

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 OS 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, Dezembro de 2012.

## ***SUMÁRIO***

INTRODUÇÃO .....	4
1. AS SOLUÇÕES .....	8
1.1. Cohen e a distinção proposição/elocução.....	8
1.2 Alexander, a autorreferência e os condicionais problemáticos .....	11
1.3. Scriven, declarações e ordenamentos.....	13
1.4. Weiss, o paradoxo da predição e os futuros contingentes .....	15
1.5. Quine e a desconfiança no professor.....	18
1.6. Shaw e a autorreferência .....	21
1.7. Lyon e o livre arbítrio.....	24
1.8. Kaplan e Montague e a formalização do problema.....	25
1.9. Sharpe.....	35
1.10. Ayer.....	36
1.11. Wright e Sudbury .....	37
1.12. Olin.....	38
1.13. Sorensen e os blindspots condicionais .....	39
1.14. Weinstraub e a solução prática .....	40
1.15. O triatlon de Levy .....	44
2. STATUS QUO DO PARADOXO .....	55

2.1. O status .....	55
2.2. Conclusão .....	55
REFERÊNCIAS BIBLIOGRÁFICAS: .....	57

## INTRODUÇÃO

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 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.

---

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

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*. 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.

O título - *Pragmatic Paradoxes* - deixa clara a intenção do autor 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 entre a apresentação e o *ato* de apresentar.

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 - uma relação do problema com o paradoxo de Moore, uma relação 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”<sup>3</sup>. 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.

---

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

3 Em DALLAGNOL, Darlei. Crer e Saber. *Dissertatio* [26], 9 – 26 verão de 2007

No entanto, outros autores buscaram resolver o problema de forma bem diversa. Avishai Margalit e Maya Bar-Hillel em sua resenha sobre o 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). Alguns autores - talvez os de opinião mais estranha - simplesmente negaram que o exame pudesse ser realizado em um dia inesperado. Outros tantos focaram no problema específico da realização da prova no último dia possível e buscaram refutar essa premissa tão intuitiva do argumento. Outros concordaram que uma prova no último dia seria impossível, mas buscaram refutar a aplicação do mesmo critério iteradamente. Outros apelaram para a impossibilidade de usar as palavras do professor como premissa, ora por negar a possibilidade de conhecermos algo pela mera afirmação de outrem, ora pela impossibilidade de conhecer o futuro, por este ser contingente. Outros, ainda, buscaram olhar o problema de outro ângulo, como um problema prático, não teórico.

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 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

### 1.1. Cohen e a distinção proposição/elocução

Para Laurence Jonathan Cohen a solução do problema apresentado por O'Connor 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 lembraríamos 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": 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 17h 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, 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.

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.

## 1.2 Alexander, a autorreferência e os condicionais problemáticos

No mesmo ano da publicação de Cohen, 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 autorreferentes, deslocando-o do campo epistemológico para o puramente lógico.

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). Pois, neste caso, para ele, a afirmação seria claramente verdadeira. Alexander apresenta duas considerações sobre este ponto: 1) Não é razoável atribuir a referência do sujeito da proposição ao cadáver. 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?”. 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 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”.

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, portanto, não desenvolverei tais ideias.

### 1.3. Scriven, 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 uma ocorrência. 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 tomá-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.

#### 1.4. Weiss, o paradoxo da predição e os 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<sup>4</sup>, a saber, “paradoxo da predição”.

O uso do título “paradoxo da predição” é premeditado. Como veremos, este título já dá um sinal de como Weiss 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 descumpra 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

---

<sup>4</sup> excluo aqui livros introdutórios de filosofia, lógica ou epistemologia, que mantêm o uso majoritário de “paradoxo do exame surpresa” ou “paradoxo do enforcado”.

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 dissossiareem 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 distinguidas 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 meu)

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.

### 1.5. Quine e a desconfiança no professor

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  $t + i - 1/2$ .”

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.

## 1.6. Shaw e a autorreferência

Cinco anos após o artigo de Quine, - o que representa um grande silêncio, afinal todos os artigos que vimos acima estão no período de 48 e 53, silêncio cujos motivos podemos apenas conjecturar: seria a autoridade de Quine que afugentou réplicas ou a força mesma de seu argumento? - 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. Assim, a regra inquebrável da escola parece ser auto-contraditória, 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 a sério. A saída de Quine – tomar saber como uma noção “vaga do senso comum” (Shaw:1958, 382, ¶4) – significa se evadir do problema e não resolvê-lo. Temos que entender “inesperado” deste modo: não deduzível 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 barrar 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 auto-contraditórias 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 arrazoados dos alunos, porém restam  $n - r$  dias onde o exame será possível. Se  $r = n$ , então as regras são simplesmente auto-contraditórias, 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 auto-referente 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 tendou 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 auto-referencial” da regra. Pois auto-referência não é condição – nem necessária, nem suficiente – para engendrar um paradoxo. Há autores<sup>5</sup> que argumentam que auto-referência não é necessária nem mesmo para definir o paradoxo do mentiroso. E proposições como “esta proposição é composta por palavras” é um exemplo evidente de uma auto-referência não paradoxal.

---

<sup>5</sup> Como Tennant 1995 e Yablo 1993. Priest 1997 busca refutar tal posição. Sorensen 1998 responde a Priest, que, por sua vez é respondido por Beall (2001). Tal resumo da ópera está em Levy 2009, p. 154, nota 25.

### 1.7. Lyon e o livre arbítrio

Provavelmente o autor que tirou conclusões mais surpreendes em relação ao problema em questão foi Ardon Lyon.

Em desenvolvimento.



## 1.8. Kaplan e Montague e a formalização do problema

Os primeiros autores a formalizarem o problema foram David Kaplan e Richard Montague. Sua solução 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;

$S$  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 “ $K$

---

<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.

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:

$$(S \wedge \neg T \wedge \neg Q \wedge \neg K_d(\underline{S})) \wedge (\neg S \wedge T \wedge \neg Q \wedge \neg K_s(\underline{T})) \wedge (\neg S \wedge \neg T \wedge Q \wedge \neg K_t(\underline{Q}))$$

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  $(Q \wedge \neg K_t(\underline{Q}))$ . Mas, não ocorrendo nem na segunda, nem na terça, o exame ocorrerá na quarta (lembramos, estamos pressupondo o cumprimento de  $D_1$ ), ou seja  $I(\underline{\neg S \wedge \neg T}, Q)$ . 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:

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

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

$A_1$  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):

$$A_3: K_d( \underline{A_1 \wedge A_2} )$$

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::

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

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* –  $K_d( \underline{\neg Q} )$ .

E se  $K_d( \underline{\neg Q} )$ , pelo princípio da retenção do conhecimento temporal (memória)  $K_s( \underline{\neg Q} )$ ,  $K_t( \underline{\neg Q} )$  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 –  $K_s( \underline{T} )$

– , o que contradiz  $\neg K_s(\underline{T})$  (do disjuncto). Assim, concluímos que  $\neg T$ . Novamente, a base para este argumento se assemelha à forma usada acima:

$$A_5: \neg S \rightarrow K_s(\neg \underline{S})$$

$$A_6: (I(\neg \underline{S} \wedge \neg \underline{Q}, \underline{T}) \wedge K_s(\neg \underline{S}) \wedge K_s(\neg \underline{Q})) \rightarrow K_s(\underline{T})$$

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  $\neg K_d(\underline{S})$ . 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,  $\neg K_s(\underline{T})$ . 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:

$$A_2: (I(\neg \underline{S} \wedge \neg \underline{T} \wedge \underline{D}_1, \underline{Q}) \wedge K_t(\neg \underline{S} \wedge \neg \underline{T}) \wedge K_t(\underline{D}_1)) \rightarrow K_t(\underline{Q})$$

Além da seguinte assunção:

$$K_t(\underline{D}_1)$$

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 auto-referencial 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 auto-referência. 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 sobrelinha 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 conseqüentemente a seguinte formulação, também auto-referente,  $D_3$ :

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

K é capaz agora de mostrar que  $D_3$  não pode ser cumprido usando um argumento análogo à A1 - A4:

$$\begin{aligned} B_1: & \neg S \rightarrow K_s(\underline{\neg S}) \\ B_2: & (I(\underline{\neg S}, \underline{D_3 \rightarrow T}) \wedge K_s(\underline{\neg S})) \rightarrow K_s(\underline{D_3 \rightarrow T}) \\ B_3: & K_d(\underline{B_1 \wedge B_2}) \\ B_4: & (I(\underline{B_1 \wedge B_2}, \underline{D_3 \rightarrow S}) \wedge K_d(\underline{B_1 \wedge B_2})) \rightarrow K_d(\underline{D_3 \rightarrow S}) \end{aligned}$$

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):  $\vdash (D_3 \rightarrow (K_d(\underline{D_3 \rightarrow S}))) \rightarrow \neg S$

Por (8) e (9):

$$B_1 \wedge B_2 \vdash D_3 \rightarrow S$$

Assim, pelo princípio usado para obter (6):

$$\vdash I(\underline{B_1 \wedge B_2}, \underline{D_3 \rightarrow S})$$

$$B_4 \vdash K_d(\underline{B_1 \wedge B_2}) \rightarrow K_d(\underline{D_3 \rightarrow S})$$

$$B_3 \wedge B_4 \vdash D_3 \rightarrow \neg S$$

Apenas para facilitar, vou usar, daqui em diante, a convenção  $P_{1..n}$  para denotar a conjunção  $P_1 \wedge \dots \wedge P_n$ . Então, por (8):

$$(11): B_{1..4} \vdash D_3 \rightarrow (\neg S \wedge \neg T)$$

$$\vdash D_3 \rightarrow (S \vee T)$$
 [(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  $\neg K_s(\underline{D_3 \rightarrow T})$ . Para aplicar a mesma linha 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:

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

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  $B_1$  -  $B_6$ :

$$C_1: K_d(\neg D_4) \rightarrow \neg D_4$$

$$C_2: \neg S \rightarrow K_s(\neg S)$$

$$C_3: K_d(C_1)$$

$$C_4: (I(C_1 \wedge \neg S, D_4 \rightarrow T) \wedge K_s(C_1) \wedge K_s(\neg Q)) \rightarrow K_s(D_4 \rightarrow T)$$

$$C_5: K_d(C_{1..4})$$

$$C_6: I(C_{1..4}, D_4 \rightarrow M) \wedge K_d(C_{1..4}) \rightarrow K_d(D_4 \rightarrow Q)$$

$$C_7: K_d(C_{1..6})$$

$$C_8: I(C_{1..4}, \neg D_4) \wedge K_s(C_{1..4}) \rightarrow K_d(\neg D_4)$$

$$(2): C_1 \vdash D_4 \rightarrow K_s(\neg D_4) \quad [(1)]$$

$$(3): C_1 \wedge S \vdash D_4 \rightarrow T \quad [(1), (2)]$$

$$(4): C_1 \vdash (D_4 \wedge T) \rightarrow \neg K_s(D_4 \rightarrow T)$$

$$(5): C_1 \vdash (D_4 \wedge T) \rightarrow \neg Q$$

$$(6): C_1 \wedge C_2 \vdash (D_4 \wedge T) \rightarrow K_s(\neg Q) \quad [(5)]$$

$$(7): \vdash I(\underline{C_1 \wedge S}, \underline{D_4 \rightarrow T}) \quad [(3)]$$

Então:

$$C4 \vdash (K_s(\underline{C_1}) \wedge K_s(\underline{\neg M})) \rightarrow K_s(\underline{D_4 \rightarrow T})$$

Por (6):

$$C_{1..4} \vdash (D_4 \wedge T) \rightarrow Ks(\underline{D_4 \rightarrow T})$$

Por (4):

$$C_{1..4} \vdash (D_4 \wedge T) \rightarrow Ks(\underline{D_4 \rightarrow T}) \wedge \neg Ks(\underline{D_4 \rightarrow T})$$

Logo:

$$(8): C_{1..4} \vdash D_4 \rightarrow \neg T$$

$$(9): C_1 \vdash (D_4 \wedge \neg T) \rightarrow S \quad [(1), (2)]$$

$$(10): C_1 \vdash (D_4 \wedge K_d(\underline{D_4 \rightarrow S})) \rightarrow \neg S$$

Por (8) e (9):

$$C_{1..4} \vdash D_4 \rightarrow S$$

E, pelo princípio invocado em conexão com (7):

$$\vdash I(\underline{C_{1..4}}, \underline{D_4 \rightarrow S})$$

Logo:

$$C_6 \vdash K_d(\underline{C_{1..4}}) \rightarrow K_d(\underline{D_4 \rightarrow S})$$

$$C_5 \wedge C_6 \vdash K_d(\underline{D_4 \rightarrow S})$$

$$C_1 \wedge C_5 \wedge C_6 \vdash D_4 \rightarrow \neg S \quad [2]$$

$$(11): C_{1..6} \vdash D_4 \rightarrow (\neg K_d(\underline{\neg D_4}) \wedge \neg S \wedge \neg T) \quad [(2), (8)]$$

$$\vdash D_4 \rightarrow (K_d(\underline{\neg D_4}) \vee S \vee T) \quad [(1)]$$

$$(12): C_{1..6} \vdash \neg D_4 \quad [(11)]$$

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:

$$C_8 \vdash K_d(\underline{C_{1..6}}) \rightarrow K_d(\underline{\neg D_4})$$

$$(13): C_7 \wedge C_8 \vdash K_d(\underline{\neg D_4})$$

$$\vdash K_d(\underline{\neg D_4}) \rightarrow D_4 \quad [(1)]$$



$$(I4): C_7 \wedge C_8 \vdash D_4$$

$$[(I3)]$$

Com essas suposições, 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>7</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 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”.

$$(I) \vdash D_6 \equiv K_d(\neg D_6)$$

O paradoxo, neste caso, é mostrado em três princípios análogos a  $C_1$ ,  $C_3$  e  $C_4$ :

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

$$E_2: K_s(E_1)$$

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

$$\vdash \neg D_6 \rightarrow K_d(\neg D_6)$$

$$[(I)]$$

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

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

$$\vdash I(E_1, \neg D_6)$$

$$[(2)]$$

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

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

$$(3) E_2 \wedge E_3 \vdash D_6$$

$$[(I)]$$

<sup>7</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).

Os autores mostram, assim, em (2) e (3), que as assunções  $E_1 - E_3$  são incompatíveis com os princípios da sintaxe elementar. Mas  $E_1 - E_3$  são ainda mais plausíveis que  $C_1 - C_8$  e ainda não contém o princípio de conhecimento por memória. Pela semelhança ao paradoxo do mentiroso [*Liar*], os autores decidem *batizar* este paradoxo –  $D_6$  – de paradoxo do conhecedor [*Knower*].

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’” seria sem muito sentido. 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 auto-referência, por exemplo.

As assunções usadas no paradoxo do conhecedor –  $E_1 - E_3$  – são instâncias dos seguintes esquemas:

$$S_1: K_d(\phi) \rightarrow \phi$$

$$S_2: K_d(K_d(\phi) \rightarrow \phi)$$

$$S_3: (I(\phi, \psi) \wedge K_d(\phi)) \rightarrow K_d(\psi)$$

Usando o “conhecedor”, os autores podem mostrar que qualquer sistema formal contendo os aparatos da sintaxe elementar e incluindo  $S_1 - S_3$  é inconsistente. Tal resultado seria análogo ao uso do mentiroso por Alfred Tarski. A relação precisa entre o resultado de Tarski e o resultado de Kaplan e Montague não é explicitada no texto, porém, sugerem eles, parece constituir um interessante campo de estudo.

## 1.9. Sharpe

Em desenvolvimento.

## 1.10. Ayer

Em desenvolvimento.

### 1.11. Wright e Sudbury

Em desenvolvimento.

## 1.12. Olin

Em desenvolvimento.

### 1.13. Sorensen e os *blindspots* condicionais

Em desenvolvimento.

### 1.14. Weinstraub e a solução prática

Ruth Weinstraub 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 Weinstraub, 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.

Weinstraub 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, Weinstraub 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: uma solução deve servir para *todas* as variações. A não ser que vejamos o problema de modo *prático*. Sorensen está errado nesta perspectiva.



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: 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: Tomar a promessa *não literalmente*. Ou seja, interpretar “surpresa” apenas como dizendo que o exame será dado em “alguma data da próxima semana”. Esta interpretação remove tanto a carga epistêmica como a carga psicológica da surpresa. Neste caso o professor a probabilidade de 1 para 5 de fazer uma prova não esperada caso escolha o dia aleatoriamente. Esta interpretação, apesar de enfraquecer o conceito de surpresa, captura a *real* intenção do professor. Não importa se o exame não for mais uma 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.

Ok, 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, “oniciência 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 de uma situação da vida real.

A solução prática que apela para a distinção entre paradoxo *atual* e *percebido* permite que expliquemos um aspecto importante do fenômeno da vida real: 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.

Mesmo concedendo como alunos ordinários agirão em relação ao anúncio, ainda precisamos entender como o anúncio afeta o aluno *sofisticado*.

Dois casos: Se o professor é ingênuo, os alunos devem responder como um aluno ingênuo. O professor 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; Se o paradoxo é conhecimento comum [common knowledge] – todos sabem que todos sabem do paradoxo – como entender o anúncio? Pode ser meramente uma piada 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 auto-contraditório, 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?

### 1.15. O triatlon de Levy

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 apresentados, o argumento do “Anúncio Improjecionável”, o argumento de “Wright e Sudbury” e o argumento dos “*Blindspots* Epistêmicos”.

O projeto de Levy então é este: apresentar como é que mesmo um anúncio feito para apenas um dia ainda será uma surpresa. Apenas esta exposição irá realmente *resolver* o problema.

Duas intuições: o aluno não pode se surpreender com um exame que ele *esperava* na noite anterior; o aluno não pode se surpreender com um exame que ele estava *certo*, na noite anterior, que ocorreria no dia seguinte; Dadas estas duas intuições, mais o fato de que um *ex hypothesi* professor confiável *prometeu* um exame até o final do período letivo, o fato de que o exame ainda não foi dado e o fato de que estamos há um dia do final do período letivo parece indisputável que o estudante não irá se surpreender com um exame no último dia.

O primeiro passo para a refutação é compreender que *surpresa*, no contexto do SEP significa isto: ausência de *conhecimento*, ou seja, ausência de uma crença verdadeira e justificada na noite anterior de que o exame ocorreria no dia posterior. E, importante, 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.

Anúncio Improjecionável. O argumento do último dia começa com uma assunção errônea: que o aluno pode, sem mais problemas, projetar o anúncio do professor para o

penúltimo dia. O que ocorre é que ele precisa, para fazer isso mudar a interpretação da regra apresentada pelo professor de “Haverá um exame surpresa até o final do período letivo” para “Haverá um exame surpresa *amanhã*”. O aluno não está, no momento do anúncio, “epistemologicamente autorizado” a supor que o anúncio é aplicável do mesmo modo no penúltimo dia como é no momento em que é feito.

Wright e Sudbury. O aluno começa o termo com o conhecimento de que um exame surpresa ocorrerá no período, porém ele *perde* esse conhecimento no penúltimo dia (se o exame não tiver sido dado). Em todo caso, o exame no último dia será uma surpresa. No entanto, este argumento não pode ser a última palavra: sua conclusão é vulnerável a um *reductio* (que Levy chama de “argumento da não-sexta renovado”). No entanto, este *reductio* também não pode ser o fim da história também, pois a conclusão dele também é vulnerável ao *reductio oposto*, ou seja, o próprio argumento de Wright e Sudbury. O resultado desse ciclo é um paradoxo lógico. *Este ciclo é o “coração” do problema do SEP.*

*Blindspot* epistêmico. O resultado circular do argumento de Wright e Sudbury é sucintamente capturado pelo conceito de Roy Sorensen de “*blindspot* epistêmico”. O “curto circuito” que resulta do círculo vicioso faz com que o aluno *não* tenha boas razões para se decidir em relação à ocorrência ou não do exame, segue-se, ele não está *justificado* em crer na sua ocorrência. E não tendo justificativa, não há conhecimento. Por fim, o exame no último dia *será* surpreendente.

Levy está interessado na versão mais forte do SEP, e estas são as assunções que ele toma como definidoras dessa versão:

- O aluno astuto deve ser considerado *máximamente* racional. Por *máximamente* racional Levy entende isto: um mestre da lógica, ele evita contradições, conhece todas as verdades lógicas e crê em todas as consequências lógicas daquilo que ele crê;
- O professor realiza o anúncio (A) numa terça-feira de que haverá um exame surpresa na quarta, quinta ou sexta. Para dividir as duas promessas (A)<sub>E</sub> é 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*”.
  - Não sexta-feira: 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. Dado (A), sexta não pode ser um dia para o exame;
  - Backtracking: 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>. Conclusão: O exame é impossível – (A)<sub>E</sub> não pode ocorrer
- A intuição:
  - 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 irá ser uma surpresa para o aluno. Então algo *deve* estar errado. Ou o argumento da impossibilidade ou a intuição estão errados. O desafio é determinar qual e porquê.

Máximamente racional: A assunção de que o aluno astuto deve ser máximamente racional é justificada pelo fato de que devemos resolver o problema em sua versão *mais forte*. Soluções para versões em que o aluno não é maximamente racional não resolvem esta versão. Devemos desconsiderar soluções em que o aluno possa esquecer o anúncio, não entender plenamente (A) ou que faça inferências equivocadas.

Distinguir “racionalidade máxima” de “infalibilidade”: Afinal o argumento aqui mostra que o argumento “da impossibilidade” *falha*. No entanto, não poderá ser dito que o argumento falha por ser *inválido*. Se fosse, *teríamos* que concluir que o aluno *não é* máximamente 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 máximamente racional. Ser máximamente racional só protege o aluno astuto de erros tolos e inferências

errôneas, porém não o impede de iniciar com premissas falsas porém inicialmente plausíveis.

Conhecimento e surpresa: (A) em termos de *conhecer*; O exame será surpresa se e somente se ele ocorre em um dia que segue uma noite em que o aluno não pode *saber* que o exame ocorrerá no dia seguinte. Conhecimento em termos de crença verdadeira e justificada (não apenas crença justificada). Levy concorda com Hall (1999) que não faz diferença formular o problema em termos de crença justificada ou crença verdadeira e justificada. Do mesmo modo, não há necessidade de considerar a questão de Gettier (Levy baseia-se em Williams 2007). (A) equivale a:

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

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.

Um argumento falho porém instrutivo contra o argumento do *backtracking*: 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*. Sexta não é um dia *impossível* de ocorrer o exame do mesmo modo que sábado o é. 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. Resposta: Ok, pode ser *fisicamente* possível dar o exame na sexta, ou em qualquer outro dia. Porém, não é *lógicamente* (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.

Porque pressupor que o professor vai cumprir (A) se puder (ou “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. 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. 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.

O argumento (1)-(7) é estranho: 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: Se o argumento da impossibilidade está correto então um exame surpresa não pode ser dado em qualquer dos dias. E isto é o mesmo que dizer:

- (a) na terça de noite astuto está justificado em crêr que haverá um exame na quarta;
- (b) se quarta passar sem exame, na quarta de noite astuto está justificado em crêr que haverá um exame na quinta;
- (c) se quinta passar sem exame, na quinta de noite astuto está justificado em crêr que haverá um exame na sexta (Austin 1969);

Intuição expandida: Imaginemos um período de 300 dias e uma prova ocorrendo no dia 155. Astuto deve estar justificado em crêr 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 crêr que ele está justificado neste dia mais do que [não] estava nos outros. Ele não estava justificado. (Cargile, 67; Gardner, 63; Halpern e Moses 1986; Williamson 92). Não há diferença fundamental entre o problema com 3 ou 300 dias.

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 *mais* do que esta mera expectativa. Esta expectativa tem de ser justificada. Ainda assim o argumento seduz: 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 e justificação uma propriedade normativa de crenças (Klein 98). Desse modo é muito importante notar isto: Pode ser o caso que astuto esteja *certo* de que o exame ocorrerá na sexta mas não esteja *justificado*. Se ele de fato ocorrer na sexta, temos que entender isso *como uma surpresa*.

Três argumentos contra a não sexta-feira. O argumento não projecionalidade do anúncio: 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*”. – (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: “Está chovendo e eu não acredito nisto”. Porém, envolve o conceito de justificação e não meramente de crença: “Haverá um exame amanhã e eu não estou justificado a crer nisso” (A\*), instanciando a forma mais geral “p e eu não creio justificadamente em p” (ou “p e eu creio justificadamente em  $\neg p$ ”).

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, não proposições “*Moore-absurd*”. Uma proposição Moore-absurd não é, de fato, auto-contraditória, 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 na cabeça do meu professor, que ele nunca tenha me dito tal coisa e que ele vá mesmo dar a prova no próximo dia. O que torna a proposição problemática é quanto *eu* a considero como uma de minhas crenças – e, no caso, justificadas.

Por distribuição lógica, 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*”.

TODO falar da 2a posição de Almeida.

Se aceitamos a apresentação das proposições Moore-absurd de Almeida não devemos considerar  $(A^*)$  como auto-contraditória, mas como fortemente insistente. 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 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 uma surpresa – como eu havia prometido”. Apesar de não cumprida, espera-se que  $(A)_s$  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.

Três argumentos contra a não sexta-feira. O argumento de Wright e Sudbury: Este argumento é um *reductio ad absurdum* do argumento da não-sexta:

(12) Assuma que o argumento da não-sexta está correto: Um exame na sexta não seria surpresa;

(13)  $\therefore$  Se o exame não foi dado até quinta, então o professor não tem como cumprir  $(A)_s$ . [(12)]

(14)  $\therefore$  Astuto não está justificado a crer em  $(A)_s$ . [(13)]

(15)  $(A)_s$  é uma parte essencial de  $(A)$ : Se alguém não está justificado a crer em  $(A)_s$  então não está a crer  $(A)$ .

(16)  $\therefore$  Astuto não está justificado a crer em  $(A)$ . [(14), (15)]

(17) 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>.

(18) ∴ Astuto não está justificado a crer em (A)<sub>E</sub>. [(16), (17)]

(19) ∴ Astuto não está justificado a crer, na quinta, que haverá um exame na sexta.  
[(18)]

(20) ∴ Um exame na sexta seria uma surpresa. [(pro definição, 19)]

(21) ⊥ [(12), (20)]

(22) ∴ (12) é falso: Um exame na sexta será uma surpresa [(12), (21)]

A não-sexta renovada: O argumento de Wright e Sudbury não é o fim do problema. Sua conclusão também é sujeita à um *reductio*:

(23) Suponha que um exame na sexta será uma surpresa (o argumento de Wright e Sudbury está correto).

(24) O professor irá cumprir (A) se puder.

(25) ∴ Astuto está justificado, na quinta, em crer que haverá uma prova na sexta.  
[(23), (24)]

(26) ∴ Um exame na sexta não será uma surpresa. [(por definição 25)]

(27) ⊥ [(23), (26)]

(28) ∴ (23) é falso: Um exame na sexta *não* será uma surpresa [(23), (27)]

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 à uma das versões do paradoxo do mentiroso, a afirmação “Esta afirmação é falsa”. No entanto, não podemos, meramente pela semelhança em relação ao círculo infinito 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 auto-referencial enquanto o do exame surpresa não. Outra diferença é que, independentemente da solução preferida para o paradoxo do mentiroso – a do autor é a de Eugene Mills, a saber, de que a proposição é *simplesmente* falsa – 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.

Três argumentos contra a não sexta-feira. O argumento de Sorensen: Sorensen defende que para elucidarmos o problema precisamos do conceito de um *blindspot* epistêmico. Isto significa que:

(30) 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 não dissolve o problema. Nem resolveria considerar astuto não maximamente racional, pois ele seria suscetível ao *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, 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*.

Conclusão: 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.

## 2. STATUS QUO DO PARADOXO

### 2.1. O status

Em desenvolvimento.

### 2.2. Conclusão

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 nos caminhos mais prováveis de uma resposta adequada.

Tal evolução se deu em cada autor – ou conjunto de opiniões – que buscou endereçar o problema, ora pelo mapeamento de uma outra alternativa de solução ora pela clarificação de algum ponto ou intuição relevante.

Dos primeiros autores 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. 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 torna “eu creio que haverá uma prova amanhã e eu não estou justificado a crer nisso”.

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. O

que esta tradição que se estendeu por diversos autores, que vimos na sua manifestação mais recente já bem diferente no artigo de Weistraub, 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 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 a filosofia, muito menos para a lógica. 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.

Desenvolver mais aqui...

Como vimos na tentativa de solução mais recente, a de Levy, praticamente todas as intuições e possibilidades apresentadas acima são levadas em consideração, mencionadas e mesmo as que o autor considera problemáticas, são endereçadas por seu complexo argumento. Sem dúvida, apesar de não possuir grandes novidades técnicas, esta é a principal virtude do artigo de Levy: buscar 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.



### REFERÊNCIAS BIBLIOGRÁFICAS:

1. ALEXANDER, Peter. Pragmatic Paradoxes. *Mind*, vol. LIX, no. 236, 1950, p. 536-8.
2. AUSTIN, A. K. On The Unexpected Examination. *Mind*, vol. LXXVIII, no. 309, 1969, p. 137.
3. AYER, A. J. *On a Supposed Antinomy*. *Mind*, vol. LXXXII, no. 325, 1973, p. 125-6.
4. BAR-HILLEL, Maya e MARGALIT, Avishai. *Expecting the unexpected*. *Philosophia*, vol. 13, no. 3, 1983, p. 263-88.
5. BURGE, Tyler. *Epistemic Paradox*. *The Journal of Philosophy*, Vol. LXXXI, no. 1, 1984, pp. 5-29.
6. CHIHARA Charles S.- *Olin, Quine, and the Surprise Examination*. *Philosophical Studies*, vol. 47, no. 2, 1985, p. 191-9.
7. CHOW, Timothy Y. *The surprise examination or unexpected hanging paradox*. *American Mathematical Monthly* 105, 1998, p. 41-51.
8. COHEN, L. Jonathan. *Mr. O'Connors Pragmatic Paradox*. *Mind*, vol. LIX, no. 233, 1950, p. 85-7.
9. FRANCESCHI, Paul. *A Dichotomic Analysis of the Surprise Examination Paradox*. *Cogprints*, 2002.
10. 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.
11. 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.

12. JANAWAY Christopher- *Knowing about Surprises: A Supposed Antinomy Revisited. Mind*, no. 391, 1989.
13. JONGELIN, Tjeerd B. e KOETSIER, Teun. *Blindspots, self-reference and the prediction paradox. Philosophia*, vol. 29, no. 1, 2002, p. 377-91.
14. KAPLAN, David e MONTAGUE, Richard. *A Paradox Regained. Notre Dame Journal of Formal Logic*, 1960, p. 79-90.
15. KIRKHAM, Richard L. *The two paradoxes of the unexpected examination. Philosophical Studies* , vol. 49, no. 1, 1986, p. 19-26.
16. ——— *On paradoxes and a surprise exam. Philosophia*, vol. 21, no. 1, 1991, p. 31-51.
17. KRIPKE, Saul. *Philosophical Troubles*. Oxford University Press, 2010, p. 27-51.
18. LEVY, Ken. *The Solution to the Surprise Exam Paradox. Southern Journal of Philosophy* , vol. 47, no. 2, 2009, p. 131-58
19. LYON, Ardon. *The Prediction Paradox. Mind*, vol. LXVIII, no. 272, 1959, p. 510-7.
20. O'CONNOR, D. J. *Pragmatic paradoxes. Mind*, vol. LVII, 1948, p. 358-9.
21. ——— *Pragmatic paradoxes and Fugitive Propositions. Mind*, vol. LX, no. 240, 1951, p. 536-8.
22. OLIN, Doris. *The prediction paradox resolved. Philosophical Studies*, vol. 44, no. 2, 1983, p. 225-33.
23. ——— *The prediction paradox: Resolving recalcitrant variations. Australasian Journal of Philosophy*, vol. 64, no. 2, 1986, p. 181-9.
24. QUINE, W. V. O. *On a So-Called Paradox. Mind* , vol. LXII, no. 245, 1953, p. 65-7.
25. SCRIVEN, Michael. *Paradoxical Announcements. Mind*, vol. LX, 1951, p. 403-7.
26. SHARPE, R. A. *The Unexpected Examination. Mind*, vol. LXXIV, no. 294, 1965, p. 255.
27. SHAW, R. *The paradox of the unexpected examination. Australasian Journal of Philosophy*, vol. 55, no. 1, 1977, p. 41-58.
28. 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.
29. ——— *Recalcitrant variations of the prediction paradox. Australasian Journal of Philosophy*, vol. 60, no. 4, 1982, p. 355-62.

30. SUDBURY, Aidan e WRIGHT, Crispin. *The paradox of the unexpected examination*. *Australasian Journal of Philosophy*, vol. 55, no. 1, 1977, p. 41-58.
31. WEINSTRAUB, Ruth. *Practical Solutions to the Surprise Examination Paradox*. *Ratio* (8), 1995, p. 161-9.
32. WEISS, Paul. *The Prediction Paradox*. *Mind*, vol. LXI, no. 242, 1952, p. 265–9.
33. WILLIAMS, John N. *The Surprise Exam Paradox: Disentangling Two Reductions*. *Journal of Philosophical Research*, vol. 32, 2007, p. 67-94.
34. WILLIAMS, Timothy. *Knowledge and Its Limits*. Oxford University Press, 2002, p. 135-46.