

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS  
CURSO DE GRADUAÇÃO EM FILOSOFIA

**EU NÃO CONHEÇO O TÍTULO:  
O PARADOXO DO EXAME SURPRESA E ALGUNS LIMITES DO CONHECIMENTO**

JUAN MAIZ LULKIN FLORES DA CUNHA

Dissertação apresentada como requisito parcial para  
a obtenção do grau de Bacharel em Filosofia

Orientador: Prof. Dr. Jaime Parera Rebello

Porto Alegre, Janeiro de 2013.

## ***SUMÁRIO***

<b>INTRODUÇÃO.....</b>	<b>4</b>
O paradoxo do exame surpresa.....	4
Uma proposta de classificação das soluções.....	6
Objetivo.....	9
<b>1. AS SOLUÇÕES PRAGMÁTICAS.....</b>	<b>10</b>
1.1. Proposições e elocuições.....	10
1.2. Futuros contingentes.....	12
1.3. Desconfiando do professor.....	14
1.4. A solução prática.....	16
<b>2. AS SOLUÇÕES LÓGICAS .....</b>	<b>21</b>
2.1 Condicionais problemáticos.....	21
2.2. Declarações e ordenamentos.....	23
2.3. Autorreferência.....	24
2.4. O conhecedor .....	27
<b>3. AS SOLUÇÕES EPISTÊMICAS .....</b>	<b>37</b>
3.1. As sete cartas.....	37
3.2. Paradoxo da loteria.....	39
3.3. Variações recalcitrantes.....	44

3.4 Os blindspots condicionais.....	53
3.5 Recalcitrâncias domadas .....	55
3.6. Triatlon .....	57
<b>5. CONCLUSÃO .....</b>	<b>66</b>
5.1. Conclusão.....	66
<b>REFERÊNCIAS BIBLIOGRÁFICAS:.....</b>	<b>70</b>

## INTRODUÇÃO

### O paradoxo do exame surpresa

*Que maravilhoso termos nos deparado com um paradoxo. Agora temos alguma esperança de progredir.<sup>1</sup> – Niels Bohr*

NO FINAL DE UMA SEMANA DE AULAS, O PROFESSOR AVISA SEUS ALUNOS: “NA PRÓXIMA SEMANA teremos uma prova surpresa”. Ao que um aluno, muito astuto, responde: “Professor, tal prova é impossível”. Surpresos, o professor e os outros alunos questionam o motivo e o aluno astuto explica:

“Se o senhor não tiver dado a prova até a quinta-feira, nós saberemos que a prova será na sexta. Sendo assim, ela não será uma surpresa. Logo, a prova não pode ser realizada na sexta. Se o senhor não tiver dado a prova até a quarta-feira, nós saberemos que a prova será na quinta, pois já vimos que não pode ser na sexta. Mas, do mesmo modo, ela não seria uma surpresa. Logo, ela também não pode ser na quinta. Se não tiver dado a prova até terça, então saberemos que a prova será na quarta, visto que não pode ser na quinta nem na sexta. Não sendo mais uma surpresa, a prova não pode ser na quarta. Se a prova não for dada na

---

<sup>1</sup> Todas as traduções são de minha autoria. Apontarei apenas aonde houver dificuldades na tradução.

segunda, por sua vez, saberemos que ela ocorrerá na terça, pois já eliminamos a quarta, a quinta e a sexta. Novamente, a terça não seria uma surpresa. Logo, a prova só pode ser dada na segunda. Mas como já sabemos que ela só pode ser dada na segunda, não será uma surpresa. Sendo assim, tal prova não é possível.”

O professor, no entanto, ignora tal argumento e aplica a prova na terça-feira, surpreendendo o aluno astuto.

O paradoxo do exame surpresa foi apresentado por Daniel John O'Connor<sup>2</sup> em um artigo publicado na revista *Mind* em 1948 com o título *Pragmatic Paradoxes* (O'Connor, 1948:358). Neste, o problema é apresentado em termos de um exercício militar - um *Class A Blackout* - apresentado à população pelo comando militar em um sábado. Tal exercício teria 2 propriedades: 1) ocorreria às 18h em um dia da semana seguinte 2) sendo um exercício *Class A*, não seria possível saber o dia de antemão.

Segundo Paul Franceschi, no entanto, tal coisa ocorreu de fato entre 1933 e 1934, quando as autoridades suecas planejaram implementar um exercício de defesa civil. A rádio, noticiou então um anúncio no qual um exercício de defesa civil ocorreria na próxima semana, porém ninguém poderia saber, de antemão, o dia da ocorrência deste. O matemático Lennart Ekbom foi então quem se deu conta do problema de tal anúncio e o expôs a seus alunos (Franceschi, 2002:1).

Voltando a O'Connor, seu título - *Pragmatic Paradoxes* - deixa clara sua intenção de enquadrar o paradoxo do exame surpresa na categoria dos paradoxos pragmáticos, por oposição aos paradoxos lógicos - como o paradoxo do mentiroso. A diferença, para O'Connor, reside no fato de que o problema apresentado no paradoxo não é um cuja definição é formalmente autocontraditória, mas que ela é “pragmaticamente autorrefutadora”. Isso significa dizer que a tensão fundamental deste problema reside, de algum modo, entre a apresentação e o *ato* de apresentar.

---

2 O problema já era conhecido antes, mas este é o primeiro artigo sobre o assunto.

O'Connor tem apenas a pretensão de chamar a atenção para o paradoxo, e não resolvê-lo. A utilidade filosófica de tal problema seria elucidar problemas relacionados, como o paradoxo suscitado pela proposição “Eu acredito que existem tigres no México, mas não há nenhum”. Este exemplo é premonitório, pois parece sugerir - apesar de não constar explicitamente no texto - que há *uma certa relação* do problema com o paradoxo de Moore, uma alternativa que é comum na bibliografia mais contemporânea e que veremos mais adiante. A formulação do paradoxo de Moore apresentada na carta de Wittgenstein tem precisamente a mesma forma: “Há um fogo nesta sala e eu não acredito que haja” (Dall'Agnol, 2007:11).

O apelo de O'Connor e os artigos que começaram a lidar com o assunto na década de 50 em diante parecem estar imbuídos do espírito do *linguistic turn*, pois o autor argumenta que o valor de tal estudo estaria em, ao menos, tornar um pouco mais claras as formas em que a linguagem ordinária pode nos “limitar” ou nos “enganar”.

### Uma proposta de classificação das soluções

Ao analisar a história do desenvolvimento do problema, vamos observar que os autores buscaram resolvê-lo de forma *bem* diversa. Avishai Margalit e Maya Bar-Hillel em sua resenha sobre o desenvolvimento do paradoxo afirmaram que “a diversidade de opiniões, abordagens e tratamentos que têm sido apresentada sobre o paradoxo é realmente impressionante” (Bar-Hillel e Margalit, 1983:263).

Curiosamente, tanto O'Connor como Cohen e Alexander, os dois autores a endereçar o problema logo a seguir, ignoraram a intuição mais evidente do problema (Bar-Hillel e Margalit, 1983:264), aquela que mostra que, simplesmente dando o exame, digamos, na terça, o professor refuta o argumento do aluno astuto. Estes poucos autores, ignorando a forte intuição de que sua ocorrência é possível, ignorando, como dizia Michael Scriven – o

primeiro a se dar conta do óbvio – “o monstro, a Realidade” (Scriven, 1951:403), disseram que o paradoxo está *inteiramente* no anúncio (Bar-Hillel e Margalit, idem).

Entre as possibilidades de refutação, a maior parte destas focou no problema específico da realização da prova no último dia possível e buscou refutar essa premissa tão intuitiva do argumento. Outros concederam que uma prova no último dia *seria* impossível, e, assim, buscaram refutar a aplicação do mesmo critério iteradamente. Independentemente do ponto onde se deve “barrar” o paradoxo, talvez a melhor classificação possível das soluções alternativas seja a partir sua “natureza”. As mais comuns, assim proponho, são a pragmática, a lógica e a epistêmica. Tal classificação também está de acordo com o trabalho de Bar Hillel e Margalit, que, em seu artigo identificam a primeira fase – mesmo que não a chamando de “pragmática”, aliás, não dando nome algum a nenhuma delas – com o trabalho de O’Connor até Quine em 1953, a fase lógica, a partir de Shaw, em 1958, e o início de uma solução carregada de noções epistêmicas em Kaplan e Montague, em 1960.

Alguns autores que defendem a natureza pragmática do problema buscam entender o anúncio do professor como um proferimento da “vida real”. O que eles querem dizer com isso, em geral, é que uma solução deve, primeiramente, focar nas *intenções* envolvidas na caracterização do problema, ou seja, que o “verdadeiro” significado do anúncio está *para além* das palavras ditas e da estrutura lógica que podemos abstrair destas. Porém, outros estão um tanto mais próximos da interpretação epistêmica, pois buscam resolver o problema apelando ao que é e o que não é razoável que o aluno pressuponha. Nesta a questão é explicitar essas pressuposições e encontrar aquelas que não podem ser sustentadas. Como algumas dessas pressuposições são claramente epistêmicas, como o princípio da retenção do conhecimento ou da memória, o princípio do fechamento epistêmico [*epistemic closure*], ou o princípio da reflexividade do conhecimento (ou “axioma KK”), podemos dizer que a linha divisória entre as posições pragmáticas e as epistêmicas não é perfeitamente nítida. Como veremos, Quine é um dos autores cuja posição é de difícil caracterização nesta classificação proposta, pois ele pretende que o aluno simplesmente não esteja justificado a tomar o anúncio do professor nem mesmo *ex hypothesi* como verdadeiro. É bem verdade, que, na *prática*, não podemos assumir as afirmações de *um professor* como perfeitamente verazes, no

entanto, ler o problema deste modo, como dirão os autores das outras posições, é diminuí-lo, pois podemos simplesmente trocar o professor por um “computador infalível”, ou um “demônio da veracidade” (o nêmesis do demônio enganador de Descartes). Esta seria a caracterização favorita dos “lógicos”. Já os “epistêmicos”, mesmo assumindo que a ideia de um ente perfeitamente veraz é algo meramente “ideal”, estes autores irão, ao menos, defender que os alunos *podem* tomar o anúncio do professor como evidência *ex hypothesi* – mesmo que isto sirva simplesmente para um *reductio ad absurdum* do anúncio.

A posição dos autores que buscaram entender o problema em termos epistêmicos tem como uma de suas principais características entender “surpresa” como “não estar justificado a crer que ocorrerá um exame antes de sua ocorrência”, onde justificação envolve todas as *evidências* disponíveis ao aluno e não meramente a possibilidade de uma dedução lógica a partir da interpretação que este faz do anúncio. O que a maioria das posições epistêmicas tem em comum com as posições lógicas é o apelo à racionalidade máxima do aluno, ou seja, que o problema não se resolva por apelo à possibilidade de uma incapacidade pessoal do aluno em entender sua posição no problema, em realizar os raciocínios corretos e não cair em contradições ou inconsistências (com algumas exceções neste caso).

Os autores que buscaram caracterizar o problema com uma natureza lógica, procuraram, primariamente, associá-lo de algum modo com o paradoxo do mentiroso, gerando formulações que envolvessem uma recursão infinita na qual o seu valor de verdade ficasse variando e não pudesse ser definido.

Existem ainda outras posições e usos bizarros na história do problema. A de Ardon Lyon, que busca, a partir de uma conclusão sobre este paradoxo, mostrar que determinismo e livre arbítrio não são incompatíveis (Lyon, 1959:514), ou a de B. Meltzer, que tenta solucionar o problema apelando ao terceiro excluído (Meltzer, 1964:432). Além destas, encontrei exposições envolvendo teoria dos jogos, lógica epistêmica dinâmica (um ramo relativamente recente da lógica epistêmica que busca formalizar as transformações de estados de conhecimento) e outras ideias ainda mais excêntricas. Sobre tais posições falarei



menos, em maior parte, por pura ignorância e total falta de tempo, e, em menor parte, por considerar crucial apresentar, ao menos, as posições mais convencionais.

## Objetivo

Além da diversidade de soluções, impressiona a diversidade de impressões que vão de “‘insignificante’ a ‘poderoso’” (idem, p. 264), e, também, como o problema se mostrou tenaz. Em 1959, Lyon afirma que a sua abordagem “tem a vantagem de ser tão simples que qualquer um que a ler irá concordar com a solução” (Lyon, 1959:510) e 50 anos depois, Levy, não menos esperançoso, diz que seu *paper* irá “terminar a intriga de uma vez por todas” (Levy, 2009:131) e que sua solução “deve ficar”.

A exposição de Margalit e Bar-Hillel foi a primeira a ter como objetivo principal (porém não exclusivo) apresentar de forma resumida a literatura sobre o problema, no intuito de eliminar a “repetição desnecessária de alguns argumentos”. No entanto, a exposição irá completar 20 anos e de lá para cá soluções mais robustas deixaram o problema mais polido. Além disso, não encontrei apresentação elementar - porém completa - e lusófona do paradoxo e seu desenvolvimento nos últimos 60 anos. Dado que este é apresentado em diversos materiais de estudo introdutórios de filosofia e lógica, creio que este fato é lastimável. Faz-se necessário um mapa atualizado do problema.

Minha pretensão é a de revisar esta literatura, apresentando nos próximos capítulos essa pluralidade de soluções com naturezas tão distintas da forma mais clara possível, para que, no fim, possamos, como disse Bohr, ter alguma esperança de progredir.

## 1. AS SOLUÇÕES PRAGMÁTICAS

### 1.1. Proposições e elocuições

A PRIMEIRA RESPOSTA AO PROBLEMA APRESENTADO POR O'CONNOR FOI FEITA POR LAURENCE Jonathan Cohen, na mesma revista *Mind*, em 1950. A sua solução do se resume a uma simples distinção entre proposição e elocução.

Em seu artigo, O'Connor tinha comparado o paradoxo ao problema sugerido pela afirmação “Eu não lembro de nada”. Tal afirmação não é nem uma contradição lógica, como  $p \wedge \neg p$ , e também não é “meramente ... falsa factualmente”, pois, se, tomamos o “nada” literalmente - absolutamente nada -, não teríamos como afirmar coisa alguma, visto que não lembrariamos nem das palavras, nem de como a linguagem funciona. O primeiro ponto é bem claro, contudo, o segundo parece carecer de mais argumento.

O ponto de O'Connor é precisamente o seguinte: ocorre que *não precisamos de um fato fora a própria afirmação* para sua falsidade. Dizer que “Joãozinho não lembra nada” é falso” depende de *verificarmos* que ele se lembra de algo. Perguntamos: “Joãozinho, você lembra o seu nome?”. Se ele responder “me lembro, é João” teremos justificativa suficiente para dizer que ele se lembra de algo, logo, não é o caso que não se lembre de nada. Já “Eu não me lembro nada” independe. Não preciso me questionar “Eu me lembro de algo?”. Daí a perplexidade de uma afirmação que já *se produz* falsa sem ser logicamente falsa.

Cohen então argumenta que tal perplexidade é desnecessária, pois O'Connor toma "afirmação" (*statement*) de modo ambíguo. Uma afirmação pode ser entendida de dois modos muito distintos. Um modo é a ação que tomamos ao afirmar, que pode ser o ato de escrever, pensar, falar, etc. Este é a sua elocução. Quando nos referimos a uma elocução, faz sentido questionar "quando foi que afirmaste X?" ou "onde estavas quando afirmaste X", do mesmo modo que "quando foi que comeste pela última vez" ou "em que restaurante comeste pela última vez". Como qualquer ato, a elocução está espaço-temporalmente localizada. Neste caso dizer "a sua elocução é verdadeira" tem o mesmo sentido que dizer "seu andar é falso"<sup>3</sup>: nenhum. O outro modo é a proposição em si. Dizer "a proposição é verdadeira", isto sim, parece apropriado. "Quando ocorre a proposição?", inadequado. Mais do que isso, como observa Cohen, a distinção parece muito plausível olhando para como a linguagem ordinária opera, pois sentenças como "Quando é que ele afirmou que 'estava a andar'?" contém essa distinção implícita. Neste exemplo, questiona-se *a posição temporal de uma elocução de uma proposição*.

"Eu, Juan Maiz, não estou escrevendo agora, dia 26 de Maio às 17 horas e 23 minutos no calendário Gregoriano, -3 horas de Greenwich" é uma proposição falsa, mas uma cuja falsidade eu já conhecia quando comecei a escrevê-la (o leitor - não sendo ele eu, Juan Maiz - não deveria estar tão seguro disso!), porque a proposição tem como objeto uma ação que supostamente não ocorre. Essa ação, no entanto, ocorre. A proposição é falsa. E é falsa justamente pelo modo como a elocução se deu, no caso, porque eu escrevi a proposição. Se eu apenas tivesse *pensado* a proposição, parando de escrever naquele minuto, ela seria verdadeira.

E porque tal distinção é útil no paradoxo em questão? O argumento de Cohen é o seguinte: Do mesmo modo que a distinção resolve o problema da proposição do parágrafo anterior, a afirmação "Ocorrerá um *Class A Blackout* na próxima semana", também é uma proposição cuja falsidade ocorre pelo ocorrência de sua elocução. De modo análogo, se o sargento tivesse apenas pensado a proposição, ela seria verdadeira.

---

<sup>3</sup> "pisar *em falso*" (traduz-se comumente por "*stumble*") é um termo comum no português, porém, obviamente, isso não quer dizer nada.

A passagem da solução inicial ao problema específico do paradoxo não é clara. Se o autor está sugerindo que o *ato* do anúncio o torna *falso*, isto significa então que o aluno deve simplesmente ignorar o anúncio. Neste caso, se o professor der uma prova no período, o anúncio se tornará verdadeiro. Assim, a distinção não resolve o problema, mas o engendra novamente. Mas mais importante ainda que esta especulação é o fato de que não está clara a relação entre estas frases que se produzem falsas (e.g. “eu não estou falando”, quando falado por um sujeito) com o anúncio do *blackout*.

Ainda assim, devemos capturar desta exposição este mérito: atentar para a *assimetria* no uso do anúncio. Esta assimetria ficará mais clara e se mostrará relevante adiante, nas soluções epistêmicas do paradoxo.

## 1.2. Futuros contingentes

EM 1952, PAUL WEISS APRESENTOU UMA SOLUÇÃO APOIADA NO CONCEITO ARISTOTÉLICO DOS futuros contingentes (*De Interpretatione*). No entanto, é bem provável que sua principal contribuição foi cunhar o nome do problema que acabou se tornando o mais utilizado no debate acadêmico, a saber, “paradoxo da predição”.

O uso do título “paradoxo da predição” é premeditado. Como veremos, este título já dá um sinal de como ele pretende resolver a questão. Weiss redefine o problema do seguinte modo: Um professor diz “é uma regra que não pode ser descumprida desta escola que haverá um exame em um dia inesperado”. Os alunos argumentam que o exame não pode ser dado no último dia de aulas, pois assim ele não será inesperado. Logo, também não pode ser dado no penúltimo dia e assim por diante. Conclusão: ou bem o professor não dá a prova ou bem o professor dá a prova em um dia esperado, o que, de qualquer forma descumpre a regra da escola.

O primeiro ponto de Weiss é buscar refutar os autores anteriores, que buscaram mostrar que o problema de alguma forma residia entre o conteúdo do anúncio e o ato do anúncio em si. Para isso ele apresenta dois exemplos: No primeiro o professor chega no

último dia de aula e anuncia “Ah, havia me esquecido: existe uma regra que não pode ser descumprida desta escola que haverá um exame em um dia inesperado. E aqui está o exame.”; No segundo o professor chega em qualquer dia letivo que não o último e dá exatamente o mesmo aviso; Em ambos os casos, fica claro que não há paradoxo algum, pois a regra não é descumprida, visto que 1) o exame ocorre e 2) ele é uma surpresa. O problema nestes dois exemplos é que não apenas a prova é uma surpresa, mas a própria apresentação da regra também é uma surpresa. Ademais, em ambos, há apenas um caso em que a regra pode ser aplicada, a saber, o próprio dia em que a regra é anunciada. É evidente que, se no segundo caso o professor não tivesse dado a prova ao mesmo tempo, o problema retornaria. Logo, para Weiss, temos que sair da perspectiva do aluno, e tomar a regra como dada, e olhar o problema de um modo mais geral. Dessa perspectiva o que é relevante é saber se é possível deduzir se haverá um exame inesperado dentro do período definido pela regra de antemão.

Para Weiss a confusão fundamental do problema é a confusão um uso “coletivo” e um uso “distributivo” do operador disjuntivo (ou). Argumenta ele que  $p$  não é dedutível de  $(p \vee q)$  mesmo que  $q$  não seja o caso. Isto porque, quando se trata de uma *predição* e consideramos um intervalo (*range*), este se apresenta apenas como uma gama de *possibilidades*, que não podem, por elas mesmas, se dissociarem umas das outras. Ou seja, quando consideramos o intervalo de dias em que é possível que ocorra o exame, não podemos excluir qualquer um destes a não ser por apelo à história ou à imaginação.

Tal apelo, feito pelos alunos ao argumentar a impossibilidade do exame no último dia, pressupõe supor que podemos, pelo apelo mesmo, tornar o intervalo de dias entre o primeiro e o penúltimo não mais passível de ocorrência de exame. É querer torná-los *atuais*, por mera imaginação. Nos termos de Weiss “é [querer] converter o ou coletivo no ou distributivo” o que não pode ser feito, pois isso só ocorre no próprio “vir a ser”.

“o último dia não pode ser deslocado dos outros dias no momento em que o anúncio é feito, exceto por um tipo de antecipação teórica da história real, afastando-se de um anúncio de uma possibilidade para um estado de coisas onde possibilidades são especificadas e distintas uma após a outra. O anúncio nos diz que algum dia vai ser selecionado como o dia do exame. Qualquer que seja, ele será inesperado, desde que

vejamos [o problema] do ponto de vista do anúncio; ele [o exame] será esperado desde que tenhamos eliminado os outros [dias] e então tornado este distinto dos outros – status que este [dia] não possui *no* anúncio.” (grifo e complementos meus)

Tendo tal solução em vista, podemos pensar no problema com apenas dois dias. Para Weiss, neste caso, o ponto é este: mesmo que o professor não dê o exame no primeiro dia e não seja mais surpresa que ele se dará no último, ele deixa de ser uma surpresa *neste momento*, mas não “do ponto de vista do anúncio”. Em resumo, no próprio anúncio não é possível prever o dia, sejam dois, seja todo o ano letivo. Mesmo no caso de um conjunto muito grande de dias e que tenhamos de fato chagado ao último sem que a prova tenha ocorrido, este dia foi inesperado até então.

Apesar da recepção geralmente ruim do trabalho de Weiss, isto se dá, assim me parece, pela sua adoção de noções aristotélicas. Porém, se lido até o fim, seu artigo apresenta esta conclusão final, que, veremos será retornada nos autores que usarão versões do paradoxo envolvendo jogos de cartas. Creio que esta parte do artigo pode muito bem ser destacada da parte inicial, e que esta conclusão não só é verdadeira, como bem fácil de se observar.

### 1.3. Desconfiando do professor

WILLARD VAN ORMAN QUINE INICIA SEU ARTIGO *ON A SO-CALLED PARADOX* AFIRMANDO conhecer o problema e estar em contenda com o mesmo há nove anos (desde 1943) e desqualificando a solução de Weiss como atitude de “extremo desespero” ao “entreter a fantasia Aristotélica de que ‘É verdade que p ou q’ é uma condição *insuficiente* para ‘É verdade que p ou é verdade que q’”.

Quine duvida do caráter paradoxal do problema. Sua apresentação se segue:

“K sabe no tempo  $t$  ... que está decretado que um evento E de um determinado tipo ocorrerá unicamente e com o conhecimento de K no tempo  $t + i$ , para um inteiro  $i$

menor ou igual que um número  $n$  especificado e que está decretado adicionalmente que  $K$  não sabe o valor de  $i$  antes passar (digamos) do tempo  $t + i - \frac{1}{2}$ .”

Seguindo a apresentação usual dos autores anteriores, neste momento  $t$ ,  $K$  argumenta que:

“ $i \leq n - 1$ ; pois, se  $i$  for  $n$ ,  $K$  saberia imediatamente após  $t + n - 1$  que  $i$  era  $n$ . Então, pelo mesmo arrazoado com ‘ $n - 1$ ’ para ‘ $n$ ’ ele argumenta que  $i \leq n - 2$ ; e assim por diante finalmente concluindo após  $n$  passos que  $i \leq 0$  e portanto que o evento não ocorrerá de modo algum.”

O erro no argumento de  $K$  consiste em considerar apenas dois cenários no tempo  $t$  ao pensar sobre o estado de coisas em  $t + n - 1$  (o penúltimo dia):

C1)  $E$  já vai ter ocorrido em algum dia que não  $t + n$ ;

C2)  $E$  vai ocorrer em  $t + n$  (o que manteria o decreto); Neste caso  $K$  vai saber em imediatamente após  $t + n - 1$  que o evento vai ocorrer em  $t + n$  (o que violaria o decreto);

Como C2 viola o decreto,  $K$  opta por rejeitar este cenário, argumentando, como vimos, que  $E$  não pode ocorrer em  $t + n$ . Quine acredita que  $K$  está errado em considerar apenas estes dois cenários, devendo incluir mais dois:

C3)  $E$  vai não vai ocorrer em  $t + n$  (o que violaria o decreto);

C4)  $E$  vai ocorrer em  $t + n$  (o que manteria o decreto) e o aluno não vai saber disso (o que manteria o decreto também);

Para Quine, como é possível que  $E$  não ocorra em  $t + n$  (mesmo não tendo ocorrido antes), o que violaria o decreto,  $K$  não tem razões suficientes em  $t + n - 1$  para estar seguro de que ocorra em  $t + n$ . Ou seja, o argumento de Quine depende que  $K$  não pressuponha que o decreto seja satisfeito por definição, pois depende que  $K$  considere a possibilidade  $E$  não ocorra em  $t + n$  mesmo imaginando-se em  $t + n - 1$ . Sendo assim o erro de  $K$  está em

associar seu argumento com um *reductio ad absurdum*, sem poder fazê-lo: em um *reductio* correto a proposição a ser refutada deve ser tomada como verdadeira. Como K não tem este direito, dado que C3 é possível, K há de considerar C4 como possível também, o que implica que seu argumento indutivo não se segue, e que, sendo assim, C1 também é um cenário que pode ocorrer.

Supor o cumprimento do decreto, e a eliminação de C4, é, como Quine apresenta, confundir duas coisas:

- i) uma hipótese, de K em  $t$ , que o decreto vai ser cumprido;
- ii) uma hipótese, de K em  $t$ , que K saberá em  $t + n - 1$  que o decreto será cumprido;

Por fim, Quine retorna ao exemplo do enforcado para mostrar como seu argumento se aplica a um cenário de um dia só. Se um juiz disser “você será enforcado amanhã, mas não sabe disso antes que ocorra” e K argumentar que o juiz está se contradizendo, logo K não deverá acreditar que será enforcado no outro dia, sendo então, o seu enforcamento uma surpresa. Tal raciocínio de K, que fique claro, é considerado errado por Quine. Se K se der conta que a conclusão de contradição leva à certeza do enforcamento, entramos em um regresso *ad infinitum* - veremos mais adiante como esse regresso de fato ocorre para outros autores e porque rejeitam a solução quineana.

O correto seria K se dar conta dos 4 cenários possíveis apresentados acima, e, é claro, sem poder saber de antemão que cenário será o caso, torcer - e com toda a reza possível! - pelo cenário em que o enforcamento simplesmente não ocorre.

#### 1.4. A solução prática

*Afirmações são feitas, palavras e sentenças são usadas. (Austin, 1950:114)*

RUTH WEINTRAUB INICIA SEU ARTIGO *PRACTICAL SOLUTIONS TO THE SURPRISE EXAMINATION Paradox*, apresentando algumas saídas já apresentadas ao problema, como a de Quine – o



aluno não pode estar seguro que ocorrerá um exame apenas pelo anúncio, logo um exame no último dia ainda seria surpreendente – , a possibilidade de que o aluno esqueça o anúncio – tornando-o verdadeiro –, ou a *reductio* mais complexa, que mostra que, como o anúncio é falsificado, a ocorrência do exame não pode ser *esperada*, e, não sendo esperada, pode se tornar verdadeira. Porém, nota Weintraub, montando seu argumento, nenhuma dessas possibilidades torna o anúncio *digno de crédito* [*credible*]. E na credibilidade é que depende a *melhor* explicação para o problema: nós atribuímos um *papel comunicativo* à promessa. O anúncio tem como objetivo passar uma informação aos alunos, que consideram o professor confiável e o professor, sabendo ser confiável, realiza o anúncio com um certo objetivo.

Weintraub considera que o objetivo do professor é fazer com que os alunos estudem diariamente, pois sabe que é mais provável que um aluno estudo caso acredite que uma prova será dada no próximo dia. Para que funcione a promessa *tem que ser* digna de crédito: qualquer análise do SEP deve permitir aos alunos saber que haverá um exame surpresa. O *insight* aqui é que devemos adotar uma posição *prática* em relação ao paradoxo; ver a promessa como um pronunciamento da *vida real* cujo *papel* – que *funciona* – uma solução deve refletir – assim como, analogamente, Weintraub apresenta o paradoxo de Zenão: o lógico pode usar sua dedução para provar a impossibilidade do movimento, porém o movimento *funciona*. Isso não significa que o paradoxo apresenta um *problema prático*, mas que a solução deste problema teórico deve incorporar uma explicação de uma certa prática.

Nesta concepção as respostas de Quine e de Olin não solucionam o paradoxo, apenas o reforçam, pois não consideram que o aluno deva acreditar na promessa. Também a solução de McLelland e Chihara – que *depende* de rejeitar a reflexividade do conhecimento, a ideia de que  $Kp \rightarrow KKp$  – é considerada errada com base no argumento de Sorensen com a versão dos alunos alinhados (ver abaixo): uma solução deve servir para *todas* as variações. Franceschi chama a tentativa de uma solução única, como visão “monista” do problema (Franceschi, 2002:2) à qual ele se opõe apresentando uma versão dicotômica, ou seja, que existem *dois* problemas envolvidos no paradoxo (de acordo com ele, os únicos autores que assim também pensaram foram Wright e Sudbury). A não ser, diz Weintraub, que vejamos o problema de modo *prático*. Neste caso, teríamos inúmeras soluções.

A consideração prática pode também atacar a solução de Wright e Sudbury: a indução reversa só pode se manter com a suposição que a promessa vai continuar sendo acreditada no penúltimo dia mesmo que o exame não tenha sido dado – a premissa da “retenção temporal do conhecimento”. Sem esta premissa o último dia *não pode* ser descartado. O exame irá surpreender os alunos pois eles não irão mais *acreditar* na promessa – um *blindspot*, nas palavras de Sorensen: uma proposição consistente mas que não pode ser acreditada.

Weintraub questiona então, que motivo alunos *ordinários* tem para desacreditar na promessa. O argumento só se segue se estes acreditarem que a promessa *irá* perder sua credibilidade.

A primeira solução então é tomar a promessa *não literalmente*. Ou seja, interpretar “surpresa” apenas como afirmando que o exame será dado em “alguma data da próxima semana”. Esta interpretação remove tanto a carga *epistêmica* como a carga *psicológica* da surpresa. Escolhendo um dos dias aleatoriamente, o professor tem probabilidade de 1 para 5 de fazer uma prova não esperada. Esta interpretação, apesar de enfraquecer o conceito de surpresa, captura, supostamente, a *real* intenção do professor. Simplesmente *não importa* se o exame não for mais uma “completa” surpresa no último dia. A intenção do professor de manter os alunos estudando que é a real motivação – e assim o “significado secreto” – para o seu proferimento.

No entanto, argumenta Weintraub quando consideramos afirmações como “ocorrerá um terremoto surpresa na Califórnia nos próximos 50 anos” a interpretação não literal não funciona: parece intuitivo crer que o autor de tal frase está mais interessado em garantir a *surpresa* do terremoto do que sua ocorrência no período em questão. Ele *provavelmente* considera apenas “muito provável” que ele ocorra neste período, mas está dizendo, ou tentando dizer, de que a população e as autoridades não estão dando atenção suficiente à esse provável desastre.

Weintraub concede que uma proposição só pode ser acreditada se não for paradoxal. No entanto, uma série de proposições soam *sutilmente* paradoxais. Sendo assim, se o ouvinte

a confundir como não paradoxal ela ainda pode ser razoável em sua crença, pois, diz ela, “onisciência lógica não é uma condição necessária para a racionalidade”. É perfeitamente razoável que estudantes ordinários acreditem na promessa *mesmo que* o paradoxo seja provado por um argumento impecável. Indução reversa não é um raciocínio que “vem” naturalmente. Olin responde a objeção de que os alunos podem não ser logicamente perfeitos como “trivial”, porém esta idealização nos remove demais de uma situação da vida real.

Na vida real, precisamos compreender a distinção entre paradoxo *atual* e *percebido*. Esta permite que expliquemos um aspecto importante do fenômeno: a diferença intuitiva entre a versão com 1 dia e a versão com 5 (ou mais) dias. Enquanto a primeira soa claramente paradoxal, mesmo para o estudante ordinário, a segunda não. Porém, mesmo concedendo como alunos ordinários agirão em relação ao anúncio, ainda precisamos entender como o anúncio afeta o aluno *s sofisticado*.

Imaginemos dois casos: 1) Se o professor é ingênuo, os alunos devem responder como um aluno ingênuo. O professor, provavelmente, irá se dar conta do paradoxo no penúltimo dia, sendo impossível prever o que ele fará no último. No entanto tal situação não faz parte do *conteúdo* da promessa original, ou seja, o professor não tinha se dado conta do problema em que se meteu ao fazer o anúncio. 2) Se o paradoxo é conhecimento comum [common knowledge] – todos sabem que todos sabem do paradoxo – como entender o anúncio? Há, segundo Weintraub, duas alternativas. Ou interpretamos o anúncio meramente como uma piada – como essas piadas “filosóficas”, retiradas dos livros do Raymond Smullyan ou das lendas do prof. Sidney Morgenbesser, que circulam pelos corredores dos institutos de filosofia –, ou os alunos devem interpretar o anúncio de acordo com a solução preferida do professor.

Em todo caso, deve-se entender a *intenção* do proferente. Se o professor crê piamente que está falando algo autocontraditório, devemos *ignorá-lo*. Se ele acha que o que diz é informativo, devemos estudar.

O argumento de Weintraub, no entanto, não precisa nem mesmo ser atacado por suas premissas metodológicas. A própria conclusão é um problema, e, sobre isso o artigo silencia:

como então deveria um aluno da própria professora Weintraub – ou de qualquer professor que concorde com ela – interpretar um anúncio dela? De acordo com a “solução prática”, “devemos interpretar o anúncio de acordo com a solução preferida do professor”. No caso, devemos pensar: qual a solução preferida por Weintraub ao paradoxo do exame surpresa? Bom, a solução preferida ou, ao menos o que é mais razoável supor, posto que ela *escreveu* um artigo sobre o tema, é que “devemos interpretar o anúncio de acordo com a solução preferida do professor”. E, novamente, *ad infinitum*, qual a solução preferida do professor?

## 2. AS SOLUÇÕES LÓGICAS

### 2.1 Condicionais problemáticos

NO MESMO ANO DA PUBLICAÇÃO DE COHEN (VER ACIMA), 1950, PETER ALEXANDER COLOCOU em dúvida a utilidade da distinção (não a validade da mesma, que é bastante evidente) de Cohen para o problema do paradoxo e apresentou outra. E com seu argumento colocou a solução do paradoxo, pela primeira vez, na trilha dos paradoxos lógicos, deslocando-o do campo pragmático.

O primeiro passo é negar a utilidade da distinção elocução/proposição. Para Cohen, não há paradoxo em “Eu não me lembro de nada”, mesmo no caso do sujeito que profere tal proposição estar morto – imagine que o sujeito gravou esta frase e solicitou que a reproduzissem em seu enterro em um *MP3 Player*. Alexander apresenta duas considerações sobre este ponto: 1) Não é razoável atribuir a referência do sujeito da proposição ao cadáver (muito menos ao *MP3 Player*!). Sendo assim proposições deste tipo bizarro não podem ser verificadas ou falsificadas. 2) Mesmo que possamos atribuir a referência ao morto é, ao menos, concebível que a frase seja falsa.

Elucidemos os dois pontos. Ponto 1: Imagine que alguém está assistindo a um vídeo no *YouTube* onde o reverendo Martin Luther King, postado em frente a uma turba, está dizendo “Eu tenho um sonho”. Esta pessoa não irá pensar “como podem tantas pessoas confiar neste mentiroso? Ele está morto!”. Ponto 2: Imagine que um médium está “recebendo” frases de

alguém imaterial que alega ser Chico Xavier. Ele diz “Eu me lembro de algo”. A referência estaria no espírito de Xavier, o que é *concebivelmente* (é possível que seja) verdadeiro.

Para Alexander, então, o problema está em uma distinção acerca do que é uma memória. Podemos ter uma memória *de uma experiência*, como em “lembro-me que ontem comi *Cheesecake*” e a memória *de uma habilidade* ou *costume*, como em “lembro-me de como andar de bicicleta” (creio que tal distinção é idêntica à que se faz em *know-how* e *know-that*, porém relacionada à memória destes conhecimentos).

Com essa distinção em mente Alexander afirma que o problema original de Cohen deve ser entendido nestes termos: “Eu ainda sei como construir proposições, porém não me lembro de nenhuma experiência”. Se assim é, não há problema algum na proposição. Já em “eu não sei como construir proposições”, para Alexander, a contradição emerge na própria proposição, independentemente da sua elocução, do mesmo modo que “eu estou mentindo agora”.

Apesar de todo esse argumento, Alexander acredita que tanto a distinção de Cohen – que continua válido para sentenças como “eu não estou escrevendo agora” –, como a sua, que procura igualar “eu não sei como construir proposições” com o paradoxo do mentiroso, não respondem ao problema do *Blackout*. Porém sua solução para este também não parece muito convincente.

Para Alexander, quando afirmo que “amanhã farei uma prova”, o que quero dizer é “se nada me impedir, amanhã farei uma prova”. Do mesmo modo, quando se apresenta o problema do *Blackout*, se pretende dizer que “se as condições do *Class A Blackout* forem satisfeitas, ocorrerá um *Class A Blackout*”. Mas como as condições são insatisfazíveis, temos uma proposição condicional cuja condição é falsa.

Este caminho deslocaria o problema para um problema da lógica dos condicionais, como no teste de Ramsey e nas teorias de cotenabilidade [*cotenability theories*]. Porém, como veremos adiante, a discussão acerca do problema não se seguiu neste caminho, mas no da autorreferência, portanto, não desenvolverei mais tais ideias.

## 2.2. Declarações e ordenamentos

*Acho que este sabor da lógica refutada pelo mundo faz o paradoxo [da predição] ser bastante fascinante. O lógico vai pateticamente por meio dos movimentos que sempre fizeram o feitiço funcionar antes, mas de alguma forma o monstro, a Realidade, não entendeu o ponto e ainda avança. – (Scriven, 1951:403)*

PARA MICHAEL SCRIVEN A DISTINÇÃO RELEVANTE PARA RESOLVER O PARADOXO É A DISTINÇÃO entre declarações (*statements*) e ordenamentos (*ordainments*). Para ele, existe uma confusão entre estes dois tipos de anúncios que engendram o problema.

Uma declaração é um anúncio regular, como “as samambaias são pteridófitas” enquanto que um ordenamento é um anúncio que “tipicamente” envolve a fixação de um local, momento ou detalhes de [promessa de, aviso de] uma ocorrência, como “o avião decolará depois dais 19 horas”. Enquanto declarações, os ordenamentos são “predições contingentes” (idem, p. 406). Sendo assim, a afirmação “ocorrerá um *Class A Blackout* no próximo sábado” é um ordenamento contraditório. Isto porque, para Scriven, o ordenamento demanda que haja um *Blackout* no sábado. Mas se o *ordenamento* é válido, então ele é inválido.

É que é importante notar aqui é como Scriven busca remover o caráter epistêmico do problema. O que ele faz é igualar a afirmação “ocorrerá um *Class A Blackout* no próximo sábado” com “se ocorrer um *Blackout* no próximo sábado, esta afirmação é falsa”. Dada esta incompatibilidade, temos duas alternativas: ou bem haverá o *Blackout* no sábado, mas não será um *Class A*, ou bem não haverá um *Blackout* no sábado (e, se ocorrer em outro dia, será *Class A*). Mas, nota Scriven, estas não são propriamente soluções, são “interpretações que poderíamos fazer em uma situação em que acreditamos que o caráter autorrefutador do anúncio ocorre por engano” (idem, p. 406).

A conclusão de Scriven é que o anúncio é contraditório *enquanto um ordenamento*. Sua simples solução é a seguinte: se concluirmos que o ordenamento é contraditório então podemos toma-lo como falso. No entanto, se tomamos como falso, estamos seguros de que não ocorrerá um *Class A Blackout* no sábado. Logo, se ocorrer um *Blackout* no sábado este será inesperado, salvando assim a *declaração* que se realizará (será verdadeira!). Como conclui Scriven “o suicídio do anúncio enquanto um ordenamento é acompanhado por sua salvação enquanto uma declaração”.

A saída de Scriven é bem simples: por meio de uma distinção tornou o problema original uma contradição, e pela contradição mostrou que podemos ser surpreendidos. Contudo, primeiramente, tal distinção é bem confusa e, segundo, se negamos a plausibilidade desta, o problema reaparece, pois, se sabemos que a afirmação é verdadeira – que o ocorrerá o *Class A Blackout* na data definida – ele não será mais *Class A*, tornando-se falso e assim por diante. Teremos, assim, não uma solução como pretende Scriven, mas um paradoxo lógico.

### 2.3. Autorreferência

CINCO ANOS APÓS O ARTIGO DE QUINE (VER ACIMA), - O QUE REPRESENTA UM GRANDE SILÊNCIO<sup>4</sup>  
- R. Shaw defende em seu artigo que o problema é de ordem lógica, ou seja, que trata-se, em essência, de um paradoxo de autorreferência.

Segue a formulação de Shaw: Um professor anuncia a seus pupilos que há uma regra inquebrável da escola de que, até o final do período letivo haverá um exame em um dia inesperado (não é possível *saber* que a prova ocorrerá no próximo dia). Os alunos argumentam que o exame não pode ser dado no último dia do período letivo, pois se este não tiver sido dado até então eles saberiam na noite anterior que ele ocorreria, não sendo mais inesperado. Do mesmo modo, não poderia ser dado no penúltimo dia e assim por diante.

---

<sup>4</sup> A final todos os artigos que vimos acima estão no período de 48 e 53. Silêncio este cujos motivos podemos apenas conjecturar: seria a autoridade de Quine que afugentou réplicas ou a força mesma de seu argumento?



Assim, a regra inquebrável da escola parece ser autocontraditória, no entanto a ocorrência *efetiva* de um exame, digamos, antes da última semana satisfaz a regra.

Para Shaw, *saber* que a prova ocorrerá no próximo dia neste contexto significa “ser capaz de prever, dado que as regras da escola não sejam quebradas”. Neste ponto Shaw se distancia de Quine, pois ele quer resolver o problema do paradoxo deixando claro que *temos* que pressupor que a regra em questão tem de ser levada à sério. A saída de Quine – tomar saber como uma noção “vaga do senso comum” (Shaw, 1958:382) – significa se evadir do problema e não resolvê-lo. Temos que entender “inesperado” deste modo: não dedutível de certas regras especificadas da escola. Assim temos uma definição do problema “puramente lógica”, com as regras servindo com axiomas e, assim, temos que qualquer problema derivado desta definição só pode ser um problema lógico e não um problema de outra natureza como queriam alguns autores anteriores à Shaw: Alexander, Scriven, Weiss e Quine.

O primeiro conjunto de regras formulado por Shaw é o seguinte:

Regra 1: Um exame ocorrerá em algum dia do período letivo;

Regra 2: O exame será inesperado, no sentido em que ele ocorrerá em um dia em que na noite anterior não será possível para os alunos deduzir que o exame ocorreria no dia posterior *a partir da regra 1*.

De acordo com Shaw, esta formulação proíbe a ocorrência no último dia, afinal os alunos podem deduzir a ocorrência da prova neste dia, caso ela não tenha sido dada, *a partir da regra 1*. No entanto, Shaw não crê que eles possam seguir o mesmo raciocínio para os outros dias – o argumento para isto não é perfeitamente claro no texto, mas podemos supor com certa confiança que se trata disso: barrar o penúltimo envolveria *violar a regra 2* – logo, não há paradoxo. Shaw supõe então uma variação do problema com uma terceira regra:

Regra 3: O exame ocorrerá em um dia em que na noite anterior não será possível para os alunos deduzir que o exame ocorreria no dia posterior *a partir das regras 1 e 2*.

Neste caso poderíamos barra o penúltimo dia, pois tal raciocínio envolveria utilizar a regra 1 e 2. No entanto, novamente, nenhum outro dia poderia ser descartado, pois isso violaria a regra 3 – e continuaríamos sem paradoxo. Entretanto, no caso de um período de apenas dois dias teríamos um cenário diferente. Neste, as regras da escola seriam simplesmente autocontraditórias e não paradoxais, pois, diz Weiss, “nenhum exame pode *de fato* ser realizado e satisfazer as regras da escola”.

Tal argumento pode generalizar o problema e vale para qualquer período de  $n$  dias, com  $r + 1$  regras (seguindo a lógica das regras 2 e 3, é claro). Se  $r < n$ , os últimos  $r$  dias poderão ser proibidos pelo arrazoadado dos alunos, porém restarão  $n - r$  dias onde o exame será possível. Se  $r = n$ , então as regras são simplesmente autocontraditórias, pois o aluno poderá dizer que a ocorrência do exame no dia  $s$  que este viola a regra  $s + 1$ . Em todo caso, diz Weiss, não temos paradoxo. Consideremos, por fim, uma última regra, e consideremos apenas esta regra somada à regra 1:

Regra 2\*: O exame ocorrerá em um dia em que na noite anterior não será possível para os alunos deduzir que o exame ocorreria no dia posterior *a partir das regras 1 e 2\**.

Neste caso, diz Shaw, fica “claro” que a origem do paradoxo é a natureza autorreferente da regra 2\*.

Três coisas podem – e devem – ser ditas neste ponto.

Primeiro, que não é claro em que sentido, no caso de  $r = n$ , um aluno possa afirmar no dia anterior – dia  $s$  – que *sabe* que haverá uma prova no outro dia – dia  $s + 1$  – com base na regra  $s + 1$ , pois, no início do período, ou mesmo em qualquer outro dia deste, como ele poderia distinguir um dia específico? Este argumento apenas poderia ser feito *após* a ocorrência do exame, onde o aluno diria que era *possível* saber da ocorrência do exame a partir das regras, porém, esta lógica também é válida para *todos os dias*.

Segundo, que a conclusão de Shaw é atroz: se a natureza paradoxal do problema está na natureza da regra 2\*, porque então ele não definiu já o problema *em termos de 2\** e tentou

dar-lhe uma solução? Ele realmente crê que *há* um paradoxo nesta definição? Parece que sim.

E, por fim, mas não menos importante, que Shaw parece afirmar que o paradoxo envolvido em 2\* se engendra *meramente* “pela natureza autorreferencial” da regra. Pois autorreferência não é condição – nem necessária, nem suficiente – para engendrar um paradoxo. Há autores<sup>5</sup> que argumentam que autorreferência não é necessária nem mesmo para definir o paradoxo do mentiroso. E, para os afeitos aos exemplos, “esta proposição é composta por palavras” é um caso evidente de uma autorreferência não paradoxal.

## 2.4. O conhecedor

A SOLUÇÃO DE DAVID KAPLAN E RICHARD MONTAGUE PARA O PROBLEMA COMEÇA COM UM assentimento com o pressuposto de Shaw – de que devemos *tomar* o anúncio do professor como verdadeiro, ao contrário da solução de Quine, e que devemos procurar encontrar a formulação mais forte e que seja genuinamente paradoxal –, cuja intenção é colocar o paradoxo *ao lado* dos paradoxos do mentiroso e o de Richard.

Porém, Kaplan e Montague afirmam ter demonstrado que a solução de Shaw não é paradoxal, ao contrário do que afirma o autor. No entanto, eles também afirmam terem encontrado variações do problema *genuinamente* paradoxais.

Vejamos a formalização proposta<sup>6</sup>, apresentada nos termos do problema do enforcado com três dias:

K é o prisioneiro;

---

<sup>5</sup> Um resumo da ópera, que ignoro cada detalhe, está em Levy 2009, p. 154, nota 25.

<sup>6</sup> Tentei manter a estrutura geral da formalização apresentada no artigo, no entanto modifiquei a sintaxe utilizada pelos autores pela sintaxe padrão que adotei para os operadores lógicos ao longo deste trabalho (e.g. & virou  $\wedge$ ,  $\sim$  virou  $\neg$ , e assim por diante). Além disso, resolvi traduzir as letras sentenciais quando achei apropriado (e.g. **W**, de *wednesday*, virou **Q**, de *quarta*). Também por motivos de consistência, optei por usar apenas parênteses para escopo e não os colchetes que os autores adotaram. Os colchetes uso à direita das expressões para informar a partir de qual expressão os autores derivam o resultado, ao contrário do modo que os autores o fazem, que é textual.

S é a proposição “K é enforcado na segunda”;

T é a proposição “K é enforcado na terça”;

Q é a proposição “K é enforcado na quarta”;

$K_d(\varphi)$  é a proposição “K sabe no domingo de noite que a proposição  $\varphi$  é verdadeira”;

$K_s(\varphi)$  é a proposição “K sabe na segunda de noite que a proposição  $\varphi$  é verdadeira”;

$K_t(\varphi)$  é a proposição “K sabe na terça de noite que a proposição  $\varphi$  é verdadeira”;

As fórmulas  $\underline{S}$ ,  $\underline{T}$  e  $\underline{Q}$  são usadas substituindo  $\varphi$  para referenciar o resultado da colocação de S, T e Q entre parênteses. Por exemplo,  $K_d(\underline{S})$  deve ser interpretado como “K sabe no domingo de noite que a proposição ‘K é enforcado na segunda’ é verdadeira”;

$D_1$  é a afirmação do juiz de que uma das três condições será satisfeita:

- K é enforcado na segunda, mas não na terça, nem na quarta, e K não tem como saber disso no domingo de noite. Formalmente:  $S \wedge \neg T \wedge \neg Q \wedge \neg K_d(\underline{S})$
- K é enforcado na terça, mas não na segunda, nem na quarta, e K não tem como saber disso na segunda de noite. Formalmente:  $\neg S \wedge T \wedge \neg Q \wedge \neg K_s(\underline{T})$
- K é enforcado na quarta, mas não na segunda, nem na terça, e K não tem como saber disso na terça de noite. Formalmente:  $\neg S \wedge \neg T \wedge Q \wedge \neg K_t(\underline{Q})$

Um outro detalhe relevante da notação de Kaplan e Montague é a substituição da notação usual de *implicação*  $\varphi \vdash \omega$  pelo a função  $I(\varphi, \omega)$ .

Assim,  $D_1$  pode ser formalizada completamente como:

$$\begin{aligned} & ( S \wedge \neg T \wedge \neg Q \wedge \neg K_d(\underline{S}) ) \wedge \\ & ( \neg S \wedge T \wedge \neg Q \wedge \neg K_s(\underline{T}) ) \wedge \\ & ( \neg S \wedge \neg T \wedge Q \wedge \neg K_t(\underline{Q}) ) \end{aligned}$$

Deste modo, K argumenta que  $D_1$  não pode ser cumprido: assumamos que ele seja; O enforcamento não pode ocorrer na quarta, pois se ocorresse, os dois primeiros adjuntos de  $D_1$  seriam falsos (clarificando: o primeiro porque S seria falso, o segundo porquê T seria falso) e o terceiro se manteria. O problema se resumiria a  $Q \wedge \neg K_t(Q)$ . Mas, não ocorrendo nem na segunda, nem na terça, o exame ocorrerá na quarta (lembramos, estamos pressupondo o cumprimento de  $D_1$ ), ou seja  $I(\neg S \wedge \neg T, Q)$ . Assim, K saberia, na terça de noite, que quarta ocorreria o enforcamento, ou seja  $K_t(Q)$ , o que contradiz obviamente  $\neg K_t(Q)$ . Logo  $\neg Q$ .

As duas assunções implícitas, porém plausíveis, de K para construir esta refutação são:

$$A_1: (\neg S \wedge \neg T) \rightarrow K_t(\neg S \wedge \neg T)$$

$$A_2: (I(\neg S \wedge \neg T, Q) \wedge K_t(\neg S \wedge \neg T)) \rightarrow K_t(Q)$$

$A_1$  é um caso do conhecimento por *memória*. Em bom português, se a prova não foi dada nem na segunda nem na terça, então K sabe, na terça de noite, que ela não foi dada na segunda nem na terça.  $A_2$  é um caso do princípio de fechamento (conhecido comumente como *epistemic closure principle*) do conhecimento: Se  $p$  implica  $q$  e eu sei que  $p$ , então eu sei que  $q$ , ou, em outras palavras, tudo que é implicado pelo meu conhecimento também é meu conhecimento. Apesar deste princípio não ser um consenso, Kaplan e Montague consideram apropriado considerar que K está em uma situação específica em que podemos tomar  $A_2$  como certo.

Podemos assumir, também, que K sabe  $A_1$  e  $A_2$  desde o início do período (domingo):

$$A_3: K_d(A_1 \wedge A_2)$$

Se ele sabe disto desde o início, podemos aplicar isto neste momento e como  $A_1$  e  $A_2$  implicam  $\neg Q$ , podemos concluir que::

$$A_4: (I(A_1 \wedge A_2, \neg Q) \wedge K_d(A_1 \wedge A_2)) \rightarrow K_d(\neg Q)$$

Usando o mesmo princípio de fechamento apresentado acima, concluímos que  $K$  não apenas que não pode ser enforcado na quarta feira –  $\neg Q$  – mas que ele *sabe* disso *já no domingo* –  $K_d(\neg Q)$ .

E se  $K_d(\neg Q)$ , pelo princípio da retenção do conhecimento temporal (memória)  $K_s(\neg Q)$ ,  $K_t(\neg Q)$  também serão verdadeiros.

Para mostrar a impossibilidade da terça, ainda assumindo  $D_1$ , o segundo disjuncto de  $D_1$  tem de ser o caso. Porém, como  $\neg S$  e  $\neg Q$  implicam  $T$ ,  $K$  sabe que  $T$  na segunda –  $K_s(T)$  –, o que contradiz  $\neg K_s(T)$  (do disjuncto). Assim, concluímos que  $\neg T$ . Novamente, a base para este argumento se assemelha à forma usada acima:

$$A_5: \quad \neg S \rightarrow K_s(\neg S)$$

$$A_6: \quad (I(\neg S \wedge \neg Q, T) \wedge K_s(\neg S) \wedge K_s(\neg Q)) \rightarrow K_s(T)$$

Por fim, basta excluir a segunda. Como  $\neg T$  e  $\neg Q$ , temos que  $S$ .  $K$  sabe  $S$ , já no domingo, o que, por sua vez, contradiz  $\neg K_d(S)$ . Logo,  $\neg S$ . Com os argumentos somados até aqui, mostramos que  $D_1$  é impossível.

No entanto, para mostrar que  $D_1$  é possível, suponhamos  $T$  (e, logo,  $\neg S$  e  $\neg Q$ ). Neste caso, a sentença  $T$ , dizem os autores, não é *analítica*. Apelando a “princípios epistêmicos intuitivos” – cuja formulação os autores omitem do texto – observamos que não é possível conhecer sentenças não-analíticas sobre o futuro.  $K$  não pode saber que  $T$  na segunda de noite, ou seja,  $\neg K_s(T)$ . Assim,  $D_1$  é perfeitamente possível. Temos portanto, uma contradição –  $D_1 \wedge \neg D_1$ .

Os autores notam que Quine acredita que o erro de  $K$  é aplicar  $A_2$ . Esta aplicação requer que  $\neg S \wedge \neg T$  implique  $Q$ . É evidente que esta implicação é falsa, a menos que, novamente, como fizemos, consideremos  $\neg S \wedge \neg T$  junto com  $D_1$ . Mas para isso, precisaríamos trocar  $A_2$  pelo seguinte análogo:

$$A_2': \quad (I(\neg S \wedge \neg T \wedge D_1, Q) \wedge K_t(\neg S \wedge \neg T) \wedge K_t(D_1)) \rightarrow K_t(Q)$$

Além da seguinte assunção:

$$A_2'' : K_t(\underline{D_1})$$

No entanto não é razoável pressupor que K *saiba* que  $D_1$  será cumprido, até porque ele está tentando *refutar*  $D_1$ ! Para Shaw o toque paradoxal do problema nasce de um elemento autorreferencial no anúncio que não está incorporado na formulação de Quine. Shaw propõe uma formulação que deixe isto explícito: Ou (1) K é enforcado na segunda e não na terça, nem na quarta, e no domingo de noite não sabe com base *neste anúncio* que 'K é enforcado na segunda' é verdadeiro, ou (2) K é enforcado na terça e não na segunda, nem na quarta, e no domingo de noite não sabe com base *neste anúncio* que 'K é enforcado na terça' é verdadeiro, ou (3) K é enforcado na quarta e não na segunda, nem na terça, e no domingo de noite não sabe com base *neste anúncio* que 'K é enforcado na quarta' é verdadeiro.

A formalização da versão de Shaw, notam Kaplan e Montague, depende da possibilidade de representar autorreferência. O mecanismo utilizado é o apresentado por Göedel e segue o que os autores já estavam adotando: ao falar *sobre* uma proposição, usa-se o “nome padrão” (*standard name*) da proposição adicionando uma sublinhado à letra sentencial. Assim os autores apresentam a seguinte fórmula  $D_2$ :

$$\begin{aligned} & ( S \wedge \neg T \wedge \neg Q \wedge \neg K_d( \underline{D_2} \rightarrow S ) ) \vee \\ & (\neg S \wedge T \wedge \neg Q \wedge \neg K_s( \underline{D_2} \rightarrow T ) ) \vee \\ & (\neg S \wedge \neg T \wedge Q \wedge \neg K_t( \underline{D_2} \rightarrow Q ) ) \end{aligned}$$

Porém os autores consideram que uma versão de apenas dois dias preserva as mesmas propriedades desta versão, portanto preferem adotar consequentemente a seguinte formulação, também autorreferente,  $D_3$ :

$$(1) \quad \vdash D_3 \equiv ( S \wedge \neg T \wedge \neg K_d( \underline{D_3} \rightarrow S ) ) \vee (\neg S \wedge T \wedge \neg K_s( \underline{D_3} \rightarrow T ) )$$

K é capaz agora de mostrar que  $D_3$  não pode ser cumprido usando um argumento análogo à  $A_1$  -  $A_4$ :

- $B_1: \neg S \rightarrow K_s(\neg S)$   
 $B_2: (I(\neg S, D_3 \rightarrow T) \wedge K_s(\neg S)) \rightarrow K_s(D_3 \rightarrow T)$   
 $B_3: K_d(B_1 \wedge B_2)$   
 $B_4: (I(B_1 \wedge B_2, D_3 \rightarrow S) \wedge K_d(B_1 \wedge B_2)) \rightarrow K_d(D_3 \rightarrow S)$

O argumento pode ser explicitado deste modo:

- (2)  $\neg S \vdash (D_3 \rightarrow T)$   
(3)  $\vdash (D_3 \rightarrow T) \rightarrow \neg K_s(D_3 \rightarrow T)$   
(4)  $\vdash (D_3 \rightarrow T) \rightarrow \neg S$   
(5)  $B_1 \vdash (D_3 \rightarrow T) \rightarrow K_s(\neg S)$  [(4)]  
(6)  $\vdash I(\neg S, D_3 \rightarrow T)$  [(2)]  
(7)  $B_2 \vdash K_s(\neg S) \rightarrow K_s(D_3 \rightarrow T)$   
(8)  $B_1 \wedge B_2 \vdash D_3 \rightarrow \neg T$  [(3), (5), (7)]  
(9)  $\vdash (D_3 \rightarrow \neg T) \rightarrow S$  [(1)]  
(10)  $\vdash (D_3 \rightarrow (K_d(D_3 \rightarrow S))) \rightarrow \neg S$   
 $B_1 \wedge B_2 \vdash D_3 \rightarrow S$  [(8), (9)]  
 $\vdash I(B_1 \wedge B_2, D_3 \rightarrow S)$   
 $B_4 \vdash K_d(B_1 \wedge B_2) \rightarrow K_d(D_3 \rightarrow S)$   
 $B_3 \wedge B_4 \vdash D_3 \rightarrow \neg$   
(11)  $B_{1..4} \vdash D_3 \rightarrow (\neg S \wedge \neg T)$  [(8)]<sup>7</sup>  
 $\vdash D_3 \rightarrow (S \vee T)$  [(1)]  
(12)  $B_{1..4} \vdash \neg D_3$  [(11)]

Assim, mostramos que, pelas suposições  $B_1$  -  $B_4$ , chegamos à conclusão (12), a saber, que o decreto não pode ser cumprido. Shaw considera o anúncio paradoxal, e não meramente impossível de ser cumprido. Para os autores, no entanto, não há boas razões para crer nisso. Suponhamos, como fizemos antes, que  $K$  seja enforcado na terça. Neste estado de coisas temos  $T$  e  $\neg S$ . Temos então que estabelecer  $\neg K_s(D_3 \rightarrow T)$ . Para aplicar a mesma linha

<sup>7</sup> Apenas para facilitar, vou usar, daqui em diante, a convenção  $P_1, \dots, P_n$  para denotar a conjunção  $P_1 \wedge \dots \wedge P_n$ .



de raciocínio, precisamos mostrar que  $D_3 \rightarrow T$ , considerada no início do período é uma sentença não-analítica a respeito do futuro. Mas  $D_3 \rightarrow T$  é analítica, pois vimos que  $\neg D_3$  se segue logicamente de princípios epistemológicos gerais, e então também  $D_3 \rightarrow T$ .

Para salvar a paradoxalidade pretendida por Shaw, propõe os autores, devemos adicionar à formulação original “a não ser que K saiba no domingo que esta afirmação é falsa”, cuja formalização seria:

$$(1) \quad \vdash D_4 \equiv K_d(\neg D_4) \wedge (S \wedge \neg T \wedge \neg K_d(D_4 \rightarrow S)) \vee (\neg S \wedge T \wedge \neg K_s(D_4 \rightarrow T))$$

Usaremos as suposições a seguir, onde  $C_1$  é o princípio de que tudo que é sabido é verdade e  $C_2$  -  $C_8$  análogos de  $B_1$  -  $B_6$ :

- $C_1: \quad K_d(\neg D_4) \rightarrow \neg D_4$
- $C_2: \quad \neg S \rightarrow K_s(\neg S)$
- $C_3: \quad K_d(C_1)$
- $C_4: \quad (I(C_1 \wedge \neg S, D_4 \rightarrow T) \wedge K_s(C_1) \wedge K_s(\neg Q)) \rightarrow K_s(D_4 \rightarrow T)$
- $C_5: \quad K_d(C_{1..4})$
- $C_6: \quad I(C_{1..4}, D_4 \rightarrow S) \wedge K_d(C_{1..4}) \rightarrow K_d(D_4 \rightarrow Q)$
- $C_7: \quad K_d(C_{1..6})$
- $C_8: \quad I(C_{1..4}, \neg D_4) \wedge K_s(C_{1..4}) \rightarrow K_d(\neg D_4)$

- (2)  $C_1 \vdash D_4 \rightarrow K_s(\neg D_4)$  [(1)]
- (3)  $C_1 \wedge S \vdash D_4 \rightarrow T$  [(1), (2)]
- (4)  $C_1 \vdash (D_4 \wedge T) \rightarrow \neg K_s(D_4 \rightarrow T)$
- (5)  $C_1 \vdash (D_4 \wedge T) \rightarrow \neg Q$
- (6)  $C_1 \wedge C_2 \vdash (D_4 \wedge T) \rightarrow K_s(\neg Q)$  [(5)]
- (7)  $\vdash I(C_1 \wedge S, D_4 \rightarrow T)$  [(3)]
- $C_4 \vdash (K_s(C_1) \wedge K_s(\neg S)) \rightarrow K_s(D_4 \rightarrow T)$
- $C_{1..4} \vdash (D_4 \wedge T) \rightarrow K_s(D_4 \rightarrow T)$  [(6)]
- $C_{1..4} \vdash (D_4 \wedge T) \rightarrow K_s(D_4 \rightarrow T) \wedge \neg K_s(D_4 \rightarrow T)$  [(4)]
- (8)  $C_{1..4} \vdash D_4 \rightarrow \neg T$

- (9)  $C_1 \vdash (D_4 \wedge \neg T) \rightarrow S$  [(1), (2)]
- (10)  $C_1 \vdash (D_4 \wedge K_d(\underline{D_4 \rightarrow S})) \rightarrow \neg S$
- $C_{1..4} \vdash D_4 \rightarrow S$  [(8), (9)]
- $\vdash I(\underline{C_{1..4}}, \underline{D_4 \rightarrow S})$
- $C_6 \vdash K_d(\underline{C_{1..4}}) \rightarrow K_d(\underline{D_4 \rightarrow S})$
- $C_5 \wedge C_6 \vdash K_d(\underline{D_4 \rightarrow S})$
- $C_1 \wedge C_5 \wedge C_6 \vdash D_4 \rightarrow \neg S$  [2]
- (11)  $C_{1..6} \vdash D_4 \rightarrow (\neg K_d(\underline{\neg D_4}) \wedge \neg S \wedge \neg T)$  [(2), (8)]
- $\vdash D_4 \rightarrow (K_d(\underline{\neg D_4}) \vee S \vee T)$  [(1)]
- (12)  $C_{1..6} \vdash \neg D_4$  [(11)]

Isso mostra que, com estas assunções  $D_4$  não pode ser cumprido. Porém, usando (12), e o princípio usado para obter (7), temos:

$$\vdash I(\underline{C_{1..6}}, \underline{\neg D_4})$$

Então:

- $C_8 \vdash K_d(\underline{C_{1..6}}) \rightarrow K_d(\underline{\neg D_4})$
- (13)  $C_7 \wedge C_8 \vdash K_d(\underline{\neg D_4})$
- $\vdash K_d(\underline{\neg D_4}) \rightarrow D_4$  [(1)]
- (14)  $C_7 \wedge C_8 \vdash D_4$  [(13)]

Com essas assunções, então, o anúncio é satisfeito. O que os autores esperam ter mostrado é que  $C_1 - C_8$  é incompatível com a sintaxe elementar<sup>8</sup>, e o interesse do problema está aí. As consequências filosóficas do problema ficarão mais claras ao considerar uma versão do paradoxo com um dia só: “A menos que K saiba no domingo de noite que este

---

<sup>8</sup> Por sintaxe elementar os autores compreendem a teoria de primeira ordem contendo, em adição às fórmulas especiais  $K_s, K_m, K_t, S, T$  e  $Q$ , todos os "nomes padrão" (de expressões), meio para expressar relações e operações sintáticas entre expressões e os axiomas apropriados envolvendo estas noções. A convenção adotada para atribuição dos nomes padrão pode ser tanto, dizem eles, a teoria P de Tarski, Mostowski, Robinson em *Undecidable Theories* (neste caso a sintaxe elementar pode ser identificada com a aritmética de Peano) ou mesmo a teoria mais fraca Q (do mesmo trabalho).

anúncio é falso, as condições a seguir serão satisfeitas: K será enforcado na segunda, mas ele não irá saber, com base neste anúncio, que ele irá ser enforcado na segunda” –  $\vdash D_5 \equiv K_d(\neg D_5) \wedge (S \wedge \neg K_d(D_6 \rightarrow S))$ . Mais do que isso, dizem eles, é possível reduzir o problema a uma versão com 0 dias! Assim: “K sabe no domingo que este anúncio é falso”.

$$(1) \quad \vdash D_6 \equiv K_d(\neg D_6)$$

O paradoxo, neste caso, é mostrado em três princípios análogos a C<sub>1</sub>, C<sub>3</sub> e C<sub>4</sub>:

$$E_1: \quad K_d(\neg D_6) \rightarrow \neg D_6$$

$$E_2: \quad K_s(E_1)$$

$$E_3: \quad (I(E_1, \neg D_6) \wedge K_s(E_1)) \rightarrow K_s(\neg D_6)$$

$$\vdash \neg D_6 \rightarrow K_d(\neg D_6) \quad [(1)]$$

$$E_1 \vdash D_6 \rightarrow \neg D_6$$

$$(2) \quad E_1 \vdash \neg D_6$$

$$\vdash I(E_1, \neg D_6) \quad [(2)]$$

$$E_3 \vdash K_s(E_1) \rightarrow K_s(\neg D_5)$$

$$E_2 \wedge E_3 \vdash K_s(\neg D_6)$$

$$(3) \quad E_2 \wedge E_3 \vdash D_6 \quad [(1)]$$

Os autores mostram, assim, em (2) e (3), que as assunções E<sub>1</sub> - E<sub>3</sub> são incompatíveis com os princípios da sintaxe elementar. Mas E<sub>1</sub> - E<sub>3</sub> são ainda mais plausíveis que C<sub>1</sub> - C<sub>8</sub> e ainda não contém o princípio de conhecimento por memória. Pela semelhança ao paradoxo do mentiroso [*Liar*], os autores decidem *batizar* este paradoxo – D<sub>6</sub> – de paradoxo do conhecedor [*Knower*]. Uma forma simplificada de apresentar este, mostrando a clara relação com o do mentiroso, é a seguinte proposição:

“K sabe que esta sentença é falsa” (Bar Hillel e Margalit, 1983:268)

Os autores resolvem então analisar as consequências epistemológicas do “paradoxo do conhecedor”. Uma série de restrições podem ser impostas em uma teoria formalizada do

conhecimento para evitar tais contradições. Uma primeira sugestão seria distinguir uma “linguagem de objetos” e uma “metalinguagem”: o termo “conhecer” ocorreria *apenas* na metalinguagem e se aplicaria apenas a sentenças da linguagem de objeto. Assim “K sabe que ‘K sabe que a neve é branca’” não seria admitido. Um recurso menos restritivo seria uma sequência de metalinguagens onde o termo “conhecer” só se aplicaria à metalinguagem “anterior” à esta na sequência. E, a mais drástica, seria negar alguma coisa que usamos na sintaxe elementar, como a autorreferência, por exemplo.

### 3. AS SOLUÇÕES EPISTÊMICAS

#### 3.1. As sete cartas

ALFRED JULES AYER INICIA SEU BREVÍSSIMO ARTIGO PROMETENDO UMA SOLUÇÃO MUITO SIMPLES ao problema. Primeiro, Ayer critica a solução de Quine de que o enforcado não pode tomar o anúncio como verdadeiro.

Tal solução – a de Quine – é “válida *ad hominem*” porém “não chega à raiz do problema”. Isto porque poderíamos, diz Ayer, adicionar a seguinte condição: entender surpresa somente no caso de haver de fato uma execução (como vimos, isto é o mesmo argumento apresentado nas cláusulas  $A_2'$  e  $A_2''$  de Kaplan e Montague). Assim o problema persiste.

Ayer desenvolve uma outra elaboração do problema que, ele pretende, irá mostrar que a resposta de Quine é insuficiente, ao mesmo tempo levando à solução do problema.

Suponhamos que um conjunto de sete cartas em sequência, incluindo o às de espadas [*Ace of Spades*] seja colocado na cela do prisioneiro em algum lugar sem a possibilidade de adulteração. Todo dia o guarda vem e saca uma carta. No dia em que ele sacar o às de espadas o prisioneiro será enforcado, desde que ele não saiba que o às será sacado naquele dia. O prisioneiro argumenta que o às de espadas não pode ser a última carta na sequência,

pois, neste caso, ele saberia que o ás seria sacado, e assim, também não poderia ser a sexta, pois com a eliminação da sétima a sequência é reduzida a seis e assim por diante.

Ayer nota que, neste caso, é irrelevante se o prisioneiro é ou não enforcado após o guarda sacar o ás ou mesmo se o guarda se lembra ou não de sacar as cartas. O que *importa* no argumento e o que é de fato a sua essência é mostrar que não pode haver um evento no qual é verdadeiro tanto que “é membro de uma dada sequência” e que “sua posição na sequência é incerta”, no sentido em que, quando alguém percorre a sequência (no caso, vai sacando as cartas), este alguém não sabe em que etapa o evento (no caso, sacar o ás de espadas) irá acontecer.

A última afirmação – sua posição na sequência é incerta – é ambígua e Ayer afirma que isto é deliberado, pois ele crê que essa ambiguidade é que engendra o *puzzle*. A ambiguidade está entre ser incapaz de prever que o evento vai ocorrer *antes que a sequência é percorrida*, e ser incapaz de prever que o evento vai ocorrer *no decorrer do percorrimeto, não importa quão longo seja*. No primeiro caso *há* incerteza, mas no segundo *pode* não haver.

Assim, se alguém é apresentado com tal jogo de um baralho de cartas (51 cartas) e é solicitado a adivinhar onde está o ás de espadas, este tem 1/51 chances de acertar (a menos que, diz Ayer, ironicamente, este tenha “percepção extra-sensorial”). Por outro lado, se alguém puder ir apostando *a medida* em que as cartas vão sendo sacadas poderá chegar um ponto em que ele irá apostar com toda a certeza – quando sobrar uma carta e o ás não tiver sido sacado.

Para “ver como isso resolve o *puzzle*”, Ayer volta ao problema com dois dias. Se a condição para o prisioneiro escapar é que ele *saiba* o dia do enforcamento, ele *não tem como* escapar, dado que ele *não sabe*, mas apenas pode adivinhar, com 50% de chances o dia. Mas se a condição for tal que haja *um certo momento* em que ele saiba o dia do enforcamento, é possível que ele escape caso não haja o enforcamento no primeiro dia. Uma falácia ocorre quando o segundo caso é projetado sobre o primeiro e é argumentado que “existem casos em que toda incerteza é removida”, não há incerteza no início. Conclui Ayer, é só pela não distinção dos dois casos que a *aparência* de uma antinomia se apresenta.

A forma como Ayer apresenta o seu argumento – por meio das cartas –, sem dúvidas é mais claro que as formulações alternativas. No entanto, este deve mesmo ser o maior mérito do artigo.

Ayer simplesmente ignora a possibilidade do problema apresentado com um dia só, ou melhor, um conjunto de cartas consistindo apenas de um às. Neste caso, o prisioneiro pode *saber* que o às será sacado, assim ele *não pode ser* enforcado.

Não há, é claro, contradição alguma desta conclusão em relação à formulação de Ayer, no entanto, parece evidente que não é bem isso que ele esperaria como resultado.

### 3.2. Paradoxo da loteria

DORIS OLIN INICIA SEU ARTIGO *THE PREDICTION PARADOX RESOLVED* DIZENDO QUE TODAS AS soluções anteriores não resolveram o problema e que a sua solução envolve uma aproximação com o paradoxo da loteria. O caminho, segundo ela, não é demonstrar que a prova é possível, pois isto é evidente. Assim, há que se “diagnosticar e isolar” os erros do argumento do aluno astuto.

Olin observa que a solução mais comumente aceita até então é a que procura construir o problema em termos puramente lógicos, onde o aluno interpreta o anúncio em termos de “dedutibilidade”. O aluno entende incorretamente o anúncio deste modo:

- (T) Haverá exatamente um exame na próxima semana e ele ocorrerá no dia D de modo que *esta proposição* e o fato que o exame não aconteceu antes do dia D não impliquem conjuntamente que o exame ocorrerá em D.

A proposição (T), nota Olin, além de ser uma forma muito pouco natural de interpretar o anúncio, de fato implica em contradição (lógica). Os defensores de tal

interpretação então apelam a versões revisadas de (T) para contornar a contradição. Porém Olin não crê que esta construção seja o caso.

O primeiro elemento para compreender o problema, nota ela, é que o [ato do] anúncio é crucial para problema. Não sabendo do anúncio, nenhum paradoxo. É vital que seja plausível supor que aluno possui boas razões para crer na ocorrência de um exame surpresa. E o único motivo para isso é que o aluno pode acreditar no professor, que é geralmente confiável [*reliable*], disse que tal exame ocorreria. Para Olin, estes conceitos epistêmicos são fundamentais para compreender o problema.

Para Olin, o que devemos entender por surpresa é que o aluno não estará *justificado* em crer, antes do dia do exame, que o exame vai ocorrer naquele dia.

A especificação do problema é dada com o seguinte anúncio S:

Um exame vai ocorrer exatamente em um dos dias de segunda até sexta; E se um exame ocorrer no dia D, você [o aluno] não estará justificado a crer nisto antes do dia D;

E, se não vamos responder à argumentação do aluno com objeções “triviais” – como a memória ou capacidade de raciocínio do aluno – temos que pressupor as seguintes coisas:

(P1) O aluno é um lógico experiente; Isto significa que se ele está justificado a crer em  $p_1, \dots, p_n$  que conjuntamente implica – ou “confirmam fortemente” –  $q$ , então o aluno é capaz de vê-lo;

(P2) O aluno lembra o anúncio durante toda a semana e lembra que o professor é confiável;

(P3) O aluno sabem sempre em que dia está e lembra se já houve ou não um exame;

Além destes pressupostos em relação às capacidades do aluno, Olin apresenta alguns pressupostos epistemológicos que devemos levar em conta:



(P4) Se A está justificado em crer em  $p_1, \dots, p_n$ , e  $p_1, \dots, p_n$  implica em  $q$  e A é capaz de vê-lo, então A está justificado em crer em  $p$ ;

(P5) Se A está justificado em crer em  $p_1, \dots, p_n$ , e  $p_1, \dots, p_n$  “confirma fortemente”  $q$ , A é capaz de vê-lo e não possui outra evidência relevante em relação a  $q$ , então A está justificado em crer em  $p$ ;

A argumentação do aluno ocorre pelo primeiro passo:

(1) Se o único exame ocorresse na sexta, então, na quinta de noite o aluno estaria justificado em crer que um exame ocorreria na sexta;

Tal conclusão é problemática, diz Olin, porque depende que ignoremos as evidências disponíveis ao aluno. O aluno ainda deve recordar, na quinta, que:

(A) Um exame vai ocorrer exatamente em um dos dias de segunda até sexta.

É de acordo com a afirmação (A) que o aluno esperava uma prova na sexta. Assim parece não razoável que ele suponha que não haverá prova na sexta, por (1). Porém, este raciocínio esquecemos de outra parte das evidências de S, a saber:

(B) Se um exame ocorrer no dia D, você [o aluno] não estará justificado a crer nisto antes do dia D.

Olin considera então que o aluno *não* estaria justificado em crer em (A) na quinta de noite. Suponhamos que ele esteja: Neste caso, ele apenas estaria justificado a crer em (A) se também estiver em crer em (B), pois “não há diferença epistemicamente relevante” entre as duas proposições. Entretanto, se S *estivesse* justificado em crer em (A) e (B), então, se dando conta da situação de quinta (nenhuma prova dada até então), ele *também* estaria justificado a crer em:

(C) Haverá um exame na sexta e eu *não* estou justificado (agora) a crer que haverá um exame na sexta.

Diz Olin que tal situação é impossível. Não é *razoável* acreditar em uma proposição com a forma ‘p e eu não estou justificado em crer em p’. Isto pois se alguém possui justificação de uma dada proposição, este não pode ser “culpado” epistemologicamente de crer em nesta proposição. No entanto, se alguém sabe que *não* possui justificação em crer nesta proposição então este estaria em “falta” ao crer.

O “desfecho”, conclui Olin, é que o aluno não estaria justificado em aceitar ambas (A) e (B) – pois isso implicaria acreditar em (C) o que estaria em uma relação “epistemologicamente problemática” com (1). E como não há motivos maiores para rejeitar uma ou outra, S não pode estar no direito de optar por aceitar apenas (A). Isto implica, diz Olin, que devemos rejeitar a premissa (P5), pois o aluno possui boas evidências que (A) é o caso, porém elas não o justificam a crer em (A).

A solução passa pelo fato de que, novamente, a *única* evidência para crer em (A) é o [ato do] anúncio. Assim, pensemos na seguinte mudança na apresentação do problema: Imaginemos que, além do anúncio do exame surpresa, o aluno possui fortes evidências independentes de que haverá um exame naquele período – como um processo automático e irreversível que foi colocado em curso por um supercomputador. Neste caso, o argumento acima não funcionaria (suponho que pelo fato de termos mais evidências de (A) do que de (B), o que problematizarei ao final da exposição). Ainda assim, poderíamos, nesta versão, “parar” o argumento do aluno no segundo estágio:

(2) Se o único exame ocorresse na quinta, então, estando na quarta de noite o aluno justificado em crer em (1), ele também estaria justificado em crer que um exame ocorreria na quinta;

Mesmo supondo que S esteja mesmo justificado em crer em (1) não podemos aceitar o resultado de (2), a saber, que seria impossível tal prova na quinta. Para estar justificado da ocorrência da prova na quinta o aluno deveria estar justificado em crer tanto em (A) como em (B), pois ele pode excluir a sexta apenas com base em (B). Mas o aluno não pode estar justificado em crer tanto em (A) como em (B), dado que isto resultaria em crer também em:

(D) Haverá um exame na quinta e eu *não* estou justificado (agora) a crer que haverá um exame na sexta.

Assim, nessas condições um exame poderia ocorrer em qualquer dos dias menos o último. Em suma, em todo estágio podemos usar o fato de que o aluno está justificado a crer em (A) para predizer a data do exame, mas, em algum ponto, para mostrar a impossibilidade de algum dos dias, temos de apelar também à crença justificada em (B). Ocorre que a crença *conjunta* em (A) e (B) sempre leva à crenças em proposições com a problemática semelhante ao paradoxo de Moore.

Por fim, Olin considera a versão “mais forte” de definição do paradoxo, a de um dia: “Haverá um exame surpresa amanhã”. Neste caso também chegaremos à proposição problemática:

(E) Haverá um exame amanhã e eu *não* estou justificado (agora) a crer que haverá um exame amanhã.

Como o estudante não pode crer em tal proposição ele não tem motivos para crer em (E), do mesmo modo que não tinha para crer em (C), e também, do mesmo modo, não tem razões para preferir (A) ou (B). No entanto, se pensamos no caso revisado, onde a evidência externa para (A) é mais forte (no caso do supercomputador) temos um resultado diferente: o aluno estará justificado em crer em (A) e não em (B), ou seja, que haverá um exame, porém que, evidentemente, este não será uma surpresa.

Em resumo o paradoxo é um paradoxo que deve ser entendido em termos epistêmicos e, ademais, a “utilidade” dele é mostrar que a premissa (P5) é problemática. Tal conclusão, aponta Olin, é análoga à conclusão de uma das maiores “escolas de pensamento” suscitada pelo *paradoxo da loteria*.

Em um jogo de loteria um bilhete é sorteado em um número de 1000 bilhetes. Digamos que **b1** é a proposição “o bilhete *i* não vai ser sorteado”. Cada uma dessas proposições é

altamente provável, ou seja, é razoável supor que cada uma delas, tomada *em separado*, é verdadeira. Assim, a conjunção das conclusões – b1 não vai ser sorteado e b2 não vai ser sorteado e ... e b1000 não vai ser sorteado – é “Nenhum bilhete será sorteado”. No entanto, a definição do problema nos diz que “Algum bilhete será sorteado”: as duas conclusões são contraditórias.

A relação do paradoxo da loteria com o do exame surpresa é então este: Mesmo que soe plausível *prima facie* estarmos justificados a crer em algo com base em um forte suporte por todas as evidências disponíveis, isto não confere o “direito” a crer.

Enfim, o resultado é a rejeição de (P5). E, se a visão de Olin está correta existe esta semelhança com o paradoxo da loteria. Ambos ensinam essencialmente a mesma lição: boas evidências não são suficientes para a crença justificada.

A análise do problema e a rejeição de (P5), assentam em duas assunções adicionais:

- (I) É impossível estar justificado em crer no par de proposições da forma “p e eu não estou justificado a crer em p”;
- (II) Se é impossível estar justificado em cada uma das proposições que são membros do conjunto  $p_1, \dots, p_n$ , e não há nenhum subconjunto de  $p_1, \dots, p_n$  em que isso seja verdadeiro, e você tem igualmente boas razões para crer em cada proposição, então você não está justificado em crer em qualquer uma das proposições.

A conclusão, nota Olin, é que “enquanto o anúncio pode ser verdadeiro, ele não pode ser aceito”. Neste sentido, é a mesma conclusão que o argumento de Wright e Sudbury. No entanto, estes, segundo ela, se utilizam de uma fundação filosófica diferente.

### 3.3. Variações recalcitrantes

EM SEU ARTIGO DE 1982, *RECALCITRANT VARIATIONS OF THE PREDICTION PARADOX*, ROY Sorensen nos diz que, como o *approach* não formal para a solução do problema não deu resultados (suponho que no período das décadas de 40, 50 e 60) a maior parte dos artigos sobre o tema voltou-se para a tentativa de solucioná-lo por meio da formalização (principalmente nas décadas 60 e 70). O objetivo da formalização é encontrar os pressupostos fundamentais que geram o paradoxo, e, subsequentemente, refutar ao menos um destes pressupostos problemáticos. O problema desse caminho, diz Sorensen, é falhar em 1) prover uma representação fiel que abarque *todas* as versões do paradoxo e 2) dar uma solução compreensiva o suficiente que resolva esta representação. As versões que apelam a uma interpretação autorreferente (iniciando por Shaw) falham especificamente no item 1. Sorensen pretende apresentar três variações do problema e, com estas, mostrar a verdadeira estrutura, antes não formulada, de uma versão generalizada do problema.

Wright e Sudbury, afirma Sorensen, buscaram rejeitar o princípio de retenção temporal do conhecimento (ou princípio da memória). Tal princípio já havia sido invocado por autores que comentam que alunos “da vida real” podem esquecer, ficar loucos, morrer... Enquanto outros autores buscaram estipular a não ocorrência de tais calamidades, para assim ter uma versão “ideal” do problema envolvendo pensadores ideais. No entanto, os autores não baseiam sua recusa nestas possibilidades não ideais. Eles defendem que o anúncio se parece com uma afirmação Mooreana  $p \wedge \neg B_a p$ , que, dizem, o agente *a* pode crer consistentemente que *foi* o caso, que *será* o caso, porém não que *é* o caso. Já o agente *b* pode crer a qualquer momento. Os autores crêem que o anúncio faz com que os alunos sejam informados de uma afirmação Mooreana, caso a prova não seja dada até o penúltimo dia.

Apesar de aceitar a rejeição do princípio de retenção temporal, Sorensen não concorda que apenas isso resolva o dilema. Para isso, vejamos sua primeira variação recalcitrante, o “paradoxo do aluno designado”:

Um exame será dado para um dos alunos: Art, Bob, Carl, Don e Eric. Estes ficam posicionados, um uma fila, de modo que o último, Eric, consegue ver as costas de todos (menos a sua), Don consegue ver as costas de todos menos Eric (e a sua) e assim por diante. O professor coloca uma estrela nas costas de cada aluno, sendo quatro estrelas prateadas e

uma dourada. O professor anuncia que a estrela dourada estará nas costas do aluno designado. Este terá de fazer a prova. Além disso, o professor diz que a prova será surpresa, no sentido em que o aluno designado não sabe que ele o é enquanto a formação original seja mantida. Um dos alunos, digamos, Art, objeta que tal exame é impossível: “Todos sabem que Eric não pode ser o aluno designado, pois, sendo, ele veria quatro estrelas prateadas e saberia que ele é o aluno designado, mas, pelo anúncio, o exame será uma surpresa, então, o aluno designado não pode saber que o é. Dado que Eric não pode ser o aluno designado, sobram Art, Bob, Carl e Don. Mas Don também não pode ser o aluno designado, por saber, baseado no argumento acima que Eric não pode ser e pelos mesmos motivos que excluem Eric. Do mesmo modo podemos eliminar Carl, Bob e Art, precisamente nesta ordem. Logo, o exame é impossível.”. Em seguida, o professor sorri e quebra a formação inicial. Carl então descobre que ele estava com a estrela dourada, sendo assim o aluno designado e, mais do que isso, sendo isto uma surpresa para ele, que tinha “comprado” o argumento.

Nesta variação o conhecimento é acumulado perceptualmente (visualmente) em vez de por memória. Poderia ser objetado que o aluno precisa reter o conteúdo do anúncio ao menos enquanto realiza sua dedução. O problema aqui é que se Eric tivesse a estrela dourada ele já saberia disso *antes* do anúncio, posto que ele vê quatro estrelas prateadas em sua frente. O anúncio, quando feito, informa Eric que a pessoa com a estrela dourada, ou seja, ele, é o aluno designado e, ao mesmo tempo, que não sabe que é o aluno designado. E como um pensador ideal não pode ser informado deste modo e se manter consistente, segue-se que um pensador ideal *não pode* ser informado deste modo. Assim, diz Sorensen, Wright e Sudbury não podem manter sua condição de informatividade.

Tal variação recalcitrante também revela errôneo o argumento de Chihara (e também J. McLelland e Craig Harrison). Estes mostram que o paradoxo só existe pelo princípio da reflexividade do conhecimento, ou “axioma KK”, ou seja, que se eu sei que *p*, eu sei que sei que *p*:

$$K_i \phi \rightarrow K_i K_i \phi$$

A rejeição se baseia no fato de que o princípio é controverso e que, dos princípios adotados ao derivar o paradoxo, este parece ser o mais descartável. No entanto o paradoxo reformulado não depende do princípio. A formalização de Sorensen do anúncio feito nesse caso variante é a seguinte:

$$\begin{aligned}
 (1) \quad & (p_1 \rightarrow \neg K_a p_1) \wedge \\
 & (p_2 \rightarrow (\neg K_b p_2 \wedge K_b \neg(p_1))) \wedge \\
 & (p_3 \rightarrow (\neg K_c p_3 \wedge K_c \neg(p_1 \vee p_2))) \wedge \\
 & (p_4 \rightarrow (\neg K_d p_4 \wedge K_d \neg(p_1 \vee p_2 \vee p_3))) \wedge \\
 & (p_5 \rightarrow (\neg K_e p_5 \wedge K_e \neg(p_1 \vee p_2 \vee p_3 \vee p_4))) \wedge \\
 & (p_1 \vee p_2 \vee p_3 \vee p_4 \vee p_5)
 \end{aligned}$$

Interpretamos do seguinte modo:  $p_1$  é 'Art é o estudante designado',  $p_2$  é 'Bob é o estudante designado' e assim por diante.  $K_a x$  significa que 'Art sabe que  $x$ ',  $K_b x$  que 'Bob sabe que  $x$ ' e assim por diante.

Os disjuntos devem ser entendidos do seguinte modo.

1<sub>1</sub>: Se for o caso que Art é o estudante designado então Art não sabe Art é o estudante designado;

1<sub>2</sub>: Se for o caso que Bob é o estudante designado então Bob não sabe que Bob é o estudante designado e Bob sabe que não é o caso que Art é o estudante designado;

1<sub>3</sub>: Se for o caso que Carl é o estudante designado então Carl não sabe que Carl é o estudante designado e Carl sabe que não é o caso que Art ou Bob sejam o estudante designado;

1<sub>4</sub>: Se for o caso que Don é o estudante designado então Don não sabe que Don é o estudante designado e Don sabe que não é o caso que Art, Bob ou Carl sejam o estudante designado;

1<sub>5</sub>: Se for o caso que Eric é o estudante designado então Eric não sabe que Eric é o estudante designado e Eric sabe que não é o caso que Art, Bob, Carl ou Don sejam o estudante designado;

1<sub>6</sub>: Um dos alunos, Art, Bob, Carl, Don ou Eric é o aluno designado;

Para satisfazer o critério de informatividade, diz Sorensen, ficamos tentados a adicionar a cada um dos disjuntos de (1) a fórmula  $\forall(x)K_x^9$ , onde  $x$  referencia cada um dos alunos. No entanto  $\forall(x)K_x(1)$  corresponde à uma variação do problema onde o anúncio do professor é feito de modo *privado* à cada um dos alunos. Neste caso, todo mundo sabe que (1), porém ninguém sabe que os *outros* sabem que (1): Don, por exemplo, não sabe que Eric sabe que (1), pois Don não sabe se o anúncio foi feito para Eric. Neste caso o anúncio, do ponto de vista de Don, parece perfeitamente plausível, se Eric for o designado.

Mas o anúncio é feito de modo *público*, assim a forma correta de representar isto é  $\forall(x)K_x(\forall(y)K_y(1))$ . Uma forma abreviada, com dois alunos apresenta a contradição em questão:

$$(2) \quad (p_1 \rightarrow \neg K_a p_1) \wedge (p_2 \rightarrow (\neg K_b p_2 \wedge K_b \neg(p_1))) \wedge (p_1 \vee p_2)$$

Deve ser possível então mostrar que é inconsistente que  $K_a(K_b(2))$ , ou seja, que Art saiba que Bob saiba da proposição (2). Sorensen se utiliza, além das regras de inferência sentenciais padrão, das seguintes regras <sup>10</sup>:

Regras de distribuição do conhecimento (RD):

$$K_\phi(A \wedge B) - K_\phi A \wedge K_\phi B$$

$$K_\phi(A \rightarrow B) - K_\phi A \rightarrow K_\phi B$$

Regra de inferência (RI):

$$\vdash A - K_\phi A$$

Regra da veracidade do conhecimento (RV):

$$K_\phi A - A$$

Regra de fechamento (RF):

---

<sup>9</sup> Por algum motivo o autor omite o quantificador universal  $\forall$  nas suas fórmulas. Eu preferi não omitir.

<sup>10</sup> Utilizo o caractere “-” para significar inferência. O *standard* é colocar as fórmulas com uma linha entre elas, verticalmente.



$K_\varphi A, K_\varphi B \rightarrow K_\varphi C$ , quando for o caso em que  $A \wedge B \vdash C$

Com estas, podemos realizar a seguinte derivação:

1.	$K_b(2)$	
2.	$p_2$	
3.	$K_b(p_1 \vee p_2)$	1, RD
4.	$(2)$	1, RV
5.	$\neg K_b p_2 \wedge K_b \neg p_1$	2, 4
6.	$K_b p_2$	3, 5, RF
7.	$\neg K_b p_2$	5
8.	$\neg p_2$	2, 6, 7, reductio
9.	$K_b(2) \rightarrow \neg p_2$	1, 8
10.	$K_a(K_b(2) \rightarrow \neg p_2)$	9, RI
11.	$K_a K_b(2) \rightarrow K_a \neg p_2$	10, RD
12.	$K_a K_b(2)$	
13.	$K_a \neg p_2$	11, 12
14.	$K_b(2) \rightarrow (2)$	1, 3
15.	$K_a(K_b(2) \rightarrow (2))$	14, RI
16.	$K_a K_b(2) \rightarrow K_a(2)$	15, RD
17.	$K_a(2)$	12, 16
18.	$K_a(p_1 \vee p_2)$	17, RD
19.	$K_a p_1$	13, 18, RF
20.	$p_1 \rightarrow \neg K_a p_1$	17, RV
21.	$p_1$	19, RV
22.	$K_a p_1 \wedge \neg K_a p_1$	19, 20, 21

Assim, com a contradição, a proposição  $K_a K_b(2) \wedge K_b K_a(2)$  não pode ser um anúncio público e informativo para *ambos* os alunos.

Tanto o paradoxo da predição como o paradoxo do aluno designado envolvem uma única e rígida ordem de eliminação. O mesmo não pode ser dito da próxima versão

recalcitrante, o paradoxo da posição indescobrível. Este paradoxo envolve um jogo com as seguintes posições:

1	2	3
4	5	6
7	8	9

O objetivo é descobrir onde o jogador foi inicialmente colocado. O jogador pode se mover nas quatro direções (para Cima, para Baixo, para a Esquerda e para a Direita), uma célula por por vez. Se o jogador se deparar com uma “parede”, seu movimento é sublinhado (e.g. se ele se moveu para cima e bateu na parede o movimento é C). Estas “batidas” ajudam o jogador a descobrir sua posição inicial. Digamos que o jogador tenha feito os movimentos C, C, E. Neste caso, fica claro que a posição inicial era 7, pois, fosse 1, 2 ou 3, os dois Cs seriam sublinhados, porém não o E no caso de 2 e 3; fosse 4, 5, ou 6 o segundo C seria sublinhado, porém não o E no caso de 5 e 6; fosse 8 ou 9, o E não seria sublinhado. Estes três movimentos, enfim, permitem a dedução de que apenas 7 é compatível com tal estado de coisas.

Se ao jogador é dado apenas dois movimentos, é possível coloca-lo numa situação em que a posição inicial seja indescobrível. Se ele é colocado inicialmente, digamos, na posição 4, para toda sequência possível de movimentos existe uma outra posição compatível (C, C e B, B são compatíveis com 5 e 6; é. E, E e D, D são compatíveis com 1 e 7; C, E é compatível com 7; C, D é compatível com 5, 7 e 8; B, E é compatível com 1 e B, D é compatível com 1, 2 e 5). Agora suponha que você proponha este jogo para um jogador e diga a este “Você foi colocado em uma posição indescobrível”. Ele devolve então o seguinte argumento:

“Suponhamos então que eu esteja em uma posição indescobrível. Segue-se que eu não posso estar em nenhuma das quinas – 1, 3, 7 e 9 –, dado que cada uma delas possui um movimento perfeitamente não ambíguo; Se estou em 3, por exemplo, C, D é um movimento que identifica perfeitamente esta célula. Tendo eliminado as quinas, podemos eliminar também as outras posições nos cantos – 2, 4, 6 e 8 –, pois estas podem ser eliminadas se o

primeiro movimento “bater” em uma parede. Se estou em 2, por exemplo, o primeiro movimento C já indicaria isto. Dado que sobra apenas a posição 5, eu descobri minha posição.”

E o problema é ainda mais complicado:

“O absurdo da suposição é tornado ainda mais manifestamente absurdo pela existência de outros oito argumentos, com outras oito conclusões em relação à conclusão inicial. Por exemplo, posso concluir que estou em 6, eliminando primeiro as quinas, depois 2, 4 e 8, e sobrando 5 e 6, eliminando 5 (as sequências E, E e D, D, eliminariam 5). E se um argumento é individuado por ter uma ordem distinta de eliminação, haveriam ainda mais de 9 argumentos, pois o 6 poderia ser eliminado nesta ordem: 7, 4 (C, E), 8, 1, 2, 5, 9 e 3. Então, eu não posso ser colocado em uma posição indescobrível.” (Sorensen, 2002:360)

Por fim, a última versão recalcitrante é “paradoxo da virgem sacrificial” que serve para mostrar que o sujeito não precisa saber quantas alternativas existem. A cada cinquenta anos os habitantes de um paraíso tropical sacrificam uma virgem para o vulcão local em uma cerimônia elaborada. Todas as virgens são trazidas com os olhos vendados para a entrada do vulcão, dão as mãos e podem apenas comunicar uma sentença: “Ninguém à sua direita é a virgem sacrificial”. Esta comunicação se dá pelo aperto forte da mão da virgem à esquerda. As virgens são confiáveis [*reliable*] e ligadas fortemente ao dever de comunicar isto se, e apenas se elas sabem que tal informação é verdadeira. Além de tudo isso, elas sabem que uma condição para ser a virgem sacrificial é que esta permanece ignorante desta “honra” até que esta é arremessada ao vulcão. O chefe deve iniciar pela virgem mais à esquerda, e, se considerar a oferta aceitável, empurra-la e dizer para as outras virgens irem para casa. Se for inaceitável ele manda esta virgem para casa e repete o procedimento com a nova virgem mais à esquerda. Este processo continua até que uma virgem é sacrificada, então é sabido que uma será sacrificada. Após ouvir o anúncio de que uma virgem será sacrificada, alguém objeta que tal cerimônia não pode ocorrer:

“A virgem mais à direita sabe que é a mais a direita, pois sua mão não está dada. Ela sabe que se ele for ‘oferecida’, então nenhuma das virgens à sua esquerda foram sacrificadas. Então, se ela é a virgem sacrificial então ela sabe que é a virgem sacrificial. Como a virgem sacrificial não pode saber do seu *status* antes de ser empurrada, a virgem mais à direita não pode ser a virgem sacrificial. Este conhecimento a obriga a apertar a mão da virgem à sua esquerda, sinalizando ‘Ninguém à sua direita é a virgem sacrificial’. Esta outra virgem, por sua vez, ou é a mais da esquerda ou está no meio. Se ela é do meio, ela pensará que, se ela for a virgem sacrificial ela saberá que nenhuma à sua esquerda foi, e, como ela sabe que nenhuma à direita será, ela pode deduzir que ela será a virgem sacrificial. Mas como esta não pode saber de sua condição, ela pode apertar a mão da virgem à sua esquerda. E podemos seguir este raciocínio para todas as virgens do meio até chegar à virgem mais da esquerda. Esta saberá que nenhuma outra virgem será a sacrificada, então ela será, porém ao deduzir isso, ela não pode ser. Logo, tal cerimônia é impossível.”

As virgens das pontas sabem que existem pelo menos duas. As virgens do meio sabem que existem ao menos três. Nas versões anteriores é fundamental que o sujeito saiba quais as alternativas em questão. No exame surpresa, os alunos precisam saber que o exame ocorrerá em algum dos dias do período. No da virgem sacrificial os participantes não tem nem mesmo uma estimativa aproximada do número de virgens em linha. As do meio apenas sabem que estão no meio de uma linha de tamanho arbitrário. Assim, *não é essencial* que o sujeito saiba a ordem em que os membros da série estejam arranjados. Ao contrário do paradoxo do aluno designado, as virgens do meio *repetem exatamente a mesma indução*, mas não replicam as deduções das outras. Nenhuma possui mais conhecimento que as outras. No caso dos alunos designados, Don só pode eliminar a si mesmo replicando o raciocínio de Eric, Carl só pode eliminar a si mesmo replicando o raciocínio de Don, e assim por diante.

Em resumo, estas três variações do paradoxo da predição mostram que o princípio da retenção temporal, o princípio KK (reflexividade), a ordem de eliminação e o número de alternativas a serem eliminadas são todos não essenciais para o paradoxo da predição. Como

as soluções passadas assumem estes como essenciais, diz Sorensen, ainda não temos uma formulação da estrutura do paradoxo da predição.

### 3.4 Os *blindspots* condicionais

JÁ EM 1984, SORENSEN, EM SEU ARTIGO, *CONDITIONAL BLINDSPOTS AND THE KNOWLEDGE Squeeze: A Solution To The Prediction Paradox*, busca dar uma solução ao paradoxo do exame surpresa, e também às versões recalcitrantes por ele elencadas.

Após argumentar que os “autorreferencialistas” distorceram o problema, Sorensen afirma que sua solução focará no passo inicial, ou seja, que mesmo um exame na sexta pode surpreender os alunos. Para ele, Wright e Sudbury, apoiados na ideia das proposições mooreanas, rejeitam o princípio da retenção do conhecimento: certas proposições deixam de ser conhecidas para certos sujeitos e em certos momentos, e isso resolve o problema. Porém, apesar de concordar com essa rejeição, Sorensen acredita ter demonstrado, com o paradoxo dos estudante designado, que tal assunção não é fundamental para a versão mais geral do problema.

Sua solução depende do que ele chama de *blindspot*<sup>11</sup> epistêmico. Uma proposição  $p$  é um blindspot se para uma pessoa  $a$ , no tempo  $t$ , se e somente se  $p$  é consistente enquanto  $K_ap$  ( $a$  sabe que  $p$  em  $t$ ) é inconsistente. Assim proposições mooreanas são *blindspots*: “Está chovendo mas Andy não sabe disso” é um blindspot para Andy, mas não para Woody. A proposição mencionada acima, formalizada como  $p \wedge \neg K_ap$  é uma proposição consistente. Porém, vejamos o que ocorre quando Andy a conhece:

- |                                |           |
|--------------------------------|-----------|
| (1) $K_a(p \wedge \neg K_ap)$  |           |
| (2) $K_ap \wedge K_a\neg K_ap$ | [(1)]     |
| (3) $K_ap \wedge \neg K_ap$    | [(2), RV] |
| (4) $\therefore \perp$         | [(3)]     |

---

<sup>11</sup> Eu poderia traduzir por ponto cego, mas me dá nos nervos.

A aceitação dos *blindspots* implica na negação de três generalizações acerca do escopo do conhecimento, a saber:

- a) Princípio do acesso irrestrito: Tudo que pode ser verdadeiro, pode ser conhecido (e.g. “ninguém sabe de nada”);
- b) Princípio do acesso intertemporal: Tudo que pode ser conhecido por alguém em um momento, pode ser conhecido por este em qualquer outro momento (e.g. João saberá que foi adotado aos 18 anos);
- c) Princípio do acesso interpessoal: Tudo que pode ser conhecido por alguém, pode ser conhecido por qualquer outra pessoa (John nunca soube, não sabe, nem saberá que foi adotado);

Uma proposição é um *blindspot* condicional para *a* em *t* se e somente se ela não é um *blindspot* mas é equivalente a um condicional cujo consequente é um *blindspot*. Um exemplo é “Se Rafael sobreviver, ele será o único que sabe disso”. Este é um *blindspot* condicional apenas para Rafael. Assim, é possível conhecer a proposição e é possível conhecer seu antecedente. Não é possível, no entanto, conhecer *em conjunto* o antecedente e o *blindspot* condicional. Voltemos ao exemplo de Andy, adicionando a proposição *r*, interpretada como “Andy está drogado”:

- (1)  $K_a r \wedge K_a (r \rightarrow (p \wedge \neg K_a p))$
- (2)  $K_a r \rightarrow K_a (p \wedge \neg K_a p)$  [(1)]
- (3)  $K_a (p \wedge \neg K_a p)$  [(1), (2)]

Para mostrar a relação dos *blindspots* condicionais com o problema em questão, Sorensen avalia uma versão com dois dias. Sendo *p*<sub>4</sub> a proposição “o exame ocorre na quinta” e *p*<sub>5</sub> “o exame ocorre na sexta”, formulamos o problema assim:

- (1)  $(p_4 \wedge \neg K_a p_4) \vee (p_5 \wedge \neg K_a p_5)$

Ou seja, “o exame ocorre na quarta e Andy não sabe disso ou o exame ocorre na quinta e Andy não sabe disso”. Mas podemos ainda modificar o exemplo deste modo:

$$(2) \neg(p_4 \wedge \neg K_a p_4) \rightarrow (p_5 \wedge \neg K_a p_5)$$

Ou seja, Andy se encontra com um *blindspot* condicional. Ele pode conhecer o anúncio do professor, porém, *ao conhecer o antecedente* (se o exame não for dado na quinta) não pode mais conhecer o anúncio. O padrão geral de um paradoxo de predição (seja o do exame surpresa, o do estudante designado, o da virgem sacrificial...) é o seguinte: Primeiro apresenta-se um caso em que um sujeito (ou grupo) *a* fica sabendo de um *blindspot* condicional para *a* (“Andy, haverá um exame surpresa na próxima semana”). Em seguida, se descreve como *a* aprende que os antecedentes do *blindspot* condicional são verdadeiros (“...o exame não foi dado até a quinta...”). Por fim, surgem duas questões distintas. A primeira é “É possível que *a* conheça tanto o *blindspot* condicional e seu antecedente?”. Como vimos, a resposta é *não*. Porém se a pergunta é “É possível que tanto o *blindspot* condicional e seu antecedente sejam verdadeiros?”, a tendência natural de responder “não” é fruto do desconhecimento das operações dos *blindspot* condicionais. A resposta é *sim*.

Assim, a solução de Sorensen para todos os casos é apontar para o erro de inferir a negação do *consequente* a partir da impossibilidade de conhecer *em conjunto* o *blindspot* condicional e o antecedente. Veremos mais sobre os *blindspots* no argumento de Levy.

### 3.5 Recalcitrâncias domadas

OLIN, EM SEU ARTIGO *THE PREDICTION PARADOX: RESOLVING RECALCITRANT VARIATIONS*, DE JUNHO de 1986, procura mostrar que o seu argumento continua válido para as versões arreadas de Sorensen.

No caso do paradoxo do aluno designado, Doris argumenta do seguinte modo: Consideremos (A1) a parte do anúncio do professor que diz que “a estrela dourada determina o aluno designado para o exame” enquanto a proposição (B1) diz que “o aluno designado não pode saber que ele é o aluno designado antes da formação ser quebrada”. Para Olin, o argumento que os alunos utilizam para demonstrar a impossibilidade dessa

designação é parado no mesmo passo em que a versão original do paradoxo da predição, a saber:

(1) Se Eric fosse o aluno designado, então ele estaria justificado em crer nisto antes de quebrar a formação (pois ele veria os outros quatro com a estrela prateada, como vimos acima).

Porém, ele forma a crença (1) com base em (A1). Porém, (A1) e (B1) possuem o mesmo peso epistêmico para ele, ou seja, não há motivo para preferir (A1) a (B1) e cada um deles só pode ser mesmo tomado como verdadeiro, caso o disjuncto de ambos seja, (A1) e (B1). Mas, neste caso, ele também estaria justificado em crer em:

(2) Eu sou o aluno designado e eu não estou justificado em crer que eu sou o aluno designado.

Sendo isto impossível para Eric crer, Eric não pode crer no anúncio [completo] do professor, *se* ele for o aluno com a estrela dourada! Assim, diz Olin, o paradoxo se resolve. Veremos adiante que tal conclusão não resolve o problema, engendrando um novo paradoxo, ao examinarmos o trabalho de Ken Levy.

De modo análogo, Olin, responde ao paradoxo da posição indescobrável. No primeiro passo do argumento, o argumento diz:

(1) Se o jogador estivesse em alguma das quinas – 1, 3, 7 ou 9 –, então ele estaria em uma posição descobrável.

Sendo (A2) as regras do jogo e (B2) a informação “Você está em uma posição indescobrável”, suponhamos que o jogador está na posição 3 e realiza a jogada C,D. Por (A2) e (B2) temos o problema análogo:

(2) A posição inicial é 3 e eu não estou justificado em crer que a posição inicial é 3.



Como é impossível que algum jogador creia em tal proposição, e o jogador tem de aceitar (A2) e (B2), ele não pode aceitar ambas. Neste ponto Olin faz uma observação em relação à uma *suposta* diferença desta variação do paradoxo: de acordo com ela o anúncio se torna incrível *apenas* quando o jogador está em alguma das quinas e *realizou* certos os movimentos. Em qualquer outro caso, diz ela, não há razão para crer que (A2) e (B2) sejam incríveis. Mas, o relevante é que mesmo estando em 3 e fazendo os movimentos certos (C,D) o jogador *não poderia* inferir sua posição inicial. Assim, os cantos não podem ser eliminados e o paradoxo também.

O mesmo também pode ser aplicado ao paradoxo da virgem sacrificial.

### 3.6. Triatlon

KEN LEVY, EM SEU ARTIGO *THE SOLUTION TO THE SURPRISE EXAM PARADOX* BUSCA REFUTAR UMA das premissas fundamentais do paradoxo, a saber, o passo inicial da negação do último dia como sendo impossível de ocorrer o exame, argumentando que a solução depende da conjunção de três argumentos já existentes na literatura, o argumento do “Anúncio Improjecionável”, o argumento de “Wright e Sudbury” e o argumento dos “*Blindspots* Epistêmicos Condicionais”.

Nossa intuição diz que o aluno não pode se surpreender com um exame que ele *esperava* na noite anterior e, do mesmo modo, o aluno não pode se surpreender com um exame que ele estava *certo*, na noite anterior, que ocorreria no dia seguinte. Estas intuições, somadas ao fato de que um *ex hypothesi* professor confiável *prometeu* um exame até o final do período letivo, ao fato de que o exame ainda não foi dado e ao fato de que estamos a um dia do final do período letivo parecem tornar indisputável que o estudante não irá se surpreender com um exame no último dia. Porém, o primeiro passo para a refutação é compreender que *surpresa*, no contexto do *SEP* significa a ausência de *conhecimento*, na noite anterior, de que o exame ocorreria no dia posterior. E, isto é fundamental importância, pois nem expectativa nem certeza psicológica são suficientes para *conhecimento*. O segundo é mostrar que a conclusão de que um exame no último dia não irá surpreender o aluno é *falsa*. Há, para Levy,

três argumentos – já citados acima – que estabelecem esta mesma conclusão. Porém, só o *alinhamento dos três* é que irá definir a questão.

Levy está interessado, como grande parte dos autores voltados às soluções epistêmicas, na versão mais forte do *SEP*. Assim, sua versão depende que tomemos o aluno astuto como *maximamente* racional, ou seja, “um mestre da lógica, ele evita contradições, conhece todas as verdades lógicas e crê em todas as consequências lógicas daquilo que ele crê” (Levy, 2004:134).

Vejamos a formulação do problema: O professor realiza, na terça-feira, o anúncio (A) de que haverá um exame surpresa na quarta, quinta ou sexta. Para dividir a conjunção das duas promessas, definimos (A)<sub>E</sub> como a afirmação “haverá um exame quarta, quinta ou sexta” e (A)<sub>S</sub> é a afirmação “o exame será uma surpresa”. Após (A) ser feito o aluno astuto apresenta o argumento “da impossibilidade” que consistem em dois subargumentos: o da “não sexta-feira” e o do “*backtracking*”. Se o exame não for dado até a quinta-feira a promessa de que o exame será uma surpresa – (A)<sub>S</sub> – será violada. E, dado (A), sexta não pode ser um dia para o exame (“não sexta-feira”). Dado o argumento da não sexta-feira, que exclui a sexta como um dia para a ocorrência do exame, se o professor não tiver dado a prova até quarta, do mesmo modo, ele não poderá ser feito na quinta, pois isso violaria (A)<sub>S</sub>. Ele só poderá ser feito na quarta, o que também violaria (A)<sub>S</sub> (*Backtracking*). A conclusão do aluno astuto é que o exame é impossível – (A)<sub>E</sub> não pode ocorrer. Ainda assim temos “a intuição”, a saber, que a conclusão do argumento da impossibilidade *não pode* ser o fim da história. Se considerarmos o argumento que se move adiante de quarta até sexta, a prova surpresa parece possível novamente. Nossa “intuição” é que uma prova na quarta ou na quinta será uma surpresa para o aluno. Então algo *deve* estar errado. Ou bem o argumento da impossibilidade ou a intuição estão errados. O desafio é determinar *qual e porquê*.

Aqui, é importante distinguir a ideia da “racionalidade máxima” de “infalibilidade”. Isto porque, como vimos, pela “intuição”, o argumento do aluno astuto *falha*. No entanto, não poderá ser dito que o argumento falha por ser *inválido*. Se fosse, *teríamos* que concluir que o aluno *não é* maximamente racional. O argumento deve ser considerado válido, porém deve ser argumentado que o aluno parte de premissas que lhe *parecem* perfeitamente plausíveis,

porém *não são*. Neste cenário o aluno poderá continuar sendo considerado maximamente racional. Ser maximamente racional só protege o aluno astuto de erros tolos e inferências errôneas, porém não o impede de iniciar com premissas falsas porém inicialmente plausíveis.

Levy não depende de que surpresa seja entendida como falta de conhecimento, mas apenas falta de justificação. Não há que se preocupar com a possibilidade de astuto estar justificado *falsamente*: Ele só se surpreende com a ocorrência do exame e não com a não ocorrência. Assim, (A) equivale a:

- (a) Haverá um exame na sexta e astuto não estará justificado em crer nisto na quinta de noite.
- (b) Haverá um exame na quinta e astuto não estará justificado em crer nisto na quarta de noite.
- (c) Haverá um exame na quarta e astuto não estará justificado em crer nisto na terça de noite.

Uma resposta ao argumento do *backtracking* que falha, porém nos é instrutiva é a seguinte: Dizer que a quinta é um dia impossível para o exame é claramente falso. Independentemente do que prova o argumento da não-sexta ainda é *fisicamente* possível que o professor dê o exame na sexta. Ocorre apenas que se ele o der ele não vai estar cumprindo a sua promessa de dar um exame *surpresa* (neste caso ele cumpre (A)<sub>E</sub> porém não (A)<sub>S</sub>). Sexta não é um dia *impossível* de ocorrer o exame do mesmo modo que sábado o é (dadas nossas assunções!). A melhor forma de apresentar a conclusão do argumento da não-sexta é: Sexta não é um dia possível para o exame *se* o professor vai cumprir (A). Assim, não podemos seguir o *backtrack* – seguir a mesma lógica para quinta e quarta – pois astuto não pode assumir que o antecedente – o professor vai cumprir (A) – é verdadeiro. A refutação deste argumento que nos ensinou algo é que pode ser *fisicamente* possível dar o exame na sexta, ou em qualquer outro dia, porém, não é *logicamente* (nem fisicamente) possível dar o exame na sexta *e ainda* cumprir (A). Astuto *deve pressupor* o cumprimento de (A) se o professor *puder*. Assim ele *pode* concluir que sexta é um dia impossível para o exame.

Mas porque pressupor que o professor vai cumprir (A) se puder (ou “argumento contra Quine”)? Primeiro porque o argumento da impossibilidade depende desta premissa; Segundo, porque rejeitar esta premissa é não atacar a versão mais forte do SEP, é atacar uma versão mais fraca, a qual Quine “já resolveu”. A versão em que astuto assume o cumprimento de (A) (se possível) é mais forte que uma versão que ele não assume isto. Nesta versão teríamos o seguinte argumento:

- (1) Assuma que o professor *não* irá necessariamente cumprir (A) se ele puder;
- (2)  $\therefore$  Astuto não está justificado em acreditar em (A). [(1)]
- (3) A única base para astuto acreditar em (A)<sub>E</sub> e (A)<sub>S</sub> é (A) mesmo.
- (4)  $\therefore$  Se astuto não está justificado em acreditar em (A), então ele não está justificado em acreditar em qualquer parte de (A) – (A)<sub>E</sub> e (A)<sub>S</sub>. [(3)]
- (5)  $\therefore$  Astuto não está justificado em acreditar em (A)<sub>E</sub>. [(2), (4)]
- (6)  $\therefore$  Astuto não está justificado em acreditar em qualquer noite que haverá um exame no próximo dia – mesmo que nenhum exame tenha sido dado até quinta. [(5)]
- (7)  $\therefore$  Um exame em qualquer dia – mesmo sexta – será uma surpresa. [(6)]

Para Levy este raciocínio de fato *resolve* uma versão do SEP – a versão em que (1) é o caso, a versão quineana. Mas ele não resolve a versão em que (1) *não* é o caso. Assim, assumir (1) não resolve o problema. Podemos, por exemplo, assumir que o professor é um computador *infalível*. Até que ponto esta abstração do problema é realmente factível é um ponto muito interessante, porém Levy não avança muito neste terreno. Certamente seria por aqui que Weintraub discordaria de sua solução, apresentando sua versão pragmática que vimos acima.

O argumento (1)-(7) é estranho, pois, assumindo que o professor não necessariamente vai cumprir (A) concluímos que ele irá cumprir (A) *desde que* dê a prova. O argumento da impossibilidade não pode estar correto, visto que se o argumento da impossibilidade está correto então um exame surpresa não pode ser dado em qualquer dos dias. Isto é o mesmo que dizer alguma das afirmações é verdadeira:

- (a) na terça de noite astuto está justificado em crer que haverá um exame na quarta;
- (b) se quarta passar sem exame, na quarta de noite astuto está justificado em crer que haverá um exame na quinta;
- (c) se quinta passar sem exame, na quinta de noite astuto está justificado em crer que haverá um exame na sexta;

Mesmo que isto não convença o leitor, Levy apela ao argumento da “intuição expandida”. Imaginemos um período de 300 dias e uma prova ocorrendo no dia 155. Astuto deve estar justificado em crer que o exame irá ocorrer no dia 155 na noite do dia 154. Como astuto teve a mesma crença nos 153 dias anteriores *não há motivo* para crer que ele está justificado neste dia mais do que [não] estava nos outros. Ele simplesmente não estava justificado, e, ademais, não há diferença fundamental entre o problema com 3 ou 300 dias.

Então temos de apreciar o argumento das “intuições enganosas”. O argumento da não sexta é muito poderoso. Sua força vem de duas intuições que nos enganam. Suponha que é quinta de noite e o exame não foi dado. Nossa primeira intuição é que ele será dado na sexta – dado (A) e a premissa de que o professor manterá (A) – logo, o exame não será surpresa para astuto. O que esta versão deixa passar é que é necessário, como já vimos, *mais* do que esta mera expectativa. Esta expectativa tem de ser *justificada*. Ainda assim o argumento seduz, afinal, nossa segunda intuição é que astuto não meramente *espera* que o exame seja dado na sexta, mas sim que ele pode estar *quase certo* disso. Mas mesmo a certeza é não suficiente para a crença justificada, visto que *certeza* é um aspecto psicológico de crenças, enquanto a justificação é uma propriedade *normativa* de crenças. Desse modo é muito importante notar que pode ser o caso que astuto esteja *certo* de que o exame ocorrerá na sexta mas não esteja *justificado*. Se o aluno estiver *apenas certo* da ocorrência do exame e ele de fato ocorrer na sexta, temos que entender isso *como uma surpresa*. Pode parecer contra intuitivo, mas o problema é que usamos surpresa de um modo bem técnico a saber “sem justificativa”. Se trocamos um pelo outro, fica mais claro: “se estou certo de algo, isto não quer dizer que estou justificado”. Assim fica mais claro! Agora o autor buscará somar seus argumentos para mostrar que astuto não está justificado.

O argumento da não-sexta repousa sobre a premissa implícita de que (A) pode ser aplicado diretamente na quinta de noite. Porém, esta aplicação representa uma mudança de “Haverá um exame surpresa até o final do período” – (A) – para “Haverá um exame surpresa *amanhã*” – usemos (A\*). (A\*)<sub>E</sub> e (A\*)<sub>S</sub> juntos parecem colocar astuto em uma situação epistemológica insustentável. (A\*)<sub>E</sub> justifica astuto a crer que o exame será dado na sexta enquanto (A\*)<sub>S</sub> afirma que astuto não está justificado a crer que o exame será dado na sexta, o que parece uma contradição. Tal afirmação é semelhante ao paradoxo de Moore: “Há um fogo nesta sala e eu não acredito que haja”. Porém, envolve o conceito de justificação e não meramente de crença: “Haverá um exame amanhã e eu não estou justificado a crer nisso”, cuja forma mais geral é “p e eu não creio justificadamente em p”.

Aqui, Levy busca alento no trabalho de Claudio de Almeida. Para ele, tanto o paradoxo de Moore, envolvendo o conceito de crença, como (A\*) envolvendo o de crença justificada, são proposições “*Moore-absurd*”. Uma proposição *Moore-absurd* não é, de fato, autocontraditória, porém é *irracional* que se creia nela. Notemos que “Haverá um exame amanhã e eu não estou justificado a crer nisso”, como já argumentaram os primeiros autores a escrever sobre o problema, é *possível* que seja verdadeiro. Para ser, basta que tal “pensamento” esteja apenas na cabeça do meu professor, que ele nunca tenha me dito tal coisa e que ele vá mesmo dar a prova no próximo dia (assim, colocando os indexicais nos devidos lugares, ele viraria “Vou dar uma aula amanhã e astuto não sabe disso”). O que torna a proposição problemática é quanto *eu* a considero como uma de minhas crenças – e, no caso, das justificadas. Por distribuição lógica, para eu crer nesta proposição, eu devo considerar que “eu *creio* que p e eu *creio* que eu não estou justificado a crer em p”. Neste caso, Almeida afirma que eu crer que eu não estou justificado a crer em p constitui uma razão para *remover* p do meu conjunto de crenças. No entanto eu creio que p. E se eu tenho motivos para não crer em parte da conjunção, isso me dá motivos para desacreditar a conjunção inteira, que é a fonte primária desta “cadeia de crenças”. Para Almeida, isto é possuir uma crença epistemologicamente “*self-defeating*”.

Se aceitamos a apresentação das proposições *Moore-absurd* de Almeida não devemos considerar (A\*) como autocontraditória, mas como *fortemente* inconsistente. Sendo assim, é irracional acreditar em (A\*). A preocupação com esta irracionalidade é que motivou,

segundo Levy, as “soluções pragmáticas”, como a de Ruth Weinstraub ao SEP. Em resumo, esta posição (a de Weinstraub) defende que (A), ao ser interpretada na quinta de noite, não se torna (A\*), porém se torna algo como “Haverá um exame amanhã porém ele não será mais uma surpresa – como eu havia prometido. Espero que tenham estudado.”. Apesar de não cumprida, espera-se que (A)<sub>s</sub> sirva seu *propósito real*. Levy concede que a interpretação pragmática dá mais sentido a (A) na quinta de noite do que (A\*), no entanto, há um alto preço a ser pago: o argumento não responde a questão do argumento da não-sexta pois este *depende* de entendermos (A) como (A\*) neste dia.

Então Levy, se volta ao argumento de Wright e Sudbury. Este é um *reductio ad absurdum* do argumento da não-sexta:

- (1) Um exame na sexta não seria surpresa (assumindo que o argumento da não-sexta está correto).
- (2) ∴ Se o exame não foi dado até quinta, então o professor não tem como cumprir (A)<sub>s</sub>. [(1)]
- (3) ∴ Astuto não está justificado a crer em (A)<sub>s</sub>. [(2)]
- (4) (A)<sub>s</sub> é uma parte essencial de (A): Se alguém não está justificado a crer em (A)<sub>s</sub> então não está a crer (A).
- (5) ∴ Astuto não está justificado a crer em (A). [(3), (4)]
- (6) Se alguém não está justificado a crer em (A), ele não está justificado em crer em qualquer parte de (A), como (A)<sub>E</sub>.
- (7) ∴ Astuto não está justificado a crer em (A)<sub>E</sub>. [(5), (6)]
- (8) ∴ Astuto não está justificado a crer, na quinta, que haverá um exame na sexta. [(7)]
- (9) ∴ Um exame na sexta seria uma surpresa. [(por definição de surpresa e 8)]
- (10) ⊥ [(1), (9)]
- (11) ∴ (1) é falso: Um exame na sexta será uma surpresa [(1), (10)]

Wright e Sudbury param por aí. Esse *reductio* parece mostrar que o exame é perfeitamente factível, mesmo na sexta feira. Porém, para Levy, ele não é o fim do problema., afinal, sua conclusão também é sujeita à um *reductio*:

- (1) Um exame na sexta será uma surpresa (assumindo que argumento de Wright e Sudbury está correto e sua conclusão em 11).
- (2) O professor irá cumprir (A) se puder.
- (3)  $\therefore$  Astuto está justificado, na quinta, em crer que haverá uma prova na sexta. [(2), (3)]
- (4)  $\therefore$  Um exame na sexta não será uma surpresa. [(por definição de surpresa e 3)]
- (5)  $\perp$  [(1), (4)]
- (6)  $\therefore$  (1) é falso: Um exame na sexta *não* será uma surpresa [(1), (5)]

Como a conclusão de cada argumento serve como *input* para o outro não podemos tomar nenhum deles como decisivo. Ficamos oscilando entre duas conclusões contraditórias. Neste aspecto o problema é *análogo* ao paradoxo do mentiroso. No entanto, não podemos, meramente pela semelhança em relação ao círculo infinito concluir que a) que um se reduz ao outro b) que ambos são suscetíveis ao mesmo tipo de solução. Um bom argumento acerca da diferença da natureza dos problemas é o fato de que o do mentiroso emerge de um problema autorreferencial enquanto o do exame surpresa não. Outra diferença é que, independentemente da solução preferida para o paradoxo do mentiroso<sup>12</sup>, o fato é que a intuição demonstra que a solução apropriada para o paradoxo do exame surpresa só pode ser uma que torne a proposição informativa – possivelmente verdadeira.

Por fim, Levy se volta à Sorensen, que defende que para elucidarmos o problema precisamos do conceito de um *blindspot* epistêmico. Isto significa que, este *loop* se resume à isto:

(EB) Astuto está justificado em acreditar em (A) *se e somente se* (A) é falso.

Considerando que astuto é maximamente racional ele está ciente de que se encontra em tal *blindspot*. Mas o mero fato desta ciência também não dissolve o problema. Nem resolveria tomar astuto como não maximamente racional, pois ele seria suscetível ao

---

<sup>12</sup> A solução preferida de Levy é a de Eugene Mills, a saber, em um resumo grotesco, de que a proposição é *simplesmente* falsa.



*blindspot* do mesmo modo. A verdade de (A) é *incompatível* com ele crer justificadamente que (A). E o mesmo é verdade para o contrário: se ele *não* crê justificadamente que (A), então será o caso que (A) – haverá um exame e será surpresa.

Em resumo, astuto *não tem* como sair de tal círculo. Ele não tem boas razões para crer, e, portanto, não pode crer justificadamente que (A). Como ele não pode crer em (A), ele não pode crer que haverá um exame na sexta-feira – (A)<sub>E</sub>. Assim, o argumento da não-sexta *falha*.

A conclusão é que o argumento da não-sexta-feira depende que reformulemos o problema de (A) para (A\*). Além disso, não basta estar seguro – psicologicamente – é necessário estar justificado que o exame ocorra na sexta-feira. O argumento de Wright e Sudbury e a renovação da não-sexta-feira, em conjunto, nos mostram que na situação (A\*) o aluno astuto se encontra em um *blindspot* epistêmico. Assim, astuto, sendo maximamente racional, deve se manter *neutro* entre acreditar em (A\*) ou não. E como astuto está neutro em relação à (A\*), logo, em relação à (A\*)<sub>E</sub> – ele não pode ter crenças em relação a ocorrência ou não do exame na sexta. Assim, o argumento da não-sexta *falha*, e falhando o resto do argumento da impossibilidade falha: como nossa intuição previa, um exame surpresa *pode* ser dado em qualquer dia da semana, incluindo a sexta-feira.

## 5. CONCLUSÃO

### 5.1. Conclusão

*A logical theory may be tested by its capacity for dealing with puzzles, and it is a wholesome plan, in thinking about logic, to stock the mind with as many puzzles as possible, since these serve much the same purpose as is served by experiments in physical science. – Bertrand Russel (Russell, 1905:484)*

ME PARECE CLARO QUE AO OLHAR PARA OS CINQUENTA E POUCOS HISTÓRIA DO desenvolvimento do problema do exame surpresa vemos uma clara evolução, senão à direção de uma resposta precisa, ao menos a um mapeamento claro das intuições envolvidas, na *explicação* de tais intuições – que creio, fazem parte de uma solução adequada: não basta apenas resolver o problema, também há que se explicar porque entramos nos caminhos errados e tortuosos do labirinto – e dos caminhos mais prováveis de uma resposta adequada. De acordo com Wright e Sudbury, as seis intuições (Sudbury e Wright, 1977:42) que uma solução compreensiva deve atender são estas:

- (A) A explicação do conteúdo do anúncio deve deixar claro que ele é *satisfazível*, dado que um exame surpresa é, *palpavelmente*, uma possibilidade lógica.
- (B) A explicação deve deixar claro que o professor pode cumprir o anúncio mesmo *depois* que ele o anunciou, dado que, *palpavelmente*, ele pode.
- (C) A explicação deve fazer justiça ao significado intuitivo do anúncio.
- (D) A explicação deve fazer justiça a plausibilidade intuitiva do arrazoado do aluno astuto.
- (E) A explicação deve tornar possível que os alunos sejam *informados* pelo anúncio.

(F) A explicação deve explicar o papel, no engendramento do problema, do anúncio ser feito *aos* alunos. Não há dificuldade se o professor realiza o anúncio para outra pessoa, ou guarda para si mesmo.

Tal lista de intuições parece ter sido compilada junto ao desenvolvimento do problema. Dos primeiros autores surge a intuição de que, em algum sentido, o problema depende do ponto de vista do próprio aluno astuto: a proposição “haverá uma prova amanhã e o aluno astuto não está justificado a crer nisso” não é problemática na “cabeça” do professor, nem de outro aluno (intuição F). Ela só é um problema quando o próprio aluno astuto *toma conhecimento* tal proposição. Neste caso, como vimos em Levy e outros, a proposição se tornaria “eu creio que haverá uma prova amanhã e eu creio que não estou justificado a crer nisso”, que é, palpavelmente, um problema, e, diferente de “eu creio que darei uma prova amanhã e eu creio que o aluno astuto não sabe disso”, que não é um problema.

De Quine vemos a primeira tentativa de resolver o problema apelando a *não* nos basearmos na premissa de que o aluno deve tomar o anúncio do professor como infalível. Esta tradição que se estendeu por diversos autores, que vimos na sua manifestação mais recente, mas já bem diferente no artigo de Weintraub, tem em comum é a de tentar ver o anúncio do professor de modo não idealizado, mas como algo que ocorre “no mundo real”, que o importante são os objetivos que *não necessariamente estão contidos no significado do anúncio*, e assim, devem ser explicitados e que isto é tudo o que precisamos, talvez não para resolver o problema, mas para entendê-lo.

Weiss, por sua vez, tem como mérito a ideia contrária: o problema *deve* ser atacado em sua forma mais forte. E esta forma *depende* de tomarmos – do aluno astuto tomar – o professor, ou as regras da escola, como *infalíveis*. Isto, não porque a solução pragmática esteja errada: ela pode explicar como *de fato* as pessoas *agem* – ou melhor, como *deveriam* agir – quando confrontadas com o problema, porém isto não parece ser de muita relevância para o “verdadeiro problema filosófico” em questão. O que Weiss, e os autores que compraram este pressuposto – de tomar as regras como infalíveis – querem é resolver o problema que surge quando refletimos detidamente no *esquema lógico relevante* do problema. Neste sentido,

Quine é um divisor de águas. Aqueles que concordaram com ele, defenderam as posições pragmáticas, os outros as posições lógicas e epistemológicas.

As variações recalcitrantes de Sorensen servem, neste quadro, o propósito de mostrar que o princípio da retenção temporal, o princípio KK (reflexividade), a ordem de eliminação e o número de alternativas a serem eliminadas são premissas inessenciais na solução do paradoxo da predição.

Porém, nenhuma das soluções que vimos parece ter dado conta dos seis critérios de Wright e Sudbury como a de Levy. E este é o seu grande mérito, apesar de não ter reinventado a roda. Ele busca conciliar todas as intuições, justificando as corretas, refutando e, mais do que isso, mostrando porque damos certo crédito às incorretas, em uma explicação unificada do problema. Minha interpretação é que o argumento de Levy falha apenas no critério (E) e apenas no caso da prova não ser dada até quinta. Ainda assim pretendo afirmar que tal “falha” é, no entanto, *fundamental* para que o problema tenha uma solução adequada e que as *outras* intuições sejam atendidas. O problema está mais em tomar (E) como uma intuição fundamental (no sentido de que precise ser cumprida) do que com o argumento de Levy.

- (A) a conclusão é que não há realmente um paradoxo envolvido;
- (B) a solução é eficaz mesmo no caso de um dia só;
- (C) o significado do anúncio é preservado em todo o argumento;
- (D) o argumento se esforça por tomar o aluno como maximamente racional (apesar de não infalível);
- (E) os alunos são, em princípio, informados, porém sua linha de raciocínio leva a um *blindspot*, do qual eles estão cientes. Neste caso, no do professor não ter dado a prova até quinta, não há informatividade: não se extrai *nada* do anúncio; Porém é por esta falta que o anúncio é cumprido quando a prova é aplicada na sexta.
- (F) o *blindspot* só ocorre ao aluno astuto;

Recusar o critério (E) não apenas resolve o *puzzle*, como queria Russell, mas mais do que isso, é da insistência em mantê-lo que brota a perplexidade do problema: temos então de

aceitar a existência de proposições verdadeiras, contingentes, não contraditórias e *perfeitamente* compreensíveis que se tornam vazias para *certos* sujeitos e por um *certo* período e que depois voltam a sua plena função.

### REFERÊNCIAS BIBLIOGRÁFICAS:

1. ALEXANDER, Peter. Pragmatic Paradoxes. *Mind*, vol. LIX, nº. 236, 1950, p. 536-8.
2. AUSTIN, J. L., *Truth. Proceedings of the Aristotelian Society*, vol., XXIV, 1950, p. 111-29.
3. AUSTIN, A. K. *On The Unexpected Examination. Mind*, vol. LXXVIII, nº. 309, 1969, p. 137.
4. AYER, A. J. *On a Supposed Antinomy. Mind*, vol. LXXXII, nº. 325, 1973, p. 125-6.
5. BAR-HILLEL, Maya e MARGALIT, Avishai. *Expecting the unexpected. Philosophia*, vol. 13, nº. 3, 1983, p. 263-88.
6. BURGE, Tyler. *Epistemic Paradox. The Journal of Philosophy*, Vol. LXXXI, nº. 1, 1984, p. 5-29.
7. CHIHARA Charles S.- *Olin, Quine, and the Surprise Examination. Philosophical Studies*, vol. 47, nº. 2, 1985, p. 191-9.
8. CHOW, Timothy Y. *The surprise examination or unexpected hanging paradox. American Mathematical Monthly*, nº. 105, 1998, p. 41-51.
9. COHEN, L. Jonathan. *Mr. O'Connors Pragmatic Paradox. Mind*, vol. LIX, nº. 233, 1950, p. 85-7.
10. DALLAGNOL, Darlei. *Crer e Saber. Dissertatio*, nº 26, verão de 2007, p. 9-26.
11. FRANCESCHI, Paul. *A Dichotomic Analysis of the Surprise Examination Paradox. Cogprints*, 2002.
12. GOOD, I. J. e MELTZER, B. *Two Forms of the Prediction Paradox. British Journal for The Philosophy of Science*, vol. XVI, nº. 61, 1965, p. 50-1.
13. HALPERN, Joseph Y. e MOSES, Yoram. *Taken by surprise: The paradox of the surprise test revisited. Journal of Philosophical Logic*, vol. 15, nº. 3, 1986, p. 281-304.
14. JANAWAY Christopher- *Knowing about Surprises: A Supposed Antinomy Revisited. Mind*, nº. 391, 1989.

15. JONGELIN, Tjeerd B. e KOETSIER, Teun. *Blindspots, self-reference and the prediction paradox*. *Philosophia*, vol. 29, nº. 1, 2002, p. 377-91.
16. KAPLAN, David e MONTAGUE, Richard. *A Paradox Regained*. *Notre Dame Journal of Formal Logic*, 1960, p. 79-90.
17. KIRKHAM, Richard L. *The two paradoxes of the unexpected examination*. *Philosophical Studies*, vol. 49, nº. 1, 1986, p. 19-26.
18. ——— *On paradoxes and a surprise exam*. *Philosophia*, vol. 21, nº. 1, 1991, p. 31-51.
19. KRIPKE, Saul. *Philosophical Troubles*. Oxford University Press, 2010, p. 27-51.
20. LEVY, Ken. *The Solution to the Surprise Exam Paradox*. *Southern Journal of Philosophy*, vol. 47, nº. 2, 2009, p. 131-58
21. LYON, Ardon. *The Prediction Paradox*. *Mind*, vol. LXVIII, nº. 272, 1959, p. 510-7.
22. MELTZER, B. *The third possibility*. *Mind*, vol. LXXIII, nº. 291, 1964, p. 430-3.
23. O'CONNOR, D. J. *Pragmatic paradoxes*. *Mind*, vol. LVII, 1948, p. 358-9.
24. ——— *Pragmatic paradoxes and Fugitive Propositions*. *Mind*, vol. LX, nº. 240, 1951, p. 536-8.
25. OLIN, Doris. *The prediction paradox resolved*. *Philosophical Studies*, vol. 44, nº. 2, 1983, p. 225-33.
26. ——— *The prediction paradox: Resolving recalcitrant variations*. *Australasian Journal of Philosophy*, vol. 64, nº. 2, 1986, p. 181-9.
27. QUINE, W. V. O. *On a So-Called Paradox*. *Mind*, vol. LXII, nº. 245, 1953, p. 65-7.
28. RUSSELL, Bertrand. *On Denoting*. *Mind*, vol. XIV, nº. 56, 1905, p. 479-93.
29. SCRIVEN, Michael. *Paradoxical Announcements*. *Mind*, vol. LX, 1951, p. 403-7.
30. SHARPE, R. A. *The Unexpected Examination*. *Mind*, vol. LXXIV, nº. 294, 1965, p. 255.
31. SHAW, R. *The paradox of the unexpected examination*. *Mind*, vol. LXVII, nº. 267, 1958, p. 382-4.
32. SORENSEN, Roy. *Conditional blindspots and the knowledge squeeze: A solution to the prediction paradox*. *Australasian Journal of Philosophy*, vol. 62, nº. 2, 1984, p. 126-35.
33. ——— *Recalcitrant variations of the prediction paradox*. *Australasian Journal of Philosophy*, vol. 60, nº. 4, 1982, p. 355-62.
34. SUDBURY, Aidan e WRIGHT, Crispin. *The paradox of the unexpected examination*. *Australasian Journal of Philosophy*, vol. 55, nº. 1, 1977, p. 41-58.

- 35. WEINSTRAUB, Ruth. *Practical Solutions to the Surprise Examination Paradox*. *Ratio* (8), 1995, p. 161-9.
- 36. WEISS, Paul. *The Prediction Paradox*. *Mind*, vol. LXI, n°. 242, 1952, p. 265–9.
- 37. WILLIAMS, John N. *The Surprise Exam Paradox: Disentangling Two Reductions*. *Journal of Philosophical Research*, vol. 32, 2007, p. 67-94.
- 38. WILLIAMS, Timothy. *Knowledge and Its Limits*. Oxford University Press, 2002, p. 135-46.