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

INTRODUÇÃO

O paradoxo do exame surpresa

Que maravilhoso termos nos deparado com um paradoxo. Agora temos alguma esperança de progredir.[1] *– 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

1 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 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 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 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). Roy Sorensen crê que contos populares germânicos do século XVIII já continham problemas de eliminação reversa, análogos ao do exame surpresa (Sorensen, 2003:267).

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.

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 de 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 fecho epistêmico [*epistemic closure principle*], 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, pois ele pretende que o aluno simplesmente não esteja justificado a tomar o anúncio do professor nem mesmo *ex hipothesi* 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 diminui-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", irão, ao menos, defender que os alunos *podem* tomar o anúncio do professor como evidência *ex hipothesi.*

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 como tendo uma natureza estritamente lógica, procuraram, primariamente, associá-lo de algum modo com o paradoxo do mentiroso, buscando formulações autorreferentes, que gerassem uma recursão infinita na qual o valor de verdade ficasse variando indefinidamente.

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 são compatíveis (Lyon, 1959:514), ou a de B. Meltzer, que tenta solucionar o problema apelando à negação do 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 elocuções

A PRIMEIRA RESPOSTA AO PROBLEMA APRESENTADO POR O'CONNOR FOI FEITA POR LAURENCE Jonathan Cohen, na mesma revista *Mind*, em 1950. A sua solução 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 de 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"[2]: 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.

---

2 "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".

Weiss define 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 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 − ½`."

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):

C_1) E já vai ter ocorrido em algum dia que não $t + n$;

C_2) 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 C_2 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:

C_3) E vai não vai ocorrer em $t + n$ (o que violaria o decreto);

C_4) 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 uma *reductio ad absurdum*, sem poder fazê-la: em uma *reductio* correta a proposição a ser refutada deve ser tomada como verdadeira. Como K não tem este direito, dado que C_3 é possível, K há de considerar C_4 como possível também, o que implica que seu argumento indutivo não se segue, e que, sendo assim, C_1 também é um cenário que pode ocorrer.

Supor o cumprimento do decreto, e a eliminação de C_4, 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 qual cenário será o caso, torcer - com toda a reza possível! -pelo cenário em que o enforcamento não ocorre.

1.4. 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 professor, supondo ser confiável, realiza o anúncio com um certo objetivo, e espera estar passando alguma informação aos alunos, que consideram, *prima facie*, o professor confiável. 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. O *insight* aqui é que devemos adotar uma posição *prática* em relação ao paradoxo e ver a promessa como um pronunciamento da *vida real*. Ao contrário dos outros autores, que buscarão uma única solução, a perspectiva prática admite uma miríade destas.

Weinstraub questiona então, que motivo alunos *ordinários* tem para desacreditar na promessa. É claro que o argumento do aluno astuto só se segue se este acreditar que a promessa *irá* perder sua credibilidade, por ser paradoxal. Porém, diz ela, é perfeitamente razoável para estes alunos acreditar na promessa mesmo que o argumento de astuto seja impecável, afinal, "omnisciência lógica não é uma condição necessária para a racionalidade". Na vida real, precisamos compreender a distinção entre paradoxo *atual* e *percebido*. Assim compreendemos 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.

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 Weinstraub 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, que a população e as autoridades não estão dando atenção suficiente à esse provável desastre.

Se o professor é ingênuo – não está ciente da existência deste problema –, e os alunos responderem de modo ingênuo – não informarem o professor do problema – tem-se uma solução, pois aí os alunos não tem como saber o dia. O professor, provavelmente, se dará conta do problema 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.

Se o problema é conhecimento comum – todos sabem que todos sabem do problema – como entender o anúncio? Há, segundo Weinstraub, 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, se o professor estiver falando sério, os aluno devem interpretar o anúncio de acordo com a solução preferida deste. 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 não há paradoxo algum, devemos estudar.

O argumento de Weinstraub, 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 Weinstraub – ou de qualquer professor que concorde com ela – interpretar um anúncio dela? De acordo com a "solução prática", quando há conhecimento comum sobre o problema, devemos interpretar o anúncio de acordo com a solução preferida do professor. No caso, devemos pensar: qual a solução preferida por Weinstraub 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 , quando há conhecimento comum sobre o problema, 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.

2.2. Declarações e ordenamentos

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.

O 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 concluímos 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, 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 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*". Assim temos uma definição do problema "puramente lógica": com as regras servindo como axiomas, qualquer problema derivado desta definição só pode ser um problema lógico e não de outra natureza.

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 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 arrazoado 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 valida 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[3] 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

DAVID KAPLAN E RICHARD MONTAGUE INICIAM SEU ARTIGO AFIRMANDO QUE A SOLUÇÃO DE Shaw não é paradoxal, ao contrário do que afirma o autor. No entanto, eles também afirmam terem encontrado uma variação do problema *genuinamente* paradoxal, com a intenção de colocar o paradoxo *ao lado* do paradoxo do mentiroso.

---

3 Um resumo da ópera, que ignoro cada detalhe, está em Levy 2009, p. 154, nota 25.

Vejamos a formalização proposta[4], 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" e Q é a proposição "K é enforcado na quarta";

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

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

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

As fórmulas $\underline{S}$, $\underline{T}$ e $\underline{Q}$ são usadas substituindo φ 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"[5].

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.
- K é enforcado na terça, mas não na segunda, nem na quarta, e K não tem como saber disso na segunda de noite.
- K é enforcado na quarta, mas não na segunda, nem na terça, e K não tem como saber disso na terça de noite.

Assim, D_1 pode ser formalizada como:

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

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

Vejamos então como procede o argumento de K que D_1 não pode ser cumprido. Primeiro, assumamos que ele seja, e que tenha se passado segunda e terça sem nenhum exame $(\neg S \land \neg T)$. O problema se resumiria a $Q \land \neg K_t(\underline{Q})$. Assim, K saberia, na terça de noite, que quarta ocorreria o enforcamento, ou seja $K_t(\underline{Q})$, o que contradiz $\neg K_t(\underline{Q})$. Logo $\neg Q$.

As assunções implícitas e plausíveis para suportar este argumento são:

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

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

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

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 fecho 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 que K está em uma situação específica em que podemos tomar A_2 como certo. A_3 assume que K conheça estes dois princípios (já no domingo). E se A_1 e A_2 implicam $\neg Q$, podemos concluir que:

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

Para mostrar a impossibilidade da terça, ainda assumindo D_1, o segundo disjunto 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 disjunto da afirmação D_1). Assim, concluímos que $\neg T$. Novamente, a base para este argumento se assemelha à forma usada acima:

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

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

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 $\neg Q$). K não sabe que T na segunda de noite, ou seja, $\neg K_s(\underline{T})$. Logo, D_1 é perfeitamente possível. Temos portanto, uma contradição – $D_1 \wedge \neg D_1$. Assim, os autores consideram que, dada esta definição inicial, o anúncio *não pode ser cumprido*, porém isto não é o mesmo que ser paradoxal.

A formalização da versão de Shaw, por outro lado, notam Kaplan e Montague, apresenta uma autorreferência. Esta seria a formulação de Shaw, com dois dias, D_2:

$$(S \wedge \neg T \wedge \neg K_d(\underline{D_2 \rightarrow S})) \vee$$
$$(\neg S \wedge T \wedge \neg K_s(\underline{D_2 \rightarrow T}))$$

Após a análise desta versão os autores concluem que esta também não representa um paradoxo[6], chegando ao mesmo resultado da versão anterior, ou seja, apenas que o anúncio não pode ser cumprido. Para salvar a paradoxalidade pretendida por Shaw, propõem os autores, devemos adicionar à formulação original "a não ser que K saiba no domingo que esta afirmação é falsa", cuja formalização seria:

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

E tal formalização poderia ser resumida, dizem eles, para 1 ou mesmo 0 dias!

6 Supondo, como fizemos antes, que K seja enforcado na terça, temos T e $\neg$S. Temos então que estabelecer $\neg K_s(D_3 \rightarrow T)$. Para aplicar a mesma linha de raciocínio, precisamos mostrar que $D_2 \rightarrow T$, considerada no início do período é uma sentença não-analítica a respeito do futuro. Mas $D_3 \rightarrow T$ é analítica, pois eles demonstram que $\neg D_3$ se segue logicamente de princípios epistemológicos gerais, e então também $D_3 \rightarrow T$.

$$(S \wedge \neg K_d(\underline{D_3 \rightarrow S})) \vee K_d(\underline{\neg D_3})$$

$$K_d(\underline{\neg D_3})$$

Pela semelhança ao paradoxo do mentiroso [*Liar*], os autores decidem *batizar* este este paradoxo 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 apresentam então uma série de "conclusões" que a existência de tal paradoxo implicaria em uma teoria formalizada do conhecimento, com o objetivo de para evitar tais paradoxos.

Contudo, mesmo aceitando a existência do paradoxo do conhecedor, não fica claro porque devemos analisar o problema tanto inserindo a autorreferência, como fez Shaw, como inserindo a cláusula que engendra o conhecedor, a saber "a não ser que K saiba no domingo que esta afirmação é falsa". Tal adição não é explícita na formulação clássica do problema e não parece nem mesmo ser necessária para engendrar a contradição que o anúncio aparentemente cria. Assim, me parece que os autores estão ignorando o problema fundamental e entendendo "paradoxo" de uma forma mais estrita, que é, no caso, uma proposição cujo valor de verdade oscila a partir da nossa relação de conhecimento com a ela mesma.

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, a de entender surpresa somente no caso de haver de fato uma execução. Ayer desenvolve uma outra elaboração do problema que, ele pretende, irá mostrar que a resposta de Quine é insuficiente, e, ao mesmo tempo, levar 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 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 às 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 às, 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 às 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 percorrimento, 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 às 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 se "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 clara que as formulações alternativas. No entanto, este deve mesmo ser o maior mérito do artigo. Ayer simplesmente ignora o problema apresentado com um dia só, um conjunto de cartas consistindo apenas de um às. Neste caso, o prisioneiro sabe que o às será sacado e assim ele não pode ser enforcado.

3.2 Os *blindspots* condicionais

ROY 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 três versões por ele elencadas em um artigo anterior, de 82, *Recalcitrant Variations of the Prediction Paradox*.

Primeiro, Sorensen busca demonstrar que os "autorreferencialistas" distorceram o problema, atacando os argumentos de Scriven, Shaw e de Kaplan e Montague. A crítica de Sorensen segue, em linhas gerais, a mesma que eu apontei no final do capítulo das soluções lógicas, que a autorreferência não está contida na interpretação mais natural do anúncio do professor e que a autorreferência não é necessária para o problema. O professor apenas fala que o aluno não poderá saber da ocorrência do exame e não que ele não poderá "de acordo com este anúncio". A estes pontos, ele acrescenta uma interessante observação: se houvesse mesmo um aspecto autorreferente no anúncio, poderíamos tomar *qualquer* anúncio de ignorância do ouvinte como tendo uma natureza autorreferente. `(B)` não se segue de `(A)`:

`(A)` Você não sabe qual a capital do Texas.
`(B)` Você não sabe qual a capital do Texas mesmo que você saiba esta sentença.

A acusação é que o autorreferencialista está interpretando `(C)` como `(C')`:

`(C)` Austin é a capital do Texas e você não sabe disso.
`(C')` Austin é a capital do Texas, mas este fato não pode ser deduzido desta sentença.

Enquanto `(C)` não é uma contradição – quiçá, mesmo que `(C)` seja dito à alguém, esta pessoa não saiba que Austin é capital do Texas: pode ser que ela esteja segura que seja Phoenix, assim a proposição é simplesmente *verdadeira* – `(C')` é.

A solução de Sorensen focará então no passo inicial, ou seja, na ideia de que mesmo um exame na sexta possa 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.

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. 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 próprio 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. Não há informatividade alguma no anúncio, para Eric, caso Eric esteja com a estrela.

Sua solução depende do que ele chama de *blindspot*[7] epistêmico. Uma proposição p é um *blindspot* para uma pessoa a, no tempo t, se e somente se p é consistente enquanto $K_a p$ (a sabe que p em t) é inconsistente. Assim proposições mooreanas – mas não apenas estas – são *blindspots*: "Está chovendo mas Andy não sabe disso" é um blindspot para Andy, mas não para Woody. A proposição mencionada acima, formalizada como $p \wedge \neg K_a p$ é uma proposição consistente, porém ela não pode ser conhecida por Andy.

(1) $K_a(p \wedge \neg K_a p)$	
(2) $K_a p \wedge K_a \neg K_a p$	[(1)]
(3) $K_a p \wedge \neg K_a p$	[(2), RV[8]]
(4) $\therefore \perp$	[(3)]

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. É possível que exista um mundo em que "ninguém sabe de nada", porém ninguém pode saber disso neste mundo);
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 pode saber que "João descobriu que foi adotado aos 18 anos" aos 19, porém não aos

7 Eu poderia traduzir por ponto cego, mas me dá nos nervos.

8 Regra da veracidade do conhecimento. Se φ sabe que p, então p é o caso. $K\varphi p \vdash p$.

17);

c) Princípio do acesso interpessoal: Tudo que pode ser conhecido por alguém, pode ser conhecido por qualquer outra pessoa (e.g. A mãe de João pode saber que "João nunca soube, não sabe, nem saberá que foi adotado", mas João não);

Já a proposição "Se Rafael sobreviveu, ele será o único que sabe disso" não é um *blindspot*, porém um *blindspot condicional* – para todos, menos Rafael. Digamos, que um meteoro gigante tenha colidido com o planeta Terra. Milagrosamente, eu e Rafael sobrevivemos. Neste cenário, eu posso saber que Rafael sobreviveu, ou seja, posso saber o antecedente. Porém, se eu sei disso, eu sei que ele não é o único. Então, se sei o antecedente, não posso saber a proposição.

É claro que a versão de um dia do problema é um *blindspot*. Mas seria a versão com mais dias um exemplo de um *blindspot* condicional? Sendo p_1 a proposição "o exame ocorre na quinta" e p_2 "o exame ocorre na sexta":

$$(1)\ (p_1 \wedge \neg K_a p_1) \vee (p_2 \wedge \neg K_a p_2)$$

Ou seja, "o exame ocorre na quinta e Andy não sabe disso ou o exame ocorre na sexta e Andy não sabe disso". Mas, por lógica elementar, podemos modificar o exemplo deste modo:

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

Visivelmente, Andy se encontra com um *blindspot* condicional. Ele pode conhecer o anúncio do professor, porém, *se conhece o antecedente* – se o exame não for dado na quinta – não pode mais conhecer o anúncio.

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

$$K_a(p_5 \wedge \neg K_a p_5)$$

$$\therefore \perp$$

O padrão geral de um paradoxo de predição – seja o do exame surpresa ou do estudante designado –, segundo Sorensen, 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 *blindspots* 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.

3.3. 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 do alinhamento de três argumentos já existentes na literatura, o argumento do "Anúncio Improjecionável", o argumento de Wright e Sudbury e o argumento dos *blindspots* epistêmicos.

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 $(A)_E$ como a afirmação "haverá um exame quarta, quinta ou sexta" e $(A)_S$ é a afirmação "o exame será uma surpresa". Após (A) ser feito o aluno astuto apresenta o argumento "da impossibilidade" que consiste 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 – $(A)_S$ – 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 $(A)_S$. Ele só poderá ser feito na quarta, o que também violaria $(A)_S$ (*Backtracking*). A conclusão do aluno astuto é que o exame é impossível – $(A)_E$ não pode ocorrer.

Ainda assim, temos "a intuição" 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*.

Levy está interessado, como grande parte dos autores voltados às soluções epistêmicas, na versão mais forte do paradoxo do exame surpresa. 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).

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*, pois ele só se surpreende com a ocorrência do exame e não com a não ocorrência. Assim, (A) equivale a disjunção de:

(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 sexta é um dia impossível para a ocorrência 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 $(A)_E$ porém não $(A)_S$). Sexta não é um dia *impossível* de ocorrer o exame do mesmo modo que sábado o é – dadas nossas assunções. A resposta a tal argumento é que a forma correta de apresentar a conclusão do argumento da não-sexta é que sexta não é um dia possível para o exame *se* o professor vai cumprir (A). 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* – logo, nem fisicamente – possível dar o exame na sexta *e ainda* cumprir (A). Astuto *deve pressupor* o cumprimento de (A) se o professor *puder*. Assim, com o que vimos até aqui, ele *pode* concluir que sexta é um dia impossível para o exame.

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" para "Haverá um exame surpresa *amanhã*". Representemos esta afirmação como (A*). $(A*)_E$ e $(A*)_S$ juntos parecem colocar astuto em uma situação epistemológica insustentável. $(A*)_E$ justifica astuto a crer que o exame será dado na sexta enquanto $(A*)_S$ 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`". Como vimos, tal proposição mooreana (também um *blindspot*) é inacreditável para astuto.

A preocupação com a irracionalidade de se crer em `(A*)` é que motivou, segundo Levy, as "soluções pragmáticas", como a de Ruth Weinstraub ao paradoxo do exame surpresa, onde `(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. Em todo caso, espero que tenham estudado!". 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, pois este 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 uma *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)`$_S$. [(1)]

(3) ∴ Astuto não está justificado a crer em `(A)`$_S$. [(2)]

(4) `(A)`$_S$ é uma parte essencial de `(A)`: Se alguém não está justificado a crer em `(A)`$_S$ 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)`$_E$.

(7) ∴ Astuto não está justificado a crer em `(A)`$_E$. [(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 uma *reductio*:

(1) Um exame na sexta será uma surpresa (assumindo que o argumento de Wright e Sudbury está correto e sua conclusão em 11).
(2) O professor irá cumprir `(A)` se puder.
(3) ∴ Astuto está justificado, na quinta, em crer que haverá uma prova na sexta. [(2), (3)]
(4) ∴ Um exame na sexta não será uma surpresa. [(por definição de surpresa e 3)]
(5) ⊥ [(1), (4)]
(6) ∴ (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[9], 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.

---

9 A solução preferida de Levy é a de Eugene Mills, a saber, em um resumo grotesco, de que a proposição é *simplesmente* falsa.

Considerando que astuto é maximamente racional ele está ciente de que se encontra em tal *blindspot*, mas esta ciência também não dissolve o problema. A verdade de (A) é *incompatível* com ele crer justificadamente que (A). E o mesmo é verdade no caso 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 – $(A)_E$. 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*). O argumento de Wright e Sudbury e a da 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 $(A*)_E$ – 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 ANOS DE HISTÓRIA DO desenvolvimento do problema do exame surpresa vemos uma clara evolução, ao menos no sentido de um mapeamento claro das intuições envolvidas, e nos 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 um aluno de outra turma – intuição F. Ela só é um problema quando o *próprio* aluno astuto toma conhecimento de tal proposição. Assim, a ideia de um problema pragmático do tipo de O’Connor (e.g. “Eu não estou escrevendo”) presta um serviço e está de acordo com às soluções epistêmicas mais contemporâneas. Afinal, um *blindspot* pode ser engendrado por uma ação. Já as soluções lógicas ignoram fortemente esta intuição, apelando a axiomas e deduções que independem das pessoas e ações envolvidas. Porém, estes primeiros autores entenderam o anúncio como *self-defeating*, não dando muito crédito às intuições A, B, C, D e E.

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 um compromisso infalível. Esta tradição que se estendeu por diversos autores, e que vimos na sua manifestação mais recente, mas já bem diferente, no artigo de Weinstraub, tem em comum tentar ver o anúncio do professor de modo não idealizado, mas como algo que ocorre “no mundo real”. Segundo esta interpretação são os objetivos que *não necessariamente estão contidos no significado do anúncio* que interessam. No entanto, o problema mais forte é um cujo aluno astuto é um “mestre da lógica”. Assim, não basta, chegada a quinta, dizer que é *mais seguro* crer que o professor irá dar a prova, mesmo que não seja mais uma surpresa, pois este parece ser o comportamento mais habitual de um professor ingênuo. Do ponto de vista do mestre lógico, a única razão para crer na ocorrência do exame é o anúncio do professor, e não se deve, portanto, tomar apenas *parte* do anúncio como satisfazível. Talvez aqui ocorra uma divergência em relação ao que estes autores entendem como “significado intuitivo do anúncio”, a intuição C, mas, em todo caso, como vimos nas objeções ao artigo de Weinstraub, a solução prática é incompleta e ainda incorre em um outro tipo de problema.

Por sua vez, Weiss tem como mérito defender a intuição contrária, a intuição D): 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 certas pessoas *agem* – ou 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 busca é resolver o problema que surge quando refletimos detidamente no *esquema lógico relevante* do problema. O mesmo vale para os outros autores "lógicos". Porém, como vimos em Sorensen, a suspeita de que a interpretação que estes fazem do anúncio do professor, tornando-o autorreferente, apesar de criar novos e interessantes *puzzles*, não é nem necessária, nem plausível quando se trata do paradoxo do exame surpresa.

Enfim, nenhuma das soluções que vimos parece ter dado conta dos seis critérios de Wright e Sudbury como as de Sorensen e a de Levy. Levy, apesar de não ter reinventado a roda, tem o mérito de buscar conciliar estas intuições. Minha interpretação é que o s argumentos de Sorensen e de Levy falham apenas no critério (E) e apenas no caso da prova não ser dada até quinta. Contudo, tal "falha" é *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 *tenha* de ser cumprida – do que com o argumento destes autores. De fato, os três princípios cuja aceitação dos *blindspots* implica na rejeição – os princípios de acesso irrestrito, interpessoal e intertemporal – podem parecer *prima facie* verdadeiros, porém, uma breve inspeção os mostra falhos.

Assim, as intuições relevantes são satisfeitas:

(A) A conclusão é que não há realmente um paradoxo envolvido. Há, no entanto, um problema complexo que requer um mestre lógico para sua solução, o que já era de se esperar;

(B) O anúncio é cumprível e essa solução é eficaz mesmo no caso de um dia só;

(C) O significado intuitivo do anúncio é preservado. Nenhuma mudança autorreferencialista é necessária;

(D) O argumento se esforça por apresentar as intuições que nos levam a dar valor ao argumento da não-sexta;

(E) Os alunos são, em princípio, informados pelo anúncio. No argumento dos *blindspots* condicionais de Sorensen, a informatividade é mantida, visto que é possível crer no anúncio, desde que não se conheça seu antecedente, como no caso em que o exame é dado em algum dia antes do último (o que também faz jus à esta intuição). No caso do professor não ter dado a prova até quinta, o anúncio não tem mais informatividade *para os alunos* – apesar de continuar informativo para outros agentes – ou seja, eles não podem estar mais, sendo mestres lógicos, justificados a crer que o anúncio será cumprido e nem o contrário. E é por isto que o anúncio é cumprido *caso* a prova seja aplicada na sexta (e, não sendo, isso não torna o anúncio paradoxal, mas mendaz).

(F) O *blindspot* só ocorre para o aluno astuto;

Recusar parte do critério (E) não apenas resolve o *puzzle*, mas mais do que isso, é da insistência em mantê-lo que brota boa perplexidade do problema e suas muitas soluções. Temos então de aceitar a existência de proposições possivelmente verdadeiras, contingentes e perfeitamente compreensíveis, porém inacreditáveis – *blindspots* – para certos sujeitos ou por um certo período.

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, p. 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*, nº. 105, 1998, p. 41–51.
8. COHEN, L. Jonathan. *Mr. O'Connors Pragmatic Paradox. Mind*, vol. LIX, nº. 233, 1950, p. 85-7.
9. DALLAGNOL, Darlei. *Crer e Saber. Dissertatio,* nº 26, verão de 2007, p. 9–26.
10. FRANCESCHI, Paul. *A Dichotomic Analysis of the Surprise Examination Paradox. Cogprints*, 2002.
11. 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.
12. 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.
13. JANAWAY Christopher- *Knowing about Surprises: A Supposed Antinomy Revisited. Mind*, nº. 391, 1989.

14. JONGELIN, Tjeerd B. e KOETSIER, Teun. *Blindspots, self-reference and the prediction paradox. Philosophia*, vol. 29, nº. 1, 2002, p. 377-91.

15. KAPLAN, David e MONTAGUE, Richard. *A Paradox Regained. Notre Dame Journal of Formal Logic*, 1960, p. 79-90.

16. KIRKHAM, Richard L. *The two paradoxes of the unexpected examination. Philosophical Studies* , vol. 49, nº. 1, 1986, p. 19-26.

17. ______ *On paradoxes and a surprise exam. Philosophia*, vol. 21, nº. 1, 1991, p. 31-51.

18. KRIPKE, Saul. *Philosophical Troubles.* Oxford University Press, 2010, p. 27-51.

19. LEVY, Ken. *The Solution to the Surprise Exam Paradox. Southern Journal of Philosophy* , vol. 47, nº. 2, 2009, p. 131-58

20. LYON, Ardon. *The Prediction Paradox. Mind*, vol. LXVIII, nº. 272, 1959, p. 510-7.

21. MELTZER, B. *The third possibility. Mind*, vol. LXXIII, nº. 291, 1964, p. 430-3.

22. O'CONNOR, D. J. *Pragmatic paradoxes. Mind*, vol. LVII, 1948, p. 358-9.

23. ______ *Pragmatic paradoxes and Fugitive Propositions. Mind*, vol. LX, nº. 240, 1951, p. 536-8.

24. OLIN, Doris. *The prediction paradox resolved. Philosophical Studies*, vol. 44, nº. 2, 1983, p. 225-33.

25. ______ *The prediction paradox: Resolving recalcitrant variations. Australasian Journal of Philosophy*, vol. 64, nº. 2, 1986, p. 181-9.

26. QUINE, W. V. O. *On a So-Called Paradox. Mind* , vol. LXII, nº. 245, 1953, p. 65-7.

27. RUSSELL, Bertrand. *On Denoting. Mind*, vol. XIV, nº. 56, 1905, p. 479-93.

28. SCRIVEN, Michael. *Paradoxical Announcements. Mind*, vol. LX, 1951, p. 403-7.

29. SHARPE, R. A. *The Unexpected Examination. Mind*, vol. LXXIV, nº. 294, 1965, p. 255.

30. SHAW, R. *The paradox of the unexpected examination. Mind*, vol. LXVII, nº. 267, 1958, p. 382-4.

31. 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.

32. ______ *Recalcitrant variations of the prediction paradox. Australasian Journal of Philosophy*, vol. 60, nº. 4, 1982, p. 355-62.

33. ______ *Paradoxes of Rationality* em *The Handbook of Rationality*. Editado por Al Mele, Oxford University Press, 2003, p. 257–75.

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, nº. 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.