios que são colocados à sua crença de volta ao falante no qual ele confiou para adquirir a crença. Mas é importante destacar que nem todos os desafios são responsabilidade do falante. Quando questionado sobre a sinceridade ou honestidade do falante, o ouvinte não pode conceder, sendo ele próprio o responsável por responder.

O ouvinte pode conceder desafios referentes ao conteúdo do testemunho do falante, e o falante é epistemicamente responsável por respondê-los, mas o ouvinte é epistemicamente responsável por responder sobre a honestidade e a sinceridade do falante. Aqui o testemunho está diretamente conectado com o tipo de responsabilidade epistêmica assumido por ambas as partes no processo de dar e receber testemunho. Testemunho é compreendido como *razão de segunda pessoa* para crer.

McMyler (2011: 94) define testemunho como a “consideração que justifica uma crença em virtude da relação de autoridade e responsabilidade existente entre um

remetente e um destinatário”, de modo que não somente o falante tem responsabilidade epistêmica quando profere um testemunho, tendo, assim, que arcar com as objeções que possam ser endereçadas à sua afirmação, como também o ouvinte tem responsabilidade quando aceita o testemunho de um falante, pois precisa arcar com as objeções sobre a sinceridade e a honestidade desse falante. O ouvinte tem que reconhecer a honestidade e a sinceridade do falante, pressupor a boa vontade deste, e assim conceder autoridade ao testemunho, formando suas crenças com base nele.

Nossas crenças testemunhais, no modelo de segunda pessoa, são justificadas através de nossa relação com outras pessoas; assim, quando questionados sobre nossa justificação, podemos simplesmente citar o falante. Isso torna o testemunho um tipo muito específico de razão para crer, uma razão de segunda pessoa. McMyler (2011) defende que essas razões de segunda pessoa têm significância epistêmica e não apenas pragmática como muitos autores têm defendido.

McMyler (2011) distingue entre (1) *confiar em* e (2) *confiar que*. *Confiar em* é ceder a alguém a responsabilidade sobre algo que é importante para nós. *Confiar que* equivale a acreditar fortemente, com base em razões, que alguém agirá de determinado modo. Em seu exemplo, um policial que está espionando Mary durante semanas *confia que*, hoje, antes do trabalho, ela deixará as crianças na escola, porque é o que ela faz diariamente antes do trabalho. Isso é diferente de afirmar que o policial *confia em* Mary para deixar as crianças na escola. Podem haver casos em que (1) e (2) coincidam, podem haver casos em que (1) seja verdadeiro e (2) não, ou que (2) seja verdadeiro e (1) não, mas definitivamente expressam coisas diferentes. E não se trata de variações contextuais; no mesmo contexto, *confiar que* e *confiar em* expressam coisas diferentes. Em (1), a confiança é dirigida a uma pessoa, enquanto em (2) é dirigida a uma proposição.

Podem-se compreender as proposições como dirigidas à segunda e à terceira pessoa, sendo que a proposição “Confio *em* Mary para pegar as crianças hoje” representa uma atitude de segunda pessoa, enquanto “Confio *que* Mary irá pegar as crianças hoje” representa uma atitude de terceira pessoa. Quando eu confio *em* Mary, estou reconhecendo que ela possui autoridade e competência, e, assim, estabeleço com ela uma relação interpessoal de segunda pessoa. Mary torna-se responsável por fazer o que lhe confiei. Caso não fizer, a consequência será uma atitude de ressentimento de minha parte.

McMyler (2011) compreende esse tipo de relação interpessoal de segunda pessoa como *relações normativas bipolares*, porque pressupõem atitudes de ambas as partes envolvidas na relação, que funcionam como respostas em relação à atitude do outro. O mesmo não ocorre em relações de terceira pessoa, ou *confiança que*. Se eu apenas *confio que* Mary pegará às crianças na escola, eu não estou cedendo autoridade e competência à Mary; ela não é responsável por realizar a ação, e tampouco eu teria direito de me ressentir com ela por não fazê-lo.

O tipo de relação de confiança no qual os teóricos da confiança estão mais interessados é o interpessoal (Cf. McLeod, 2011), que envolve fiar-se na boa vontade do confiado, mais do que formas de confiança diretamente cognitivas. Esse tipo de relação é essencialmente interpessoal, principalmente porque envolve o risco de traição. O confiado coloca-se em uma posição vulnerável, na qual pode ser traído por aquele em quem confia, tornando-se assim dependente do outro.

Já para McMyler (2011), confiar é uma atitude cognitiva de segunda pessoa. Confiança envolve o reconhecimento do confiado pelo confiante em uma relação interpessoal, mas envolve também tomar algo como verdadeiro com o objetivo de acertar. Nesse sentido, confiar é muito semelhante a crer; a diferença está apenas

no que justifica a crença. A confiança é justificada pela relação interpessoal entre confiante e confiado. Quando confio em alguém, acredito que essa pessoa fará algo que espero que ela faça. No exemplo anterior, se confio em Mary para pegar as crianças na escola, é porque acredito que ela pegará as crianças na escola. O que justifica minha confiança é a relação interpessoal que estabeleço com Mary. Essa relação interpessoal dá suporte epistêmico à minha crença. Confiança interpessoal é crença embasada em razões fornecidas por outra pessoa.

Confiar tem sido distinguido de crer por muitos filósofos, ou ainda compreendido como crer sem razões. Na perspectiva de McMyler (2011), não se trata de crer sem razões, senão de crer com base em razão testemunhal. Quando confiamos em alguém, acreditamos no que a pessoa diz com base em seu testemunho; trata-se de uma razão de segunda pessoa para crer. Pode ser que não tenhamos nenhuma razão impessoal ao acreditar com base em testemunho, mas isso não quer dizer que não tenhamos razões. Nessa concepção de testemunho, confiança é embasada em razões de segunda pessoa e pressupõe uma dependência interpessoal irreduzível na pessoa confiada.

É importante distinguir entre o que justifica a atitude de confiar em alguém para algo e o que justifica a crença envolvida em confiar em alguém para algo. Se consideramos que a crença envolvida em confiar em alguém é justificada apenas através de nossas evidências para a fidedignidade da pessoa, então é difícil ver como isso difere de uma crença que não envolve confiança. Mas McMyler (2011) sugere que devemos conceber a crença envolvida em confiar em alguém como justificada através do relacionamento interpessoal entre falante e ouvinte. Nesse caso, a fidedignidade do falante serviria como um pano de fundo que coloca o ouvinte em uma posição de considerar o relacionamento interpessoal como capaz de gerar razões de segunda pessoa, ao menos implícitas, para

---

confiar. Confiança é, deste modo, uma atitude irredutivelmente interpessoal.

Quando confio em alguém para crer que *p*, essa pessoa assume parte da responsabilidade sobre minha crença. Eu não estou indo às minhas próprias razões sobre *p*, apenas tenho minhas próprias razões para crer na fidedignidade da pessoa. O relacionamento interpessoal que mantemos gera razões de segunda pessoa para crer que *p*. Se desafiado sobre minha crença, posso conceder o desafio a outra pessoa, que tem a responsabilidade epistêmica de dar conta dele.

Confiança é compreendida tanto como cognitiva, quanto como irredutivelmente interpessoal. É cognitiva na medida em que confiar em alguém para algo envolve acreditar na pessoa com base em razão genuinamente epistêmica. E é interpessoal na medida em que a razão que justifica a crença é irredutivelmente de segunda pessoa. Isso faz com que, ao confiar em alguém, eu não seja completamente epistemicamente responsável pela minha crença, já que posso conceder as razões para crer à pessoa em quem confiei. A pessoa em quem confio é parcialmente epistemicamente responsável pela justificação de minha crença, de modo que, ao ser desafiado sobre a crença, posso conceder as razões para crer a ela.

O modelo de segunda pessoa considera que o testemunho pode ser fonte de conhecimento e/ou justificação porque é uma prática generalizada na vida diária de todos os seres humanos graças a qual tornamos possível a distribuição e a partilha do conhecimento. Essa prática implica em um tipo de atitude cooperativa entre agentes.

Testemunho é uma relação social comunicativa, que envolve reconhecimento. Por parte do ouvinte, o reconhecimento da autoridade do falante; por parte do falante, o reconhecimento da precariedade epistêmica do ouvinte. Seria esse reconhecimento mútuo suficiente para tornar a resposta do falante uma razão epistêmica para o

ouvinte<sup>8</sup>? McMyler oferece uma boa contribuição ao salientar a presença de deveres e direitos na relação de testemunho, mas não dá conta de todas as preocupações epistemológicas. Vejamos a seguir as objeções que atingem diretamente sua proposta, bem como aquelas que se aplicam às visões interpessoais como um todo.

### **Objeções às Visões Interpessoais: A Impossibilidade de uma Analogia**

Uma objeção que se pode levantar à proposta de McMyler é a de que o falante e o ouvinte podem cumprir todos os requisitos para a relação interpessoal de testemunho sem que nada epistemicamente relevante esteja envolvido. Podemos pensar em um caso onde o falante assume a responsabilidade sobre o que ele testifica, e assim o ouvinte possui o direito de deferir. O falante é sincero<sup>9</sup>, e existe uma relação de confiança entre ambos, de modo que o ouvinte avaliou a fidedignidade do falante, e o falante tem boa vontade para com o ouvinte. Todas as características interpessoais para o conhecimento testemunhal descritas por McMyler estão preenchidas, e ainda assim o falante não transmite conhecimento testemunhal porque nada implica que seu testemunho seja favelmente conducente à verdade (Hinchman, 2012).

No último capítulo de seu livro, McMyler insere mais dois requisitos que o falante deve cumprir em uma relação testemunhal, que até então não haviam sido tratados como

---

<sup>8</sup> Esses questionamentos são fruto dos diálogos com o Professor Dr. Fernando Broncano, que suscitaram tais inquietações.

<sup>9</sup> McMyler considera que não se pode adquirir conhecimento testemunhal a partir de testemunho insincero, pois, quando há insinceridade, não há responsabilidade em relação ao que se afirma, e assim não se cumpre uma das condições necessárias para o conhecimento testemunhal tal como proposto pelo autor.

condições para o conhecimento e/ou justificação testemunhal. Além de sincero, o falante deve ser competente e fiável. Ele parece lançar mão dessas exigências ao notar que os requisitos da relação interpessoal de testemunho não são suficientes para haver transmissão de conhecimento e/ou justificação.

Até o momento, a argumentação afirmava que o ouvinte deveria *julgar* o falante como epistemicamente competente ou uma autoridade em relação à p. No capítulo 5 de *Testimony, Trust and Authority*, ele afirma que o falante não deve apenas ser *julgado* como autoridade relevante pelo ouvinte, mas *deve realmente ser* uma autoridade relevante. Passa, então, a defender como condição necessária que o falante esteja “[...] em uma posição para si mesmo gerar a razão da crença fornecida por ela ao vir e dizer ao público que p” (McMyler, 2011: 162).

Epistemólogos tendem a considerar que uma razão para crer é algo que torna uma crença provavelmente verdadeira. Não parecem dispostos a admitir que crenças podem estar justificadas, ou que possa haver transmissão de conhecimento simplesmente em decorrência da dependência epistêmica entre falante e ouvinte. Tudo indica que McMyler apela para os requisitos de competência e fiabilidade do falante ao perceber a forte objeção que pode ser direcionada à sua proposta, a saber, que as condições apresentadas para o conhecimento testemunhal podem ser cumpridas mesmo que nenhum objetivo epistêmico esteja envolvido.

No capítulo 5, ele assume que “uma consideração é evidência de uma proposição p apenas no caso em que conta a favor da verdade, confirma, ou torna provável p” (McMyler, 2011: 153). Entretanto, durante todo o desenvolvimento da proposta, os únicos pressupostos para que houvesse justificação e/ou conhecimento testemunhal eram a responsabilidade epistêmica do falante sobre a afirmação, o direito epistêmico do ouvinte deferir ao

falante, e a presença de sinceridade e confiança. Mas, mesmo que todas essas condições estivessem cumpridas, o falante poderia não ser conducente à verdade, por ter um histórico de erros em testemunhos desse tipo, por exemplo (Hinchman, 2012). Para resolver esse problema, McMyler apela aos novos requisitos de competência e fiabilidade. Se o falante for competente e fiável, ele é conducente à verdade. Anteriormente, bastaria que o ouvinte o julgasse competente e fiável, não que ele realmente o fosse.

Segundo Hinchman (2012), essa inserção de novos requisitos de competência e fiabilidade para o falante está em profundo desacordo com o modelo de segunda pessoa proposto pelo autor. Inicialmente, o modelo de segunda pessoa era sustentado pela tese de que as relações interpessoais de autoridade e responsabilidade existentes entre o falante e o ouvinte contavam em favor da proposição acreditada como razões evidenciais. McMyler (2011: 155) formula seu modelo de segunda pessoa como segue:

Assim, não só são razões testemunhais probatórias/evidenciais, não só elas genuinamente contam em favor da verdade da proposição acreditada, mas fazem isso precisamente em virtude das relações interpessoais de autoridade e responsabilidade existente entre remetente e destinatário. O sentido no qual razões testemunhais são evidenciais, não pode, assim, ser dissociado do sentido no qual elas são de segunda pessoa. O testemunho equivale à evidência precisamente em virtude de suas características de segunda pessoa, ou assim parece.

Nota-se que razões testemunhais são consideradas como equivalentes a razões evidenciais, dando conta da conducência à verdade, exatamente por suas características interpessoais. O que as torna razões de segunda pessoa as



torna evidenciais. Os conceitos de autoridade e responsabilidade cumpririam um papel evidencial. Posteriormente, aumenta-se a exigência evidencialista solicitando que o falante seja competente e fiável.

Em um primeiro momento, no qual o foco da justificação testemunhal é a relação interpessoal e não a competência do falante, é o conceito de responsabilidade que deve dar conta das exigências epistêmicas. O conceito de autoridade apenas expressa a capacidade do falante de fornecer razões, não explicando o caráter evidencial dessas razões. Assim, o que torna o testemunho uma razão para crer é a relação de responsabilidade instituída entre falante e ouvinte. Mas Hinchman (2012) questiona como essa relação de responsabilidade poderia contar em favor da verdade da proposição se o falante não for conducente à verdade. Por isso, nas últimas páginas do livro, o autor apela para as condições de competência e fiabilidade. O problema é que essas condições parecem desempenhar todo o papel de garantir que o testemunho seja adequadamente epistêmico, e toda a descrição da relação testemunhal desenvolvida no decorrer do livro parece perder importância.

McMyler tenta resolver a falha epistêmica de sua teoria afirmando que falantes que não são fiáveis e competentes não podem fornecer razões evidenciais através de testemunho. Razões evidenciais são razões conducentes à verdade. O que ocorre é que toda a caracterização da relação interpessoal de testemunho que focava na relação entre ouvinte e falante, bem como na responsabilidade do falante, não faz qualquer trabalho estritamente epistêmico.

Hinchman (2012) observa que confiança e autoridade muitas vezes estão no cerne da relação através da qual um falante torna o conhecimento disponível a um ouvinte, mas questiona se esse relacionamento interpessoal pode ele próprio explicar a relação epistêmica. A proposta de McMyler, a despeito de fornecer uma boa explicação das

relações interpessoais de testemunho, falha em apontar o caráter epistêmico dessa relação.

Thomas Simpson (2013) também salienta a falta de objetivos epistêmicos compondo a relação interpessoal de testemunho proposta por McMyler. A suposição da responsabilidade do falante pode fornecer ao ouvinte razões pragmáticas ou morais para crer, mas nada na natureza da relação interpessoal indica que ela seja conducente à verdade. A mesma objeção pode ser dirigida a todas as contribuições interpessoais de confiança, conforme salienta Lackey em *Learning from words*. Para Lackey (2008: 222) as visões interpessoais de testemunho, quando genuinamente interpessoais são epistemologicamente impotentes.

Lackey (2008) analisa as visões interpessoais de testemunho e apresenta algumas objeções cruciais a estas. Mesmo concepções não evidenciais precisam admitir que crenças justificadas por testemunho têm que ser resistentes a contraevidências. A justificação testemunhal não pode ser mantida diante da presença de contraevidência não derrotada. Mas Lackey vai argumentar que, se, como afirmam as concepções interpessoais, o valor epistêmico das crenças testemunhais é definido por certas características não evidenciais da relação interpessoal entre falante e ouvinte, não fica claro como essas características poderiam derrotar uma contraevidência para o testemunho<sup>10</sup>. Assim,

[...] se o defensor de VIT<sup>11</sup> faz um sacrifício e afirma que a razão não evidencial fornecida pelo seu testemunho não pode simplesmente funcionar como contraevidência à contraevidência, então isto equivale a uma concessão para a irrelevância

---

<sup>10</sup> Ver o exemplo do ornitologista (Lackey, 2008: 225).

<sup>11</sup> Abreviação de ‘visões interpessoais de testemunho’.

---

epistemológica das razões fornecidas por testemunho (Lackey, 2008: 226).

Se as características não evidenciais da relação interpessoal de testemunho, que justificam a crença do sujeito, não são capazes de funcionar como derrotadores de contraevidências, então elas não têm qualquer significância epistemológica, o que torna as teorias interpessoais epistemologicamente impotentes.

Ainda, a contraevidência pode ser fornecida pelo testemunho de outra pessoa<sup>12</sup> com quem o ouvinte também possui uma relação interpessoal. Nesse caso, qual testemunho irá prevalecer? Como uma relação interpessoal não evidencial poderia servir como contraevidência para outra relação interpessoal não evidencial? Por outro lado, como uma razão evidencial poderia, de algum modo, derrotar crenças testemunhais não evidenciais? Lackey (2008: 226) pontua a dificuldade inerente às visões interpessoais não evidenciais:

[...] se as razões apresentadas pelo testemunho ficam inteiramente fora do reino das evidências, como todos os proponentes da VIT mantêm, então esse ponto de vista enfrenta o difícil desafio de explicar a relação entre tais razões não evidenciais e as contraevidências que são capazes de derrotar tais razões.

McMyler (2011) admite que a responsabilidade do falante dá ao ouvinte razões para acreditar no testemunho. Entretanto, Lackey (2008) questiona que assumir responsabilidade possa ter qualquer valor epistêmico, porque não é conducente à verdade. Pode ter valor moral ou pragmático, mas não epistêmico. Um defensor da concepção de segunda pessoa poderia argumentar que a

---

<sup>12</sup> Tal como no exemplo do ornitologista formulado por Lackey.

suposição de responsabilidade do falante pelo ouvinte o torna epistemicamente responsável, não culpável ou racional ao crer que *p*. Entretanto, isso é um erro, pois o fato de o falante assumir a responsabilidade sobre a verdade de *p* não torna o ouvinte epistemicamente responsável. O fato de que um mentiroso compulsivo assume a responsabilidade pela verdade de uma afirmação não me torna não epistemicamente censurável ou epistemicamente responsável ao acreditar em seu testemunho (Lackey, 2008: 229).

Assumir a responsabilidade sobre a verdade das afirmações só possui valor epistêmico quando o falante o faz de forma apropriadamente epistêmica, sobre circunstâncias epistemicamente apropriadas. Por exemplo, eu posso assumir a responsabilidade sobre a afirmação de que me lembro de que, aos três anos, sofri um acidente de carro, mas minha memória não é uma fonte epistêmica fiável nesse caso, mesmo que eu faça essa afirmação de forma sincera, com boa vontade e assumindo a responsabilidade. “Por isso, *é somente quando um falante está se responsabilizando pela verdade de sua asserção, que ela tem a capacidade de conferir valor epistêmico sobre crenças testemunhais*” (Lackey, 2008: 229, grifo da autora). Ou a responsabilidade assumida pelo falante é conectada com a verdade (não vem ao caso de que modo, desde que seja), ou ela não é propriamente epistêmica, e, se esse for o caso, não é do interesse da Epistemologia do Testemunho.

Deste modo, embora genuinamente interpessoal, o modelo de segunda pessoa é epistemologicamente impotente, porque, da forma como é proposto, mesmo na ausência de condições epistêmicas no falante e/ou no ouvinte, a relação interpessoal de testemunho seria capaz de transferir justificação. Podemos supor um caso em que o falante seja completamente não fiável para aqueles casos de testemunho, mas, se ele assumir a responsabilidade, for

sincero e possuir boa-vontade, as condições para justificação testemunhal se cumprem – ou um caso no qual o ouvinte possa adquirir crença justificada mesmo diante de contraevidência relevante.

Lackey (2008: 235) apresenta, ainda, outro exemplo que atinge as visões interpessoais de testemunho e consequentemente o modelo de segunda pessoa. Suponhamos um contexto no qual duas pessoas, João e Marcelo, estão conversando, e João *conta* a Marcelo que seu chefe está tendo um caso com uma funcionária. Uma terceira pessoa escuta a afirmação, que não foi dirigida a ela. Esse “intruso” não faz parte do relacionamento interpessoal, não há boa vontade do falante para com ele, não há responsabilidade e, por isso, não há direito do ouvinte de deferir. Entretanto, não parece haver nenhuma diferença epistemológica significativamente relevante entre Marcelo, que participa da relação interpessoal, e o intruso. Isso significa que não há nenhuma diferença epistemológica relevante entre *contar* e dizer, pois ambos estão ouvindo a mesma afirmação com o mesmo grau de fiabilidade e a mesma condutibilidade à verdade.

Os dois também possuem a mesma base para a fiabilidade do falante, todos trabalham juntos no mesmo escritório; desse modo, não é mais racional para um do que para outro crer no testemunho, pelo simples fato de que ele foi direcionado a Marcelo e não ao “intruso”. Lackey (2008: 236) ainda varia o exemplo, supondo um caso no qual Marcelo pudesse ter com João uma amizade desde a infância, o que forneceria mais razões para confiar no testemunho. Mas, segundo a filósofa, isso não serviria para suportar as contribuições interpessoais de testemunho, porque mostraria apenas que Marcelo tem mais razões para crer no testemunho do que o intruso que escuta atrás da porta, não que as características do relacionamento podem afetar o valor epistêmico do testemunho. Isso mostra apenas que um ouvinte que possui mais informações sobre

a fidedignidade do falante pode adquirir crença testemunhal mais justificada do que aquele que não possui essas informações, não que a relação interpessoal gera mais justificção.

A diferença estabelecida entre *contar* e dizer é que, quando alguém *meramente* ouve outra pessoa dizer que *p*, ela não pode conceder ao falante que *p*; já quando o falante *conta* ao ouvinte que *p*, ele assume a responsabilidade, o que dá ao ouvinte direito de conceder desafios sobre *p* ao falante. Mas Lackey afirma que essa diferença não apresenta nenhuma significância epistemológica relevante para o conhecimento testemunhal. Ela questiona ainda se as relações testemunhais são estabelecidas dessa maneira.

Suponha que João diz a Marcelo que a irmã de Ana tem problemas com abuso de álcool, e Ana escuta essa afirmação. Sendo que se trata de sua irmã, parece que Ana deveria ter o direito de exigir a responsabilidade de João sobre a afirmação. Mas, como a afirmação foi dirigida a Marcelo, conforme as visões interpessoais apenas ele poderia exigir responsabilidade. Contudo, se ambos ouviram a mesma afirmação, proferida pela mesma pessoa, por que apenas Marcelo teria o direito de conceder os desafios e exigir a responsabilidade de João? Considerando que a afirmação diz respeito à irmã de Ana, parece que ela tem ainda mais direito de exigir que João se responsabilize sobre o que diz<sup>13</sup>.

Características interpessoais não acrescentam valor epistêmico às crenças testemunhais para além do valor epistêmico que elas já possuem por sua condutibilidade à verdade; isso torna as concepções interpessoais de testemunho epistemologicamente impotentes. As características que atribuem valor epistêmico às crenças independem das relações interpessoais, e isso pode ser notado nos exemplos apresentados por Lackey. Podemos

---

<sup>13</sup> Exemplo da Lackey (2008: 236).

---

retirar todas as características da relação interpessoal, mantendo somente a condutibilidade à verdade, e teremos o mesmo resultado em termos de justificação testemunhal.

Assim, quando as contribuições interpessoais buscam inserir condições epistêmicas na relação testemunhal, tal como McMyler (2011) no capítulo 5 insere as condições de competência e fiabilidade, de modo a tornar a crença testemunhal conducente à verdade, elas acabam por perder suas características essencialmente interpessoais. As características evidencialistas introduzidas por McMyler acabam por desempenhar todo o papel epistêmico, fazendo com que as características interpessoais careçam de relevância na relação testemunhal. Por isso, Lackey (2008) afirma que ou as contribuições interpessoais são epistemicamente impotentes e genuinamente interpessoais, ou genuinamente epistêmicas mas não interpessoais. A inserção de condições evidencialistas, que trazem para a relação testemunhal relevância epistêmica, torna as exigências interpessoais dispensáveis.

Contribuições interpessoais fazem boas caracterizações da relação de testemunho, entretanto não levantam considerações úteis para a Epistemologia do Testemunho, por carecerem de relevância epistêmica. Não é possível dar conta desse problema sem comprometer a abordagem interpessoal e, nesse caso, não oferecer nenhuma posição nova para o debate em Epistemologia do Testemunho.

Axel Gelfert (2013) considera que *Testimony, Trust and Authority* amplia o debate sobre a importância das relações interpessoais no conhecimento testemunhal, além de possibilitar reflexões acerca da dimensão ética de nossa dependência em outras pessoas para obter conhecimento. McMyler define confiança do modo como foi

compreendida na Ética<sup>14</sup>, realizando, desse modo, uma analogia conceitual, transpondo o conceito da perspectiva ética para uma perspectiva epistemológica. O filósofo realiza, assim, algo bastante inovador no debate sobre a confiança no testemunho de outras pessoas, tentando unir a dimensão epistêmica e a dimensão ética das relações testemunhais. Entretanto, não é possível notar como essas considerações de âmbito moral ampliam ou contribuem diretamente na discussão epistêmica sobre testemunho.

A analogia falha. Confiança em Ética pode ser definida como *fiar-se na boa vontade de outra pessoa para fazer p*. Transposto para Epistemologia, seria *fiar-se na boa vontade de outra pessoa para crer que p*. Concepções interpessoais têm assumido esse conceito e tentado aplicá-lo em Epistemologia do Testemunho. O problema é que a boa vontade da outra pessoa não otimiza fins epistêmicos, e sim fins morais. Um defensor da concepção interpessoal poderia tentar defender que *fiar-se na boa vontade* pode otimizar fins epistêmicos, mas qual o papel da boa vontade em crer em verdades e evitar falsidades?

Um falante pode ter boa vontade mas não ser competente no assunto do qual trata, não otimizando fins epistêmicos. O defensor da concepção interpessoal poderia inserir a exigência de competência. Mas, de qualquer forma, a boa vontade continuaria a não desempenhar papel epistêmico e poderia ser descartada, o que possibilitaria uma economia de conceitos no debate, preservando apenas um conceito já consagrado na tradição epistemológica – *fiar-se*.

O conceito de confiança é rico e esclarecedor quando tratamos de relações interpessoais, e é verdade que o testemunho é uma relação interpessoal. Entretanto, confiança não carrega em si conducência à verdade, para

---

<sup>14</sup> O conceito de confiança utilizado pelo autor é o mesmo proposto por Anette Baier.



otimizar fins epistêmicos; deste modo, não parece contribuir para o debate em Epistemologia. As tentativas de inserir no conceito condições que a tornem epistêmica fracassam, pois, inseridas as condições, dispensam o viés interpessoal descaracterizando o conceito.

## Referências

- Baier, Annette C. (1986). 'Trust and Antitrust', *Ethics*, 96(2): 231–260.
- Firth, Roderick (1978). 'Are Epistemic Concepts Reducible to Ethical Concepts?'. In: Goldman, Alvin; Kim, Jaegwon (Eds.). *Values and Morals: Essays in Honor of William Frankena, Charles Stevenson, and Richard Brandt*, Springer Science & Business Media, Dordrecht: Reidel, 215-229.
- Gelfert, Axel (2013). 'Review of Benjamin McMyler: Testimony, Trust and Authority', *Journal of Applied Philosophy* 30(1): 101-103.
- Hinchman, Edward S. (2012). *Review Testimony, Trust, and Authority*. Publicado na internet, disponível através da seguinte URL: <<http://ndpr.nd.edu/news/29388-testimony-trust-and-authority/>>.
- Lackey, Jennifer (2008). *Learning from Words: Testimony as a Source of Knowledge*, Oxford: Oxford University Press.
- Locke, John (1975). *An Essay concerning Human Understanding*, Oxford: Clarendon Press.
- McLeod, Carolyn (2015). 'Trust'. In: Zalta, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL:

<<http://plato.stanford.edu/archives/spr2011/entries/trust/>>.

McMyler, Benjamin (2011). *Testimony, Trust and Authority*. Oxford: Oxford University Press.

Simpson, Thomas (2013). 'Review of Benjamin McMyler, 'Testimony, Trust, and Authority' and Paul Faulkner 'Knowledge on Trust'', *Mind*, 122(485): 305-311.

DEONTOLOGISMO  
EPISTÊMICO, CETICISMO E  
EPISTEMOLOGIA MORAL,  
AUTOCONHECIMENTO E  
FALIBILISMO  
INTROSPECTIVO: UMA  
TENTATIVA DE  
APROXIMAÇÃO

---

*Ricardo Rangel Guimarães<sup>1</sup>*

O Deontologismo Epistêmico pode, de uma maneira bastante resumida e superficial, ser definido, em parte, e entendido, minimamente, como uma espécie de justificação internalista em Epistemologia ou Teoria do Conhecimento: de maneira geral, uma justificação dita deontológico-internalista defende que um sujeito cognoscente S tem o direito de dar assentimento apenas às proposições para as quais o mesmo dispõe de uma boa garantia ou aval epistêmico explícito (“*warrant*”)<sup>1</sup>. De forma análoga, tal garantia poderia também ser interpretada como um critério epistêmico que permitiria a S eliminar o componente da dúvida ao crer em uma proposição P qualquer<sup>2</sup>. A dúvida

---

<sup>1</sup> Email: rirangel@gmail.com

<sup>1</sup> Conforme salienta, por exemplo, Alvin Plantinga (1993), na sua epistemologia, onde este conceito de “*warrant*” é elemento chave para uma compreensão mínima da mesma.

<sup>2</sup> René Descartes, em “Meditações Metafísicas”, por exemplo, utiliza esse critério após introduzir no seu sistema filosófico o método da

hiperbólica, metódica e radical, em uma perspectiva cartesiana, por exemplo, tem aplicação na investigação científica e do senso comum, do assim denominado “saber ordinário”, mas, nas nossas práticas e condutas, em relação a determinadas crenças que formamos, do ponto vista interno a evidência que as justifica, digamos assim, é, até certo ponto, suficiente para não suspendermos o juízo sobre as mesmas, bem como também sobre alguns dados externos imediatos que recebemos através dos sentidos, que, como Descartes (1991 [1641]: 179) coloca, “seria ridículo colocá-los em dúvida”. Tal atitude também seria doxasticamente válida e racionalmente defensável para outras crenças mais fundamentais, tanto da perspectiva de adquirir conhecimento certo e indubitável, quanto para determinadas atitudes proposicionais que norteiam e balizam certas ações e comportamentos básicos. Se nós deixássemos de crer ou suspendêssemos o juízo acerca das mesmas, correríamos o risco de pormos nossas próprias vidas em risco, o que seria uma conduta passível de alguma repreensão moral e de irresponsabilidade epistêmica, no sentido de não se crer responsabilmente, vindo, com isso, a abalar e violar certas regras básicas do deontologismo e da justificação internalista que tal concepção carrega consigo.

Ao adquirirmos crenças acerca do mundo exterior, se pretendemos ter conhecimento dos objetos deste mundo, uma das exigências é que estas crenças sejam justificadas epistemicamente, e que haja um componente de irrepreensibilidade nessa justificação. Para os objetos externos à mente, se há uma correspondência da verdade desse objeto com a crença justificada no mesmo, podemos

---

dúvida hiperbólica, onde o sujeito começa por duvidar de tudo para, no prosseguimento da ordem das razões, dirimir muitas dúvidas que haveria no início de sua investigação. Certos pensamentos e crenças, entretanto, na metafísica/teoria do conhecimento cartesianas não são passíveis de dúvida metódica e racional, sendo o mais indubitável de todos o famoso *cogito ergo sum* (“penso, logo existo”).

---

dizer que conhecemos o objeto, em um paradigma epistemológico de conhecimento como crença verdadeira justificada. O mesmo raciocínio poderia ser feito em relação às crenças da nossa vida mental, crenças introspectivas: na natureza da justificação internalista de uma crença introspectiva, poder-se-ia falar em um componente deontológico desta justificação, em que entram em jogo no ato de crer componentes éticos de responsabilidade moral, como deveres e direitos epistêmicos, por exemplo, bem como os conceitos de “obrigação”, “permissão”, “requerimento”, dentre outros semelhantes, conforme chama a atenção William Alston (1989: 115)<sup>4</sup>.

Alvin Plantinga (1993: 13-14), por sua vez, chega a afirmar que “toda a noção de justificação epistêmica tem a sua origem e residência nesse território deontológico do dever e da permissão, e é somente por meio da extensão análoga que o termo ‘justificação epistêmica’ é aplicado em outros sentidos”.

A normatividade e a objetividade e subjetividade da justificação é um problema que o Deontologismo Epistêmico procura um ponto de equilíbrio: alguns filósofos identificam a noção deontológica com justificação epistêmica subjetiva, enquanto outros procuram fornecer um caráter de objetividade a esta justificação. No primeiro caso, a subjetividade não parece oferecer uma excelência epistêmica normativa que o conceito de justificação exigiria para uma crença ser um caso de conhecimento, ao passo que, no segundo, o risco seria pela perda de um caráter supostamente irrepreensível no seguir uma regra com vistas a uma condução à verdade; por exemplo, que tal

---

<sup>4</sup> Alguns epistemólogos contemporâneos internalistas são partidários e defensores do deontologismo epistêmico, como por exemplo Laurence Bonjour, Roderick Chisholm, Carl Ginet e Matthias Steup, dentre outros.

justificação deveria possuir como norma epistêmica adequada.

A discussão sobre essa distinção, bem como sobre a normatividade da justificação no que diz respeito a tal traço deontológico, que exigiria um tratamento de princípios éticos neste contexto, é extremamente fundamental dentro do domínio investigativo proposto neste trabalho, e sua relevância e importância não é menor do que todas as outras questões aqui expostas para análise. Uma das pretensões aqui presentes é abordar a relação do caráter deontológico da justificação epistêmica internalista com a natureza das crenças introspectivas: a questão de se um sujeito cognoscente S pode e tem o direito de exercer alguma espécie de ceticismo, particularmente um ceticismo de natureza moral em relação a tais crenças, ou, pelo menos algumas delas, que, em alguma medida podem ser consideradas básicas, no sentido de poderem ser, em alguma medida, fundamento para outras crenças, é uma das questões a serem incansavelmente perseguidas e minimamente tratadas aqui.

Em vista deste fim, faz-se alusão novamente a Descartes, cuja distinção entre o domínio da conduta da vida e da investigação científica na busca da verdade funcionará como pano de fundo e baliza para a presente análise. Segundo o filósofo racionalista, a razão pela qual a dúvida metódica aplica-se legitimamente ao domínio da investigação científica e não da conduta da vida é que as verdades da ciência, ancoradas no conhecimento matemático das verdades necessárias, parecem, em alguma medida, serem certas e indubitáveis, ao passo que, nas nossas ações relativas às condutas da vida, as “certezas absolutas”, além de serem contingentes, podem também se mostrarem perniciosas e danosas em muitos casos. Faz-se a referência, aqui, a um exemplo dado pelo próprio autor em uma carta sua de agosto de 1641 como resposta a um correspondente desconhecido, que ficou conhecido como

---

“*Hyperaspistes*”, bem como a um correlato a ser desenvolvido na sequência. O exemplo é da abstinência da ingestão de um alimento supostamente envenenado, que alguém evita comer em detrimento de dar cabo de sua própria vida. Segundo Descartes,

Seria desejável tanta certeza nas coisas que dizem respeito à conduta da vida quanto é requerido para adquirir a ciência; mas é muito fácil demonstrar que não se deve buscar nem esperar uma tão grande. [...] Como, por exemplo, se alguém quisesse se abster inteiramente de ingerir algum alimento por tanto tempo que morresse de fome, sob o pretexto de que não estava certo que não havia veneno algum misturado, e acreditasse não estar obrigado a comer, por não saber clara e distintamente que tinha ante de si do que sustentar sua vida, e que mais vale esperar a morte abstendo-se de comer do que matar a si mesmo ingerindo tais alimentos. Certamente esse sujeito deveria ser qualificado de louco e acusado de ser o autor da sua morte. E se ao contrário supuséssemos que esse homem só pudesse ter alimentos envenenados, os quais, todavia, não os pareceria como tais, mas ao contrário muito saudáveis; e supuséssemos também que ele recebeu tal temperamento da natureza, que a completa abstinência serve à sua saúde, ainda que lhe pareça que ele não deva se alimentar menos do que os outros homens, é certo, apesar disso, que esse homem será obrigado a usar esses alimentos, e assim fazer antes o que parece útil do que o que o é efetivamente. E isso é de si tão manifesto que me espanto que o contrário possa ter vindo ao espírito de alguém (Alquié, 1996: 359-360).

Nessa passagem, podemos constatar que existem, para Descartes, diferentes tipos de certeza relativos aos diferentes domínios, tanto o da investigação da verdade

quanto o da vida: enquanto o primeiro requer a indubitabilidade das nossas crenças, o segundo, não, onde o guia motor aparentemente é a utilidade das nossas ações. Traz-se à baila este exemplo não tanto para se discutir esta distinção entre investigação da verdade e conduta da vida, mas um ponto correlacionado, a saber: que seria uma possível dúvida que este sujeito poderia ter em relação à sua própria crença de estar sentido fome, por exemplo, resultante de seu desejo apetitivo. Esta espécie de ceticismo introspectivo, com certo cunho moral subjacente, haja vista a questão essencial aqui ser de natureza ética, é acerca de uma crença da própria vida mental do sujeito cognoscente S, e este ponto pode ser relacionado com a questão do autoconhecimento e a incompatibilidade que este teria com uma perspectiva externalista acerca do conteúdo mental, questão esta já levantada por Paul Boghossian (1989)<sup>7</sup>, por exemplo, em que poderia admitir-se a tese absurda da perda da autoridade da primeira pessoa ao postular que uma outra mente extra-epidérmica a S poderia saber mais sobre os próprios pensamentos do que a mente que os habita, na autoridade desta (o famoso experimento mental de pensamento da Terra-Gêmea de Hilary Putnam ilustra a problemática; tal perda da autoridade da primeira pessoa também é uma questão que suscita desdobramentos éticos relevantes).

Embora esta questão da compatibilização de uma perspectiva anti-individualista (ou externalista) acerca do conteúdo mental com o conhecimento dos próprios estados mentais não vá ser diretamente tratada aqui, haja vista os tópicos de pesquisa centrarem-se na perspectiva da possibilidade deste conhecimento, tal discussão é relevante no que tange à questão do chamado autoengano, por exemplo; ponto este sobre o qual um breve

---

<sup>7</sup> Questão esta especialmente analisada em seu artigo ‘Content and Self-Knowledge’ (1989).



desenvolvimento faz-se pertinente, mesmo que tangencialmente, a fim de delinear um pouco as razões pelas quais podemos e/ou devemos crer irracionalmente, e também se podemos e temos alguma espécie de autoridade moral para duvidar de algumas de nossas crenças introspectivas. Na sequência, buscar-se-á relacionar minimamente tais questões com o aspecto deontológico da justificação epistêmica e da epistemologia moral de um modo geral, na tentativa de encontrar conexões entre estes tópicos de pesquisa aqui expostos.

O autoengano está presente na vida mental de um sujeito cognoscente S quando tal sujeito dá assentimento (assentimento este dado com determinada motivação para crer) a uma ou mais crenças que, de acordo com a sua própria normatividade racional, são falsas ou infundadas. Tal fenômeno possui uma natureza psicológica, e caracteriza-se por uma falha de racionalidade de S segundo alguma motivação, motivação esta que o fez crer de forma irracional. A gênese de tais crenças irracionais pode ser oriunda da fantasia ou da “imaginação indisciplinada”, conforme observa, por exemplo, Robert Audi (1998: 55)<sup>8</sup>.

O que motiva o interesse pelo problema do autoengano no contexto da discussão presente diz respeito a uma questão fundamental, a saber: como pode um sujeito crer contrariamente às evidências que possui?<sup>9</sup> No caso, tal

---

<sup>8</sup> O exemplo de Audi refere-se a uma crença retida sobre o passado que não possui uma natureza puramente memorial: não vejo determinada pessoa há um ano, mas creio irracionalmente tê-la visto há cerca de um mês usando o mesmo cinto que está usando no instante quando a encontro e percebo presentemente a presença de tal cinto.

<sup>9</sup> Há muitas definições de evidência em epistemologia realizadas por vários autores. É importante referirmos aqui alguns deles, pois o conceito de evidência é fundamental e muito utilizado no contexto da teoria do conhecimento como um todo. Feldman (2003), por exemplo, fala que a posse de evidência é a marca de uma crença justificada, e a isto ele denomina ‘teoria evidencialista da justificação’, ou ‘evidencialismo’. Já Chisholm (1974: 18) afirma que ‘[...] *evidência*

sujeito terá outra crença que não a suportada pela evidência de que ele dispõe poderia (e deveria também?) causar em sua vida mental (supondo-se, nesse tratamento, uma teoria da causação de crenças pelas evidências disponíveis), uma crença contrária às suas evidências<sup>10</sup>. O ceticismo moral introspectivo implicaria na suspensão do juízo da crença introspectiva, não se seguindo desta atitude proposicional de dúvida uma crença contrária à pretensa crença racionalmente fundamentada em questão. Mas, uma vez que eu suspenda o juízo acerca de uma crença racional, posso crer irracionalmente na crença contrária: por exemplo, S pode estar acometido por uma doença terminal, e aparentemente recusa-se a crer nisso. S tem tal crença, e a mesma é racional, mas tem dúvida sobre esta racionalidade, e suspende o juízo sobre a mesma. Embora não seja uma atitude necessária em decorrência desta sua atitude cética, S pode passar a crer irracionalmente que não está doente, recaindo assim no autoengano, equívoco cognitivo este originado, em alguma medida, pela suspensão do juízo da crença racional que teve antes, a de estar doente: a dúvida cética de sua crença racional pode o levar a crer irracionalmente no contrário do que deveria crer nesse caso, de que estava doente.

Crer falsamente não é uma atitude irracional, e não é desta atitude que se está falando aqui: errar no sentido de nossas crenças não atingirem a marca da verdade é algo normal e frequente, sendo muitas vezes o caso. O que estaria vedado epistemicamente seria a possibilidade de crer

---

*adequada* é aquela que, quando adicionada à opinião verdadeira, produz conhecimento'. A evidência, em Chisholm, pode ser entendida em termos do quão razoável é para S crer na proposição P.

<sup>10</sup> O que se denomina de 'subdeterminação' em epistemologia tem como princípio que E é boa evidência para a crença de que P se e somente se E não for boa evidência para a crença de que não P; algo necessário para a admissão de alguma espécie de ceticismo, por exemplo.

sem assentir uma conducência à verdade inerente ao próprio ato de crer, algo como o paradoxo de Moore, do tipo “É o caso que P, mas eu não creio que P”, ou não seguir, por exemplo, o objetivo epistêmico norteador da administração de nossos sistemas crenças de William James (1897), que estabelece como critério normativo a regra “crer verazmente e não crer falsamente”. Também não se trata de alguma espécie de *wishful thinking*, de crer naquilo que se deseja que seja ou que venha a ser o caso, porque, no pensamento desejoso, não está pressuposto alguma espécie de conflito com o conjunto de evidências disponíveis para se crer. Sob uma perspectiva epistemológica, psicológica e, sobretudo, ética, no sentido de se formular um juízo moral acerca de questões desta natureza, o fenômeno do autoengano pode ser passível de censura pela presença de uma justificação contrária para se crer, em detrimento da suposta justificação racional que uma crença bem formada, digamos assim, teria. O mesmo questionamento desta censura poderia ser aplicado não para esta crença contrária à evidência em questão, o que caracterizaria propriamente o autoengano, mas a ausência de evidência e racionalidade para crer justificadamente ou estar justificado em crer: assim como S pode crer contrariamente, por exemplo, de que sente fome (S crê que não sente fome), quando de fato sente fome, se autoenganando, S poderia nem crer nem descrever (ou não crer) de que sente fome, exercendo aqui alguma forma de atitude cética introspectiva? Em ambos os casos, independente da atitude a ser tomada, a consequência trágica final seria a morte por inanição, se S abstivesse-se de ingerir alimentos.

Nosso senso moral atestaria que, numa perspectiva epistêmico-deontológica, tanto em um caso como em outro, S teria o dever de crer que sente fome, ou de que está doente e precisa de tratamento adequado, no caso do outro exemplo, e não tem o direito de crer contrariamente

nisso, não podendo dar-se ao luxo de ser irracional a ponto de pôr a própria vida em risco, e também não tendo o direito a ser cético nesses casos, duvidando de tais crenças introspectivas. Nesse ponto, o autoengano e o ceticismo moral introspectivo nos conduziriam a resultados equivalentes, e, no domínio modal das possibilidades, o desconforto de uma suposta falibilidade acerca das nossas crenças introspectivas não seria menor ou menos oneroso do que a possibilidade do autoengano, a que todos podem ser suscetíveis, e isso não ter relação alguma com alguma espécie de defeito cognitivo ou de ausência de racionalidade por parte do agente.

Ao contrário da perspectiva anti-individualista acerca do conteúdo mental e da questão da incompatibilidade de um autoconhecimento básico, como observou Tyler Burge (1993), com esta perspectiva, em que teríamos de recorrer ao mundo exterior e aos objetos empíricos para conhecer os nossos próprios pensamentos e nossas crenças, em vista de perder a autoridade da primeira pessoa e destes pensamentos, de a mente conhecer o mundo e ter acesso cognitivo a outras mentes (no que Wittgenstein já rotulava como sendo o “problema das outras mentes”), a pergunta fundamental aqui é: que condições necessárias e suficientes devem ser satisfeitas para o conhecimento da própria mente e dos seus estados mentais e suas crenças introspectivas. Tal indagação é complexa e extrapola, em certa medida, o objetivo delimitador desta pesquisa em termos da resolução de um problema filosófico que se mostra quase que insolúvel: a intenção não é pretender dar conta deste tópico tão amplo e espinhoso por si só, mas sim apontar e relevar esta questão no âmbito da investigação presente, e tê-la também como balizadora de uma discussão com os direitos e deveres epistêmicos nesta busca por esta relação que é um dos objetivos primários e fundamentais a serem perseguidos aqui.

---

A conexão destas questões com o Deontologismo Epistêmico e com a Epistemologia Moral, como estão colocadas neste ensaio, é complexa e nada trivial, e uma discussão mais aprofundada sobre a justificação internalista faz-se pertinente e mais do que necessária neste contexto a fim de lançar luz a esta questão. Em Epistemologia, a justificação doxástica forma-se juntamente com a crença: no momento em que S crê que P, S está justificado em crer que P (da distinção relevante a ser feita entre ter justificação para crer que P e crer justificadamente que P, uma proposição qualquer: a terceira condição da Análise Tradicional do Conhecimento (ATC) trata da justificação para crer que P <sup>12</sup>). Tal justificação é uma justificação relativa a um determinado conteúdo proposicional, no sentido de ser uma evidência justificadora para crer, noção de evidência para se estar justificado a crer mais forte que o conceito de evidência de Chisholm (1974), por exemplo, que exige um determinado grau de presunção para se estar justificado. Mas nem sempre a justificação é formada por crenças: dados dos sentidos, por exemplo, fatos do mundo, ou mesmo fatos da vida mental de um sujeito cognoscente S podem causar crenças, e o conteúdo mental nesses casos seria não proposicional. Como estamos interessados em discutir e investigar as crenças introspectivas, faz-se pertinente perguntar pela natureza e origem da justificação introspectiva: esta teria como causa as crenças introspectivas, ou a mesma seria formada por estes fatos da

---

<sup>12</sup> Conforme Richard Feldman (2003), por exemplo, que desenvolve e expõe a noção canônica de conhecimento proposicional como crença verdadeira justificada. Não apenas este autor, mas praticamente toda a tradição da epistemologia analítica contemporânea considera como condições necessárias e suficientes para o conhecimento proposicional por parte de um sujeito cognoscente qualquer S que S tenha crença verdadeira justificada, e mais ainda, que esta crença justificada não seja acidentalmente verdadeira a fim de a mesma ser uma forte candidata a ser um caso de conhecimento proposicional.

vida mental de S? Quanto ao caráter da inferencialidade, a justificação introspectiva parece não ter uma natureza inferencial: se sinto dor, fome, ou tenho alguma espécie de *wishful thinking*, por exemplo, a justificação em crer nisso não parece ter sido formada através de um processo inferencial em minha mente, e intuitivamente resulta ser mais razoável conceber um processo direto e imediato na formação desta justificação, sem a intermediação de nenhum outro componente cognitivo.

Alguns epistemólogos internalistas, que, no geral, requerem alguma espécie de razões para justificarem suas crenças de que P, além das evidências já referidas, admitem um fundacionismo no seu sistema de crenças (em oposição ao infinitismo e diferentemente também da concepção coerentista canônica), dentre as quais há algumas crenças consideradas básicas: se a justificação é internalista, e se existem algumas destas crenças básicas no sistema para um sujeito S, então todas as outras crenças são causadas por estas crenças, bem como as respectivas justificações para se crerem nas mesmas<sup>13</sup>. No caso das crenças introspectivas, podemos tratar de uma espécie de fundacionismo, e determinar para esse sujeito S formador de crenças quais são as suas crenças básicas e quais não são, bem como a justificação para tais crenças: tal procedimento leva a uma consequência lógica de infalibilidade acerca destas crenças básicas e da justificação para as mesmas, implicando, com isso, o fato de que não posso enganar-me acerca dos conteúdos de meus pensamentos e estados mentais, e sobre estas crenças não sou passível de atribuir-lhes alguma espécie de falibilidade epistêmica e erro cognitivo. A

---

<sup>13</sup> Cabe aqui ressaltar o exemplo clássico de Descartes, um internalista e fundacionista clássico, a ponto de estabelecer em sua teoria do conhecimento uma crença básica da qual todas as outras crenças dependeriam causalmente, a saber, o “cogito ergo sum” (“penso, logo existo”).

---

introspecção produz crenças justificadas acerca da vida mental do sujeito S numa analogia com a percepção, por exemplo: enquanto através desta temos contato com o mundo exterior recebendo do entorno dados dos sentidos (*sense datum*), a introspecção funciona como uma espécie de percepção “interna”, que autodireciona crenças para nossa vida mental e fornece justificação epistêmica para estas poderem, inclusive, tornarem-se conhecimento. Para ilustrar e trazer alguma luz a estas questões, cabe trazer à baila uma noção de Robert Audi, por exemplo, sobre crenças serem classificadas como estados mentais disposicionais, e não necessariamente ocorrentes e conscientes: não é requerido de S que este pense auto-reflexivamente em P para, deste processo mental ocorrente e consciente, seguir-se uma atitude proposicional de crença, a saber, a crença de que P. Para Audi (1998: 75), ter uma propriedade mental disposicional ou estar em estado mental disposicional significa estar ou tender a “fazer ou submeter-se a alguma coisa sob certas condições, mas não necessariamente estar realmente fazendo, submetendo-se ou experienciando alguma coisa ou mudando de alguma maneira”, do que se segue disto que se poderia crer não reflexivamente. Para Audi (1998: 74), pensar é um processo que envolve uma sequência de eventos na mente de um sujeito S<sup>15</sup>, diferente de imaginar algo, que pode tanto ser um processo, se envolver uma sucessão de imagens, quanto um estado mental, onde a imagem permanece inalterada. Nem todo imaginar é um processo, pois alguns destes são estáticos, e processos mentais ocorrentes, como o pensamento, por exemplo, são chamados de *propriedades de processo-experienciais*, e propriedades mentais ocorrentes com imagens mentais estáticas na mente designam-se *propriedades de estado-experienciais*, onde ambas distinguem-se das

---

<sup>15</sup> “Thinking is a kind of *process* and involves a sequence of events [...]”.

propriedades disposicionais<sup>16</sup>. Estas distinções audianas são relevantes para um entendimento mínimo da introspecção enquanto faculdade cognoscente, pois a analogia que o mesmo faz com a percepção e do pensamento introspectivo ser esta espécie de “percepção interior” é válida enquanto a crença introspectiva for imagética e ter associada a si uma representação mental que utilize o recurso pictórico, de associação da imagem mental com o objeto percebido. No caso da percepção, o objeto é externo à mente, e o que causa a crença perceptual ou é este objeto ou os dados dos sentidos, que são os *inputs* que os sujeitos cognoscentes recebem, são as informações sensoriais que o aparato cognitivo processa. Nesse ponto, a analogia não resulta nada óbvia em relação à introspecção, não sendo também para a própria percepção: há uma espécie de zona cinzenta na definição e delimitação do objeto da introspecção, e disto resultam muitas dificuldades para uma epistemologia da introspecção que identifique onde se formam as crenças introspectivas, bem como a justificação para estas crenças, e a possibilidade do conhecimento introspectivo, o autoconhecimento, ou conhecimento de segunda ordem, segundo Tyler Burge, por exemplo, que utiliza esta nomenclatura a fim de oferecer uma solução para o famigerado clássico problema da incompatibilidade do anti-individualismo acerca do conteúdo mental, com a autoridade da primeira pessoa em relação aos seus próprios pensamentos e crenças. Como na percepção, não há um realismo direto que associe a experiência sensorial com a imagem mental do objeto do mundo a ser “copiado” do mundo, digamos assim: na introspecção, é nebulosa a ideia de que haja um objeto introspectado, nos moldes do objeto da percepção, que

---

<sup>16</sup> Audi (1998: 74-75) utiliza, para ambos os casos, as expressões “experiential process properties” e “experiential state properties”, respectivamente.



sirva de contrapartida deste. O que o “olho introspectivo” vê é diferente do que o olho da percepção enxerga, o primeiro necessitaria de uma espécie de imagem de “segunda ordem” para enxergar “mentalmente”, digamos assim.

Dando-se seguimento a esta análise do aspecto cognitivo da introspecção, a abordagem da tese do acesso privilegiado faz-se necessária e pertinente, uma vez que está intimamente relacionada ao modo como é dado o contato com os conteúdos de nossos estados mentais. A tese do acesso privilegiado sustenta que um sujeito S é infalível e onisciente em relação a estes conteúdos mentais, não havendo espaço lógico e epistêmico para o erro e o engano, pois esta condição cognitiva de S, a saber, sua infalibilidade e onisciência, mostrariam que ninguém mais está em condições de saber de sua vida mental do que o próprio S, condição esta dita privilegiada em relação aos objetos do mundo exterior e aos conteúdos mentais de outros sujeitos, por exemplo. Esta tese pode ser contestada e negada em parte, apesar de ser contra-intuitiva esta possível refutação: se condicionarmos esta infalibilidade e onisciência apenas à chamada introspecção *atenta*, podemos aceitar tal asserção, pois o que seria infalível, de fato, são os estados e/ou fenômenos mentais ocorrentes (conscientes), excluindo a esfera dos eventos mentais disposicionais e não ocorrentes que S pode ter sem ter consciência e autorreflexão de que os possui.

É exata e especificamente neste ponto nevrálgico da discussão que se pode trazer novamente à baila o problema do autoconhecimento e de uma suposta incompatibilização deste com uma perspectiva anti-individualista (ou externalista) acerca do conteúdo mental e da justificação epistêmica: na concepção incompatibilista, para S saber o que pensa, ele deve, em última instância, fazer uma investigação empírica do mundo exterior a fim de investigar os conteúdos dos seus pensamentos, hipótese esta

inverossímil, pois o levaria à perda da sua autoridade de primeira pessoa como aparato cognitivo privilegiado destes pensamentos que são próprios da vida privada e mental de S. Não é o objetivo aqui tratar dessa deferência ao conteúdo mental putmaniano, por exemplo, da *slow switching* (mudança não informada de ambiente externo) da Terra para a Terra-Gêmea, e das consequências de sua famosa afirmação “significados simplesmente não estão na cabeça!” (Putnam, 1992), bem como de outros experimentos mentais de pensamento presentes na filosofia analítica de modo geral, mas sim fazer uma conexão com o conhecimento de crenças de nossa vida mental que não formamos ao fazer referência direta aos objetos exteriores. Numa analogia forçada, seria epistêmica e moralmente legítimo nos enganarmos acerca de se a dor que sinto, que creio que sinto, ou a crença de que estou faminto não ser de fato dor ou fome, mas uma falácia decorrente de um modo de raciocinar equivocadamente, e eu não poder saber disto, não ter este acesso cognitivo privilegiado adequado na formação desses silogismos introspectivos, e com isto ser falível acerca destas crenças de minha vida mental?

Há dois domínios cognitivos aqui, o raciocínio e a introspecção: no primeiro, posso supor uma relação inferencial, e nesta relação fazer inferências inválidas, mas com transmissão de racionalidade nesta relação, e cometer um equívoco falacioso sem saber que o cometi, ao passo que, no segundo domínio, mesmo com uma suposta infalibilidade na formação da crença introspectiva, e desta não ter uma natureza inferencial, poderia haver alguma espécie de “equívoco semântico” como no raciocínio, e, com isso, se crer irracionalmente e recair em algum tipo de autoengano. Boghossian, por exemplo, em “Externalism and Inference” (1992), na crítica que faz ao externalismo e à tese de Burge do autoconhecimento como um conhecimento de segunda ordem, oferece não apenas um incompatibilismo com o conhecimento do mundo exterior

com a perspectiva anti-individualista, mas também em relação ao caráter a priori das nossas capacidades lógicas, numa espécie de “infortúnio lógico” advindo daí, que decorreria em falácias de equivocação em relação a raciocínios mal feitos com o consentimento do sujeito cognoscente, mas sem o seu conhecimento disso: não poderia haver também um “infortúnio epistêmico” para esses casos, bem como para nossas crenças introspectivas, de ignorância epistêmica, de não saber que P? Ou S tem o dever e tem de saber que P, não podendo não saber, não tendo o direito de não saber? Parece clara e expressa aqui uma problemática de natureza moral nesta questão nevrálgica de fronteira, e sua análise e discussão no contexto dos presentes tópicos de pesquisa fazem-se mais do que necessários, pertinentes e altamente relevantes.

Nos casos de *slow-switching* do externalismo, do conteúdo mental em relação a pensamentos passados, por exemplo, até se poderia, com algum esforço intelectual, talvez, mostrar que o sujeito S estaria em alguma medida isento, digamos assim, de algum compromisso moral de dever crer, estar justificado que e saber que P. Mas, e na ausência destas trocas não informadas de ambiente externo, não supondo um entorno a S, do ponto de vista intraepidérmico, de sua vida mental, este tem esse direito de não saber, ou o dever de saber, supondo uma falibilidade acerca do conhecimento de suas crenças e estados mentais básicos nos moldes semelhantes a uma perspectiva anti-individualista, por exemplo? E se agregarmos a toda esta gama variada de tópicos de interesse para discussão os conceitos de “sorte epistêmica” (“*epistemic luck*”), de Duncan Pritchard, e de “sorte lógica” (“*logical luck*”) e “sorte moral” (“*moral luck*”), de Roy Sorensen, por exemplo, teríamos nós introspectivamente o direito de saber, quando podemos estar “gettierizados” e atingirmos a marca da verdade por mero acaso, tendo crença verdadeira justificada, mas não tendo conhecimento (a assim chamada

“gettierização” em Epistemologia consiste justamente no acidente cognitivo por parte do agente epistêmico que crê justificadamente em uma verdade por sorte)? Ou temos o *dever* de saber, mesmo acertando nesta marca por mera crença verdadeira justificada, onde não saberíamos, no caso, por termos atingido a verdade por acaso, não se constituindo tal crença verdadeira justificada em conhecimento? Talvez para a atitude da vida cartesiana que se fez referência anteriormente, chegar à verdade “por linhas tortas”, digamos assim, no sentido de não constituir-se tal trajeto uma rota epistemicamente adequada, correta e irrepreensível, seja suficiente, mas, e para determinados fatos de nossas vidas mentais que exigiriam um rigor cognitivo maior, talvez não o seja. Como *saber* tudo isto, se é possível, ou o mais sensato seria, de fato, a suspensão do juízo acerca dessas questões, e disto resultar um ceticismo introspectivo ou autoconsciente como especulação teórica? A mera possibilidade do autoengano e de nos equivocarmos em relação aos nossos próprios pensamentos, crenças e mesmo crenças justificadas, segundo uma perspectiva deontológica insuficientemente respondida aqui, mas de máximo interesse e relevância para toda a discussão posterior e ulterior que será necessária no aprofundamento destas questões, causa-nos uma espécie de ojeriza, e é uma questão epistemológica e moral de difícil e complexo tratamento. Não está em questão se podemos ou não ter acesso ao conteúdo de nossos pensamentos: uma vez pressupondo-se que podemos, a pergunta é se temos justificção adequada para crermos introspectivamente, ou se estamos sendo racionais ao crer que P, não nos sendo passível a possibilidade do erro e da falibilidade e, ao fim e ao cabo, se temos ou não o direito e/ou o dever de assim crer. E, uma vez rejeitando um infalibilismo acerca de nossa vida mental, das crenças verdadeiras supostamente justificadas introspectivamente, temos o direito de sermos falíveis, ou é nosso dever, epistemicamente, crer

---

justificadamente da perspectiva introspectiva de forma irrepreensível e sem a possibilidade do autoengano? Há espaço conceitual para crer irracionalmente, e o podemos fazê-lo sem saber, julgando sermos racionais: podemos nos dar a esse luxo, ou temos que ter consciência reflexiva dessa irracionalidade, mesmo que, de uma perspectiva de atitudes e ações em relação à nossa vida e nossos comportamentos, como observou Descartes, crer irracionalmente conforte-nos acerca de determinadas vicissitudes de nossa existência? A verdade, muitas vezes, *dói*, diz o ditado: é nosso dever sempre atingir a sua marca, independentemente de por crença acidentalmente verdadeira justificada e/ou conhecimento, nesse caso, ou podemos ter ou nos darmos o direito de não sabê-la, de não chegar à mesma, em detrimento de nossa satisfação pessoal e de interesses epistêmicos e morais, dentre outros, em alguma medida, nos autoenganando inconscientemente? E de que modo questões práticas, pragmáticas e prudenciais poderiam ser consideradas dentro deste contexto, no sentido de considerar o que seriam virtudes epistêmicas e morais no âmbito das atitudes mentais do sujeito cognoscente?

Em consonância e na relação direta ou indireta com o que foi exposto até este momento e dando prosseguimento a estas discussões, é que determinados aspectos conceituais que constituem as bases gerais da epistemologia moral podem e devem ser desenvolvidos e investigados. O cognitivismo moral, por exemplo, é a concepção dentro da interface ética/epistemologia que defende que existam fatos morais objetivos no mundo, as crenças morais dos sujeitos cognoscentes são verdadeiras ou falsas, e podem as mesmas ser justificadas objetivamente, seja por critérios internalistas já referidos e discutidos anteriormente, como evidências e/ou razões para crer, por exemplo, ou externalistas, como a conexão confiável e estável com a verdade, o fato do mundo. Já o

não cognitivismo, a tese moral contrapositiva ao cognitivismo, defende que não há objetividade nos juízos morais no sentido destes serem verdadeiros ou falsos, apresentarem valores de verdade objetivos, embora uma ética não cognitivista, que caracteriza-se basicamente pela impossibilidade do estabelecimento de crenças morais verdadeiras e justificadas, não implique que tais crenças morais não sejam passíveis de graus justificacionais, que não necessariamente envolvem atitudes cognitivas quanto a fatos ou valores morais, embora envolvam atitudes cognitivas quanto a fatos “*tout-court*”, como em qualquer prática guiada pela ação humana. Dentro desta dicotomia entre as posições cognitivistas e não cognitivistas na epistemologia moral, e fazendo a conexão com a questão do autoconhecimento e a defesa racional ou não de alguma forma plausível de ceticismo moral e de aspectos deontológicos, como já foi colocado anteriormente, a grande dificuldade do estabelecimento das bases conceituais mínimas de uma epistemologia moral é a definição e a garantia da existência do seu objeto de estudo, basicamente. Duas tarefas primordiais e fundamentais colocam-se de antemão para o epistemólogo moral neste contexto investigativo, a saber: a do mesmo justificar a possibilidade e a eficácia da busca por este objeto, o conhecimento moral, e determinar positivamente o que deve ser aceito como critério de justificação para esse conhecimento moral, definindo com isso o seu objeto. Uma epistemologia moral mínima, de modo geral e canônico, pode ou não estar comprometida com as teses do cognitivismo e do não cognitivismo morais, bem como com algumas formas de realismo e relativismo morais e culturais, sob a égide de pluralismos éticos diversos, e evidentemente com a análise tradicional do conhecimento proposicional como crença verdadeira justificada e a possibilidade de justificação das crenças morais, seja por critérios internalistas e/ou externalistas, conforme já

---

exposto e discutido anteriormente, e o comprometimento ontológico com a existência ou não de fatos morais objetivos e uma resposta convincente aos argumentos céticos e relativistas, como os clássicos problemas epistemológicos do regresso ao infinito, o fundacionismo com justificação de crenças básicas na cadeia doxástica e o coerentismo, seja ele epistêmico ou não, no qual também temos a teoria dos juízos morais de John Rawls, por exemplo, e a sua tese do equilíbrio reflexivo para fundamentar uma justificação de crenças morais, bem como o intuicionismo moral de David Ross ou o utilitarismo de John Stuart Mill, por exemplo. Enfim, são todas questões de alta relevância e pertinência no domínio amplo, irrestrito e ilimitado de pesquisa nesta área, com muitas possibilidades de conexões e relações entre estas diversas temáticas que se relacionam entre si, tendo o grande campo da epistemologia moral e do deontologismo epistêmico como pano de fundo e substrato teórico/conceitual, onde muita pesquisa ulterior e futura é requerida, a fim de atingir os objetivos mínimos aqui propostos diante da complexidade de tais temas, seja nas suas particularidades ou no conjunto dos tópicos minimamente expostos e discutidos no presente contexto investigativo.

## Referências

Alquié, Ferdinand (1996). *Oeuvres philosophiques de Descartes II (1638-1642)*, Paris: Dunod.

Alston, William (1989). 'The Deontological Conception of Epistemic Justification'. In: Alston, William. *Epistemic Justification: Essays in the Theory of Knowledge*, Ithaca: Cornell University Press, 1-354.

- Audi, Robert (1998). *Epistemology: a Contemporary Introduction to the Theory of Knowledge*, London and New York: Routledge.
- Boghossian, Paul A. (1989). 'Content and Self Knowledge', *Philosophical Topics* 17(1): 5-26.
- Boghossian, Paul A. (1992). 'Externalism and Inference', *Philosophical Issues* 2:11-28.
- Burge, Tyler (1993). 'Content Preservation', *Philosophical Review* 102(4): 457-488.
- Chisholm, Roderick (1974). *Teoria do Conhecimento*, Rio de Janeiro: Zahar Editores.
- Descartes, Rene (1991 [1641]). *Meditações Metafísicas*, Coleção Os Pensadores. São Paulo, Editora Nova Cultural, 5 Ed.
- Feldman, Richard (2003). *Epistemology*, Upper Saddle River, New Jersey: Prentice Hall.
- James, William (1897). 'The Will to Believe'. In: William, James. *The Will to Believe and Other Essays in Popular Philosophy*, New York: Longmans, Green, and Co, 1-332.
- Plantinga, Alvin (1993). *Warrant: the Current Debate*, Oxford: Oxford University Press.
- Putnam, Hilary (1992). *Raizão, Verdade e História*, Lisboa: Dom Quixote.



# ANTI-ANTI-REDUCTIONIST CONSIDERATIONS ABOUT THE JUSTIFICATION OF TESTIMONIAL BELIEFS

---

*Luis Rosa & André Neiva*<sup>1</sup>

## 1. A dilemma

Anti-reductionists about the epistemic status (e.g. justification or knowledge) of testimonial beliefs<sup>2</sup> assume that some paradigmatic cases of testimony—more precisely, cases where the hearer has no precise information about the speaker—motivate their position. Consider the following quote from Lackey's *Learning from Words*:<sup>3</sup>

When I ask a stranger on the street for directions to the Empire State Building, do I have enough information about her to justify my accepting her testimony that it is six blocks north? Or, traveling to London for the first time, do I have enough evidence about a random British newspaper to adequately justify the beliefs I acquire while reading it?

This, to my mind, is a very compelling objection to reductionism.

---

<sup>1</sup> Emails: [fsopho@gmail.com](mailto:fsopho@gmail.com) (Luis Rosa), [al.neiva@gmail.com](mailto:al.neiva@gmail.com) (André Neiva).

<sup>2</sup> The epistemic status we are interested in here is *justification*, not *knowledge*.

<sup>3</sup> Lackey (2008: ch. 5).

Let us ask whether Lackey's type of objection against reductionism is in fact compelling. In particular, let us ask if it is compelling when it comes to epistemic justification. The objection would be, then: there are cases in which the hearer clearly lacks justifying information about the testifier and, yet, the hearer forms an apparently justified belief on the basis of the relevant testimony. Let us leave aside for a moment cases of testimony from *experts*—testimonial beliefs that are held in virtue of an attribution, on the part of the hearer, of expertise to the speaker on the topic of her assertion. The cases we are interested in (and the ones that are present in Lackey's objection) are cases of *testimony from strangers*—a more dangerous territory.

Consider the following disjunction concerning beliefs formed on the basis of testimony from strangers, in the type of scenario envisioned by Lackey: either (a) those beliefs are justified or (b) they are not. If (a) is the case (first *lemma*), then Lackey's case is compelling *only if* those beliefs are justified non-inferentially or, for that matter, only if those beliefs are justified in a way that differs from the one required by the reductionist about testimonial beliefs. If (b) is the case (second *lemma*), however, then Lackey assumes something that is actually false to be true. The anti-reductionist argument about testimonial justification does not get off the ground without the assumption that the target beliefs are justified.

Our considerations here will run as follows. The cases that are supposed to motivate anti-reductionism about the justification of testimonial beliefs are either cases of beliefs that are justified, but this can be properly explained by making reference to reasons available to the hearer, or they are not cases of epistemically justified beliefs at all. If that is the case, then maybe we have no good reasons to believe that the anti-reductionist view is true. That is, none of the disjunctions presented above, (a) or (b), seems to favour anti-reductionism.

---

Let us explore *lemmas* (a) and (b) in turn.

## 2. *Lemma* (a)

Here is a straightforward explanation why (a) is true (assuming that it is true): in the relevant cases of testimony from strangers, hearers form their testimonial beliefs *on the basis of good reasons*.<sup>4</sup>

Suppose Mabel is in Porto Alegre for the first time. She is hungry and she asks a stranger where she can find a restaurant. ‘See that stairs over there?’, the stranger asks Mabel, ‘You go up and you’ll see a restaurant on your left’. Mabel goes up the stairs and she finds a restaurant on her left—just as the stranger said. (Here is one way we can explain Mabel’s behaviour: after hearing what the stranger said, she trusted the stranger’s testimony; that is, she formed a belief that *There is a restaurant upstairs on the left*, or something along these lines and acted on it).

Now, it seems clear that Mabel has no specific information about that stranger’s authority on the location of restaurants in Porto Alegre—she does not know that the stranger is a reliable source of information about that issue. They never met before. For all she knows, the stranger can be a sincere, well-informed citizen as well as a liar, or maybe an evil epistemologist who is trying to fool people in order to lower the reliability of their testimonial beliefs in the actual world. But is Mabel totally deprived of good reasons to believe what the stranger said?

It is doubtful that this is the case: maybe as a result of being aware of the way the stranger answered her question (*plus* other background beliefs), Mabel believes that *It is highly unlikely that the stranger would answer her question so readily and with such a conviction if he did not know whether there is a restaurant upstairs*, and/or Mabel believes that *Probably that*

---

<sup>4</sup> For a similar point, see Kenyon (2013).

*stranger lives in Porto Alegre given his accent*, etc. So Mabel's belief about what the stranger said may not be the only one that she forms as a result of being in the testimonial situation. And those beliefs plus Mabel's belief about what the stranger said may be good reasons for her to believe that *There is a restaurant upstairs on the left*.

Do these considerations indicate that Lackey-type considerations against reductionism are not compelling? It will depend on how we interpret the position called 'reductionism' here. Lackey herself seems to acknowledge that both can be true at the same time: (i) the justification of testimonial beliefs does not necessarily depend on inferential processes from other beliefs (more specifically, beliefs whose justification are non-testimonial) and (ii) the justification of testimonial beliefs depends on at least some positive reasons available to the believer. So, as long as we take reductionism to be a thesis according to which the justification of testimonial beliefs necessarily depends on *processes of inference* from other beliefs (beliefs whose justification are non-testimonial), then the Lackey-type objection to reductionism may look compelling.

However, reductionists need to require hearers to perform specific pieces of inference, maybe overly demanding ones, when they form their testimonial beliefs. It is sufficient for a reductionist to claim that the justification of the relevant testimonial beliefs *is not basic* (whatever that means), or that it depends on the epistemic status of further doxastic attitudes. If that is the case (under assumption (a)), we can conclude that although Lackey makes a compelling point against a certain interpretation of the reductionist view, perhaps she does not make a compelling point against *a better interpretation* of the reductionist view.

A good way to start fleshing out a better version of reductionism of this kind is to consider a Humean take on the issue of testimony. Adler (2012), along with a number

---

of epistemologists compares this view with a Reidian, anti-reductionist view: “Reid's position is that any assertion is creditworthy until shown otherwise; whereas Hume implies that specific evidence for its reliability is needed.” As Adler himself notes, we need not interpret the Humean position as stating that hearers must go out and check the credentials of the speaker, inquiring into his or her reliability concerning the relevant topic. Rather, it can be best interpreted as stating that the *presence* or *availability* of evidence or reasons for thinking that the speaker is reliable is necessary for one to be justified in believing the proposition expressed by the speaker's utterance. What seems to make for the contrast between Humeans and Reidians is that the latter ones, but not the former ones, claim that the absence of defeating reasons or counter evidence for testimonial beliefs is *sufficient* to make them justified.

(At this point someone could point out that the ‘better version’ of reductionism simply consists in a very trivial point: that justified beliefs have some sort of dependence on ‘background evidence’, or non-explicit reasons. But we should not think of background evidence as something of less importance, as if its presence was some sort of trivial condition for justification, on a par with conditions such as the one expressed by: *S* has a justified belief at *t* only if *S* is alive at *t*. The claim that *S* is alive at *t* does not raise the probability that *S* is justified in believing that *p* at *t* to a significant level, as does the claim that *S* has background reasons for believing that *p* at *t*.)

But even if inferences on the part of the hearer were required by the reductionist (in order for us to attribute justification to his/her testimonial belief), maybe we would have no good objection of the Lackey-type at the end of the day. Indeed, some authors<sup>5</sup> would like to describe the

---

<sup>5</sup> See Fricker 1995.

reductionist view as making exactly such a requirement. A reductionist need not assume that the relevant inferences are processes in which one *explicitly* entertains premises and derives a conclusion from them (perhaps also taking the premises to give support to the conclusion). Suffice it that the hearer has good reasons available (maybe not explicitly held ones) that in some sense ground one's belief in the proposition expressed by the speaker's utterance. (We can make sense of the relevant grounding relation in a number of ways—as a counterfactual dependence, as a non-deviant causal sustaining connection, etc.—but we need not delve into the details about this now.)

But even if the reductionist position requires some sort of inference on the part of the hearer, it would appear that some sort of *explanatory inference* may be plausibly assumed to take place in some of the relevant cases. The idea is that the hearer reasons roughly as follows when she accepts the testimony of a stranger:<sup>6</sup>

Stranger says '*p*' *because* she knows that *p* and she cared to inform me that *p*.

Again, this type of inference—like many 'automatic' or 'System 1' cognitive processes<sup>7</sup>—need not require the hearer to consciously and explicitly go through the relevant inferential steps. They can be instantiated effortlessly by the hearer, just like many other ordinary cases of reasoning by explanation (e.g., when one reasons as follows: Amanda is crying *because* Amanda is sad).

---

<sup>6</sup> See similar proposals in Fricker (1995), and Malgren (2006).

<sup>7</sup> See Kahneman (2011).

---

### 3. *Lemma* (b)

Blindly assuming that paradigmatic beliefs formed on the basis of testimony from strangers are always justified is not a responsible philosophical attitude. So let us submit that assumption to critical scrutiny.

Here is one consideration in support of the idea that many testimonial beliefs of the relevant kind are not epistemically justified: in many standard cases of testimony from strangers, the hearer actually has no good reasons or evidence for holding the testimonial belief, although that belief has some (yet to be specified) practical value to her. Here is another consideration in support of that idea: in many standard cases of testimony from strangers, the hearer does not actually form a *belief* toward the proposition expressed by the speaker's claim, but rather some (yet to be specified) weaker type of propositional attitude toward it. Let us consider each of these possibilities in turn.

We saw that Mabel believes that *The stranger said that there is a restaurant upstairs*. Call this her 'testimonial evidence'. We also saw that we can plausibly assume that Mabel also forms other beliefs as a consequence of hearing what the stranger said (and *how* he said it): beliefs whose contents may constitute good explanations why the speaker said what he said, or why the speaker said what he said in the way he did. If that is the case, it makes sense to suppose that Mabel uses those explanations to draw further inferences—e.g., that *There is a restaurant upstairs*.

It may be the case, however, that Mabel's testimonial evidence *plus* her explanatory beliefs constitute less than good reasons for her to believe that *There is a restaurant upstairs*. For one thing, the explanatory beliefs she forms may themselves be only weakly justified. For another, the support that her testimonial evidence *plus* her explanatory reasons confer upon the proposition that *There is a restaurant*

*upstairs* is possibly below the threshold required for epistemic justification (whatever that is).

If one were to follow this line of argument, one would need to explain why it is intuitive for most of us to think that paradigmatic testimonial beliefs are epistemically justified, despite the fact that they are not (that is, some kind of error-theory would be required here). Are there good explanations to that effect? Here is one candidate explanation: it is *practically rational* for hearers to behave as if what the stranger said is true—what the stranger said is what the hearer ‘has to go about’ in her practical decisions. E.g., as far as Mabel can tell, it is more likely that she will satisfy her practical goals (e.g., eating in a restaurant) conditional on her going upstairs—that is, conditional on her behaving as if what the stranger said is true (that *There is a restaurant upstairs*)—than it is conditional on her not so behaving. But this does not necessarily mean that she *believes* what the stranger said. She may rather believe it is likely or probable that the stranger is correct. Furthermore, it would be intuitive for us to attribute epistemic justification to Mabel’s testimonial belief because it is actually practically rational for her to behave as if what the stranger said is true.

It is not at all implausible to assume that in most cases of testimony from strangers the hearer does not actually form a *belief* toward the proposition that was testified by the stranger. Maybe the attitude the hearer takes toward that proposition in the relevant cases is that of an *assumption*: she just assumes that what the stranger said is true and starts operating under that assumption. Or maybe the attitude the hearer takes toward that proposition in the relevant cases is that of a not-so-high credence. Or, as it was hinted above, maybe it is a belief about what is likely or not.

Here is a mark of belief: in general, when one believes that  $p$  one also has a disposition to answer ‘Yes’ to



---

a person that is interested in knowing whether  $p$  is the case. In most paradigmatic cases of testimony from strangers, however, the subject has only a disposition to answer: ‘Well, someone out there just told me that  $p$ —maybe he is right...’. Imagine that soon after hearing what the stranger said, Mabel is asked by another stranger whether there is a restaurant nearby. Would she firmly answer: ‘There is a restaurant upstairs’, as if she believed that proposition to be true? Probably not. She would rather tell the person that *someone just told her that there is a restaurant upstairs*. That much is believed by her—but that does not mean that she also believes that *There is a restaurant upstairs*.

There is another striking difference between attitudes of belief and whatever weaker sorts of attitude one may form in a situation of testimony from strangers. Although beliefs are revisable, they mostly persist when confronted with not-so-good reasons to believe the contrary, or even to suspend judgment about the relevant subject matter. If you actually believe that *The Comic Book Store is in the 7th Avenue*, say, because you have been there before and you know the names of the avenues, etc., then you will not stop believing such a thing when a stranger tells you that *The Comic Book Store is in the Road St*. Compare this case, however, to the one in which a stranger tells you that *The Comic Book Store is in the 7th Avenue* and, right after that, another stranger tells you that *The Comic Book Store is in the Road St*. In the latter but not in the former case, you change your attitude toward the proposition *The Comic Book Store is in the 7th Avenue*. That is because it is only in the former case that you really believe that proposition.

## Conclusion

We conclude that, as long as anti–reductionists take those paradigmatic cases of testimony that we have been dealing with here to be their main motivation, they are in

trouble. *Lemma* (a) can be properly explained by a reductionist view and *Lemma* (b) is a dead-end for anti-reductionists.

## References

Adler, Jonathan (2012). 'Epistemological Problems of Testimony'. The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), accessible in the following URL: <<http://plato.stanford.edu/archives/spr2014/entries/testimony-episprob/>>.

Kahneman, Daniel (2011). Thinking Fast and Slow. Farrar, Straus and Giroux.

Kenyon, Tim (2013). 'The Informational Richness of Testimonial Contexts', *Philosophical Quarterly* 63(250): 58-80.

Lackey, Jennifer (2008). Learning from Words: Testimony as a Source of Knowledge. New York: Oxford University Press.

Hume, David (1977 [1748]). An Enquiry Concerning Human Understanding. Eric Steinberg (ed.), Indianapolis: Hackett Publishing Company.

Malmgren, A. S. (2006) 'Is There A Priori Knowledge By Testimony?', *The Philosophical Review*, 115: 199–241.

# A AGÊNCIA EPISTÊMICA REFLEXIVA DE ERNEST SOSA<sup>1</sup>

---

*Thiago Rafael Santin*<sup>2</sup>

## **Introdução**

Neste artigo, exporemos e examinaremos a proposta de agência epistêmica reflexiva de Ernest Sosa, principalmente a exposta no livro *Knowing full well* (2011), capítulo dois, *Epistemic agency*, e no artigo *Epistemic agency* (2013)<sup>3</sup>, a partir da sua tradução publicada, *Agência epistêmica* (2012).<sup>4</sup>

“Agência epistêmica” surge como um termo análogo ao termo “agência”. Assim como em ética prescrevem-se deveres éticos, fala-se em epistemologia em deveres epistêmicos, em relação à tomada de decisão doxástica sincrônica – instanciada na pergunta: “o que devo crer agora?”. A partir do princípio da tradição kantiana de que *dever implica poder*, tem-se como pressuposto a capacidade de controle voluntário do agente sobre suas crenças, tal qual há controle voluntário sobre suas ações – posição padrão em filosofia da ação. Em epistemologia, porém, a posição

---

<sup>1</sup> Este artigo é uma versão melhorada de parte da minha dissertação de mestrado.

<sup>2</sup> Email: [santin.thiago@gmail.com](mailto:santin.thiago@gmail.com).

<sup>3</sup> A versão utilizada é o rascunho disponível na página pessoal do autor (link: <<http://www.ernestsosa.com/>>); portanto, a paginação citada não corresponde à versão publicada.

<sup>4</sup> Em alguns trechos, em função de erros de tradução, usaremos o artigo original, em inglês, que teve sua publicação posterior. No restante, seguiremos a tradução publicada, para padronização terminológica em português.

*standard* é a falta de controle voluntário sobre os estados doxásticos, chamada de involuntarismo doxástico. A partir da falta de controle direto, buscaram-se controles indiretos, como a deliberação e a reflexão.

Abrimos mão da agência doxástica, controle direto sobre os estados doxásticos, em função da falta de controle voluntário sobre nossos estados doxásticos, o involuntarismo doxástico. Investigaremos, então, agência epistêmica através de desempenhos de um crente. Examinaremos se os desempenhos são do tipo epistêmico e se possibilitam agência epistêmica.

Desempenhar implica risco, o risco de falhar, pois é sempre possível obter sucesso ou fracassar no que se desempenha. A formação de crença (e aquisição de conhecimento) pode, então, falhar. A falha epistêmica pode ter consequências práticas, como quando, por exemplo, eu falho em saber os horários de ônibus da minha região e me atraso para um compromisso, ou quando falho em acreditar nas evidências que me são apresentadas e creio com base no meu preconceito, agindo sobre uma falsidade.

A teoria de Sosa tem se desenvolvido como uma epistemologia da virtude de dois níveis<sup>5</sup>. Ela abarca tanto os processos cognitivos de busca da verdade quanto a consideração reflexiva das evidências na decisão doxástica, e tem como noção central a avaliação de desempenhos de agentes motivados que visam a alcançar um fim, o conhecimento, manifestando virtudes ou competências, que são disposições de um agente. Exporemos a epistemologia da virtude de Sosa, que é um caso especial de normatividade e agência em geral, para, depois, abordar o domínio epistêmico – normatividade e agência epistêmica (Sosa, 2012).

---

<sup>5</sup> Para uma caracterização e crítica da teoria de Sosa, ver *Bi-level virtue epistemology*, de John Turri (2013).

---

## Estrutura e normatividade de crer

### I) Estrutura

Os desempenhos epistêmicos têm duas ordens. Como um caso especial, os termos de desempenho e aptidão são aplicados à formação de crença e, conseqüentemente, à aquisição de conhecimento. Porém, surgem algumas especificidades e assimetrias, como a suspensão de juízo.

Crer é um desempenho que tem duas ordens: execuções e avaliações. Na primeira, é um desempenho que tem por objetivo, em geral, a verdade<sup>6</sup>, e conta com os três níveis da estrutura de funcionamento: a crença é precisa se é verdadeira (atinge/alcança a verdade); é hábil se é formada de maneira competente (mira adequadamente a verdade); e é apta se demonstra a competência do crente (atinge a verdade porque mirou adequadamente e obteve a verdade porque a buscou)<sup>7</sup>. Os três níveis indicam, respectivamente, se a crença alcançou a verdade sobre a questão, se buscou alcançá-la e se obteve a verdade justamente porque buscou alcançá-la.

---

<sup>6</sup> Seguiremos Sosa (2011: 36), falando somente em “crer verdadeiramente”, comportando, assim, as respostas positiva e negativa de crer que P e descrever que P, entendido como crença na contraditória (crer que  $\sim$ P), referente à questão de *se* P é o caso ou não é o caso. A suspensão de juízo, não crer, é um caso similar à suspensão da ação, abstenção, dado o risco presente, seja pela falta de perícia do agente epistêmico (p.e.: é daltônico), por condições inadequadas (p.e.: está drogado), ou por situações inadequadas (p.e.: a luz é muito fraca para distinguir a cor de um determinado objeto).

<sup>7</sup> Sosa (2011: 64) distingue ‘atingir/alcançar’ (*to reach*) de ‘obter’ (*to attain*). O primeiro indica apenas a posse de sucesso (crença verdadeira), enquanto o segundo indica que o sucesso (crença verdadeira) foi alcançado pela manifestação da habilidade, ou seja, pelo desempenho apto (crença apta).

Na segunda ordem, as metacrenças avaliam a primeira ordem, pela competência e pelo ambiente. Elas constituem diligências para buscar a verdade sobre a questão considerada. Nelas, o agente baseia-se para decidir crer, descreer ou suspender o juízo. As metacrenças são originadas da metacompetência do agente, sua capacidade de avaliação da sua competência de primeira ordem dentro do seu ambiente atual.

A metacrença apta não deve ser confundida com a crença de primeira ordem, que é meta-apta. No primeiro caso, há uma crença de segunda ordem, que é resultado da metacompetência epistêmica – ou seja, uma metacrença com aptidão de segunda ordem, que é conhecimento reflexivo. No último, há uma crença de primeira ordem guiada por uma metacrença adequada.

No domínio prático, o agente é guiado por sua meta-aptidão para desempenhar ou abster-se. Já no domínio epistêmico, não há abstenção *simpliciter*, em primeira ordem<sup>8</sup>. Para Sosa, o agente ou crê que “P” ou crê que “não P”, que é equivalente a descreer. A abstenção no domínio epistêmico é a suspensão de juízo, em segunda ordem. Ao suspender o juízo, o agente não crê, não desempenha. Assim, a meta-aptidão epistêmica guia o agente no domínio epistêmico para desempenhar ou suspender o juízo.

Mas a suspensão de juízo, ao contrário da abstenção, precisa ser motivada por razões. A abstenção prática ocorre em primeira ordem, enquanto a suspensão de juízo ocorre somente na segunda ordem. A meta-aptidão, nesse caso, é a avaliação das razões que não levaram à crença ou à descrença e que conduziram à suspensão de juízo.

---

<sup>8</sup> O mais próximo de abstenção seria não considerar uma questão ou proposição. Assim, não haveria oportunidade de julgamento. Esses casos são possíveis (alguém que foge de considerar questões), mas não são interessantes, contudo, pois eliminam a situação de desempenhos.

---

Para mostrar a estrutura de crer, passemos ao seguinte exemplo:

### O Julgamento Olímpico

Imaginemos que há uma prova de arco e flecha em um campo com alvos, na qual há duas pessoas competidoras, um arqueiro homem e uma arqueira mulher. Para avaliar, temos uma banca com uma juíza mulher e dois juízes homens.

O primeiro juiz, Sexista, é um juiz que tem forte preconceito a favor de seu gênero, masculino, contra o gênero feminino. Ele faz seus julgamentos reiteradamente com base no preconceito, ao invés de fazer julgamentos que busquem a verdade. Assim, ao avaliar a prova de arco e flecha, ele julga que o arqueiro homem teve os melhores tiros e que, por isso, merece nota mais alta em relação à arqueira mulher. Ocorre que a crença do juiz (proferida em seu julgamento, pela nota) é verdadeira: o arqueiro homem de fato teve os melhores tiros. Contudo, o juiz Sexista tem crença precisa, pois ele atingiu a verdade, mas não tem crença hábil, pois não mirou adequadamente a verdade, nem crença apta. Embora tenha crença precisa, ela não é precisa em virtude de ser hábil, ele não obteve a verdade, apenas alcançou-a acidentalmente.

O colega do juiz Sexista é outro juiz homem, Distráido. Ele não está enviesado, como o colega, e quer buscar a verdade acerca da prova. Porém, ele fica muito próximo do local onde ocorre a sinalização sonora para autorizar os tiros das atletas. Ele é distraído pelo barulho do sinalizador do tiro, não olha para as competidoras e perde o momento do tiro. Contudo, ele espia o guia de anotações de sua colega, a juíza mulher, Justa, e marca tal qual ela.

Ocorre que a juíza Justa observou adequadamente e tem crença apta, pois sua crença é verdadeira porque ela

observou adequadamente. Nessa situação, o juiz Distráido tem uma crença precisa, pois é crença verdadeira, ele buscava alcançar a verdade ao crer, era hábil, mas não tem uma crença apta, porque atingiu a verdade acidentalmente, considerando que sua habilidade não foi suficiente para seu ambiente naquela ocasião: ele distraiu-se, não viu os tiros e também não sabe se a colega realmente viu, nem se o julgamento dela é verdadeiro. O juiz Sexista não apresenta um caso de conhecimento, assim como o juiz Distráido. Somente a juíza Justa tem conhecimento, como veremos adiante.

## II) Normatividade

Passemos à normatividade de desempenho, através dos três componentes: competência, objetivo e motivação, e sua aplicação ao domínio epistêmico.

### i) Competência

A competência epistêmica contém o conjunto de todas as subcompetências epistêmicas envolvidas (perceptiva, mnemônica, inferencial, introspectiva, etc.). Elas são as faculdades cognitivas e fontes de conhecimento que baseiam um desempenho para alcançar a verdade sobre a questão que se apresenta. Mas a competência epistêmica também inclui nossa capacidade racional:

Observe que nenhum humano abençoado com a razão tem meramente conhecimento animal do tipo obtido por feras. Pois mesmo quando uma crença perceptual deriva tão diretamente quanto sempre faz dos estímulos sensoriais, ainda é relevante que o sujeito não tenha percebido os sinais do testemunho em contrário. [...] Mesmo quando a resposta ao estímulo é muito direta, se o sujeito também ouvisse ou visse os sinais de testemunho



---

em contrário crível, isso mudaria a sua resposta. As crenças de um animal racional parecem, portanto, nunca provir de introspecção, memória ou percepção sozinhas. Pois a razão é sempre pelo menos uma parceira silenciosa na vigília de outros dados relevantes, uma parceira silenciosa cujo próprio silêncio é uma causa contribuinte da crença resultado (Sosa, 1991: 240)<sup>9</sup>.

Essa competência epistêmica, baseada na razão, é complexa e tem vários níveis:

Eu falo de “mecanismos” ou processos de formação de crença e algumas vezes de “mecanismos de entrada/saída”, mas eu quero negar explicitamente qualquer implicação que eles são simples ou modulares. [... Um] mecanismo pode ser algo próximo de um reflexo, ou pode ser uma habilidade de processamento central de nível muito alto, do tipo que permite uma crítica sensível a “decidir” como avaliar um trabalho, baseada em ponderação complexa e hábil (Sosa, 2007: 87, nota de rodapé 9).<sup>10</sup>

---

<sup>9</sup> No original: “Note that no human blessed with reason has merely animal knowledge of the sort attainable by beasts. For even when perceptual belief derives as directly as it ever does from sensory stimuli, it is still relevant that one has not perceived the signs of contrary testimony. [... E]ven when response to stimuli is most direct, if one were also to hear or see the signs of credible contrary testimony, that would change one’s response. The beliefs of a rational animal hence would seem never to issue from unaided introspection, memory, or perception. For reason is always at least a silent partner on the watch for other relevant data, a silent partner whose very silence is a contributing cause of the belief outcome”.

<sup>10</sup> No original: “I speak of “mechanisms” or processes of belief formation, and sometimes of “input/output mechanisms”, but I want to disavow explicitly any implication that these are simple or modular. [... A] mechanism can be something close to a reflex, or it can be a very high-level, central-processing ability of the sort that enables a

Conforme a estrutura de desempenhos epistêmicos, temos que crença precisa é crença verdadeira. Crença hábil é crença que busca obter a verdade, podendo ou não ser precisa (verdadeira). Há crenças que podem ser falsas, mas ainda assim hábeis. Esses são os casos fortuitos onde o agente manifesta adequadamente suas competências epistêmicas, mas ainda assim não consegue alcançar a verdade – o mundo pode não ter contribuído ou a ocasião exige competências que ele não possui.

A metacompetência consiste na avaliação dessas competências em nosso ambiente atual. Ela é o uso da razão para refletir acerca delas:

Tendo compreendido como é ao nosso redor, através de conexões envolvendo a percepção/memória/raciocínio exigidos para o conhecimento animal, continuar o raciocínio nessas bases pode aumentar nossa integração consciente e coerência explicativa, emprestando virtude epistêmica às nossas crenças (Sosa, 2010a: 187)<sup>11</sup>.

Essas competências terão relevância sobre como buscamos e alcançamos a verdade, sobre as questões que consideramos, pois constituem nossos processos de formação e avaliação de crenças. Assim, voltemos a considerar a finalidade dos desempenhos.

## ii) Objetivos

---

sensitive critic to “decide” how to assess a work, based on complex and able pondering”.

<sup>11</sup> No original: “Having grasped how it is around us through connections involving the perception/memory/reasoning required for the animal knowledge, further reasoning on that basis may enhance our conscious integration and explanatory coherence, lending epistemic virtue to our beliefs”.

---

Crer pode ter outros objetivos além da verdade, mas a verdade é o único objetivo epistêmico e é inerente à crença. Objetivos práticos são possíveis, tais como buscar conforto através de uma crença, o pensamento positivo (*wishful thinking*), e falhas e vieses de raciocínio. Mas isso não é relevante para a formação do conhecimento – pelo contrário, como veremos adiante, esses fatores não devem atrapalhar o objetivo epistêmico, para que haja conhecimento:

Uma crença pode assim derivar de causas ou mesmo razões que não são epistêmicas. Tome tal crença, motivada por razões pragmáticas. Mesmo se razões epistêmicas também motivam-na, entretanto, essa crença ainda pode falhar em constituir uma busca da verdade, se o sujeito crê como alguém em desconsideração extrema da evidência (Sosa, 2011: 20)<sup>12</sup>.

A crença apta é aquela que obteve a verdade devido ao crente tê-la buscado, ficando manifestada a competência epistêmica na sua aquisição, que resulta em conhecimento. A finalidade inerente, ainda que não única, da crença é a verdade.

Sosa defende a relação específica das crenças com a verdade, mas essa relação não é exclusiva. Ou seja, crenças podem ter outros fins e serem baseadas em razões práticas. Mas, nessas situações, não serão manifestações de competências epistêmicas do agente, falhando em constituir conhecimento: “Crenças que visam somente o conforto ou outro objetivo pragmático não são apropriadamente

---

<sup>12</sup> No original: “A belief can thus derive from causes or even reasons that are not epistemic. Take such a belief, one motivated by pragmatic reasons. Even if epistemic reasons also motivate it, however, that belief might yet fail to constitute pursuit of truth, if one believes as one does in extreme disregard of the evidence”.

guiadas por uma competência epistêmica para obter a verdade” (Sosa, 2011: 16)<sup>13</sup>.

iii) Motivação

Os desempenhos são baseados em razões motivadoras, assim como as crenças tradicionalmente são baseadas em razões epistêmicas. Crer é um desempenho também baseado em razões, podendo ser motivado psicologicamente por razões que não sejam epistêmicas. As razões para crer podem, então, ser do tipo epistêmico, em crenças que visam à verdade, ou do tipo prático, nas que têm múltiplos alvos.

Sosa considera que as razões epistêmicas são um subconjunto das razões práticas. A aquisição da verdade (alvo epistêmico) visa à obtenção de conhecimento, um dos possíveis fins de crer e o único fim epistêmico. A verdade é o valor epistêmico, e a avaliação epistêmica dá-se por ela: se foi buscada, se foi alcançada e se há relação entre a busca e a consecução.

Como exemplo, elaboramos um caso a partir de Sosa (2011: 18):

Primo Vinny na Prisão

O primo da juíza olímpica, Justa, chamado Vinny, está preso por um crime terrível. Ela não suporta a ideia de que ele seja culpado e quer confortar a si mesma, e forma a crença falsa de que ele não é culpado. Ela, então, pede para Vinny dizer que não cometeu o crime, mas ele responde confessando que é culpado. Justa pode manter a crença de que ele não é culpado, ou crer no testemunho de Vinny e ter uma crença verdadeira, apesar do seu forte desejo de autoconforto.

---

<sup>13</sup> No original: “Beliefs aimed only at comfort or some other pragmatic objective are not properly guided by an epistemic competence for attaining truth”.

---

Se mantiver a crença de que ele não é culpado, terá um mau desempenho, por formar crença com base no autoconforto e mantê-la, ignorando as evidências contrárias. Mas, se considerar o testemunho e deixar seu desejo de autoconforto, buscará a verdade e poderá ter conhecimento. Então, se ela deixar de lado seu autoconforto e acreditar no testemunho de Vinny, será meta-apta, pois se absteve de considerar o desejo de sua inocência, e poderá ter crença apta e um desempenho plenamente apto, que manifesta sua competência completa. Por outro lado, se Justa falhar em afastar seu desejo de conforto, terá mantido uma crença falsa com o fim de autoconforto. Essa crença poderá ser apta do ponto de vista prático, se ela se confortar, mas Justa não será apta epistemicamente. Já se ela permanecer em dúvida e abster-se, por seu desejo de autoconforto conflitar com a confiança no testemunho do primo, estará exercendo bem sua metacompetência epistêmica, apesar de não ser apta epistemicamente – pois não produziu desempenho. Ela reflete sobre o risco de crer conforme seu conforto e suspende o juízo, dadas as dúvidas sobre sua competência de formar a crença naquela ocasião, que exige pesar corretamente as evidências contra o seu desejo prático. Ao duvidar da manifestação da sua competência em primeira ordem e abster-se, está sendo louvável na segunda ordem, por avaliar bem o risco presente, dada a situação e suas competências epistêmicas. Justa seria epistemicamente meta-apta sem ser apta, ao suspender o juízo e evitar uma crença falsa.

### **Estrutura epistêmica e ordens de conhecimento**

Sosa define duas ordens de desempenho epistêmico em sua estrutura normativa, a fim de mostrar

conhecimento de um modo mais pleno, através das relações entre os níveis epistêmicos e o seu resultado.

O conhecimento é adquirido pelo agente que desempenha uma diligência em busca da verdade. Ele conta com uma estrutura própria, mas há uma relação do conhecimento com o agente. Os processos de formação de crença, e conseqüente aquisição de conhecimento, têm sua aptidão derivada do agente que desempenha esses processos:

Somente crenças na diligência de obter a verdade se qualificarão [como candidatas ao conhecimento]. E elas são aptas somente indiretamente. Elas derivam sua aptidão da aptidão da diligência correspondente, a diligência de obter assim a verdade. É essa diligência que mais diretamente se enquadra na estrutura AAA<sup>14</sup> (Sosa, 2011: 17)<sup>15</sup>.

Isso implica na avaliação não ser de uma crença ou de um caso de conhecimento, de maneira isolada, mas da formação e da crença de alguém, do conhecimento adquirido por um agente. Isso não exclui a avaliação sobre esses processos e as crenças, que também pode ser feita:

---

<sup>14</sup> A estrutura de virtude em três níveis “AAA”: *accuracy*, *adroitness*, *aptness* (precisão, habilidade, aptidão). Ela estipula o sucesso do desempenho (precisão), a competência do agente (habilidade) e a relação entre esses dois (aptidão). A execução de um desempenho encaixa-se na estrutura AAA se ele for preciso, ou seja, se atingiu seu objetivo; se for hábil, ou seja, se mirou adequadamente seu objetivo; e se for apto, ou seja, se o primeiro derivou do segundo, se atingiu o objetivo por tê-lo mirado adequadamente.

<sup>15</sup> No original: “Only beliefs in the endeavor to attain truth will qualify. And they are apt only indirectly. They derive their aptness from the aptness of the corresponding endeavor, the endeavor thereby to attain truth. It is this endeavor that most directly falls under the AAA structure”.

---

São somente agentes e sujeitos que são avaliados como justificados ou não? Não, ações e crenças são avaliáveis também. De fato, quando alguém é avaliado como justificado, ele é avaliado como justificado *em agir ou acreditar de certa maneira*. Não obstante, a avaliação de atos particulares pode implicar em uma avaliação indireta do agente ou sujeito (grifo original, Sosa, 2010a: 186)<sup>16</sup>.

Dada a relação entre conhecimento e agente conhecedor, temos a relação da avaliação de crença/conhecimento com a avaliação do agente, e vice-versa. Não é uma relação isomórfica, mas é próxima o bastante para usarmos os termos de forma intercambiável, por simplicidade. O importante é que não se trata de elementos separados, agente e crença/conhecimento, nem de uma avaliação normativa que se dê separadamente, apenas fazemos um exame separado, por questões de investigação teórica.

Temos até aqui que conhecimento é crença apta. Para termos conhecimento, é preciso crença verdadeira que seja alcançada pela manifestação da competência epistêmica relevante. Ou seja, o primeiro nível de sucesso epistêmico é a aptidão de primeira ordem e consiste em conhecimento animal.

S sabe que p se, e somente se, a correção da crença de S que p manifesta, parcialmente ou plenamente, uma competência epistêmica de formação ou sustentação de crença de S, que, ao fazê-lo, também manifesta plenamente uma competência

---

<sup>16</sup> No original: “Is it only agents and subjects that are assessed as justified or not? No, actions and beliefs are assessable, too. Indeed, when one is assessed as justified, one is assessed as justified *in acting or believing in a certain way*. Nevertheless, the evaluation of particular acts might imply an indirect evaluation of the agent or subject”.

(possivelmente mais complexa) assentada pelo menos parcialmente em S (Sosa, 2011: 88)<sup>17</sup>.

Assim, a competência do agente necessita ser manifestada para que ele tenha conhecimento, incluindo a condição e a situação que, ao menos parcialmente, envolvem fatores externos. O que nos leva à questão da sorte epistêmica, pois o agente pode estar em um ambiente adequado e ter conhecimento por sorte ou acidentalidade, como nos casos do tipo Gettier<sup>18</sup>.

Então, há também a necessidade de eliminar a sorte e a acidentalidade, dada na relação do agente com o seu ambiente. Isso ocorre com a avaliação das competências de primeira ordem, ou seja, uma metacompetência. Há crenças reflexivas (ou crenças de segunda ordem, ou, ainda, metacrenças – todas são sinônimas), sobre a competência de primeira ordem e o ambiente, o que resulta em conhecimento de segunda ordem, ou conhecimento reflexivo. Para que haja conhecimento reflexivo, é preciso competência de segunda ordem.

Seguindo a estrutura normativa, o conhecimento de segunda ordem é aquele que avalia a pertinência da crença em relação às competências do agente dentro do seu ambiente. Esse metaconhecimento é necessário para que o agente epistêmico oriente a competência epistêmica, avaliando o risco de falha em crer, podendo decidir crer ou suspender o juízo:

---

<sup>17</sup> No original: “S knows that p if, and only if, the correctness of S's belief that p manifests, partially or fully, an epistemic belief-forming or belief sustaining competence of S's, in doing which it also manifests fully a (possibly more complex) competence seated at least partially in S”.

<sup>18</sup> Para uma introdução, ver *Gettier problem*, em *Analysis of Knowledge*, da SEP, disponível em: <<http://plato.stanford.edu/entries/knowledge-analysis/#GetPro>>.



---

Para uma epistemologia da virtude baseada em competência, então, conhecimento animal exige crença apta, mas conhecimento reflexivo exige mais. Conhecimento reflexivo exige crença apta que é metacompetente (e, de fato, meta-apta). O conhecedor deve ter uma compreensão de segunda ordem – uma crença ou pressuposição – que sua crença de primeira ordem é ou seria apta (Sosa, 2010: 475)<sup>19</sup>.

A avaliação adequada sobre crer ou suspender o juízo é resultado da meta-aptidão. O agente é guiado pela metacompetência ao ponderar adequadamente a resposta, conforme sua competência de primeira ordem e seu ambiente na ocasião, levando em conta o risco de falhar e formar crença falsa. A falha em segunda ordem é compreendida como subavaliação ou sobreavaliação do risco em primeira ordem: ou o agente epistêmico forma crença quando é arriscado demais e deveria suspender o juízo, ou suspende o juízo quando o risco é baixo e deveria formar crença (seja em P ou  $\sim$ P, conforme o caso). Em um caso no qual o agente desempenha bem em segundo nível, temos a meta-aptidão. Há, aí, o segundo nível de sucesso epistêmico, de segunda ordem, que é a metacrença apta (ou crença de segunda ordem apta), e consiste em conhecimento reflexivo.

O sucesso completo não é apenas precisão e metaprecisão, mas aptidão plena, aptidão derivada da meta-aptidão. Isso significa que o agente manifesta sua

---

<sup>19</sup> No original: “For a competence-based virtue epistemology, then, animal knowledge requires apt belief, but reflective knowledge requires more. Reflective knowledge requires apt belief that is meta-competent (and, indeed, meta-apt). The knower must have a second-order grasp — a belief or presupposition — that his first order belief is or would be apt”.

competência que é guiada por sua metacompetência naquela ocasião.

Assim, conhecimento humano é constituído de conhecimento animal e conhecimento reflexivo, que, juntos, constituem o conhecimento pleno, que veremos abaixo.

Conhecimento humano está, então, em dois níveis. Todo conhecimento animal humano exige um grau de meta-aptidão, de algum nível mínimo. (Cognição animal bruta exige um nível substancialmente mais baixo). Conhecimento reflexivo, por contraste, exige a operação das capacidades reflexivas humanas. Estas, nas capacidades superiores, envolvem a reflexão consciente sofisticada demandada, em última instância, pelo exercício da sabedoria racional (Sosa, 2011: 93, nota de rodapé 17)<sup>20</sup>.

Há relações específicas entre conhecimento e metaconhecimento, bem como entre aptidão e meta-aptidão epistêmica. Temos, como visto anteriormente, uma independência entre aptidão e meta-aptidão, que se segue na relação epistêmica.

A crença apta pode ocorrer sem meta-aptidão; o agente pode crer verdadeiramente sem considerar o risco envolvido. Também o agente epistêmico pode ser meta-aptos sem ter crença apta. Ele pode falhar no desempenho da formação de crença, manifestando mal a competência epistêmica de primeira ordem e crendo em uma falsidade.

---

<sup>20</sup> No original: "Human knowledge is then on two levels. All human animal knowledge requires a degree of meta-aptness, of some minimal level. (Brute animal cognition requires a substantially lower level.) Reflective knowledge by contrast requires the operation of human reflective capacities. These, at the higher reaches, involve the sophisticated conscious reflection demanded ultimately by the exercise of rational wisdom".

Alternativamente, pode suspender o juízo adequadamente e não seguir com a formação da crença, não havendo desempenho em primeira ordem. A suspensão de juízo mostra como o segundo nível pode interferir no primeiro, até mesmo suspendendo sua atividade adequadamente – assim como visto no exemplo “primo Vinny na prisão”.

A suspensão de juízo adequada é resultado da relação na qual a competência de segunda ordem interfere na competência de primeira ordem, suprimindo sua manifestação, a fim de que o agente epistêmico evite erros. Há um resultado mais positivo dessa relação, na qual a meta-aptidão guia a aptidão, e que leva ao terceiro nível de sucesso. A relação entre conhecimento animal e conhecimento reflexivo, que resulta no conhecimento pleno, é estabelecida através da manifestação de aptidão derivada da meta-aptidão:

Conhecimento animal é crença apta de primeira ordem. Conhecimento reflexivo é conhecimento animal afirmado aptamente pelo sujeito. Podemos ver agora que saber algo plenamente requer que o sujeito tenha conhecimento animal e reflexivo sobre isso, mas também que o saiba com aptidão plena. Requer, por assim dizer, que a correção das crenças de primeira ordem do sujeito manifeste não apenas o animal, competências de primeira ordem que, de modo suficientemente confiável, produzem a correção das crenças formadas. As crenças de primeira ordem do sujeito ficam aquém, se não são guiadas apropriadamente pela metacompetência relevante do sujeito. Essa metacompetência governa se o sujeito deveria em absoluto formar ou não uma crença sobre a questão em disputa ou deveria ao invés suspender o juízo completamente. É somente se essa metacompetência está operante na formação de crença do sujeito sobre esse assunto que a sua crença pode alcançar as alturas epistêmicas. A crença de primeira ordem do sujeito

é apta na proporção de quão confiável é a competência de primeira ordem manifestada em seu sucesso. Além disso, ela é mais plenamente apta na proporção de quão confiável é a metacompetência que seu sucesso também manifesta. Essa metacompetência é manifestada em uma distância, entretanto, porque o metaconhecimento de que uma crença é suficientemente provável para ser apta no nível de base é constituído pelo fato de que a correção da metacrença correspondente manifesta, ela mesma, a metacompetência relevante do sujeito (Sosa, 2011: 11s)<sup>21</sup>.

O conhecimento pleno, ou conhecer plenamente (*knowing full well*) é, para Sosa, o exemplar de conhecimento humano<sup>22</sup>. Ele é resultado da relação entre o conhecimento

---

<sup>21</sup> No original: “Animal knowledge is first-order apt belief. Reflective knowledge is animal belief aptly endorsed by the subject. We can now see that knowing something full well requires that one have animal and reflective knowledge of it, but also that one know it with full aptness. It requires, that is to say, that the correctness of one's first-order belief manifest not only the animal, first-order competences that reliably enough yield the correctness of the beliefs produced. One's first order belief falls short if it is not appropriately guided by one's relevant meta-competence. This meta-competence governs whether or not one should form a belief at all on the question at issue, or should rather withhold belief altogether. It is only if this meta-competence is operative in one's forming a belief at all on that subject matter that one's belief can reach the epistemic heights. One's first-order belief is apt in proportion to how reliable is the first-order competence manifest in its success. What is more, it is more fully apt in proportion to how reliable is the meta-competence that its success also manifests. This meta-competence is manifest at a remove, however, because the meta-knowledge that it is a belief likely enough to be apt on the ground level is constituted by the fact that the correctness of the corresponding meta-belief itself manifests the subject's relevant meta-competence”.

<sup>22</sup> Sosa (2011) tem uma nota de rodapé sobre a escolha do termo “knowing full well”, cf. nota 16, p. 92s. Nele, aceita o uso de “conhecimento pleno”, bem como outros usos de pleno (*fully*), como

---

de segunda ordem, reflexivo – uma manifestação bem-sucedida de metacompetência epistêmica, que avalia como adequada a formação de crença em um ambiente –, e o conhecimento de primeira ordem, animal – que é a formação de crença verdadeira, que manifesta a competência epistêmica –, aquela crença que é verdadeira porque buscou a verdade corretamente, guiada pela metacrença adequada. Essa relação é a do conhecimento animal com o conhecimento reflexivo, que gera o conhecimento pleno. Esse é o terceiro nível de sucesso epistêmico: conhecimento pleno é o conhecimento animal que é guiado pelo conhecimento reflexivo.

A aptidão da crença é gradual, é produto da confiabilidade da competência manifestada. De igual modo, ocorre com a meta-aptidão, conforme a confiabilidade da metacompetência manifestada. A avaliação de sucesso no desempenho de crer, para Sosa, está ligada à obtenção de crenças verdadeiras. O objetivo de crer verdadeiramente é atingir conhecimento animal, em primeira ordem, e conhecimento reflexivo, em segunda ordem. A relação adequada faz com que o resultado seja conhecer plenamente – o valor epistêmico mais alto. O conhecimento pleno é aquele plenamente apto, que resulta do desempenho apto guiado pelo desempenho meta-apto.

Voltemos ao “Julgamento Olímpico”, para exemplificar agora toda a estrutura epistêmica e os níveis de conhecimento.

O juiz Sexista não tem conhecimento sequer em primeira ordem, pois falha em buscar a verdade devido ao viés sexista. Apesar de ter crença verdadeira, ele não é hábil por não buscar a verdade, portanto não é apto, tampouco meta-apto por não avaliar as competências de seus julgamentos e corrigir o viés (falta de habilidade para

---

metafórico, no sentido de que é possível ter conhecimento ainda melhor, de que há espaço para melhorias.

judgar). Atribuímos seu fracasso à falta de habilidade na formação da crença, uma falha na execução, que é seguida de falha na avaliação, na qual ele não percebe seu viés. Ele constantemente falha por falta de habilidade, não sendo sequer funcional.

O juiz Distráido também não tem conhecimento animal; apesar de buscar a verdade e de atingir crença verdadeira, ele falha em obter a relação adequada, pois não obtém crença verdadeira por buscar a verdade, mas sim por acidente – seu julgamento é cópia da colega, juíza Justa. Assim, ele tem precisão e habilidade, mas não aptidão. Seu fracasso é devido à insuficiência da sua habilidade naquele ambiente, que se origina na falha da metacompetência, por não avaliar o risco do ambiente (distrair-se com o sinal sonoro e perder a visualização do tiro). Ele acerta eventualmente, é parcialmente funcional, porque não avalia adequadamente os riscos quando desempenha, e não é reflexivo, ficando suscetível a desempenhar em ambientes onde sua competência não será suficiente.

A juíza Justa é a única que obteve a verdade (crença precisa) porque buscou-a ao julgar (crença hábil), sendo apta, ou seja, tendo conhecimento animal. Ela avaliou bem a ocasião, sem cair em viés (falha pela competência) nem se distrair (falha pelo ambiente), sendo meta-apta, tendo conhecimento reflexivo. Como o conhecimento reflexivo guiou o conhecimento animal, podemos afirmar que juíza Justa tem conhecimento pleno ou conhece plenamente.

Por fim, cabe explicitar que esses níveis constituem uma escala de valor referentes à crença e ao conhecimento:

Crença apta observada aptamente, conhecimento reflexivo, é melhor do que mera crença apta ou conhecimento animal, especialmente quando o conhecimento reflexivo ajuda a guiar a crença de primeira ordem, de modo que ela é apta. Em tal

---

caso, a crença é plenamente apta e o sujeito sabe plenamente (Sosa, 2011: 12-13)<sup>23</sup>.

A disputa do valor do conhecimento não é o enfoque aqui, mas a observação é pertinente para o objetivo de agência epistêmica de Sosa. Essas definições formam o quadro de normatividade epistêmica. A satisfação dessas normas requer que o agente tenha uma série de desempenhos e avaliações sobre eles, manifestando adequadamente sua competência epistêmica. Vejamos a seguir a concepção de agência epistêmica, a função que tem dentro desse cenário, bem como se ela é robusta.

### **Ordens de agência epistêmica**

Seguindo a estrutura em dois níveis, há também duas ordens de agência epistêmica.

#### **I) Primeira ordem**

Sosa apresenta uma concepção de agência epistêmica relacionada aos desempenhos que visam a objetivos, através da manifestação de competências, motivados por razões epistêmicas. O ganho teórico em deixar para trás problemas tradicionais da agência doxástica, como a ação e o voluntarismo, permite superar a restrição de uma concepção mais fraca.

Ser um perseguidor da verdade quanto a uma questão exige o exercício de agência epistêmica e competência na formação de sua crença. E alguém pode, assim, formar crenças racionalmente, mesmo

---

<sup>23</sup> No original: “Apt belief aptly noted, reflective knowledge, is better than mere apt belief or animal knowledge, especially when the reflective knowledge helps to guide the first-order belief so that it is apt. In such a case the belief is fully apt, and the subject knows full well”.

que não deliberadamente, nem voluntariamente, nem mesmo conscientemente (Sosa, 2011: 19)<sup>24</sup>.

A agência epistêmica de primeira ordem consiste em desempenhos do tipo *funcionamento*: formar crença conforme as evidências disponíveis, com a manifestação das competências epistêmicas de primeira ordem. As competências epistêmicas de primeira ordem são constituídas das diversas subcompetências que conduzem os processos doxásticos do sujeito. Seu exercício adequado é a aptidão e leva o agente a adquirir conhecimento animal: “Uma crença é um caso de conhecimento se e somente se é apta à verdade, isto é, se e somente se sua precisão manifesta a habilidade epistêmica do crente” (Sosa, 2011: 24)<sup>25</sup>.

Ela é uma agência epistêmica funcional, na qual o desempenho é assim avaliado: “Nesse quadro funcional, distinguimos entre desempenhos que satisfazem ao menos os padrões mínimos de operações epistêmicas apropriadas e aqueles que não os satisfazem” (Sosa, 2012: 24). Igualmente, os agentes têm uma avaliação funcional nessa ordem de agência epistêmica: “Agentes que meramente *funcionam* estão sujeitos a falhas ou responsabilizações, mais do que a pecados ou infrações, ou a outras violações que pressupõem liberdade” (grifo original, Sosa, 2012: 24). A avaliação de um é indiretamente a avaliação de outro, mas ambas se dão em termos funcionais.

Essa ordem de agência não comporta espaço para liberdade nem para a suspensão de juízo. É a capacidade de

---

<sup>24</sup> No original: “To be a pursuer of the truth on a question requires the exercise of epistemic agency and competence in one’s belief formation. And one can thus form beliefs rationally even if not deliberately, nor voluntarily, nor even consciously”.

<sup>25</sup> No original: “A belief is a case of knowledge if and only if it is truth-apt, i.e., iff its accuracy manifests the believer’s epistemic adroitness”.



---

colocar-se em posição de formar crenças conforme a ocasião. Emprega o princípio de racionalidade meios e fins<sup>26</sup>, mas tem seu fim estabelecido como a verdade: “Finalmente, os funcionamentos, na nossa região intermediária [entre os sofrimentos e meros fazeres e a liberdade das diligências], são avaliáveis num aspecto epistêmico específico, o que significa que a sua avaliação deve ser feita com relação à verdade” (Sosa, 2012: 25).

Nossa competência padrão nos conduz racionalmente à verdade. Mas podemos ser desviados da verdade, através de vieses, ou motivados por interesses práticos. Nos casos em que podemos não estar sujeitos à força da razão, necessitamos da agência de segunda ordem, a fim de preservarmo-nos de falhas.

## II) Segunda ordem

A agência epistêmica de segunda ordem guia o exercício adequado de primeira ordem. Ela é um desempenho do tipo *diligência*, em busca da verdade. Acompanha a agência de primeira ordem constantemente, mas é disparada por problemas ou desvios na formação de crenças. Esses problemas podem ter três origens: i) riscos de engano por um ambiente inadequado; ii) redução ou impedimento da competência; e iii) conflitos de razões (práticas e epistêmicas). Podemos afirmar, portanto, que Sosa atribui três papéis distintos para sua atuação:

i) avaliação do ambiente e da competência epistêmica de primeira ordem;

---

<sup>26</sup> O princípio é assim formulado por Sosa: “S visa a fazer X, ao fazer Y, SSE S faz Y no esforço para fazer X” (2011: 18). No original: “S aims to X, by Y'ing, IFF S Y's in the endeavor to X”.

ii) decisão de formar crença ou suspender o juízo (desempenhar ou abster-se);

iii) separação entre razões práticas e epistêmicas, através da reflexão.

Cada um desses papéis liga-se a um nível de sucesso epistêmico e implica um tipo de conhecimento (animal, reflexivo e pleno). Examinemos cada um deles.

i) Meta-avaliação ambiental e funcional

O primeiro papel da agência de segunda ordem consiste na avaliação do risco de falha pelo ambiente e das competências de primeira ordem. Ele é um nível não manifestado, silencioso ou implícito.

Os dois primeiros problemas que desencadeiam a agência reflexiva são relativos aos componentes da competência, à situação e à condição do agente, respectivamente. O primeiro é acerca de situações enganadoras, como os exemplos céticos e os casos do tipo Gettier trazem-nos. O segundo é sobre o agente estar impedido ou incapaz, ainda que temporariamente, de manifestar sua habilidade. Vejamos esses pontos.

Sosa afirma que as condições contextuais *podem* exercer um papel comprometedor em relação à competência epistêmica. Com isso, está afirmando a falibilidade das nossas competências – podemos estar errados, mesmo em situações em que não percebamos motivos para pensar que estejamos, pois podemos falhar em ambas as ordens (é preciso que *haja*, de fato, esses motivos, e que seja falha do agente não percebê-los, caso contrário, estaremos em um cenário cético).

Porém, relembando os condicionais disparadores de manifestações, o que interessa é a ocasião atual. Sosa descarta as demais ocasiões espaço-temporais e modais, por

---

serem irrelevantes. O necessário é que o agente tenha manifestado sua competência plena na situação na qual se encontra:

Raramente, se alguma vez, uma reconhecida disposição do senso comum exige que suas condições de desencadeadoras desencadeariam suas manifestações resultantes não somente no próprio tempo e lugar em que o sujeito da disposição está localizado, mas também ao longo de uma vizinhança mais ampla de tempos e espaços (Sosa, 2011: 80)<sup>27</sup>.

E essa restrição vai além do mundo físico atual, abrangendo mundos possíveis (modais)<sup>28</sup>:

Quer definamos a vizinhança por proximidade física ou por proximidade modal, não é plausível que uma competência, perícia ou disposição seja manifestada em um certo local somente se o agente tivesse sucedido similarmente em qualquer lugar em geral na vizinhança (Sosa, 2011: 91)<sup>29</sup>.

A agência de segunda ordem avalia constantemente nosso ambiente e funcionamento, seja implicitamente,

---

<sup>27</sup> No original: “Rarely if ever does a recognized, commonsense disposition require that its triggering conditions would trigger its resultant manifestations not only at the very place and time where the host of the disposition is located, but also throughout some wider neighborhood of places and times”.

<sup>28</sup> Para uma introdução, ver *Possible worlds*, da SEP, disponível em: <<http://plato.stanford.edu/entries/possible-worlds/>>.

<sup>29</sup> No original: “Whether we define the neighborhood by physical proximity, or by modal proximity, it is not plausible that a competence, skill, or disposition is manifest at a certain location only if the host would have similarly succeeded elsewhere generally in the neighborhood”.

quando não há sinais de inadequação, seja reflexivamente, quando há. Nos primeiros, não há disparador, então ela não é manifesta, ainda que presente. Esses casos são os que desempenhamos ordinariamente, nossa agência funcional desempenha bem e a agência reflexiva fica ociosa. Não há complicações ou necessidade de ir a outros níveis de conhecimento. É o papel silencioso da razão, pelo qual adquirimos conhecimento animal – cuja execução é frequente, constantemente formamos conhecimento sobre o mundo no qual estamos e agimos.

ii) Decisão de desempenho ou abstenção

Quando há sinais de risco ambiental ou falta de competência funcional, é preciso ir para o papel reflexivo de decisão manifestada. Esse segundo papel implica, portanto, o primeiro papel cumprido – percebemos os riscos porque os monitoramos.

Quando a agência de segunda ordem é manifestada, forma-se a metacrença de aceitação ou rejeição do risco presente na ocasião e decide-se por manter ou formar<sup>30</sup> as crenças-alvo ou por suspender o juízo.

Porém, mesmo essa ponderação de razões não precisa ser consciente, nem ligada à vontade do agente, portanto não precisa ser controlada positivamente. São os casos em que os processos de formação de crenças seguem sem sinais de problemas ou desvios, há somente deliberação de razões epistêmicas.

---

<sup>30</sup> Há casos nos quais a crença foi formada, e a metacompetência indica que há risco de ela ser falsa. Nesses casos, é adequado falar em sustentar/manter ou revisar/rejeitar a crença. Porém, esses casos são análogos aos de formação para o que nos interessa. Assim, por uma questão de economia, iremos nos referir somente à formação, pois o que se aplica a uma, aplicar-se-á a outra.

Alegadamente, a competência relevante de segunda ordem é uma competência padrão que assume a competência de primeira ordem a menos que haja sinais de alerta para o contrário. Tais sinais podem derivar do testemunho contrário, ostensivamente confiável, ou de alguma outra fonte de incoerência. Alguém submetido ao teste Muller-Lyer começa a suspeitar de sua competência de primeira ordem, por exemplo, quando a medição colide com a aparência visual<sup>31</sup>.

Nossa competência de segunda ordem tem, então, um componente constitucional que responde com suspeita quando impelido por sinais de alerta e, com confiança, na luz de sua ausência. A competência interior completa exigirá, é claro, que o sujeito esteja em boa forma. Ele não pode estar bêbado, drogado, dormindo ou incapacitado de outra forma. Além disso, e aqui é o ponto crucial, a competência de segunda ordem plena exigirá também que ele esteja adequadamente situado. Assim como a competência plena de primeira ordem de visão de cor exige boa luz, também a competência plena de segunda ordem exige que o sujeito esteja adequadamente situado. Com o intuito de ser competente na segunda ordem, o sujeito deve estar situado de tal forma que, mais cedo ou mais tarde, haveria sinais de alerta se a competência de primeira ordem relevante estivesse ausente (Sosa, 2010: 473)<sup>32</sup>.

---

<sup>31</sup> A ilusão de ótica Muller-Lyer consiste em duas linhas paralelas do mesmo tamanho, mas com ângulos invertidos em suas extremidades. Isso leva à percepção de que são de tamanhos distintos.

<sup>32</sup> No original: “Arguably, the relevant second-order competence is a default competence that assumes first-order competence unless there are tell-tale signs to the contrary. Such signs might derive from contrary, ostensibly reliable testimony, or from some other source of incoherence. A Muller-Lyer subject begins to suspect his first-order competence, for example, when measurement clashes with visual appearance.

Há uma reflexão sobre as razões, ainda que implícita, que leva à decisão de crer, descrever ou suspender o juízo, bem como sobre a maneira pela qual nós chegamos a esse julgamento ou suspensão:

Ao raciocinar a partir de tais comprometimentos implícitos [de julgamentos de primeira ordem], nós podemos, ao fim, ganhar atenção consciente (*conscious awareness*) de nossas faculdades e subfaculdades. Podemos talvez assim ganhar conhecimento consciente da sua natureza e de como elas tornam possível nosso sucesso cognitivo. E isso pode ajudar a prover uma perspectiva que subscreve, com entendimento coerente, nosso uso daquelas faculdades (Sosa, 2010a: 189)<sup>33</sup>.

Estar bem situado e em boas condições em segunda ordem é o que nos possibilita atingir esses níveis de raciocínio e até mesmo atenção consciente. Esse segundo

---

Our second-order competence has thus a constitutional component that responds with suspicion when prompted by tell-tale signs, and with trust in light of their absence. The fuller inner competence will of course require that one be in good shape. One must not be drunk or drugged or asleep or otherwise disabled. In addition, and here is the crucial point, the complete second-order competence will require also that one be properly situated. Just as the first-order complete color-vision competence requires good light, so the complete second order competence requires that the subject be properly situated. In order to be competent on the second-order, one must be so situated that sooner or later there would be tell-tale signs if the relevant first-order competence were absent”.

<sup>33</sup> No original: “By reasoning from such implicit commitments we may eventually gain conscious awareness of our faculties and sub-faculties. We may thus gain conscious knowledge of their nature and of how they enable our cognitive success. And this can help provide a perspective that underwrites, with coherent understanding, our use of those faculties”.

---

papel da agência reflexiva está ligado ao conhecimento reflexivo, que implica no conhecimento da competência de primeira ordem, no metaconhecimento ambiental e na suspensão de juízo racional.

### iii) Separação das razões

O terceiro papel de agência reflexiva requer o sucesso relacional nos dois primeiros. É preciso que haja avaliação ambiental e funcional adequada e que se tenha decidido por formar crença por causa disso.

O nosso arrazoadado sugere uma distinção entre tal crença animal e a crença reflexiva (*reflective*). A crença animal é constituída por um estado armazenado que pode guiar a conduta subconscientemente. A crença reflexiva, por contraste, é uma disposição para julgar afirmativamente em resposta a uma pergunta – se sob a influência de nenhum objetivo além daquele de responder corretamente – com verdade. É esse “juízo” que a pessoa está disposta a emitir é um ato consciente distinto ou um estado conscientemente mantido. É esse ato ou estado que está no âmbito do nosso controle livre quase tão frequentemente e plausivelmente quanto o estão as escolhas que fazemos ordinariamente e as intenções conscientes que mantemos (Sosa, 2012: 34).

Seu exercício adequado leva o agente a não se guiar por razões práticas, nem permitir que elas interfiram na consideração das razões epistêmicas na formação da crença.

Quando há sinais de risco, o agente avalia que ele é baixo e decide pela formação de crença; podemos exercer a agência reflexiva nesse terceiro papel. O agente, então, permanece no padrão de formação de crença em primeira ordem e tem de manter a primazia das razões epistêmicas

para ser bem-sucedido e obter conhecimento, ou falha em separar as razões e é guiado pelas razões práticas (autoconforto, pensamento mágico, etc.).

Esse é o papel no qual a agência epistêmica usa o conhecimento reflexivo para guiar o conhecimento animal, cujo exercício resulta no conhecimento pleno. A agência epistêmica de segunda ordem, assim manifestada, pode resultar em controle sobre as crenças. Vamos examinar mais profundamente essa possibilidade.

### **Agência epistêmica reflexiva**

A agência epistêmica consiste em desempenhos epistêmicos; ela tem duas ordens, correspondentes aos funcionamentos e às diligências. Ela tem três papéis na vida cognitiva: monitoramento de sinais de risco, decisão de aceitação do risco e controle negativo sobre influências práticas.

As diligências podem ser livremente determinadas, mas nem sempre o são. Nos casos em que o são, temos a agência epistêmica livre manifestada, segundo Sosa (2012: 32s). Ela é a mais forte, mas ainda não envolve nenhuma ação ou controle direto:

Além do mais, competência epistêmica [no sentido de agência] não é acional. Não é instituída através de raciocínio direto, positivo, prático-silogístico. [...] Em vez disso, competência epistêmica resulta em crenças, mas não através de decisões práticas ou silogismos. Então, em um sentido, somos passivos, não diretamente ativos, na nossa formação de crenças. Apesar disso, nossas crenças estão sob nosso controle (Sosa, 2011: 32)<sup>34</sup>.

---

<sup>34</sup> No original: “Moreover, epistemic competence is not actional. It is not instituted through positive, direct practical-syllogistic reasoning. [...] Instead, epistemic competence results in beliefs, but not through



---

Os dois primeiros papéis da agência reflexiva são constituídos por metacompetência, e derivam seu sucesso dela: monitorar o ambiente e decidir sobre a confiabilidade das competências de primeira ordem, que pesam as evidências e formam as crenças. Porém, Sosa define um terceiro papel para a agência reflexiva: a separação das razões. Como dá-se esse controle sobre as razões? De onde vem a capacidade de evitar razões práticas e resguardar as razões epistêmicas, a fim de que nossa agência de primeira ordem manifeste nossa competência? Como podemos controlar as crenças sendo passivos na sua formação?

Sosa parece apelar para uma concepção de agência supressiva. Vejamos a definição de Peter Hacker:

[...] um agente pode também fazer algo que impede ou cessa a mudança que, de outro modo, ocorreria. Em tais casos de atuação impeditiva ou supressiva, o paciente [objeto da agência] é impedido de mudar ou mudar ainda mais. Isso pode ser feito pela ação no paciente, interferindo ou agindo em alguma outra coisa que, de outro modo, teria agido sobre o paciente (Hacker, 2010: 135).

Vejamos também a caracterização negativa:

Observe-se que o resultado da agência impeditiva ou supressiva é a não ocorrência de um evento que, de outro modo, teria ocorrido, o não início ou o término de um processo que, de outro modo, teria começado ou continuado, e a falha na obtenção de um estado que teria sido o caso de outro modo (Hacker, 2010: 151).

---

practical decisions or syllogisms. So, in one sense we are passive, not directly active, in our belief formation. Nevertheless, our beliefs are under our control”.

Assim, a agência reflexiva tem seu último papel como impeditiva de razões práticas na agência epistêmica funcional. Ela atua somente impedindo interferência nos casos em que há esse risco. O domínio epistêmico é resguardado dos interesses práticos e o conhecimento é adquirido ao evitar-se vieses e outros problemas de origem prática. Mas como a reflexão pode caracterizar esse tipo de agência? Qual a base para a manutenção de nossa agência epistêmica funcional?

Quando temos bons funcionamentos, somos bons agentes, consideramos as razões epistêmicas e formamos crenças adequadamente; não necessitamos da manifestação da agência reflexiva, ela permanece em um estado de monitoramento, sem necessitar de decisão alguma, pois temos bons desempenhos até que haja sinais para pensarmos o contrário.

Mas, quando esses sinais são percebidos, os desempenhos funcionais já estarão comprometidos com as razões epistêmicas que os motivam em busca da verdade, contrapostas a quaisquer razões práticas que os possam desviar. Então, seria inviável usar essas mesmas razões epistêmicas e contrapô-las às razões práticas, para permanecer na busca da verdade. É preciso que haja um segundo nível de motivação para a busca da verdade, que possibilite ao agente fundamentar sua agência epistêmica reflexiva:

Sendo assim, dois tipos de fundamentação motivacional podem subjazer a uma atitude que mantemos. A primeira é uma fundamentação constituída por razões baseadas nas que o sujeito sustenta a atitude. A segunda é uma fundamentação constituída por razões baseadas nas que nós nos abtemos de impedir (ou de mudar ou afetar de outro modo) a atitude, onde a atitude é então sustentada por padrão. O desejo puro,

---

desinteressado, pela verdade pode motivar racionalmente a crença de alguém da segunda maneira. Pode ser uma razão na qual alguém se baseia para abster-se de afetar sua crença na busca de objetivos pragmáticos, como o conforto. É claro, tudo isso é compatível com a noção de racionalidade epistêmica pura e com a possibilidade de que uma crença seja epistemicamente irracional, embora racional todas as coisas consideradas, onde essa última deve ser entendida como racional todas as considerações práticas feitas, inclusive o desejo pela verdade na questão em voga (Sosa, 2011: 33-34)<sup>35</sup>.

É o desejo de busca da verdade, livremente escolhido, em segunda ordem, que orienta o agente para resguardar sua agência primária. Porém, a liberdade na definição dos fins de uma diligência poderia levar à arbitrariedade. E, no domínio epistêmico, arbitrariedade não pode ser aceita; nossas atitudes e processos têm de ser motivados racionalmente: “*Diligências fundacionais* estariam desprovidas de bases motivadoras racionais para explicar o seu estatuto como racionais. Tais diligências pareceriam, assim, arbitrárias, e, portanto, irracionais” (grifo original,

---

<sup>35</sup> No original: “All that being so, two sorts of motivational rationale can underlie an attitude that we hold. One is a rationale constituted by reasons based on which one holds the attitude. The second is a rationale constituted by reasons based on which we forbear from precluding (or from changing or otherwise affecting) the attitude, where the attitude is then held by default. The pure, disinterested desire for truth can rationally motivate one's beliefs in the second way. It can be a reason based on which one forbears from affecting one's beliefs in the pursuit of pragmatic goals such as comfort. Of course, that is all compatible with a notion of purely epistemic rationality, and with the possibility that a belief be epistemically irrational though rational all things considered, where this last is to be understood as rational all practical considerations considered, including the desire for the truth on the question at hand”.

Sosa, 2012: 26). Como é possível, então, para uma diligência epistêmica livre não ser acusada de irracionalidade?

Sosa responde que “*Diligências* epistêmicas livremente racionais podem basear-se em desempenhos *funcionalmente* racionais” (grifo original, *ibidem*). Ou seja, a racionalidade pode basear-se no funcionamento, que é motivado pelas razões epistêmicas próprias daquela ocasião: “*Funcionamentos fundacionais* não partilham do mesmo estatuto problemático. Não envolvendo escolhas de todo, não envolvem escolha arbitrária” (grifo original, *ibidem*).

Além disso, formamos compromissos através das nossas disposições baseadas na vontade. Esses compromissos orientam nossas escolhas, manifestados nas ocasiões adequadas, em que devemos implementar atitudes específicas, a partir das disposições voluntárias. Igualmente, temos disposições voluntárias livres para implementar tais e tais desempenhos em ocasiões adequadas:

Aqui não adquirimos uma crença animal por meio do funcionamento apropriado “automático”, não livre, dos nossos mecanismos cognitivos. O compromisso, ao invés disso, é escolhido livremente, tal que o “mecanismo” é ele mesmo voluntário. O que esse compromisso governa, contudo, não é crença animal armazenada automaticamente por meio do funcionamento apropriado das nossas faculdades animais. Antes, os nossos compromissos livremente escolhidos governam os nossos juízos livres quando sopesamos conscientemente uma questão à luz de todas as razões sincronicamente em vista (Sosa, 2012: 34).

Uma vez que tenhamos o compromisso de buscar a verdade sobre a questão, orientamos nossa agência reflexiva para atuar supressivamente e resguardar a agência

epistêmica padrão (primeira ordem). Através dessa derivação, é que escolhemos buscar a verdade, deliberamos nosso ambiente e competências de primeira ordem e, assim, controlamos nossas crenças animais: “Agora [após a deliberação reflexiva], podemos sopesar novamente a crença no nosso resultado. E aqui podemos de novo exercer liberdade, à medida que está em nossas mãos acreditar no resultado. Dessa maneira, nós podemos controlar as nossas crenças animais” (Sosa, 2012: 33).

A agência reflexiva nesse último papel é livre, racional e garante nossos funcionamentos epistêmicos (agência de primeira ordem), através do impedimento de razões práticas.

## Referências

Hacker, Peter M. S. (2010). *Natureza Humana: Categorias Fundamentais*, tradução de José Alexandre Durrty Guerzoni, Porto Alegre: Artmed.

Ichikawa, Jonathan J.; Steup, Matthias (edição da primavera de 2014). The Analysis of Knowledge. In: Zalta, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL:

<<http://plato.stanford.edu/archives/spr2014/entries/knowledge-analysis/>>.

Menzel, Christopher (edição do verão de 2015). Possible Worlds. In: Zalta, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Publicado na internet, disponível através da seguinte URL:

<<http://plato.stanford.edu/archives/sum2015/entries/possible-worlds/>>.

Sosa, Ernest (1991). *Knowledge in Perspective: Selected Essays in*

*Epistemology*, New York: Cambridge University Press.

Sosa, Ernest (2007). *A Virtue Epistemology: Apt Belief and Reflective Knowledge*, vol. I, New York: Oxford University Press.

Sosa, Ernest (2010). 'How Competence Matters in Epistemology', *Philosophical Perspectives*, 24, Epistemology: 465-475.

Sosa, Ernest (2010a). 'Ernest Sosa'. In: Dancy, Jonathan; Sosa, Ernest; Steup, Matthias (Eds.). *A Companion to Epistemology*, 2nd ed., Malden MA: Wiley-Blackwell, 185-190.

Sosa, Ernest (2011). *Knowing Full Well*, Princeton: Princeton University Press.

Sosa, Ernest (2012). 'Agência Epistêmica', tradução de Rafael Lopes Azize, *Caderno CRH*, 25(2): 23-37.

Sosa, Ernest (2013). 'Epistemic Agency'. *The Journal of Philosophy*, CX(11): 585-605.

Turri, John (2013). Bi-Level Virtue Epistemology. In: Turri, John (Ed.). *Virtuous Thoughts: the Philosophy of Ernest Sosa*, Philosophical Studies Series 119, Dordrecht: Springer, 147-164.