

# DO DESACORDO AO PARADOXO DA AUTORIDADE: UMA ANÁLISE DA CONCEPÇÃO DE SERVIÇO DE AUTORIDADE À LUZ DA TEORIA DO “PONTO-CEGO”, DE R. SORENSEN

*Ramiro de Ávila Peres*

PPGFil - Universidade Federal do Rio Grande do Sul

**Resumo:** A partir de uma revisão crítica da literatura, estuda-se um desafio do anarquismo filosófico à teoria da autoridade jurídica de J. Raz: seguir uma ordem de que se discorda seria irracional, pois equivaleria a agir contra o que se considera mais justificado. Utilizando referências atinentes à teoria da decisão e à epistemologia, além de exemplos sobre ferramentas de auxílio à tomada de decisão do cotidiano (como *experts* e programas), esboçamos duas possíveis respostas, que correspondem a duas concepções instrumentais distintas de autoridade: numa, é justificável para o agente colocar-se numa situação que o leve a agir de forma inconsistente com suas crenças – a ideia de “submeter-se à autoridade”, afim a uma “confiança cega”; noutra, utilizando a teoria dos “pontos-cegos”, delineiam-se restrições que uma teoria da racionalidade impõe à teoria da autoridade.

**Palavras-chave:** Autoridade, racionalidade, paradoxo.

**Abstract:** Using a critical review of the literature, we study a challenge from philosophical anarchism to J. Raz's theory of legal authority: it would be irrational to follow an order with which one disagrees, since it would mean acting against what is considered more justified. Through references from decision theory and epistemology, and deploying examples about tools for assisting in routine decision-making, we sketch two possible answers: first, it may be justifiable to put yourself in a situation that leads to irrationally acting; second, according to the “blindspots” theory, we outline some restrictions from a theory of rationality to the authority.

**Keywords:** Authority, rationality, paradox.

## 1- Introdução

Uma teoria da autoridade política precisa dar conta da conexão entre a questão da obrigação política, a da legitimidade da autoridade e o problema da lei injusta. Em outras palavras, ela precisa responder por que razão estamos *justificados* em obedecer a entidades como governos e leis, mesmo quando consideramos que as normas reconhecidas por essas entidades não são justas. Afinal, se a autoridade deve ser capaz de fazer uma *diferença prática* na deliberação do sujeito, ele deve estar obrigado a segui-la mesmo quando ela esteja errada, ou ele não concorde com o teor de suas diretivas – no vocabulário corrente de filosofia política, deve segui-las independentemente de seu *conteúdo*.

O anarquismo<sup>1</sup> fornece uma resposta cética a esse problema: não existe obrigação política – nem, portanto, o problema da lei injusta. Uma das principais inspirações filosóficas dessa teoria são as situações paradoxais a que somos levados pela noção de independência de conteúdo – a exemplo do que Shapiro chama *paradoxo da racionalidade*: “quando as autoridades estão erradas, elas não podem ter o poder de obrigar os outros – quando estão certas, o seu poder de obrigar não tem sentido” (SHAPIRO, 2000, p.2-3, trad. livre)<sup>2</sup>.

Na seção a seguir, estudaremos uma das principais teorias desenvolvidas para responder a esse paradoxo – a “concepção de serviço” (*service conception*) de J. Raz. Após, apresentaremos uma reformulação do problema, para a qual a teoria *raziana* não fornece uma resposta clara. Sugerimos a possibilidade de uma “concessão mínima” ao paradoxo – que um sujeito pode ser vítima de uma forma de irracionalidade ao obedecer a uma ordem da qual discorda, sem que isso implique que a própria aceitação da autoridade é irracional.

Essa solução, contudo, não faz jus às aspirações da concepção de serviço. Para elucidar melhor o paradoxo, vamos compará-lo ao que Roy Sorensen (1988) chama de “pontos-cegos” (*blindspots*) – proposições afins a sentenças *mooreanas* (do tipo “*p*, mas não acredito que *p*”), como, e.g., “é racional fazer o que não considero racional fazer”. Caso a comparação seja correta, a conclusão do paradoxo é não uma proposição absurda, mas uma sentença em relação à qual o sujeito racional está impedido de adotar certas atitudes proposicionais. Esclarecemos a teoria de Sorensen na terceira seção.

Por fim, encerramos com uma análise mais detalhada sobre o que está envolvido quando um sujeito discorda de uma ordem da autoridade aceita. Conquanto observemos que o paradoxo da autoridade seja muito mais inofensivo do que a maneira como a literatura filosófica costuma apresentá-lo, sua discussão revela um limite para a aceitação racional de ordens.

## 2 - O caráter paradoxal da autoridade

### 2.1 A independência de conteúdo e a “concepção de serviço”

---

<sup>1</sup> Para os de índole pragmática, que consideram pouco relevante ou urgente responder aos desafios do ceticismo filosófico, lembro que o anarquista se distingue de suas contrapartes epistemológicas (como os filósofos que põem em dúvida a realidade do movimento, do tempo, da percepção, de outras mentes, etc.) na medida em que sua teoria – para usar a pecha que Dworkin (2003, 292) imputa ao que ele chama “ceticismo externo” – não “deixa tudo como está”: ela serve de fundamento geral para o libertarianismo político e tem implicações práticas genuínas.

<sup>2</sup> “[...] when authorities are wrong, they cannot have the power to obligate others — when they are right, their power to obligate is meaningless”.

Pode-se começar o estudo filosófico da autoridade com o reconhecimento de um paradoxo:

[...] a razão requer que o sujeito sempre deve agir sob a ponderação das razões de que está ciente. É da natureza da autoridade que esta demanda submissão mesmo quando o sujeito pensa que o que é demandado é contra a razão. Portanto, a submissão à autoridade é irracional (RAZ, 1979, p. 3, trad. livre).

Para responder a esse desafio, J. Raz desenvolve sua concepção de serviço da autoridade, segundo a qual ordens válidas são razões protegidas: elas são ao mesmo tempo razões para realizar uma ação (razões de primeira ordem) e razões para excluir a deliberação sobre outras razões (razões excludentes, de segunda ordem). Para ser dotada de autoridade, segundo Raz, uma ordem deve mediar entre pessoas e razões; ela deve ser apta a guiar a conduta do agente, sem que ele precise deliberar sobre as razões que justificariam a ação – para isso, ela substitui as razões que o agente já tinha. Essa é a *tese da preempção*: “o fato de que uma autoridade requer a realização e uma ação é uma razão para sua realização, a qual não deve ser adicionada às outras razões relevantes quando se pensa sobre o que fazer, mas deveria substituir algumas delas” (RAZ, 1994, p.214, trad. livre). Mas, para efetuar tal substituição efetivamente, a norma deve ser baseada nas razões que o sujeito já tem – i.e., sua justificação *depende* das razões para agir de primeira ordem do sujeito, ou a autoridade não seria distinta de simples coerção (ibid.). Essa tese da *dependência* implica a da *preempção*: uma vez que a autoridade já teria considerado algumas das razões aplicáveis ao sujeito, deliberar sobre estas junto com o comando da autoridade seria “contá-las” duas vezes (RAZ, 1986, p. 58)<sup>3</sup>.

A concepção de serviço tornou-se a teoria da autoridade mais difundida entre filósofos do direito – ao menos no ambiente acadêmico anglo-saxão (HERSHOVITZ, 2010). A noção de preempção explica os requisitos de ‘peremptoriedade’ e de ‘independência de conteúdo’ usualmente associados aos comandos de uma autoridade. Junto com a tese da dependência, as duas teses favorecem uma teoria da legitimidade, centrada na ‘tese da justificação

<sup>3</sup> Nesse ponto, discordamos de Raz: qual o problema de contar duas vezes? Se eu obedecer à autoridade, a dupla contagem não prejudica nossa análise (na verdade, essa parece ser a situação do sujeito que segue a regra, a despeito de seu próprio julgamento – e que Raz considera normal); por outro lado, se, apesar da dupla contagem, eu desobedeço à diretiva, isso significa que as razões em contrário eram muito significativas, pois superaram as razões (superavaliadas) para agir de acordo com a ordem. O ponto é que são justamente as razões em contrário que o comando busca excluir: “These exclusionary reasons do not, of course, exclude relying on reasons for behaving in the same way as the directives requires. [...] Hence the pre-emption excludes only reasons that conflict with the authority’s directive” (RAZ, 2009, p.144). Logo, o argumento da dupla contagem é uma petição de princípio: é a tese da preempção que torna a dupla contagem desnecessária (porque a preempção exclui as razões contrárias à ordem), e não a irracionalidade da dupla contagem que implica a preempção.

normal<sup>7</sup> – segundo a qual uma autoridade é legítima quando o indivíduo a ela sujeito tem maior probabilidade de atender às razões que se aplicam a ele aceitando as diretrizes dela e tentando segui-las do que se buscar seguir essas razões diretamente (RAZ, 1986, p.53). Essa concepção permitiria dissolver o paradoxo: embora a razão exija que o agente aja de acordo com as razões de que ele está ciente, isso não significa que ela exija que ele aja de acordo com todas essas razões; uma ordem legítima substitui as razões de primeira ordem. Ela também permite fundamentar a tese de que toda autoridade *de facto* reivindica (podemos também dizer que *pretende*) ter autoridade *de jure* – i.e., autoridade legítima (RAZ, 1986, p.26-7).

Se a autoridade é legítima, isso significa que o sujeito tem obrigação de segui-la, ainda que julgue de modo contrário (RAZ, 1986, p.62). É importante diferenciar isto da conclusão de, e.g., Hart (1982, p.253) de que a autoridade, ao menos em algum ponto, pretende “cortar” a deliberação sobre os prós e contras da ação – o que Raz chama de “submissão do juízo” (*surrender one’s judgement*). Como os comandos da autoridade prática são razões para agir, seria irrelevante (para caracterizar alguém como autoridade) se o sujeito julga de acordo com eles ou se se abstém de fazer um julgamento (RAZ, 1986, p.39). Em conjunto com a tese da *preempção*, porém, isso parece implicar que o sujeito que aceita uma autoridade legítima e que não se abstém de fazer esse juízo está agindo de forma irracional – pois estaria “reabrindo” uma deliberação encerrada.

Por outro lado, se o comando da autoridade legítima é, além de uma razão prática, uma razão epistêmica (e parece ser esse um dos pontos cruciais do caráter instrumental da concepção de serviço), então o sujeito age de forma irracional ao discordar dela – de forma semelhante a um investigador que ignora evidência relevante. Não posso beneficiar-me do conhecimento especializado de meu corretor se não aceitar seus conselhos de investimentos, assim como não posso beneficiar-me dum programa que efetue a mesma tarefa. É claro, há situações em que não devo aceitá-los – e.g., o instrumento pode deixar de ser confiável (se meu corretor tem conflito de interesses, ou se o programa apresenta *bugs*...). Essas são as situações que Raz (1986, p.62) chama de erros *claros* (e.g., se o resultado da soma de números inteiros for apresentado com uma fração decimal), em contraposição a um erro grande (se o mesmo resultado for um erro de milhares de unidades), porém difícil de detectar; é compatível com a concepção de serviço que ela seja condicionada à ausência de erros claros.

Isso aparentemente inverte o paradoxo: sendo  $\alpha$  uma ordem legítima, não deveríamos perguntar “como pode ser racional fazer  $\alpha$ , se não o considero racional”, mas “como posso não considerar racional fazer  $\alpha$ ?”. Submeter-se à

autoridade, na concepção de serviço, não equivale a abdicar de julgar racionalmente, mas a fazê-lo de acordo com a evidência mais relevante.

## 2.2 Novos paradoxos

Na Introdução, citou-se o paradoxo da racionalidade, que questionava como alguém pode submeter-se a uma autoridade se, quando discordar dela, a submissão é injustificada, e, quando concordar com ela, a submissão é desnecessária. Essa formulação sugere uma analogia com o que podemos chamar de “falácia do estudante”: “se eu estudar e não passar no teste, é porque sou ignorante e o estudo é inútil; se passar, é porque sou inteligente, e o estudo é desnecessário”. A falácia é rapidamente exposta porque é evidente que o estudo aumenta a probabilidade de passar no exame; ele tem um efeito epistêmico benéfico, na medida em que, no momento do teste, permite fazer um julgamento melhor.

Algo semelhante valeria para a autoridade, segundo a concepção de serviço: de acordo com a tese da justificação normal, o sujeito que obedece a autoridade legítima tem maior probabilidade de agir racionalmente. Mas tentemos reformular o paradoxo de uma forma (talvez) ainda mais primitiva. Suponha que F (uma autoridade) ordene a R que faça  $\alpha$ ; este, porém, *acredita* que não deve fazer  $\alpha$ . Sabendo disso, F expõe a teoria da autoridade *razziiana*, que lhe permite enunciar a R<sup>4</sup>: (i) “você deve fazer  $\alpha$ , ainda que acredite que não deve fazer  $\alpha$ ” e (ii) “você deve fazer  $\alpha$ , mesmo que eu esteja errado”. R, para concordar com (i), deveria concluir que (iii) “devo fazer  $\alpha$ , ainda que acredite que não devo fazer  $\alpha$ ”. O problema aqui surge da manutenção de duas *atitudes* incompatíveis (aceitar uma ordem e discordar dela); embora nenhuma dessas sentenças seja contraditória, a proposição (ii) não deve ser enunciada por F, e R não pode crer em (iii).

## 3 - Paradoxos e pontos-cegos

### 3.1 Racionalidade e justificação

O problema em questão refere-se à relação entre a formação de juízos justificados e ações. Uma possível solução para o paradoxo seria o que Derek Parfit (1984, p.12) chama de “irracionalidade racional”: situações em que, paradoxalmente, um agente racional, visando ao melhor estado de coisas para si, preferiria agir de maneira inconsistente com suas crenças ou com a própria racionalidade. O exemplo mais extremo é a “resposta de Schelling ao assalto à

---

<sup>4</sup> Apelar para o conteúdo ilocutório da ordem não resolve o problema. Se F enunciar o imperativo “faça  $\alpha$ ” com qualquer uma das sentenças supra estaremos no mesmo paradoxo, na medida em que a autoridade reivindica legitimidade.

mão armada”<sup>5</sup> (PARFIT, 1984, p.11), que demonstraria que qualquer teoria da racionalidade deve aceitar que, em algumas situações, pode ser racional causar a si mesmo irracionalidade, ainda que temporariamente.

O objetivo de Parfit é estudar a tensão entre a concepção de racionalidade como auto-interesse e o cumprimento de promessas e ameaças por um indivíduo racional. Imagine que ordeno que você me preste um serviço, sob a ameaça de explodir uma bomba que nos matará a ambos; supondo que sou racional, você decide ignorar a ordem – afinal, não ganharei nada (e perderei muito) cumprindo a ameaça. Nesse caso, é racional para o agente formar a intenção de que cumprirá suas ameaças e promessas, a fim de que outros indivíduos acreditem nisso; mas, considerando que seu cumprimento é custoso, por que ele deveria cumpri-las, se essas ações não terão qualquer consequência desejada? Ao invés de um “compromisso crível”, essa ameaça é o que teóricos dos jogos chamam de *cheap talk*.

Isso pode lembrar situações polêmicas entre os estudiosos da teoria da decisão causal e os da teoria dos jogos, como o paradoxo da toxina de G. Kavka (1983, p.33-4), o “dilema do prisioneiro” iterado um número determinado de vezes, ou situações afins ao jogo do ultimato, em que as estratégias adotadas por jogadores em testes de laboratório (e recomendadas pelo senso comum, na medida em que maximizam a utilidade esperada) contrastam com “equilíbrios perfeitos sub-jogo”<sup>6</sup>. Se um agente só é racional

---

<sup>5</sup> *Schelling's Answer to Armed Robbery*: referência a algumas das estratégias para vencer situações de barganha estudadas por Thomas Schelling no clássico *Strategy of Conflict*. A ideia é que frequentemente pode ser uma vantagem limitar as próprias opções. Suponha que um assaltante invadisse minha casa, ouvisse-me chamar a polícia (a qual demora alguns minutos para chegar) e, por fim, exigisse o ouro de meu cofre, sob ameaça de morte; o problema é que não tenho garantia alguma de que ele não me matará, já que posso identificá-lo – aparentemente, isso seria o mais racional (cf. o jogo *kidnap* em Binmore, 2007, p. 44). Nesse caso, seria racional tomar uma “poção” que me deixasse temporariamente irracional, sob cujo efeito eu adotaria crenças e atitudes manifestamente inconsistentes, e agiria visando a consequências indesejadas – como dizer “por favor, torture-me; odeio dor”. Segundo Parfit (1984, p. 11) “In the case just described, while this man is still in my house, it would be irrational for me to cause myself to cease to be very irrational. [...] During this period, my acts are very irrational. [...] We should claim instead that, since it was rational for me to cause myself to be like this, this is a case of rational irrationality”. O exemplo de Parfit pode parecer assaz estranho; mas isso não é tão relevante. Se não estiver convencido, suponha que um sujeito “confiável” ameace matar minha família a menos que eu tome a poção. O cenário é basicamente o mesmo: tomar a poção é uma decisão racional; as ações cometidas sob o efeito da poção, não. Afinal, se pode ser racional causar a si mesmo um estado de privação temporária de qualquer capacidade cognitiva (é isso o que analgésicos e anestésicos fazem), por que não o poderia ser privar-se da própria racionalidade?

<sup>6</sup> Um dilema do prisioneiro é um jogo simétrico entre dois jogadores ( $x$  e  $y$ ), que devem escolher (unilateralmente) uma das seguintes estratégias: cooperar (*dove*) ou não (*hawk*). O jogador  $x$  prefere: a) que  $x$  não coopere e que  $y$  coopere a b) que ambos cooperem, e isto a c) que ninguém coopere, e isto a d) que  $y$  coopere e  $x$  não coopere. Um *nash-equilibrium* é a situação em que a estratégia escolhida por cada jogador é a melhor resposta à estratégia escolhida por outro; já uma estratégia dominante é a que

quando age da maneira que ele acredita que possa causar a consequência desejada, é concebível que, em algumas situações estratégicas ou nas quais haja “falhas de racionalidade”, pode ser melhor para o agente *não agir de forma racional*.

Um exemplo pode ser a teoria *causal* da autoridade esboçada por Shapiro (2000, p.53), voltada a situações de racionalidade limitada – como seria o caso de alguém que contratasse um *personal trainer* rigoroso para obrigá-lo a exercitar-se mesmo quando, por preguiça, preferisse não o fazer. Submeter-se a outrem para evitar *akrasia* seria análogo a instituir uma regra para evitar tomar decisões em contextos nos quais a probabilidade de erro é grande, ou mesmo aceitar um código de conduta a fim de tornar-me confiável e entrar num grupo<sup>7</sup>:

Quando alguém se compromete a seguir uma autoridade para beneficiar-se de suas diretivas, essa pessoa restringe seu “eu” futuro de forma a agir segundo as ordens da autoridade, quaisquer que sejam. O sujeito não escolhe obedecer quando a ordem é emitida – quando ordenado, ele não tem escolha, exceto obedecer<sup>8</sup>.

---

maximiza o *pay-off* do agente, não importa o que o outro faça (Binmore, 2007, p.20). No dilema do prisioneiro, só há um *nash-equilibrium*, em que ambos *não cooperam*. O que há de paradoxal é que isso implica uma situação *pior* do que se ambos cooperassem.

Não devemos, porém, confundir esse jogo com sua versão iterada indefinidamente – na qual a cooperação mútua é um *nash-equilibrium*: se *x* não cooperar num *round*, *y* o “punirá” não cooperando no *round* seguinte (Binmore, 2007, p.21). O problema é quando o jogo tem um número *definido* de *rounds*: nesse caso, a indução retrospectiva (*backward induction*) mostra que os jogadores não deveriam cooperar. Então, jogar *hawk* em todos os casos é uma estratégia dominante, e a situação em que ambos sempre jogam *hawk* é um “equilíbrio perfeito sub-jogo”. Prova: suponha que o jogo tem 2 *rounds*; nesse caso, os jogadores sabem que não cooperarão no 2o (porque se reduz ao problema original). Mas isso implica que eles não têm por que cooperar no 1o *round*. Mas se isso vale para dois *rounds*, também vale para três, e assim sucessivamente. Mas imagine que o jogo tem 100 *rounds*: a indução retrospectiva diz que os jogadores vão adotar *hawk* em todos os *rounds*, embora uma estratégia que misture *dove* e *hawk* de acordo com a disposição dos demais de cooperarem retornaria um *pay-off* melhor (Binmore, 1997).

Nesse caso, poderíamos sentir-nos tentados a concluir que se *x* e *y* cooperassem no primeiro *round*, eles estariam adotando uma decisão irracional do ponto de vista do jogador em *cada round*, mas acertada quando se considera o jogo de forma geral. Mas, se essa possibilidade não pode ser excluída previamente, então um jogador não sabe se o outro agirá de forma racional e o equilíbrio sub-jogo não estará acessível – donde não temos motivo para chamar de irracional a decisão de cooperar no primeiro *round*.

<sup>7</sup> Interessante notar que o argumento do estado de natureza de Hobbes (2002, p.94-97) pode ser visto dessa forma: cada indivíduo renunciaria a seu direito a tudo, aceitando um soberano, a fim de deixar de ser uma ameaça para os demais. Isso não é coincidência, já que Parfit apresenta a noção de “racionalidade irracional” ao discutir a teoria da racionalidade como auto-interesse – a qual teria Hobbes como um dos principais expoentes.

<sup>8</sup> When a subject commits himself to following an authority in order to benefit from its directives, that person constrains his future self to act on the demands of the authority, whatever they may be. The

Nestes casos, o ato de submeter-se é racional (da mesma forma que, nos exemplos de Parfit, tomar a poção é uma ação racional, ou fazer uma ameaça autodestrutiva é racional), mas isso não significa que cada ação subsequente o seja. Aplicando o mesmo raciocínio à teoria da autoridade, a preempção se aplicaria, pois, não às *razões para decidir*, mas à própria decisão, uma vez que é o próprio ato de decidir (aqui considerado como deliberação e escolha) que é substituído pelo comando da autoridade. Se isso estiver correto, podemos dizer que R age racionalmente ao submeter-se ao conjunto de ordens de F (podemos dizer que esse conjunto é um “*global optimum* racional”), ainda que isso não implique que ele aja racionalmente ao submeter-se a cada ordem (pois esta não é compatível com outras de suas atitudes proposicionais). Pode haver uma ordem em relação à qual não é racional submeter-se – mas se R investigar, para uma determinada ordem, se é esse o caso, deixará de submeter-se ao conjunto de ordens de F.

Embora isso forneça uma solução para os paradoxos de F e R, na medida em que aceita um dos “chifres do dilema” – i.e., que R está agindo com uma forma de inconsistência no *caso particular de a* (mas não, de forma geral, por aceitar F como autoridade) – ela é difícil de extrapolar para todos os casos, pois frequentemente pensamos em regras e autoridades como algo semelhante a *instrumentos* para tomar decisões melhores, como parece ser a proposta da concepção de serviço. Além disso, essa solução ignora um aspecto fundamental de nosso paradoxo: no exemplo de F e R, não é a obediência à ordem em si que é irracional, mas a enunciação da conjunção das proposições que expressam que F deve ser obedecido e que a ação ordenada não é devida. Veremos uma alternativa que faz jus a esses pontos.

### 3.2 Pontos-cegos

Todos os vertebrados têm “pontos cegos” no campo de visão, causados pela passagem do nervo ótico pela retina; esses pontos são “preenchidos” na formação da imagem pela paisagem (é o *scotoma* do campo de visão). De forma análoga, há sentenças consistentes em relação às quais não podemos ter certas atitudes proposicionais – incluindo crenças e conhecimento. Por isso, essas sentenças foram batizadas de “pontos-cegos” por Roy Sorensen (1988).

Pontos-cegos decorrem do caráter reflexivo do auto-conhecimento – i.e., do fato de que temos crenças que têm por referência outras crenças ou outras atitudes proposicionais. O exemplo clássico seriam as *sentenças mooreanas*

---

subject does not choose to obey when the directive is issued – when ordered, the subject has no choice but to obey” (Shapiro, *ibid.*, trad. livre)

– assim chamadas em homenagem a George Moore e a sua descoberta da sentença paradoxal “chove, mas não acredito nisso”. Ela não é uma contradição: posso perfeitamente saber que “chove, mas  $x$  não acredita nisso”, desde que  $x$  não seja substituído por “o autor deste artigo” (i.e., eu). Sorensen (2011) apresenta uma ilustrativa prova (baseada na “prova da cognoscibilidade” de F. Fitch) de que é impossível ter conhecimento proposicional de um “ponto-cego”, pois isso implicaria uma contradição:

1.  $K(p \leftrightarrow \sim Kp)$  [leia-se: “sei que ‘ $p$ , mas não sei se  $p$ ”]
2.  $Kp \leftrightarrow K\sim Kp$  [(1), uma vez que o conhecimento se distribui pela conjunção]
3.  $\sim Kp$  [consequência de (2), considerando que conhecimento implica verdade]
4.  $Kp \ \& \ \sim Kp$  [(2), (3)]

Considerando que conhecimento é crença verdadeira e justificada<sup>9</sup> (*Gettier cases* à parte), e que (1) pode ser verdadeira (afinal, há várias verdades que desconhecemos;  $p$  pode ser qualquer uma delas), então o que me impede de conhecer (1) é o fato de que **não pode ser objeto de uma crença justificada**.

Note-se como, quando um sujeito não tem opinião sobre o valor de verdade de uma proposição, ele pode crer numa disjunção verdadeira de “pontos-cegos”; e.g., sei que ou “o número de estrelas é par, e não acredito nisso” é verdadeiro, ou “o número de estrelas é ímpar, e não acredito nisso” é verdadeiro: embora não possa crer em nenhum dos disjuntos especificamente, devo crer na disjunção.

Dessa forma, a noção de “ponto-cego” permite explicar como seres humanos normais podem ter *crenças falibilistas*: se eu fosse um lógico perfeito, deveria acreditar na conjunção de todas as minhas crenças (só assim poderia dizer que meu sistema de crenças é *completo*) – i.e., deveria acreditar que *todas* as minhas crenças são verdadeiras; mas isso contradiz a crença falibilista e modesta de que tenho ao menos uma crença falsa<sup>10</sup>. Se  $\beta$  representa o operador “creio que”, e as letras minúsculas, proposições em que creio, então a conteúdo da “crença modesta” é a seguinte disjunção de pontos-cegos: “( $p$  é

<sup>9</sup> Apesar do que esta passagem pode sugerir, nada no presente argumento depende da adoção de uma “concepção internalista” de conhecimento – i.e., da ideia de que um agente só detém conhecimento de uma proposição quando está justificado em crer nela e está ciente disso.

<sup>10</sup> Se o escopo dessa crença (chame-a  $\beta^\infty$ ) é *todo* o meu conjunto de crenças (permitindo, pois, auto-referência), então podemos ter um “paradoxo do mentiroso”: ou há ao menos uma sentença do conjunto complementar de crenças que é falsa, ou  $\beta^\infty$  o é; mas, nesse caso, ela é verdadeira, e assim *ad infinitum*.

*falso*  $\leftrightarrow \beta p$ ), ou (*q é falso*  $\leftrightarrow \beta q$ ), (*r é falso*  $\leftrightarrow \beta r$ )...”. Como já observamos, porém, não posso crer em nenhum dos disjuntos. Em bom português, podemos dizer que, embora eu creia que *cada uma* de minhas crenças é verdadeira, isso não significa que creia que *todas* (i.e., seu conjunto) sejam verdadeiras.

### 3.3 Pontos-cegos da razão prática

Um caso de ponto-cego de relevância prática é a “prova da auto-aniquilação” do utilitarismo, por Bernard Williams (2005, p.161 e s.), que ora resumimos:

- a) “um utilitarista deve buscar a atitude menos danosa possível que seja capaz de evitar consequências piores” (WILLIAMS, 2005, p.161);
- b) “assim, ele será frequentemente impelido a fazer coisas que, consideradas em si mesmas, são francamente abjetas” (como o que Williams chama de “intervenções preventivas”) (*ibid.*);
- c) se um número muito grande de pessoas for utilitarista, será empiricamente provável uma escalada de “intervenções preventivas” (até porque nem todas as pessoas são utilitaristas; do contrário, viveríamos numa utopia, onde o pensamento moral seria em boa medida desnecessário);
- d) esse cenário é pior do que se ninguém for utilitarista.

#### Conclusão de Williams:

se o utilitarismo estiver certo, e com ele algumas proposições empíricas bastante prováveis, será melhor que as pessoas não creiam no utilitarismo. Se, por outro lado, ele for falso, então certamente será melhor que as pessoas não creiam nele (2005, p.165).

De forma análoga, Sorensen (1988, p.382) menciona que um utilitarista pode ter dificuldades auto-referenciais ao aprender que um juízo falso sobre uma situação *S* pode levar a melhores consequências; isso proscreeveria julgamentos racionais a respeito de *S*. O exemplo de Williams apenas aplica a mesma lógica ao próprio utilitarismo; essa teoria implicaria, nesses casos, uma tese do tipo: “devo agir de modo a maximizar a utilidade geral, mas não posso acreditar nisso (pois implica agir contrariamente à maximização da utilidade geral)”. Essa “fragilidade à auto-referência” não é exclusiva do utilitarismo, mas se aplica a qualquer teoria da decisão que pretenda aplicar-se à *própria atividade de tomar decisões*, já que implica a possibilidade de “x deve fazer  $\alpha$ , mas não deve pensar nisso”.

Algo semelhante pode ocorrer com teorias da justificação da decisão jurídica. Seria esse o caso da “pretensão de correção” de R. Alexy, bem como

da reivindicação de legitimidade de Joseph Raz<sup>11</sup> por parte da autoridade jurídica; embora toda autoridade seja falível e esteja em condições de reconhecê-lo<sup>12</sup>, ela não pode dizer, ao enunciar certa decisão, que *esta* decisão é incorreta. Uma “autoridade falibilista” está numa situação análoga à da pessoa que admite que tem uma crença falsa – mas que é incapaz de apontá-la.

O que não significa que não possamos fazê-lo por ela: posso enunciar um de seus pontos-cegos (“*p*, mas você não acredita que *p*”), mas *você* não pode acreditar na proposição enunciada (ao menos não sem a tornar falsa). Sorensen (1988, p.161) considera que isso explica a discórdia mesmo em situações idealizadas; certamente, isso implica a ausência de conhecimento comum (SORENSEN, 1988, p.320) no que for relacionado ao ponto-cego – uma vez que não posso ter conhecimento de algo que é um ponto-cego para mim<sup>13</sup>.

#### 4 - Dissolvendo o paradoxo

Retomando o paradoxo inicial: F não pode enunciar “você deve fazer  $\alpha$ , ainda que eu esteja errado” porque esse é um “ponto-cego da autoridade legítima”: sua pretensão de refletir corretamente as razões do sujeito é incompatível com a admissão conjunta de erro. Isso não proscreve a modéstia ao “estilo Ayres Britto”: F pode admitir que existe *alguma decisão sua*  $x$  que é errada, mas que, apesar disso, deve ser seguida. Há uma diferença lógico-formal entre os dois casos: enquanto “alguma decisão  $x$ ” aqui é uma variável ligada por um quantificador existencial (“existe um  $x$  qualquer tal que  $x$  é uma decisão e...”), “ $\alpha$ ” denota, em nosso exemplo, o que seria uma proposição determinada (i.e., opera substituindo o ‘nome’ de um objeto – trata-se *desta* decisão); i.e., quando dizemos “alguma decisão  $x$ ”, não precisamos *apontá-la, discriminá-la*. A admissão de F de que há alguma decisão  $x$  que não é correta é análoga ao caso da “crença falibilista”: equivale a uma disjunção de pontos-cegos, do tipo “ou a decisão 1 é errada, e tenho a pretensão de que não seja errada; ou a decisão 2 é errada, e tenho a pretensão de que não seja errada...”.

F também deve abster-se de enunciar “você deve me obedecer, ainda que não acredite nisso”, mas por um motivo diferente (conquanto conexo): ele pretende que sua ordem seja uma razão para R, e essa é uma proposição que R *não pode aceitar*. O problema maior está em R: é mais fácil abster-se de produzir

<sup>11</sup> Para uma comparação entre as duas pretensões, ver Gardner (2012, p.125).

<sup>12</sup> E.g., a afirmação de Carlos Ayres Britto (2001, p.10), então Ministro do Supremo Tribunal Federal, de que essa corte tem o direito de errar por último.

<sup>13</sup> De acordo com o Teorema do Acordo de Robert J. Aumann (1976), dois agentes racionais (mais precisamente, *bayesianos*) com conhecimento comum de suas crenças iniciais (seus *priors*) e conhecimento comum das respectivas implicações não podem discordar em suas conclusões (seus *posteriors*).

um enunciado que uma crença; é por isso que os paradoxos da autoridade costumam ser enunciados do ponto de vista de R, o ‘súdito’. Se este acredita que F é uma autoridade legítima, mas discorda de F, pode não ser tão simples abster-se de formar uma crença que ele já tem; vale a pena analisar cada uma das possibilidades lógicas em que isso pode ocorrer:

1. F é uma autoridade legítima, e  $\alpha$  é devido: se a concepção de serviço estiver correta, R está sendo irracional em discordar. Ele não está desobedecendo à autoridade (i.e., não viola uma obrigação), mas seria racional (a fim de manter a consistência entre suas crenças e atitudes) rever sua segunda crença (de que  $\alpha$  não é devido).
2. F é uma autoridade legítima, mas  $\alpha$  não é devido: podemos dizer que F está dentro de sua “margem de erro” usual. Nesse caso, F não deixou de refletir as razões de primeira ordem de R. Ainda assim, R está sendo, para tomar de empréstimo outra expressão de Parfit (2011, p. 34), “menos que racional”<sup>14</sup> ao discordar. Embora suas duas crenças sejam verdadeiras (na legítimidade de F e no erro de  $\alpha$ ), ele está mais justificado em manter a primeira – donde ele “deve racionalmente” (parafraseando Parfit, *ibid.*) segui-la. É possível discutir se F cria uma obrigação nesse caso; se seu comando realmente substitui as razões de primeira ordem (i.e., se faz “diferença prática”), mas isso é irrelevante para o paradoxo (e para a perspectiva de R)<sup>15</sup>.

---

<sup>14</sup> É interessante distinguir duas noções: a) “R tem razões para fazer  $\alpha$ ”, de b) “R deve racionalmente fazer  $\alpha$ ”. Suponha que R se depara com uma perigosa serpente no deserto; ele acredita que o melhor a fazer é sair correndo – e, dada essa crença, isso é o que ele *deve, racionalmente, fazer*. Infelizmente, R ignora que as serpentes dessa espécie são muito rápidas, mas não atacam indivíduos que ficam parados; nesse caso, embora ele tenha uma razão para não correr, R a desconhece – donde ele deve, racionalmente, correr (v. PARFIT, 2011, p.34 e s.).

<sup>15</sup> É irrelevante porque o paradoxo não deveria ocorrer: não estou justificado em pensar que a ordem é errada. Apenas sugiro que R pode não ter obrigação porque, se ele não cumprisse a determinação de F, é discutível até que ponto seria passível de crítica – se, além de ter falhado em fazer o que racionalmente deveria fazer, ele teria violado uma obrigação. Suponhamos que meu general me ordene prestar apoio a um ataque – o qual, não sabemos, seria um desastre. Por minha preguiça, intuição, ou pacifismo, meu regimento se atrasa, o que (depois descobrimos) nos poupa de um vexame, corte marcial ou morte. Posso ser passível de alguma forma de repreensão, mas é discutível se violei uma obrigação para com meu general, já que ele estaria justificado em agradecer-me.

Na realidade, uma das críticas que se faz à concepção de serviço é que mostrar que um sujeito tem uma *razão* especial para seguir outrem não é suficiente para mostrar que tem uma *obrigação* de fazê-lo (HERSHOVITZ, 2010). Seria necessário mostrar que ele tem uma *relação* especial com quem emite a ordem. Se a relação entre comandante-subordinado exige obediência cega, então violei essa obrigação; mas pode haver outras relações de autoridade/subordinação que não sejam assim.

Em certo sentido, podemos dizer que a autoridade não comete um erro nesse caso – assim como R não erra ao segui-la (i.e., ele age de acordo com a razão). Se é possível chegar a conclusões teóricas falsas mesmo depois de reunir toda a informação disponível e raciocinar da maneira correta (isso é, sem cometer nada que possa ser considerado um erro), também o é para conclusões práticas; nesse caso, não há falha no próprio ato de julgar. Enquanto instrumento, a autoridade acerta ao desempenhar adequadamente sua função: o sujeito age melhor ao aceitar seu comando do que ao ignorá-lo.

1. F não é uma autoridade legítima,  $\alpha$  não é devido, mas R está *epistemicamente* mais justificado em crer que F é uma autoridade legítima: um exemplo comum disso é quando a informação é limitada; nesse caso, a probabilidade *a priori* de R respeitar suas razões de primeira ordem é maior se obedecer ao comando – e, na ausência de informação relevante (que pode não estar acessível a R), ele não pode extrair a respectiva probabilidade *a posteriori*. Esse caso é semelhante ao de (2), e R está sendo “menos que racional”, mas com a diferença de que, agora, é mais difícil sustentar que F faça “diferença prática” – já que sua ilegitimidade implica que a preempção das razões de primeira ordem *não ocorre*.

2. F não é uma autoridade legítima,  $\alpha$  não é devido, e R está *epistemicamente* mais justificado em crer que F não é uma autoridade legítima: nesse caso, a Tese da Justificação Normal de Raz não vale, e a autoridade não é legítima. Ela é como um instrumento não é útil – ainda que isso possa ocorrer apenas excepcionalmente (e.g., quando tenho informação que a autoridade não tem, de forma que a probabilidade *a posteriori* de que eu esteja certo, e ela errada, é maior)<sup>16</sup>.

Cabe discutir um ponto levantado por (3) e (4), ao qual se poderia objetar, inspirado pelos exemplos de Raz. A ideia é que parece difícil aceitar que uma autoridade possa ser “transitoriamente” ilegítima – i.e., que ora posso decidir que ela é legítima, ora posso decidir que não. Afinal, a peremptoriedade do comando parece implicar que ele deve sempre ser contado não como um *prior*, mas como um *posterior* – como uma razão conclusiva.

O estranho é que isso parece quebrar a analogia com instrumentos: ninguém se priva *a priori* da possibilidade de julgar quando eles são úteis ou não; posso rever minha submissão a um agente de investimentos, assim como

<sup>16</sup> Claro, ainda assim, pode ser justificável cumprir a ordem por outros motivos; e.g., por apego, medo ou admiração por F – alguma forma de “relação especial” com F.

posso deixar de usar um aplicativo, em ocasiões em que eles não se mostrarem úteis – o que não implica que os deixe de usar completamente. Isso parece uma caracterização mais adequada da distinção *razziãna* entre erros “claros” e erros “grandes”, vista na seção 2; erros claros são aqueles em que sabemos que a autoridade (ou qualquer outro instrumento) não são úteis. Afinal, se a legitimidade da autoridade deriva de seu caráter instrumental, por que quebraríamos a analogia nesse ponto?

Não há problema algum em acreditar que outrem tem maior probabilidade *a priori* de estar certo; o problema é acreditar que isso também vale sempre para a probabilidade *a posteriori*, no caso de um juízo determinado. Suponha que R saiba que F acerta com uma frequência 50% maior, e que F envie um e-mail dizendo que R deve comprar ações de certa companhia; no dia seguinte, pouco antes de confirmar a operação, R descobre que a companhia está quase quebrada. Não seria racional desconsiderar o e-mail neste caso?

Se R acreditar que, não importa quanta informação possua a mais, ainda assim o juízo de F será melhor do que o dele, R estará se comprometendo com um tipo estranho de dogmatismo (recusando evidência acessível – i.e., a informação a mais) e se considerando um *anti-expert* em relação ao caso. Essa é uma possibilidade lógica, e as outras pessoas poderiam, quiçá, constatá-lo, mas ele, não; pois, como aponta Sorensen (1988, p. 394-6), considerar a si mesmo um *anti-expert* implica crer num *ponto-cego*, do tipo “julgo que meus juízos não são confiáveis”.

Isso nos levaria de volta ao paradoxo de F e R; no caso, R estaria próximo de assumir que “devo fazer  $\alpha$  & não posso saber que devo fazer  $\alpha$ ”<sup>17</sup>. Esse problema vindica os que consideram que há algo de estranho na ideia de “submeter” sua capacidade de julgamento à autoridade. Se R deseja tal submissão, precisa de uma teoria afim à da “racionalidade irracional” para justificá-la.

## Conclusões

Na primeira parte do artigo, apresentamos a concepção de serviço da autoridade de J. Raz, desenvolvida como uma resposta ao paradoxo com que começamos este artigo; Raz busca mostrar que seguir uma ordem pode ser justificado, desde que satisfeitas as teses da preempção, da dependência e da justificação normal: nesse caso, a ordem substitui as razões primárias, embora dependa (ainda que parcialmente) delas.

---

<sup>17</sup> Isso é uma instância de  $(p \& \neg Kp)$ , que, como já vimos, não pode ser conhecida – justamente por não ser passível de crença justificada.

Como vimos em seguida, o paradoxo retorna quando a autoridade ou o súdito pronuncia um *ponto-cego* - um limite para aquilo em que estamos justificados em crer ou defender. A primeira resposta ao paradoxo assume que o súdito “submete o próprio juízo” à autoridade, que se abstém de julgar por si mesmo; mas, na medida em que discorda da ordem, ele incorre em algo afirmativo em uma inconsistência – embora se possa cogitar justificar a manutenção disso por meio de uma teoria da “racionalidade irracional”.

Alternativamente, assumindo a existência de falhas de racionalidade, pode-se apontar que um sujeito que segue uma ordem da qual discorda não está sendo irracional apenas *por seguir a ordem*, mas por manter *duas* atitudes incompatíveis entre si; sua falha pode ser antes epistêmica, se ele está deixando de acreditar naquilo em que tem mais razões para acreditar (i.e., no comando da autoridade legítima). Portanto, como em outras situações em que as evidências parecem conflitar com as crenças já mantidas por um agente, ele precisa definir, por si mesmo, qual é a atitude que ele está justificado em manter.

## Referências

- AUMANN, R. J.. “Agreeing to Disagree”. In: *Annals of Statistics*, v.4, n.6, 1976, p.1236-1239. Disponível em: <<https://projecteuclid.org/euclid.aos/1176343654>>, acesso em 24.7.2016.
- BRITTO, C. A. “O Regime Constitucional dos Tribunais de Contas”. In: *Revista Diálogo Jurídico*, Salvador, CAJ - Centro de Atualização Jurídica, v.I, n.9, (Dezembro) 2001. Disponível em: <<http://www.direitopublico.com.br>>.
- BINMORE, K. G. “Rationality and backward induction”. In: *Journal of Economic Methodology*, v.4, n.1, 1997, p.23-41. Disponível em: <<http://www.tandfonline.com/doi/citedby/10.1080/13501789700000002>>, acesso em 24.7.2016.
- BINMORE, K. G. *Game theory: A very short introduction*. New York: Oxford University Press, 2007.
- DWORKIN, R. *O Domínio da Vida: aborto, eutanásia e liberdades individuais*. Trad.: Jefferson Luiz Camargo. São Paulo: Martins Fontes, 2003.
- GARDNER, J. *Law as a leap of faith: essays on law in general*. Oxford: Oxford University Press, 2012.
- HART, H. L. A. *Essays on Bentham: Studies in Jurisprudence and Political Theory*. Oxford, UK: Clarendon Press, 1982.

HERSHOVITZ, S. “The Role of Authority”. *Philosophers' Imprint*, University of Michigan Public Law Working Paper, n.201, 2010. Disponível em: <<http://ssrn.com/abstract=1615675>>, acesso em 24.7.2016.

HOBBS, T. *Os Elementos da Lei Natural e Política*. Trad. Fernando Dias Andrade. São Paulo: Ícone, 2002.

KAVKA, G. “The Toxin Puzzle”. In: *Analysis*, v.43, n.1, 1983, p.33-36. Disponível em: <<http://analysis.oxfordjournals.org/content/43/1/33.extract>>, acesso em 24.7.2016.

PARFIT, D. *Reasons and Persons*. Oxford: Clarendon Press, 1984.

\_\_\_\_\_. *On What Matters. Vol. 1* Oxford: Oxford University Press, 2011.

RAZ, J. *The Authority of Law: Essays on law and morality*. Oxford: Clarendon Press, 1979.

\_\_\_\_\_. *The Morality of Freedom*. Oxford: Clarendon Press, 1986.

\_\_\_\_\_. *Ethics in the public domain: Essays in the morality of law and politics*. Oxford: Clarendon Press, 1994.

\_\_\_\_\_. *Between Authority and Interpretation: on the theory of law and practical reason*. Oxford: Oxford University Press, 2009.

SHAPIRO, S. J. “Authority”. In: *Stanford/Yale Jr. Faculty Forum Research Paper 00-05*; Cardozo Law School, Public Law Research Paper, n.24, 2000.

Disponível em SSRN: <<http://ssrn.com/abstract=233830> ou <http://dx.doi.org/10.2139/ssrn.233830>>, acesso em 24.7.2016.

SORENSEN, R. A. *Blindspots*. Oxford: Clarendon Press, 1988.

\_\_\_\_\_. “Epistemic Paradoxes”. In: *The Stanford Encyclopedia of Philosophy* Spring 2014 Edition, Edward N. Zalta (ed.). Última atualização em 2011. Disponível em <<http://plato.stanford.edu/archives/spr2014/entries/epistemic-paradoxes/>>, acesso em 24.7.2016.

WILLIAMS, B. *Moral: uma introdução à ética*. Trad. Remmo Manarino Filho. São Paulo: Martins Fontes, 2005.

Email: bolivarcambara@yahoo.com.br

Recebido: 6/2018

Aprovado: 3/2019