





Revista de Pesquisa em Filosofia

# FUNDAMENTO

Ouro Preto - Minas Gerais - Brasil



**UFOP**

Universidade Federal  
de Ouro Preto

Fundamento Revista de Pesquisa em Filosofia / Universidade Federal de Ouro Preto / Grupo de Pesquisa em Filosofia Contemporânea. n. 5, (jul-dez. 2012) – Ouro Preto:

Ed. UFOP, 2012

Semestral – Tiragem: 200 exemplares.

ISSN: 2177-6563

1. Filosofia - Periódicos. I. Universidade Federal de Ouro Preto. II. Instituto de Filosofia, Artes e Cultura. III. Grupo de Pesquisa em Filosofia Contemporânea.

Endereço para correspondência:

**Revista Fundamento**  
**Instituto de Filosofia, Artes e Cultura – IFAC**  
**Rua Coronel Alves, 55, Centro, Ouro Preto, MG, Brasil**  
**CEP: 35400-000**

**Endereço eletrônico: [revistafundamento@ufop.br](mailto:revistafundamento@ufop.br)**

## **Revista Fundamento**

UNIVERSIDADE FEDERAL DE OURO PRETO

**Reitor**

Prof. Dr. Marcone Jamilson Freitas Souza

INSTITUTO DE FILOSOFIA, ARTES E CULTURA/IFAC

**Diretor**

Prof. Dr. Guilherme Paoliello

DEPARTAMENTO DE FILOSOFIA/DEFIL

**Chefe de Departamento**

Prof. Dr. Olímpio José Pimenta

**Bibliotecária**

Neide Nativa

FUNDAMENTO

REVISTA DE PESQUISA EM FILOSOFIA

**Diretores Gerais**

Prof. Dr. Mário Nogueira (UFOP)

Prof. Dr. Sérgio Miranda (UFOP)

**Diretores Executivos**

Aline Monteiro (UFOP)

Lívia Reis (UFRJ)

Roberto Wagner de Carvalho Júnior (UFOP)

**Editor**

Prof. Dr. Mário Nogueira (UFOP)

**Diagramação**

Aline Monteiro

**Revisão**

Lívia Reis

**Assistente de Revisão**

Aline Monteiro

**Comitê Editorial**

Prof. Dr. Alexandre Noronha Machado (UFPR)

Prof. Dr. Hélio Lopes da Silva (UFOP)

Prof. Dr. Danilo Marcondes (PUC-Rio)

Prof. Dr. Frank Thomas Sautter (UFSM)

Prof. Dr. Guido Imaguire (UFRJ)

Prof. Dr. Marco C. Ruffino (UFRJ)

Prof. Dr. Nelson Gonçalves Gomes (UNB)

Prof. Dr. Túlio Xavier de Aguiar (UFMG)

## SUMÁRIO

### I - CONVITE À PESQUISA EM FILOSOFIA

- 11 AMARTYA SEN E AS SOCIEDADES MAIS JUSTAS COMO IDEIA E REALIDADE  
*AMARTYA SEN AND FAIR SOCIETIES AS IDEA AND AS REALITY*  
Mário Nogueira de Oliveira - UFOP
- 23 O QUE QUEREMOS DE UMA TEORIA DA JUSTIÇA?  
*WHAT DO WE WANT FROM A THEORY OF JUSTICE?*  
Autor: Amartya Sen  
Tradução: Mário Nogueira de Oliveira - UFOP

### II - ARTIGOS INÉDITOS

- 40 SERÁ QUE, COMO DEFENDEM OS EPICURISTAS, ESTAR MORTO NÃO É UM MAL PARA QUEM MORRE?  
*AS ARGUED BY THE EPICUREANS, IS BEING DEAD NOT BAD TO THOSE WHO DIE?*  
Domingos Faria - Universidade de Lisboa
- 61 A MORAL SENSUALISTA DE HELVÉTIUS E O PRINCÍPIO DO INTERESSE  
*THE SENSUALIST MORAL OF HELVETIUS BY THE PRINCIPLE OF INTEREST*  
Camila Ferraz - UFPR
- 77 ROBÔS PODEM SER MORALMENTE RESPONSABILIZADOS? REFLEXÕES À LUZ DO PRAGMATISMO DE JOHN DEWEY  
*CAN ROBOTS BE MORALLY RESPONSIBLE? REFLECTIONS FROM JOHN DEWEY'S PRAGMATISM*  
Fernando Cesar Pilan - UENP
- 97 SISTEMA POLÍTICO COMUNAL, LIBERDADE, IGUALDADE E RELIGIÃO ENTRE OS NORTE-AMERICANOS SEGUNDO TOCQUEVILLE  
*COMMUNAL POLITICAL SYSTEM, FREEDOM, EQUALITY AND RELIGION IN NORTH AMERICA ACCORDING TO TOCQUEVILLE*  
Cícero J. Silva Oliveira - PUC-Rio
- 109 COMO TORNAR A BIVALÊNCIA COMPATÍVEL COM A INTUIÇÃO DE QUE O FUTURO É ABERTO?  
*HOW TO MAKE BIVALENCE COMPATIBLE WITH THE INTUITION THAT THE FUTURE IS OPEN?*  
Matheus Silva - UFRGS

## **II - ARTIGOS INÉDITOS**

- 127 QUEM SABE OUTRA HORA, OUTRO MUNDO  
*MAYBE ANOTHER TIME, ANOTHER WORLD*  
Fernando Furtado - UFMG

## **III - TRADUÇÃO**

- 141 NOTAS SOBRE O PRECONCEITO  
*NOTES ON PREJUDICE*  
Isaiah Berlin  
Tradução de Mário Nogueira de Oliveira - UFOP

## **IV - RESENHA**

- 147 OBRA: *PHILOSOPHICAL DEVICES, PROBABILITIES, POSSIBILITIES, AND SETS*  
Autor: David Papineau  
Resenha por: Matheus Martins Silva - UFRGS

**Convite à Pesquisa em  
Filosofia**



# Amartya Sen e as sociedades mais justas como ideia e realidade

## *Amartya Sen and fair societies as idea and as reality*

Mário Nogueira de Oliveira <sup>1</sup>  
Universidade Federal de Ouro Preto

*“Vivemos em um mundo de opulência sem precedentes, mas  
também de privação e opressão extraordinárias. O  
desenvolvimento consiste na eliminação de privações de  
liberdade que limitam as escolhas e as oportunidades das  
pessoas de exercer ponderadamente sua condição de cidadão”*  
Amartya Sen

Amartya Sen possui pelo menos três obras *principais* traduzidas no Brasil: *Desigualdade reexaminada*, *Desenvolvimento como Liberdade* e *A ideia de Justiça*. Nelas ele apresenta sua tese sobre a abordagem das capacidades (*capabilities*), originalmente formulada junto a Martha Nussbaum.<sup>2</sup>

---

1 - Gostaria de agradecer especialmente a K. C. Tan e G. Rainbolt, meus supervisores de estágio pós-doutoral, que propiciaram os meios para que a minha pesquisa - da qual este texto é uma parte - pudesse ser realizada na GSU e UPENN. Sou grato a eles por isso e pelas conversas sempre estimulantes. Desse período de pesquisa derivaram textos, palestras, conferências, publicações e diversas orientações de alunos além de cursos de Filosofia Política Contemporânea. Agradeço também ao DE-FIL e à UFOP pelo período de pesquisa que me foi concedido. Também sou grato às revisões da língua portuguesa feitas por Livia Reis e Aline Monteiro.

2 - Cf. Nussbaum, Martha C. and Sen, Amartya (Eds). *The Quality of Life*. Oxford Univeristy Press: Oxford, 1993.

Em *Desigualdade Reexaminada*, Amartya Sen escreve que ser igualitário, no sentido de buscar algum tipo de igualdade para todos os cidadãos de uma determinada sociedade ou mesmo para toda a humanidade não é realmente uma característica “unificadora” (2001:45) e atesta isso apresentando teses de reconhecidos autores de diferentes perspectivas éticas e políticas, tais como J. Rawls, R. Dworkin, T. Nagel, T. Scanlon, R. Nozick, J. Buchanan e W. Letwin em seu artigo “The case against equality” (“O argumento contra a igualdade”) na obra *Against Equality (Contra a Igualdade)*. Sen mostra como todos esses autores assumem uma forma igualitária para abordar os ordenamentos sociais, embora a igualdade seja posta em “lugares” diferentes para cada um dos autores mencionados, a saber: igualdade de tratamento por merecimento (Letwin), igualdade de direitos libertários (Nozick) - isto é, nenhuma pessoa tem mais direito à liberdade que qualquer outra -, igual liberdade e igualdade na distribuição de “bens primários” (Rawls), “tratamento como iguais” e “igualdade de recursos” (Dworkin) ou “igualdade” (Scanlon) (Cf. Sen, 2001: 44-5).

O autor que marcou definitivamente a obra de filosofia política de Sen e a quem ele reconhecidamente dedica seu livro *A ideia de justiça* foi John Rawls. Sen é um leitor dedicado de Rawls desde 1968, quando foi *visiting scholar* em Harvard vindo de Nova Deli e recebeu de Rawls os originais de *Uma teoria da justiça*. Nos dois anos seguintes, lecionou junto ao autor dos manuscritos e a Ackerman a teoria que viria a ser um divisor de águas na filosofia política mundial e o maior legado de Rawls.

Segundo K. C. Tan, a importância de Rawls para toda a filosofia política reside justamente no fato de ele ter sido responsável por um novo modo de se trabalhar a filosofia política e a ética na segunda metade do séc. XX. Tan mostra que por muito tempo a filosofia analítica anglo-americana assumiu que o papel da filosofia seria exatamente o de fazer a análise da linguagem e dos conceitos. As filosofias moral e política, especificamente, eram estudadas exclusivamente como metaética, ou seja, clarificando o significado de termos como “bem”, “correto”, “justo”, entre outros. E “as questões normativas, como visionar uma sociedade justa, estariam fora do alcance da filosofia”.

Rawls deixou de aplicar a análise no sentido estrito às investigações éticas e procurou construir uma teoria. Lembrando que “análise no sentido estrito” foi aquela nascida com Moore e Russel, que marca o início da aplicação da análise da proposição como método de esclarecimento filosófico. Rawls enquadra-se mais ao sentido amplo de “filosofia analítica” pelo qual entendemos um modo de fazer filosofia centrada na busca da clareza conceitual e no rigor argumentativo. Sobre esse ponto, Daniels escreveu, em acordo com K. C. Tan:

*Ao contrário da ênfase dominante no século XX da Ética analisar a linguagem moral e investigar tópicos de metaética, Uma Teoria da Justiça argumenta com extremo rigor sobre princípios morais substantivos e discute suas implicações para a construção de instituições sociais básicas. Esse padrão normativo estabelecido por Rawls encorajou outros trabalhos sobre “justiça” e em outras áreas da ética aplicada e da filosofia política. Essa nova perspectiva da filosofia analítica explica a relevância do trabalho de Rawls sobre muitas gerações.* (Daniels, 2005, p. 361).

Somente ao analisar e encontrar pontos na teoria de Rawls com os quais não concorda ou pensa em aperfeiçoá-los, Sen pode delimitar claramente sua singularidade na defesa do ideal de igualdade liberal como sendo a igualdade das capacidades básicas de cada indivíduo e a defesa da ideia de que os limites apropriados para a igualdade fundamental entre cidadãos devem considerar as suas “capacidades básicas”.

Amartya Sen define uma capacidade como a habilidade de uma pessoa realizar atos valiosos ou alcançar estágios de vida considerados valiosos por ela própria. Uma capacidade representa uma série de combinações alternativas de estados físicos e mentais que uma pessoa é capaz de fazer, de se tornar ou de ser (Cf. Sen, 1993; Sen, 2001). Portanto, as capacidades são *oportunidades* ou *liberdades* (*freedoms*) para alcançar aquilo que um indivíduo considera valioso.

A abordagem das capacidades advoga que cada pessoa não deve ver o outro como um meio para o crescimento econômico ou para a estabilidade social. Devemos levar em conta *as liberdades* que propiciem que as pessoas se tornem capazes de escolher aquilo que elas valorizam e também devemos levar em conta as liberdades para que as pessoas se tornem capazes de trabalhar na remoção dos obstáculos a essas liberdades (“remoção de injustiças”, como diz Sen). Em *Desenvolvimento como Liberdade*, fica claro que ampliar as liberdades de uma pessoa implica ampliar suas capacidades. Uma sociedade desenvolvida seria aquela em que seus membros possuem liberdade para escolher a vida que desejam e também gozam de liberdade para “funcionar” bem, entendendo-se “funcionar” como poder realizar concretamente suas capacidades.

Em seu texto intitulado *Desigualdade Reexaminada*, Sen explica que a ampliação das capacidades humanas envolve as *liberdades* (*freedoms*) que as pessoas de fato vivenciam ao escolher as vidas que elas têm razão para valorizar. As pessoas devem ser capazes de escolher aquilo que importa para uma vida valiosa. Não há como compreendermos o conceito capacidade sem o conceito liberdade. As capacidades são o conjunto de oportunidades e opções que uma pessoa dispõe quando elege a vida que deseja.

Quando analisamos os arranjos sociais tendo em vista a justiça e a igualdade, são as capacidades das pessoas que devem guiar a nossa análise e nossa avaliação, ao invés de certa quantidade de dinheiro ou recursos ou mesmo qualificações que elas possam gerir. Os recursos são os meios, mas não os fins intrínsecos do bem-estar humano. A abordagem trata de um vínculo entre a oportunidade e seus efeitos e consequências, ou seja, as capacidades e os funcionamentos.<sup>3</sup>

Amartya Sen chama de funcionamentos (*functionings*), por exemplo, ao ato de poder ler, a compreensão do que é lido, a participação na vida social de uma comunidade (associações de moradores, partidos políticos, sindicatos e demais esferas de uma sociedade democrática) e estar calmo ao realizar tudo isso (não estar submetido cargas excessivas de estresse, vergonha, cansaço ou baixa autoestima). Já as capacidades, como condições para os funcionamentos, constituem um tipo de “potencial” para adquiri-los. Sen cita como exemplos: ter sido *ensinado* a ler, possuir livros ou outros materiais disponíveis para a leitura, viver em um contexto onde as pessoas da sua classe social, do seu gênero, da sua etnia estão permitidas a falar sobre os assuntos que lhes interessam e podem fazer isto sem estresse ou nervosismo exagerado.

O vínculo e a diferença entre esses conceitos podem ser sintetizados assim: a capacidade está para a oportunidade assim como o funcionamento está para a realização concreta de uma potencialidade. Há uma parte que Sen chama de transcendental, a ideia ou a possibilidade, e outra que chama de empírica, a realização concreta.

## 1. Além de Rawls:<sup>4</sup>

Como visto, Amartya Sen parte da pergunta “o que é uma sociedade justa?”, que originou vários trabalhos a partir de *Uma teoria da justiça* de J. Rawls, para elaborar sua tese própria sobre a justiça social. Essa pergunta é considerada central para a maioria das teorias da justiça e Sen denomina o modo como Rawls responde à questão como abordagem “transcendental” da justiça, uma vez que enfoca a identificação dos arranjos sociais *totalmente* justos. Esse ponto será amplamente discutido no artigo “O que queremos de uma teoria da justiça?”, traduzido na íntegra, com autorização do próprio Amartya Sen, especialmente para este número da *Revista Fundamento*.

---

3 - Um estudo ao qual devo muito para minha pesquisa é a coletânea de artigos editada por Melanie Walker e Elaine Unterhalter, *Amartya Sen's capability approach and social justice in education*, New York, Palgrave-Macmillan, 2007. Diversos ensinamentos dos doze artigos de diferentes autores aumentaram meu entusiasmo pela obra de Sen e passaram a fazer parte do que compartilho aqui.

4 - O segundo capítulo de *A ideia de justiça* é intitulado “Rawls e mais além”.

Segundo Sen, Rawls explora com profundidade a natureza de uma sociedade inteiramente justa, inclusive economicamente, a partir da equidade contratual.<sup>5</sup> Os problemas relacionados à justiça social e econômica são apresentados quando o debate enfoca os objetivos e as ramificações da sua teoria de justiça. Sen busca aprimorar o trabalho de Rawls, ampliando seu caráter igualitário e a defesa da igualdade de recursos para cada um dos cidadãos. Para isso, como mencionamos anteriormente, ele define o ideal de igualdade liberal como sendo a igualdade das capacidades básicas de cada indivíduo. Seu argumento reside na defesa da ideia de que os limites apropriados para a igualdade fundamental entre cidadãos devem considerar as suas capacidades básicas.

Sen afirma que a abordagem do problema da desigualdade feita por Rawls não é “suficientemente igualitária” porque ela não leva em conta a diversidade e a heterogeneidade das capacidades de cada agente que, a rigor, sobrepõem-se a eles e a suas reais condições de possuir os meios e identificarem os fins para converter suas capacidades em fontes de liberdade de fato. Segundo Sen, é possível listarmos alguns fatores que afetam as vidas dos agentes tornando a disparidade social um fato real. Alguns exemplos desses fatores causadores de desigualdade social e econômica vão desde diferenças climáticas na região de domicílio, uma distribuição arbitrária e desigual de bens, diferenças de ordem pessoal (diferenças pessoais sejam físicas ou socioeconômicas), a vulnerabilidade a doenças endêmicas ou parasitárias (como a constante exposição à malária, febre amarela, dengue, etc.). De acordo com Sen, a tese rawlsiana da igualdade de recursos não dedica atenção suficiente a esses fatores que, na realidade cotidiana, implicam disparidades no que concerne à liberdade e capacidade das pessoas para que alcancem as metas que consideram valiosas (Sen, 1985, p.81). *As teorias anteriores à tese de Sen falham porque acatariam apenas determinados bens e por que ignorariam aquilo que pessoas diferentes precisam: ou seja, pessoas diferentes precisam de meios diferentes para que consigam bens diferentes* (Sen, 1988 p. 153).

A falha também pode ocorrer porque as pessoas podem ajustar seus fins, suas expectativas e preferências às suas condições iniciais (e injustamente desiguais), como já alertara Martha Nussbaum<sup>6</sup>. Isto quer dizer que a igualdade social e econômica básica é afetada porque alguns setores da população ignoram uma diversidade de aspectos que determinam e influenciam a qualidade de nossas vidas. Em outras palavras, o meu desconhecimento de

---

5 - Para uma síntese dos fundamentos da tese rawlsiana, Cf. Sen, A. “What Do We Want from a Theory of Justice?”. *The Journal of Philosophy*, Vol. 103, No. 5 (May, 2006), pp. 215-238 e, como mencionado, em tradução para o português neste número de *Fundamento: Revista de Pesquisa em Filosofia*.

Todas as menções, daqui em diante, feitas ao artigo “What do we want from a theory of justice” estão citadas na paginação do texto original encontrada em *The Journal of Philosophy*.

6 - Cf. Nussbaum, M., *Women and Human Development: The Capabilities Approach*, Cambridge: Cambridge University Press, 2000 e Nussbaum, M., “Capabilities as Fundamental Entitlements: Sen and Social Justice,” *Feminist Economics*, 9(2/3): 33–59, 2003.

um modo de vida possível e melhor faz-me aceitar e buscar somente os modos de vida que já conheço e acredito serem os únicos possíveis.

## **2. Liberdade e seres sociais mais completos**

Como Sen deixa claro em *Desenvolvimento como Liberdade*, a expansão das liberdades substantivas desfrutadas pelos indivíduos é a finalidade a qual sua teoria visa. O próprio Sen resume bem essa tese ao afirmar que, se não assumirmos que desenvolvimento implica exclusivamente crescimento do Produto Interno Bruto, aumento de rendas pessoais (com ou sem bolsas e ações paternalistas), industrialização ou avanço tecnológico, poderemos entender que o desenvolvimento de uma sociedade demanda a remoção das “principais fontes de privação da liberdade”, que ele lista como: “pobreza, tirania, carência de oportunidades econômicas e destituição social sistemática, negligência nos serviços públicos, intolerância ou interferência excessiva de Estados repressivos” (Sen, 2000, p. 18).

No mesmo livro, um pouco mais adiante, Sen escreve que o desenvolvimento de uma sociedade deve estar “relacionado sobretudo com a melhora de vida que levamos e com as liberdades que desfrutamos”. E complementa: “expandir as liberdades que temos razão para valorizar não só torna nossa vida mais rica e mais desimpedida, mas também permite que sejamos seres sociais mais completos, pondo em prática nossas volições, interagindo com o mundo em que vivemos e influenciando esse mundo” (Sen, 2000, p. 29).

Mas, ao considerarmos que há riqueza a ser distribuída em uma dada sociedade, como tornar a sua repartição justa com todos os membros dessa sociedade? Esse é um problema que Sen percebe na teoria de Rawls, e pretende contribuir para sua resolução.

A conversão dos bens primários ou dos recursos pode trazer problemas advindos das diferenças entre as pessoas e suas diferentes escolhas sobre o que deve ser feito em relação aos seus bens. De acordo com Sen, uma teoria que busque ser justa deve ser sensível às condições nas quais a pessoa vive no momento em que faz suas escolhas na vida: em uma sociedade justa, é preciso que cada cidadão seja o único responsável por suas escolhas, e para isso, ele deve estar livre de fatores arbitrários impostos pelos aspectos natural e social, que limitam suas opções e suas oportunidades.

Portanto, devemos enfrentar os fatores arbitrários que podem ser de ordem natural: certas restrições impostas pela natureza, como ter nascido com doenças congênitas, por exemplo, ou ser nascido muito pobre. Tampouco podem ser negligenciados os fatores arbitrários

causados pelos fatos sociais, tal como Rawls exemplifica com a “loteria do nascimento”. Essa “loteria” afirma que se alguém tiver a má sorte de nascer em uma família miserável, não terá muitas chances de realizar uma vida que possa ser considerada boa, pois as condições básicas para o desenvolvimento da sua liberdade estará obstado ou dificultado pelas privações que sofre desde o útero materno (por exemplo, má nutrição da mãe durante a gestação). Enquanto uma pessoa que nasce em uma família mais abastada terá muitas oportunidades para se desenvolver e para ser responsável por suas próprias escolhas, pois terá diante de si os meios para que possa alcançar e ser responsável sobre o que fizer com sua vida.

A consideração feita por Sen mostra que os bens primários de Rawls (que para este são meios para se alcançar a liberdade efetiva, e que tal liberdade merece um lugar central na política e distribuição justa da sociedade) *não* consistem em liberdade efetiva, e sim em *meios* para se alcançar tal liberdade. Para Sen é importante focar a liberdade como um fim em si mesmo, e que isso seja a finalidade buscada pela distribuição da riqueza em um arranjo político.

Como Michael Sandel explica com muita clareza, Sen muda a perspectiva rawlsiana da seguinte forma: do enfoque em bens, passa-se ao enfoque em pessoas e suas condições de vida. Sen trabalha com as capacidades das pessoas em transformar os bens em liberdade efetiva para se buscar e concretizar sua própria concepção de vida boa.<sup>7</sup>

Sen aborda a perspectiva da capacidade sobre dois aspectos, o bem-estar e a liberdade, para realizar ou buscar o bem-estar (Cf. Sen, 2000, p. 90-ss). No que diz respeito ao primeiro aspecto, ele argumenta que o bem-estar de uma pessoa está ligado aos seus funcionamentos. E tais funcionamentos podem ser tanto mais simples como estar adequadamente nutrido, gozar de boa saúde, estar livre de doenças que podem ser evitadas e também da morte prematura, entre outros fatores, até realizações mais complexas, tais como ser feliz, gozar de autorrespeito e tomar parte na vida da comunidade sem se sentir discriminado.

Para que uma pessoa esteja apta a realizar sua própria concepção de vida boa, é preciso que ela esteja funcionando bem em suas mais variadas formas. Ela deve ter a capacidade de realizar os funcionamentos que achar necessário em busca de seu bem-estar.

Quanto maior a capacidade de uma pessoa para realizar os mais diversos funcionamentos, maior será sua liberdade efetiva de escolher como deve funcionar.

Certa vez, foi-me perguntado, em um debate sobre o trabalho de Sen, se, ao colocarmos os membros de uma sociedade em uma fila, qual deles deveria estar no início da fila? Para

---

7 - Cf. SANDEL, Michael. *O liberalismo e os limites da justiça*. Trad. Carlos Pacheco do Amaral. Lisboa: Fundação Calouste Gulbenkian, 2005. Este livro também foi especial para a pesquisa geral e deixo registrada a minha dívida ao trabalho de Sandel, não só para compreensão e redação de textos, mas nas minhas aulas e orientação de alunos.

mim, a resposta de Sen parece ser bastante clara: são as “capacidades básicas” que devem fundamentar a distribuição de bens em um determinado arranjo político que busque a justiça. E, com base nas capacidades, ele afirma quem são *os menos favorecidos da sociedade, ou seja, aqueles que tiverem menor capacidade para converter bens distribuídos em liberdade efetiva para realização de funcionamento*. Portanto, estes devem estar no início da fila.

### 3. A arbitragem equitativa

Em seu artigo “What do we want from a theory of justice?” (2006), Sen apresenta um outro problema da teoria de Rawls. Para ele, a tese não é exequível porque a conquista de uma sociedade plenamente justa exigiria uma superabundância de instituições, incluindo as operações irrestritas de um Estado soberano. Algumas dessas instituições são fortemente maculadas por falhas e defeitos ou são inexistentes em muitos países do mundo, principalmente em países como o Brasil e em grande parte do nosso hemisfério (por exemplo, instituições tomadas pela corrupção). Tais países poderiam estabelecê-las mesmo sem a possibilidade de organizar algumas dessas instituições, e, assim, seria até possível fazer avançar a justiça – ou reduzir a injustiça – a um grau considerável, mas, embora isso possa talvez ser adequado ou suficiente para aplicar a abordagem comparativa da justiça (que é a abordagem de Sen), ainda assim não aconteceria a conquista da justiça transcendental (ou a sociedade plenamente justa) como argumenta Rawls. Se tal justiça sem mácula (a justiça transcendental rawlsiana, tal como explicada em *A ideia de Justiça*) fosse o único foco de atenção em uma teoria da justiça, então as condições institucionais formariam um tipo de “barreira de entrada”, levando-nos a uma impossibilidade de aplicação da teoria da justiça em situações nas quais essas mesmas demandas institucionais são inexistentes e não podem ser estabelecidas em um futuro previsível (em nosso país, parece-me ser o caso da nossa imprevisibilidade de um poder legislativo não corrupto e, ainda, a imprevisibilidade de um ensino fundamental de qualidade para todos).

Sen lembra também que outro problema reside na necessidade de um ponto de vista comum alcançado a partir da posição original, por meio do esquema conceitual do véu da ignorância, também explicitado por Rawls. Esse esquema implica que nossos princípios e convicções tornam-se objetivos na medida em que são aceitos e testados ao se assumir o ponto de vista geral e pela avaliação dos argumentos a favor deles, por meio das restrições expressas pela concepção da posição original (Cf. Rawls, 1991, p. 517; 1999, p. 453)

Em contraposição à justificação negociada, que fica confinada a um único aspecto dos

pontos de vista ou a um consenso sobreposto, usando a expressão de Rawls, existe uma abordagem diferente para a imparcialidade que produz diferentes vozes, possivelmente mesmo a partir de uma distância que o recurso argumentativo do véu da ignorância não dá conta. Isso pode ocorrer precisamente porque essas vozes múltiplas iluminam as decisões públicas e ajudam a torná-las imparciais. Na terminologia da resolução de conflitos que Amartya Sen analisou,<sup>8</sup> isso é mais parecido com a arbitragem do que com a negociação, já que os árbitros não necessitam tomar parte na disputa. Sen afirma que a interpretação da equidade e da imparcialidade, por meio do entendimento de uma “arbitragem equitativa”, é uma tese rival séria à via da “negociação equitativa”, que é a direção exclusiva na qual a característica contratualista da abordagem transcendental de Rawls procede (Cf, Sen, 2002, “Open and Closed Impartiality”). O artigo “What do we want from a theory of justice?” trata amplamente dessa questão, mas apresento apenas breves linhas gerais aqui: existem duas fontes de diferenças substanciais entre os procedimentos de Smith (assumida por Sen) e de Rawls. Primeiro, a abordagem contratualista rawlsiana direciona-se à identificação das demandas da transcendência (o principal questionamento na posição original tem por objetivo as demandas de uma sociedade plenamente justa), enquanto os espectadores imparciais de Smith, tal como interpretados por Sen, são invocados tipicamente para alternativas diferentes com objetivo de esclarecer e elucidar problemas específicos do avanço ou do retrocesso da justiça em uma abordagem comparativa entre sociedades e arranjos políticos. Outra diferença decorre do fato de que os observadores imparciais podem ser imaginados como vindos de longe ou de perto, com perguntas feitas sobre como o problema da decisão pareceria para aqueles que podem ter tido experiências social e institucional de modos diferentes. Enquanto os juízes imparciais ideais podem considerar útil perguntar, por exemplo, quais possibilidades de resolução de conflitos políticos parecem com aquelas que eles encontram quando são partes envolvidas no processo político (aqui o exercício seria bastante semelhante ao exercício rawlsiano do momento do estabelecimento do contrato político), com a perspectiva de Sen, os árbitros poderiam ser considerados como portadores de perspectivas políticas diferentes daquelas geralmente aceitas na sociedade e na cultura na qual o problema político precisa ser resolvido, trazendo pluralidade de valores e de perspectivas ao processo político.

Sobre esse ponto, Sen reitera sua interpretação de Smith em uma palestra no Brasil, em 2012, ao afirmar que precisamos levar em consideração os sentimentos sobre como se comportar e o que uma pessoa tem razões para fazer, ou seja, “sentimentos intermediados por raciocínio, um raciocínio interativo que faça as pessoas avaliarem como a mesma situação

---

8 - Estudo realizado antes da elaboração do artigo “What Do We Want from a Theory of Justice?” e do seu livro principal *A ideia de justiça*.

se apresentaria para outras pessoas distantes”. Para Sen, seria uma abordagem que leva em consideração a “justiça global” na qual constantemente se pede para pensar nas pessoas de outros lugares. Não se trataria apenas de “considerar os efeitos de nossa ação sobre outras pessoas, mas de nos libertarmos do nosso modo de pensar, no qual frequentemente nos enclausuramos”, e, para tanto, Sen retorna a Smith ao afirmar que o melhor caminho para assim agirmos é invocando outras culturas, outras sociedades. Sen exemplifica ampliando o argumento de Smith ao narrar que, enquanto na Grécia Antiga pensava-se ser impossível construir uma sociedade sem infanticídio, na Índia, no mesmo período, o infanticídio era completamente proibido. Assim, percebemos a importância de “olhar como outras pessoas pensam, incorporar outras perspectivas ao nosso raciocínio, para então traduzirmos isso naquilo que devemos fazer”. (p. 23-4)

Possivelmente, se não fossem os textos de Rawls não existiriam pelo menos as obras de Sen mencionadas aqui e tal como estão redigidas: com suas críticas de anos e anos até chegar ao ponto em que ele considera desnecessária uma abordagem transcendental que apresente um esboço teórico de uma sociedade plenamente justa; esforço que remonta a obras como as de Platão e Kant. Sen percebe que, a partir do empirismo de Bentham, Smith, Stuart Mill, entre outros, ele poderá elaborar um trabalho com a profundidade da pesquisa filosófica que caracterizou John Rawls, sem abandonar a aplicabilidade de um trabalho em Ciências Sociais e sua própria marca de pesquisador, especialmente se lembrarmos que, junto com Mahbub ul Haq, também é criador do Índice de Desenvolvimento Humano (IDH).

Amartya Sen nos mostra como o respeito à diversidade e à pluralidade é importante para a remoção da injustiça. Como mencionei, ele afirma que vozes distantes também estão submetidas à disciplina do escrutínio crítico, incluindo a convocação aos espectadores imparciais de longe e de perto. Segundo ele, a disciplina da arbitragem justa no contexto da justiça global também pode ser vista como interativa: vozes de longe e de perto implicam a razão pública interativa e aberta que é crucialmente importante para o entendimento das demandas que constituem os direitos humanos, a despeito das diferenças e práticas dentro de cada país e entre países.

## **Referências Bibliográficas:**

DANIELS, Norman. “John Rawls”. *A Companion to Analytic Philosophy*. Ed. P. Martinich and D. Sosa. Blackwell: Oxford, 2005, p.351-370.

K. C. Tan. "Liberal Equality: What, Where and Why". The Oxford Handbook of American Philosophy. Oxford: OUP, 2008 (p. 516)

NUSSBAUM, M e SEN, A. The Quality of Life, Oxford: Oxford University Press, 1993.

NUSSBAUM, M. "Capabilities as Fundamental Entitlements: Sen and Social Justice," Feminist Economics, 9(2/3), 2003, p. 33–59.

\_\_\_\_\_. Women and Human Development: The Capabilities Approach, Cambridge: Cambridge University Press, 2000 e Nussbaum, M., "Capabilities as Fundamental Entitlements: Sen and Social Justice," Feminist Economics, 9(2/3): 33–59, 2003

SEN, Amartya. A ideia de justiça. São Paulo: Companhia das Letras, 2011.

\_\_\_\_\_. "Capability and Well Being" in NUSSBAUM, M e SEN, A. The Quality of Life, Oxford: Oxford University Press, 1993, p.30-53

\_\_\_\_\_. Commodities and Capabilities. North Holand: Amsterdam, 1985.

\_\_\_\_\_. Desenvolvimento como Liberdade. São Paulo, Companhia das Letras, 2000.

\_\_\_\_\_. Desigualdade Reexaminada. Rio de Janeiro: Record, 2001.

\_\_\_\_\_. "Igualdade de que?" in S.M. Murrin (Ed.) Liberdade, igualdade e direito. Barcelona: Ariel, 1988.

\_\_\_\_\_. "Open and Closed Impartiality", Journal of Philosophy, XCIX, 9, 445-69, Setembro de 2002.

\_\_\_\_\_. "Justiça, esperança e pobreza" - conferência proferida em 25 de abril de 2012. Trad. Luis Felipe Garcia. Pensar a Filosofia. Org. E. Wolf. Porto Alegre: Arquipélago Editorial, 2013, p. 12-33

\_\_\_\_\_. "What Do We Want from a Theory of Justice?." The Journal of Philosophy, Vol. 103, No. 5, May, 2006.

RAWLS, John. A theory of justice. Cambridge, MA: Harvard, 1999.

\_\_\_\_\_. A Theory of Justice. Revised Edition. Cambridge, MA: Harvard, 1999

\_\_\_\_\_. Political Liberalism. New York: Columbia, 1993.

SMITH, Adam. Theory of Moral Sentiments. (Introduction by Amartya Sen) Londres: Penguin, [1790] 2009.



# O que queremos de uma teoria da justiça?\*

*Autor: Amartya Sen  
Harvard University*

*Tradução: Mário Nogueira de Oliveira<sup>1</sup>  
UFOP*

Eu começo a partir da posição geral rawlsiana na qual a interpretação da justiça está vinculada com a razão pública. O enfoque tem que ser, nas palavras de John Rawls, em uma “estrutura pública de pensamento” que propicie “uma noção de acordo nos juízos entre agentes razoáveis.”<sup>2</sup> Rawls ressalta essa exigência para evitar o que ele chama de “uma inclinação pessoal”:

*Não olhamos para a ordem social a partir da nossa situação, mas assumimos um ponto de vista que todos possam adotar em pé de igualdade. Nesse sentido, olha-*

---

\* Pelas úteis discussões, o autor do artigo é grato a Sabina Alkire, Kenneth Arrow, Akeel Bilgrami, Sissela Bok, Joshua Cohen, Peter Hammond, Isaac Levi, Thomas Nagel, Prasanta Pattanaik, Mozaffar Qizilbash, Ingrid Robeyns, Emma Rothschild, Carol Rovane, Maurice Salles, Thomas Scanlon, Patrick Suppes, Kotaro Suzumura, Philippe Van Parijs, e aos participantes dos seminários na Harvard University, Stanford University e Yale University, onde versões anteriores deste artigo foram apresentadas. Ver Rawls, *Political Liberalism*. New York: Columbia, 1993. pp. 110-13.

1 - O tradutor agradece ao Prof. Amartya Sen e *The Journal of Philosophy* pela permissão da tradução em Língua Portuguesa em *Fundamento: Revista de Pesquisa em Filosofia* da UFOP. O artigo foi publicado em inglês em *Journal of Philosophy*, Vol. 103, No. 5 (May, 2006), pp. 215-238. O tradutor agradece às diretoras executivas da revista, Livia Reis e Aline Monteiro, pelas excelentes contribuições ao trabalho.

2 - Ver Rawls, *Political Liberalism*. New York: Columbia, 1993. pp. 110-13.

*mos para a sociedade e para o nosso lugar nela objetivamente; compartilhamos um ponto de vista comum com os outros e não elaboramos nossos juízos a partir de uma inclinação pessoal.*<sup>3</sup>

A sustentação da *razão pública* na teoria da justiça leva a duas questões posteriores: *qual é o público relevante?* e *em quais questões a razão deve se concentrar?* A segunda questão diz respeito à variedade de pontos de vista que deve contar na razão pública (por exemplo, se elas todas têm que vir de dentro de um dado estado político), enquanto a primeira questão relaciona-se ao assunto de que trata a razão pública, em particular quais são as perguntas a serem respondidas para uma teoria da justiça satisfatória? Argumentarei que os dois pontos estão vinculados e que, juntos, levam-nos à questão fundacional: *o que queremos de uma teoria da justiça?*

Comecei valendo-me da orientação de Rawls acerca da conexão básica entre objetividade, razão pública e teoria da justiça. Todavia, tenho que argumentar a favor de um modo diferente de buscar essa conexão, partindo não apenas do conteúdo substantivo da teoria da justiça de Rawls, mas também do diagnóstico de Rawls acerca dos requisitos de uma teoria da justiça, incluindo a questão da razão pública e do alcance e cobertura da participação pública.

## O transcendental *versus* o comparativo

Começemos com a questão de uma teoria da justiça satisfatória. Na sua análise da “justiça como equidade”, Rawls toma como questão principal: *o que é uma sociedade justa?* De fato, na maioria das teorias da justiça na filosofia política contemporânea, essa pergunta é considerada central. Isso leva ao que pode ser chamado de uma abordagem “transcendental” da justiça, uma vez que enfoca a identificação dos arranjos sociais perfeitamente justos. Contrariamente, o que pode ser chamado de uma abordagem “comparativa” estaria concentrado em um ranking<sup>4</sup> alternativo de arranjos sociais (se alguns arranjos são mais ou menos justos que ou-

3 - Rawls, *A Theory of Justice*. Cambridge: Harvard, 1971. pp. 516-17. Corresponde a página 453 da edição revista do livro (Cambridge: Harvard, 1999). Sobre a relação entre objetividade política e razão pública, ver também Bruce A. Ackerman, *Social Justice in the Liberal State*. New Haven: Yale, 1980; Joshua Cohen, “An Epistemic Conception of Democracy”, *Ethics*, XCVII, 26-38, 1986-87; Seyla Benhabib (Ed.), *Democracy and Difference: Contesting the Boundaries of the Political*. Princeton: University Press, 1996; Amy Gutmann e Dennis Thompson, *Democracy and Disagreement*. Cambridge: Harvard, 1996; Thomas Scanlon, *What We Owe to Each Other*. Cambridge: Harvard, 1998.

4 - N. T. A palavra “ranking” já se encontra nos dicionários de Língua Portuguesa, no Brasil, que apresentam como significado a “formação ou listagem (de pessoas, órgãos etc.); classificação ordenada de acordo com critérios determinados”. Cf. *Grande Dicionário Houaiss da Língua Portuguesa*.

tros), ao invés de focar exclusivamente na identificação de uma sociedade completamente justa. As abordagens transcendental e comparativa são muito distintas e, como será discutido aqui, nenhuma das abordagens em geral subsume ou engloba a outra.<sup>5</sup>

A abordagem transcendental da justiça não é nova (pode remontar no mínimo a Thomas Hobbes), mas as contribuições recentes fizeram muito para consolidar a confiança nessa abordagem. Na sua investigação sobre “justiça como equidade”, Rawls explora com profundidade a natureza de uma sociedade inteiramente justa entendida na perspectiva da equidade contratual. A investigação de Rawls inicia-se com a identificação dos requisitos da equidade ao explorar uma imaginada “posição original” na qual os membros da sociedade ignoram suas respectivas características individuais, incluindo suas próprias preferências abrangentes. Os princípios de justiça que emergem na posição original passam a ser considerados imparciais porque são escolhidos pelas pessoas envolvidas sob um “véu da ignorância”, sem conhecimento das suas identidades individuais na sociedade com interesses específicos e prioridades particulares.

Adiante neste artigo, discutirei algumas limitações dessa compreensão dos requisitos da equidade (e perguntarei se os pontos de vista a serem considerados devem todos vir da população de um dado Estado), mas o ponto imediato a notar no contexto do entendimento da abordagem transcendental é que o exercício da equidade tem como objetivo total identificar os princípios apropriados para uma sociedade completamente justa e isolar as necessidades institucionais para a estrutura básica de tal sociedade. Por sua vez, o trabalho dessas instituições conduz a mais decisões sociais nos estágios posteriores no sistema rawlsiano, por exemplo, por meio da legislação apropriada (que Rawls chama de “o estágio legislativo”). A sequência avança passo a passo em linhas especificadas firmemente, com o desdobramento muito elaborado de arranjos sociais justos.

Apesar do uso constante e generalizado da abordagem transcendental, o interesse intelectual e a relevância prática das questões comparativas sobre justiça são difíceis de negar. A investigação dos diferentes modos de avançar a justiça na sociedade (ou no mundo), ou de reduzir injustiças manifestas que possam existir, exige juízos comparativos sobre a justiça, para os quais a identificação de um arranjo social completamente justo não é necessária nem suficiente. Para ilustrar a oposição envolvida, pode acontecer que, em uma perspectiva comparativa, a introdução de políticas sociais que abolem a escravidão, ou que eliminem a fome generalizada, ou que removem o analfabetismo excessivo, pode atestar a produção

---

5 - A Teoria da Escolha Social, iniciada na forma moderna por Kenneth Arrow (*Social Choice and Individual Values*. New York: Wiley, 1951), baseia-se em uma abordagem fundamentalmente comparativa. A prioridade analítica dessa abordagem também pode ser vista nas formulações iniciais de escolha social no século XVIII, particularmente na obra de Condorcet, *Essai sur l'application de l'analyse à la Probabilité des Décisions rendues à la Pluralité des Voix*. Paris: L'Imprimerie Royale, 1785.

de um avanço da justiça. Porém, a implementação de tais políticas poderia ainda deixar as sociedades envolvidas longe dos requisitos transcendentais de uma sociedade completamente justa (uma vez que a transcendência teria outros requisitos em relação às liberdades igualitárias, equidade distributiva, etc.).

A grande divisão entre o “justo” e o “injusto”, que é aquilo que uma teoria transcendental da justiça fornece, deixaria a sociedade no lado “injusto”, mesmo depois de reforma, a despeito do que pode ser visto em uma perspectiva comparativa como uma mudança que reforça a justiça. Alguma articulação não transcendental é claramente necessária. Tomando um outro tipo de exemplo, a instituição de um sistema público de segurança de saúde nos Estados Unidos que não deixe dezenas de milhões de americanos sem qualquer garantia de atendimento médico pode ser julgada como sendo um avanço da justiça, mas tal mudança institucional não tornaria os Estados Unidos uma “sociedade justa” (uma vez que ainda restaria uma centena de outras transgressões a serem remediadas).

Uma abordagem transcendental não pode, por si própria, abordar as questões sobre o avanço da justiça e comparar propostas alternativas para se ter uma sociedade mais justa, faltando a proposta de um salto radical para um mundo perfeitamente justo. De fato, as respostas que uma abordagem transcendental da justiça dá – ou pode dar – são muito distintas e distantes do tipo de interesses que engajam as pessoas em discussões sobre a justiça e a injustiça no mundo, por exemplo, iniquidades como fome, analfabetismo, tortura, prisão arbitrária ou falta de assistência médica, como características sociais particulares que necessitam ser remediadas. O foco desses engajamentos tende a ser os modos e meios de fazer avançar a justiça – ou reduzir a injustiça – no mundo, remediando essas iniquidades, ao invés de apenas procurar a satisfação simultânea de um conjunto inteiro de arranjos sociais perfeitamente justos, requisitados por uma teoria transcendental particular.

## **A defesa possível de uma abordagem transcendental**

A argumentação tem sido muito fácil até aqui em uma acepção importante. Certamente, as respostas transcendentais não podem ser tudo o que queremos de uma teoria da justiça. Mas poderá existir – esse é um assunto a ser investigado – alguma conexão menos óbvia, alguma relação entre o transcendental e o comparativo, que poderia tornar a abordagem transcendental o caminho certo de proceder nas avaliações comparativas. O afastamento formal da abordagem transcendental em relação à invocação da ideia de justiça nos debates e discussões sobre assuntos práticos não indica, por si mesmo, que a abordagem transcen-

dental não possa ser a abordagem correta.

Assim, pelo menos mais duas questões têm de ser tratadas com relação à possibilidade, respectivamente, (1) da suficiência e (2) da necessidade da abordagem transcendental para se elaborar juízos comparativos sobre a justiça. Primeiro, as respostas para os questionamentos transcendentais podem também levar-nos, indiretamente, às avaliações comparativas da justiça, em particular por meio de comparações entre as “distâncias” do transcendente e a posição em que os conjuntos particulares de arranjos sociais encontram-se respectivamente? Segundo, pode ser o caso que a pergunta transcendental (“o que é uma sociedade justa?”) tenha de ser respondida primeiro, como um requisito essencial para que uma teoria da justiça comparativa seja cogente e bem fundada, que, de outro modo, a sua fundação seria disjuntiva e frágil? Para uma crítica adequada da abordagem transcendental da justiça, a partir da perspectiva comparativa, temos que avaliar essas possibilidades criticamente.

As crenças implícitas na suficiência ou necessidade (ou em ambas) de uma abordagem transcendental para a avaliação comparativa claramente desempenharam um papel poderoso na crença generalizada de que a abordagem transcendental é crucial para toda a teoria da justiça. De fato, mesmo em uma teoria da escolha social, onde a estrutura analítica é firmemente relacional e, ao mesmo tempo, baseada em comparações entre pares, as investigações da justiça, em particular, têm sido repetidamente ampliadas a fim de que saiam da base comparativa para chegar a uma identificação da justiça transcendental (frequentemente no molde rawlsiano).<sup>6</sup> Ao argumentar a favor de uma abordagem comparativa mais forte para a justiça, com a qual este artigo está interessado (e para qual sugiro que a teoria da escolha social pode desempenhar um importante papel), seria necessário examinar se as conclusões comparativas *seguem-se de* ou *precisam de* alguma identificação transcendental.

## A especificação transcendental fornece rankings comparativos?

Início com a questão da *suficiência*. A abordagem transcendental produz, como um sub-

---

6 - As conclusões comparativas têm sido tipicamente reunidas para chegar a algumas reivindicações definitivamente transcendentais (uma tradução que é possível com pressupostos bastante restritivos e limitantes, como será discutida aqui). Veja, por exemplo, Sen, *Collective Choice and Social Welfare*. San Francisco: Holden-Day, 1970; Kotaro Suzumura, *Rational Choice, Collective Decisions, and Social Welfare*. New York: Cambridge, 1983; Peter J. Hammond, “Equity, Arrow’s Conditions and Rawls’ Difference Principle”, *Econometrica*, XLIV, 793-804, 1976; Claude d’Aspremont e Gevers Louis, “Equity and the Informational Basis of Collective Choice”, *Review of Economic Studies*, XLVI, 199-209, 1977; Arrow, “Extended Sympathy and the Possibility of Social Choice” *American Economic Review*, LXVII, 219-25, 1977. A influência da perspectiva rawlsiana sobre a justiça tem sido muito forte na aplicação da Teoria da Escolha Social, apesar da distância fundamental entre as respectivas abordagens comparativa e transcendental.

produto, conclusões relacionais que estão prontas para serem classificadas, de modo que a transcendência termine por nos propiciar muito mais do que as suas formas manifestas articulam? Em particular, a especificação de uma sociedade totalmente justa é suficiente para nos dar rankings de pontos de partida de justiça em termos de “distâncias” comparativas em relação à perfeição, a fim de que uma identificação transcendental possa, também, vincular imediatamente as graduações comparativas?

Aqui a resposta é um firme não. A principal dificuldade repousa no fato de que existem características diferentes envolvidas na identificação de distância relacionadas, entre outras distinções, com (1) diferentes pontos de partida, (2) variação nas dimensões das transgressões dentro do mesmo ponto de partida geral e (3) modos diversos de comparar as infrações individuais. A identificação da transcendência não fornece qualquer meio de abordar esses problemas para se chegar a um ranking que relacione os desvios em relação à transcendência.

Por exemplo, no contexto da análise rawlsiana da sociedade justa, os desvios podem ocorrer em muitos espaços diferentes. Podem incluir o rompimento da liberdade que, ademais, pode envolver diversas violações de liberdades distintas (muitas das quais figuram na ampla cobertura da liberdade e sua prioridade em Rawls). Podem, também, existir violações – novamente, em formas possivelmente discrepantes – dos requisitos da equidade na distribuição dos bens primários (podem existir desvios diferentes das exigências do Princípio da Diferença que forma uma parte do segundo princípio de Rawls). De modo similar, diversas transgressões podem ocorrer em outras teorias transcendentais da justiça (por exemplo, aquelas que substituiriam o enfoque rawlsiano nos “bens primários” no Princípio da Diferença por “capacidades”, “recursos” ou “oportunidades” ou mesmo outro modo de formular as necessidades de distribuição e alocação da justiça transcendental).

Existem também modos diversos de avaliar a extensão de cada discrepância e de avaliar o afastamento comparativo das distribuições reais daquilo que os princípios da justiça plena exigiriam. Além disso, temos que considerar os desvios na equidade procedimental (tais como as infrações da justa igualdade de meios e oportunidades públicas), que figuram dentro do domínio dos requisitos rawlsianos da justiça (na primeira parte do segundo princípio). Mensurar esses desvios procedimentais *contra* impropriedades dos padrões emergentes de distribuição interpessoal (por exemplo, distribuições de bens primários), que também figuram no sistema rawlsiano, exigiria especificação distinta – possivelmente em termos axiomáticos – de importância ou significância relativa (ou *trade-offs*<sup>7</sup> como

---

7 - **N.T. Trade-off** é uma expressão usada em Economia que qualifica situações de escolha entre opções que estão em oposição. Por exemplo, quando elaboradores de determinada política pública encontram-se em uma situação na qual necessitam optar uma das alternativas à disposição abdicando das demais, trata-se de um *trade-off*.

são chamados às vezes no vocabulário de avaliação multidimensional). Mas essas extensões, embora muito úteis, repousam muito além do exercício específico de identificação da transcendência, e são, de fato, os ingredientes básicos de uma abordagem “comparativa” da justiça, ao invés de uma abordagem “transcendental”. A caracterização de uma justiça sem mácula não traz consigo qualquer delineamento sobre como os diversos desvios da perfeição podem ser comparados e ranqueados.

Certamente, a falta de tais desdobramentos comparativos não é um problema para a teoria transcendental da justiça vista como uma conquista autônoma. O silêncio quanto às relações não é em qualquer sentido uma dificuldade interna de uma teoria da justiça transcendental. De fato, alguns transcendentalistas puros seriam totalmente opostos, mesmo ao flerte com as gradações e as avaliações comparativas, e podem se esquivar muito provavelmente das conclusões advindas dos confrontos de relações. Eles podem apontar, em particular, o seu entendimento de que um arranjo social “correto” não deve, de forma alguma, ser entendido como um arranjo social melhor, o que poderia abrir a porta para o que, às vezes, é visto como o mundo intelectualmente piegas de avaliações gradativas na forma de “melhor” ou “pior” (vinculado com o superlativo relacional “melhor”). O estado absoluto de o “correto” transcendental – contra expressões relacionais como “melhor que” e “o melhor” – pode ter uma posição própria fortemente fundamentada, mas isso não ajuda em nada nas avaliações comparativas de justiça.

Sem dúvida, os membros de qualquer comunidade politicamente organizada podem contemplar como uma reorganização enorme e totalmente abrangente pode ser realizada, movendo-nos *de uma só vez* para o ideal de uma sociedade plenamente justa. Uma teoria transcendental que não seja contrária ao bom senso pode servir, nesse sentido, como algo tal qual “o manual completo do grande revolucionário”. Mas esse manual não seria muito invocado nos debates sobre a justiça nos quais constantemente participamos e que enfocam as múltiplas maneiras para reduzir as injustiças que caracterizam o mundo.<sup>8</sup>

Mesmo se pensamos na transcendência fora dos termos “sem classificação” dos arranjos sociais “certos”, mas nos termos de classificação gradual dos arranjos sociais “melhores”, a identificação do melhor, em si mesma, não nos diz muito sobre a classificação total, tal como a maneira adequada para comparar duas alternativas que não são as melhores. A identificação do que seria o melhor não especifica um ranking único com respeito ao qual o melhor situa-se no cume. Considerando uma analogia, o fato de uma pessoa considerar a *Monalisa*

---

8 - É importante notar aqui que o diagnóstico de injustiça não exige uma identificação única da “sociedade justa”, uma vez que muitas identificações diferentes de arranjos sociais perfeitamente justos podem concordar com o diagnóstico de uma deficiência sanável de um arranjo social particular (por exemplo, como no caso de fome comprovada, analfabetismo ou negligência médica).

como a melhor pintura do mundo, não revela como ela classificaria um Gauguin em relação a um Van Gogh. A busca pela justiça transcendental é um exercício envolvido em si mesmo, mas independente de pensarmos a transcendência em termos de o “certo” sem graduação ou se pensamos na transcendência em uma estrutura de o “melhor” graduado, ela não nos diz muito sobre os méritos comparativos de muitos – de fato, tipicamente a maioria – dos diferentes arranjos sociais.

## **Uma teoria transcendental é necessária para as comparações de justiça?**

Dedico-me agora à segunda questão, sobre a hipótese de que a identificação do melhor é necessária, mesmo se não for suficiente, para graduar duas alternativas quaisquer em termos de justiça. No sentido usual de necessidade, esta seria uma possibilidade de alguma forma estranha. Na disciplina dos juízos comparativos em qualquer campo, as avaliações que relacionam duas alternativas em geral tendem a ser uma questão entre elas, sem existir a necessidade de pedir a ajuda de uma terceira alternativa “irrelevante”. Realmente, não é totalmente óbvia a razão pela qual ao fazermos o julgamento de que algum arranjo social  $x$  seja melhor do que o arranjo alternativo  $y$ , tenhamos que invocar a identificação de que alguma alternativa muito diferente  $z$  seja o arranjo social “melhor” ou “certo”. Ao defendermos um Picasso em detrimento de um Dali, não precisamos nos preocupar com a identificação da pintura mais perfeita do mundo, que superaria os Picassos e os Dalis e todas as outras pinturas no mundo.

No entanto, é possível pensar que a analogia com a Estética é problemática, uma vez que uma pessoa não poderia ter a ideia de uma pintura perfeita, tal como a ideia de uma sociedade perfeitamente justa aparece como identificável em teorias transcendentais de justiça. Argumentarei, depois, que a existência de uma alternativa melhor, ou transcendente, não está de fato garantida mesmo no campo da justiça. Porém, estou pronto para prosseguir, momentaneamente, na suposição de que tal identificação pode, de alguma forma, ser feita. No entanto, apesar dessa aceitação preliminar, a existência de uma alternativa totalmente identificável, ou uma alternativa melhor, não indica que ela seja necessária (ou mesmo útil) para se referir a ela quando julgamos os méritos relativos de duas outras alternativas. Por exemplo, podemos desejar aceitar de fato, com grande convicção, que o Everest é a montanha mais alta do mundo, completamente imbatível quanto à altura por qualquer outro pico; entretanto, esse entendimento não é necessário nem particularmente útil na comparação de

alturas entre, por exemplo, o Kanchenjunga e o Mont Blanc. Existiria algo profundamente estranho em uma crença geral de que uma comparação entre duas alternativas não pode ser sensatamente feita sem a identificação anterior de uma alternativa suprema.

Assim, a hipótese da necessidade em um sentido padrão seria difícil de sustentar. Contudo, existe uma forma mais fraca da hipótese da necessidade, que meramente assegura que, se avaliações comparativas podem ser sistematicamente feitas, então, essa disciplina também deve ser capaz de identificar a melhor. Nesse caso, a reivindicação não seria tanto que duas alternativas não podem ser comparadas quanto à justiça sem primeiro saber-se qual é a alternativa melhor ou perfeita, mas que o ranking comparativo das diferentes alternativas deve *inter alia* também ser capaz de identificar a resposta para a questão transcendental sobre a sociedade perfeitamente justa. Ou, em outras palavras, se a questão transcendental não pode ser respondida, então, tampouco pode ser respondida a questão comparativa. Esse entendimento da necessidade não defende o caráter imprescindível de seguir *via* a abordagem transcendental para a avaliação comparativa, mas isso daria, pelo menos à identificação transcendental, uma presença necessária na teoria da justiça. Temos, igualmente, que examinar essa afirmação consideravelmente mais fraca de “necessidade”.

## Comparativos sem transcendência

Uma sequência de comparações de pares de alternativas nos levaria invariavelmente à melhor alternativa? Essa pressuposição possui algum atrativo, uma vez que o superlativo poderia de fato parecer ser o ponto final natural de uma comparação forte. Mas essa conclusão, em geral, é um *non sequitur*. De fato, é apenas com um ranking “bem ordenado” (por exemplo, uma ordenação completa e transitiva de um conjunto finito) que podemos ter certeza de que o conjunto de comparações de pares deve, também, identificar uma alternativa “melhor”.

Portanto, devemos perguntar: quão completa deve ser a avaliação para que ela seja uma disciplina sistemática? Na abordagem “totalista”, que caracteriza as teorias de justiça padrão (incluindo a teoria de Rawls), a incompletude tende a aparecer como uma falha, ou, pelo menos, como um sinal da natureza incompleta do exercício. Realmente, a sobrevivência da incompletude é algo visto como um defeito de uma teoria da justiça, o que traz a questão das afirmações positivas que tal teoria elabora. Entretanto, uma teoria da justiça que dá espaço de modo sistemático para a incompletude permite-nos chegar a julgamentos possivelmente muito fortes (por exemplo, sobre a injustiça da fome continuada em um mundo de

prosperidade ou sobre a persistente subjugação grotesca das mulheres), sem ter que encontrar uma avaliação altamente diferenciada para cada arranjo político e social na comparação com todos os outros arranjos (por exemplo, abordando questões como: a taxa de imposto máximo de 45% é mais ou menos justa que a taxa máxima de 46%?).

Em outro lugar, eu argumentei por que uma teoria sistemática e disciplinada da avaliação normativa, incluindo a avaliação da justiça social, não precisa assumir uma forma “totalista”<sup>9</sup>. A incompletude pode ser do tipo permanente por várias razões diferentes, incluindo as lacunas intransponíveis em informações e a insolubilidade de juízos envolvendo considerações discrepantes. Por exemplo, pode ser difícil resolver o equilíbrio geral das demandas comparativas de considerações de equidade que repousam por trás do *maximin* lexicográfico rawlsiano, comparado com, digamos, os *sum-rankings*<sup>10</sup> em uma forma ajustada para a equidade<sup>11</sup>. Ainda assim, apesar de tal incompletude durável, podemos ainda ser capazes de concordar prontamente que existe uma clara injustiça social envolvida na persistência da fome endêmica ou na exclusão do acesso aos cuidados médicos, o que exige um atenuante bem específico para o avanço da justiça (ou redução da injustiça), mesmo depois de levar em consideração os custos envolvidos. De modo similar, podemos reconhecer a possibilidade de que as liberdades de diferentes pessoas podem, até certo ponto, estar em conflito umas com as outras (de modo que qualquer ajuste fino das demandas por igual liberdade pode ser difícil de se realizar) e, ainda assim, concordamos fortemente que torturar pessoas acusadas seria uma violação injusta da liberdade e que essa injustiça exige uma retificação urgente.

9 - Sen, “Consequential Evaluation and Practical Reason”, *Journal of Philosophy*, XC VII, 9, 477-572, Setembro de 2000. Veja também Sen, *Collective Choice and Social Welfare*; “Maximization and the Act of Choice”, *Econometrica*, LXV, 745-79, 1997; e “The Possibility of Social Choice” *American Economic Review*, LXXXIX, 349-78, 1999; e também “The Possibility of Social Choice”, *Synthese*, CXL, 43-59, 2004. Veja também a resposta de Isaac Levi para o último artigo, em “Amartya Sen”, no mesmo número de *Synthese*, pp 61-67.

10 - N.T.: Definição de sum-ranking pelo próprio A. Sen: um conjunto de vantagens individuais é, pelo menos, tão bom quanto outro se, e somente se, ele possui, pelo menos, uma soma total de vantagens tão grande quanto o outro. “Utilitarianism and Welfarism” em *The Journal of Philosophy*, Vol. 76, No. 9 pp. 463-489, 1979.

11 - A vasta literatura sobre este tema inclui, entre outras contribuições, S.-Ch. Kolm, “The Optimum Production of Social Justice”, em J. Margolis e H. Guitton, (Eds.) *Public Economics*. London: Macmillan, 1969. pp. 145-200; A.B. Atkinson, “On the Measurement of Inequality”, *Journal of Economic Theory*, II, 244-63, 1970; James Mirrlees, “An Exploration of the Theory of Optimal Income Taxation”, *Review of Economic Studies*, XXXVIII, 175-208, 1971; Sen, *On Economic Inequality*. New York: Oxford, 1973 (edição ampliada com um novo adendo redigido em conjunto com James Foster, 1997); Claude d’Aspremont e Gevers Louis, “Equity and the Informational Basis of Collective Choice”, XLIV, 199-209, 1977; Eric Maskin, “Decision-making under Ignorance with Implications for Social Choice”, *Theory and Decision*, XI, 319-37, 1979; Kevin W.S. Roberts, “Interpersonal Comparability and Social Choice Theory”, *Review of Economic Studies*, XLVII, 421-39, 1980; Charles Blackorby, David Donaldson, e John Weymark, “Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction”, *International Economic Review*, XXV, 327-56, 1984; d’Aspremont, “Axioms for Social Welfare Ordering”, em Leonid Hurwicz, Schmeidler David e Hugo Sonnenschein (Eds.), *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*. New York: Cambridge, 1985. pp 19-76; e d’Aspremont e Gevers, “Social Welfare Functionals and Interpersonal Comparability”, Blackorby, Walter Bossert, e Donaldson, “Utilitarianism and the Theory of Justice”, e Bhaskar Dutta, “Inequality, Poverty and Welfare”, os três últimos em Arrow, Sen, e Suzumura (eds.), *Handbook of Social Choice and Welfare*. Amsterdam: North-Holland, 2002. Capítulos 10-12, Vol. 1.

Existe uma consideração adicional que pode funcionar efetivamente na abertura de espaço político para a incompletude de decisões sobre a justiça social, mesmo se for o caso que toda pessoa possua um poder completo de decidir nos arranjos sociais possíveis. Uma vez que uma teoria da justiça implica o acordo entre diferentes partes (por exemplo, na “posição original” na abordagem rawlsiana), a incompletude também pode vir a partir da possibilidade de que diferentes pessoas possam continuar a ter algumas diferenças (de modo consistente com o acordo em muitos dos juízos comparativos). Mesmo depois que interesses já assentados e prioridades pessoais tenham sido, de alguma forma, “retirados” de consideração por meio de recursos como o “véu da ignorância”, podem restar visões possivelmente conflitantes sobre as prioridades sociais, por exemplo, ao mensurar as reivindicações de necessidade sobre o direito aos frutos do trabalho de alguém.

Os conflitos dos princípios distributivos que são difíceis de erradicar podem ser ilustrados com um exemplo, que discuti em outro contexto. O exemplo está relacionado com o problema de se decidir qual entre três crianças deve ganhar uma flauta que disputam. A criança A é a única das três que sabe tocar flauta (as outras não negam isso); a criança B é a única que não possui nenhum brinquedo (as outras duas concedem que são muito mais ricas e com boas condições para diversões); a criança C trabalhou duro para construir a flauta sozinha (as outras confirmam isso). Teóricos de diferentes perspectivas – utilitarianistas, igualitaristas ou libertarianistas – podem acreditar que uma resolução justa pode ser prontamente percebida aqui, embora eles vissem, respectivamente, resoluções totalmente diferentes como sendo exatamente certas. O objetivo principal a notar no contexto presente é que todas as resoluções diferentes possuem argumentos sérios a apoiá-los, e podemos não ser capazes de identificar exatamente um dos argumentos alternativos como sendo o único (para invocar o critério de Thomas Scanlon) que “poderia ser justificado aos outros com bases que, se apropriadamente motivados, eles não poderiam rejeitar razoavelmente”<sup>12</sup>.

Mesmo quando cada parte envolvida possui a sua própria especificação completa de justiça, a “intersecção” entre os rankings – que é as crenças *compartilhadas* das diferentes partes – pode fornecer um ranking parcial, se os juízos não forem todos congruentes<sup>13</sup>. A aceitabilidade da incompletude avaliativa é realmente um assunto central na escolha social em geral, e também é relevante para teóricos da justiça, ainda que teóricos rawlsianos e outros afirmem (e essa é uma afirmação e não algo que é estabelecido em um percurso livre qualquer) que um acordo total emergirá definitivamente na “posição original” e em outros

---

12 - Scanlon, *What We Owe to Each Other*, p. 5; ver também seu artigo “Contractualism and Utilitarianism,” em Sen e Bernard Williams (Eds.), *Utilitarianism and Beyond*. New York: Cambridge, 1982. pp. 103-28.

13 - As características formais dos “ordenamentos parciais de intersecção” são discutidas em Sen, *On Economic Inequality* e “Maximization and the Act of Choice”.

formatos semelhantes<sup>14</sup>.

De fato, por razões tanto de avaliações individuais incompletas e de congruência incompletas das avaliações individuais, incompletude pode ser uma característica forte dos juízos de justiça social. Isso pode ser problemático para a identificação de uma sociedade perfeitamente justa e torna as conclusões transcendentais difíceis de derivar<sup>15</sup>. Ainda assim, tal incompletude não evitaria elaborar juízos comparativos de justiça em muitos casos, em que poderia existir um acordo justo em rankings particulares em relação a pares, sobre como fortalecer a justiça e reduzir a injustiça. Uma ordem parcial pode ser muito útil sem ser capaz de levar a qualquer identificação transcendental de uma sociedade plenamente justa.

A pergunta “o que é uma sociedade justa?” não é, portanto, um bom *ponto de partida* para uma teoria da justiça útil. Para isso, deve ser acrescentada a conclusão adicional de que tampouco pode ser um *ponto final* plausível. Uma teoria da justiça comparativa não precisa, nem tampouco necessariamente fornece, uma resposta para a pergunta “o que é uma sociedade justa?”.

## Barreiras institucionais e silêncio transcendental

Dedico-me agora a uma característica diferente – embora também relacionada com as anteriores – da abordagem da justiça transcendental, em particular os requisitos institucionais extremamente exigentes da conquista da justiça sem máculas. A conquista de uma sociedade plenamente justa exigiria uma superabundância de instituições, incluindo as operações irrestritas de um Estado soberano. Algumas dessas instituições são defeituosas ou inexistem em muitos países do mundo; nem tais países podem estabelecê-las prontamente. Mesmo sem a possibilidade de organizar algumas dessas instituições, é, de fato, possível fazer avançar a justiça – ou reduzir a injustiça – a um grau considerável, mas, embora isso seja suficiente para aplicar a abordagem comparativa da justiça, não se propicia aí a conquista

14 - Para uma primeira expressão de ceticismo sobre a plausibilidade das decisões unânimes na “posição original”, ver meu ensaio em conjunto com W.G. Runciman, “Games, Justice and the General Will” *Mind*, LXXIV, 554-62, Setembro de 1965.

15 - Do ponto de vista matemático, deve-se reconhecer que um ordenamento transitivo, embora incompleto, sobre um conjunto finito, invariavelmente fornecerá um ou mais elementos “maximais”, no sentido de não haver qualquer elemento que seja melhor que um elemento maximal. No entanto, a maximalidade não garante a existência de um melhor elemento. A natureza fundacional da distinção matemática envolvida e suas implicações significativas são investigadas em N. Bourbaki, *General Topology*, partes I e II, versão em Inglês (Reading, MA: Addison-Wesley, 1966), e *Theory of Sets*. Reading, MA: Addison-Wesley, 1968. A presença de elementos maximais é suficiente para a escolha razoável de uma alternativa que não seja pior do que nenhuma outra. Contudo, ela não permitirá, em geral, o desdobramento de um arranjo social perfeitamente justo, nem mesmo permitirá um arranjo único de justiça perfeita (como no sistema rawlsiano e outros sistemas transcendentais). Sobre a relevância de longo alcance da distinção entre maximalidade e otimalidade (e transcendência), ver Sen, “Internal Consistency of Choice”, *Econometrica*, LXI, 495-521, 1993; e “Maximization and the Act of Choice”, *Econometrica*, LXXV, 745-79, 1997.

da justiça transcendental. Se tal justiça sem mácula fosse o único foco de atenção em uma teoria da justiça, então as condições institucionais formariam um tipo de “barreira de entrada”, levando-nos a uma falta de aplicação da teoria da justiça em situações nas quais essas mesmas demandas institucionais são inexistentes e não podem ser estabelecidas em um futuro previsível.

As condições institucionais seriam particularmente difíceis de estabelecer ao lidarmos com, digamos, problemas de justiça global. A alegação de que precisamos de um Estado soberano para aplicar os princípios de justiça – uma reivindicação que foi bem articulada por Thomas Hobbes – está substancialmente conectada com as elaboradas demandas institucionais de um entendimento transcendental de justiça. A destituição da relevância da “ideia de justiça global”, fortemente sustentada por Thomas Nagel, vale-se do seu entendimento de que essas demandas institucionais extensas não podem ser realizadas em nível global nesse tempo. Como ele diz, “parece-me muito difícil resistir à alegação de Hobbes sobre a relação entre justiça e soberania” e “se Hobbes está certo, a ideia de justiça global sem um governo mundial é uma quimera”<sup>16</sup>. No contexto global, Nagel concentra-se em clarificar outras demandas distinguíveis das demandas da justiça, tal como a “moralidade humanitária mínima” (que “governa nossas relações com todas as outras pessoas”), e também distinguível de estratégias de longo prazo para mudanças radicais nas possibilidades institucionais (“acredito que o caminho mais provável para alguma versão de justiça global seja por meio da criação de estruturas de poder notoriamente injustas e ilegítimas que são toleráveis com os interesses das nações-estados atuais mais poderosas”)<sup>17</sup>.

Também na abordagem rawlsiana, a aplicação de uma teoria de justiça exige uma quantidade de instituições que determinam a estrutura básica de uma sociedade plenamente justa. De modo não surpreendente, Rawls, de fato, abandona os seus próprios princípios da justiça quando se trata de avaliar como pensar sobre a justiça global. Em uma contribuição posterior, *O direito dos povos*, Rawls invoca uma “segunda posição original”, com uma negociação justa que envolve representantes de diferentes arranjos políticos – ou diferentes “povos”, como Rawls os chama – que funcionam como partes sob esse segundo véu da ignorância<sup>18</sup>. No entanto, Rawls não tenta derivar princípios de justiça que possam emanar dessa segunda posição original e, ao invés disso, concentra-se em certos princípios gerais do comportamento humanitário. O veículo rawlsiano de justiça que nos levaria rapidamente

---

16 - Nagel, “The Problem of Global Justice”, *Philosophy and Public Affairs*, XXIII, 113-47, 2005; *Journal of Philosophy*, Vol. 103, No. 5, p.115, Maio de 2006.

17 - Nagel, “The Problem of Global Justice”, pp. 130-33, 146-47. Uma crítica importante ao que Joshua Cohen e Charles Sabel descrevem como o “estatismo” de Nagel pode ser encontrada no artigo de ambos, “Extra Rempublicam Nulla Justitia?”, *Philosophy and Public Affairs*, XXXIV, 147-75, Primavera de 2006. Nessa mesma edição, ver também A.L. Julius, “Nagel’s Atlas”, pp. 176-92.

18 - Rawls, *The Law of Peoples*. Cambridge: Harvard, 1999.

adiante na busca de alguma justiça dentro de um mundo sem justiça permanece enguiçado e estacionado em uma manhã gelada de um mundo sem um Estado global.

Decerto, Rawls não precisa concordar que o mundo seja realmente injusto se ele permanece plenamente vinculado à crença de que o conceito de justiça não se aplica ao nível global. Por outro lado, Nagel parece definitivamente convencido que “não vivemos em um mundo justo” (a sentença de abertura de “O problema da justiça global” (*op.cit.*, p.113)). Enquanto eu concordo firmemente com essa conclusão (dado o que eu vejo como o papel de uma teoria da justiça, o tema deste artigo), não fica inteiramente claro para mim como Nagel pode estabelecer argumentos desse tipo, dada sua convicção de que a ideia de uma justiça global é “uma quimera”.<sup>19</sup>

Então, o desafio de tornar o mundo menos injusto permanece sem ser tratado dentro de uma abordagem transcendental. Contudo, não fica totalmente claro por que devemos nos recolher ao silêncio, no que diz respeito à justiça, meramente porque o alcance das possibilidades institucionais não nos prepara para uma justiça transcendental. Apesar disso, a pergunta sobre como a justiça global pode ser impulsionada mantém-se pertinente, a não ser que nós a removamos à força do território da justiça, com base em que a justiça transcendental seja a única ideia satisfatória de justiça – ou a única ideia compreensível. De fato, essa pergunta sobre o progresso da justiça (ou sobre a redução da injustiça global) pode ser uma parte fecunda do tema da abordagem deliberativa da razão pública. Uma “abordagem pública do pensamento”, que Rawls nos ensinou a valorizar e a usar, não deve tornar-se completamente inoperante somente porque as demandas institucionais de uma sociedade perfeitamente justa demonstraram-se impraticáveis.

Nesse caso, existe uma tensão real entre fazer um bom uso da razão pública (aproximando os argumentos gerais poderosos de Rawls nessa direção) e manter o silêncio quando algumas condições institucionais exigentes que são necessárias para a justiça transcendental não possam ser inteiramente satisfeitas (uma conclusão que emerge da forma transcendental especial que Rawls atribui à sua teoria da justiça). Há, aqui, uma séria perda. A deliberação pública pode ser importante e útil, tanto na busca da reforma institucional (mesmo quando a totalidade das reformas institucionais necessárias para a perfeição não pode ser executada) quanto ao examinar o que pode ser feito para reduzir a injustiça – uma questão básica na abordagem comparativa da justiça –, mesmo quando nem todas as instituições certas estão

---

19 - Se Nagel está contando com uma possível interpretação formal (ao prolongar um argumento) de que, na falta de uma teoria da justiça aplicável, pode de fato ser dito que “nós não vivemos em um mundo justo”, então Nagel poderia, com a mesma facilidade, dizer também “nós não vivemos em um mundo injusto”. Todavia, isso não pode ter servido tão bem para a sentença de abertura do artigo refinadamente argumentado por Nagel, e eu diria, existe razão para perguntarmos a nós próprios como essa assimetria ocorre se a ideia de justiça global é realmente uma quimera.

organizadas.<sup>20</sup> Uma abordagem comparativa não rawlsiana (ao invés de uma abordagem transcendental) para a justiça pode ser uma boa base conceitual para tal razão pública – essencialmente rawlsiana.

## O localismo transcendental e interesses globais<sup>21</sup>

A importância da razão pública para lidar com problemas globais de justiça é um tema relevante por si próprio. Todavia, deve ser reconhecido que, embora Rawls fosse um líder visionário do pensamento sobre a importância da razão pública, ele foi um cético considerável acerca do uso da razão pública no nível global. É importante distinguir dois fundamentos possíveis para a relutância de Rawls. Um problema, que já foi mencionado, diz respeito à inaplicabilidade da abordagem bastante exigente da justiça transcendental no nível global por causa das limitações institucionais. Uma segunda razão para a relutância de Rawls é sua insistência em vincular a razão pública com o formato contratualista da “posição original”. Isso envolve um exercício deliberativo planejado que pareceria ser difícil de aplicar além dos limites de uma sociedade particular (ou dos limites de um “povo” particular, como Rawls define essa coletividade em seus trabalhos posteriores).<sup>22</sup>

A afirmação de Rawls sobre a necessidade de um “ponto de vista comum” que foi citado (de *Uma teoria da justiça*) no início deste artigo é seguida imediatamente pela invocação desse esquema conceitual particular:

Assim, nossos princípios e convicções morais são objetivos na medida em que são aceitos e testados ao se assumir esse ponto de vista geral e por avaliar os argumentos a favor deles, por meio das restrições expressas pela concepção da posição original.<sup>23</sup>

Desse modo, a deliberação toma forma de uma negociação equitativa na qual a equidade da negociação razoável baseia-se na exigência de que o raciocínio aconteça sob o véu da ignorância especialmente concebido. Mas os participantes na deliberação são exatamente as partes do contrato social para a sociedade em questão. A voz de uma pessoa conta porque

---

20 - Sobre este assunto, ver meu artigo “Elements of a Theory of Human Rights”, *Philosophy and Public Affairs*, XXXIV, 315-56, Outono de 2004.

21 - N.T. O termo ‘localismo’ é comumente usado no Português do Brasil como bairrismo ou regionalismo.

22 - Entretanto, é importante notar que, na versão do “contratualismo” que Scanlon desenvolve e investiga, que exige um juízo sobre o que “seria razoável para aqueles afetados por um princípio a ser rejeitado”, o confinamento do cidadão não é necessário nem sensato, uma vez que as políticas públicas e instituições de uma nação, ou Estado, ou “povo”, como Rawls chama essa coletividade, podem afetar de modo significativo outros povos em outros lugares. Ver Scanlon em *What We Owe to Each Other*, e também a sua contribuição anterior, “Contractualism and Utilitarianism”.

23 - Rawls, *A Theory of Justice* (1971), pp. 517; e (1999), p. 453; itálico acrescentado por mim.

ela está diretamente envolvida no contrato social que “regulará as instituições” da sociedade da qual ela é um membro.<sup>24</sup>

Em contraposição à justificação negociável confinada a pontos de vista locais, existe uma abordagem diferente para a imparcialidade, que produz diferentes vozes, possivelmente mesmo a partir de “uma distância” (para usar a inteligente frase de Adam Smith), precisamente porque essas vozes iluminam as decisões públicas e ajudam a torná-las imparciais. Na terminologia da resolução de conflitos, isso é mais parecido com a arbitragem que com a negociação; os árbitros não necessitam ser partes da disputa. Argumentei em outro lugar que a interpretação da equidade e da imparcialidade por meio do entendimento de uma “arbitragem equitativa” é um rival sério para a via da “negociação equitativa”, que é a direção exclusiva na qual a característica contratualista da abordagem transcendental de Rawls – Justiça como equidade – procede.<sup>25</sup>

A abordagem da arbitragem equitativa está bem exemplificada pela invocação das perspectivas dos “espectadores imparciais” de Smith. Os espectadores imparciais são observadores ideais que não precisam ser membros da sociedade e suas imparcialidades não são advindas, como no sistema rawlsiano, exclusivamente – ou mesmo primariamente – por meio do experimento pensado de um véu de ignorância acerca das circunstâncias pessoais dos membros individuais de uma dada sociedade. Ao invés disso, o experimento pensado pelos membros da sociedade no sistema de arbitragem equitativa de Smith invoca os juízos dos observadores desinteressados que não são, eles próprios, partes das decisões sociais que devem ser tomadas.

Em si mesmo, isso pode não parecer uma grande diferença já que ambos são experimentos meramente pensados que devem ser empreendidos dentro dos respectivos formatos pelas pessoas da sociedade real. Também não há qualquer coisa que evite os árbitros equitativos ideais de empreenderem o exercício de se colocarem na posição das partes envolvidas sob o planejado véu da ignorância, de modo que a arbitragem equitativa possa fazer bom uso dos *insights* que podem advir da negociação equitativa.

Entretanto, existem duas fontes de diferenças substanciais entre os procedimentos de Smith e de Rawls. Primeiro, a abordagem contratualista procede, por padrão, em direção à identificação das demandas da transcendência (o principal questionamento na posição original tem por objetivo as demandas de uma sociedade justa), enquanto os espectadores imparciais de Smith são invocados tipicamente para alternativas diferentes, a fim de esclarecerem problemas específicos do avanço ou do retrocesso da justiça em uma abordagem comparativa. Uma segunda diferença decorre do fato de que os observadores imparciais

24 - Rawls, *A Theory of Justice* (1971), p. 23.

25 - Sen, “Open and Closed Impartiality”, *Journal of Philosophy*, XCIX, 9, 445-69, Setembro de 2002.

podem ser imaginados como vindos de longe ou de perto, com perguntas feitas sobre como o problema da decisão pareceria para aqueles que podem ter tido experiências social e institucional diferentes (uma questão de importância, como Smith argumentou). Enquanto os juízes imparciais ideais podem considerar útil perguntar *inter alia* quais coisas parecem com aquelas que eles realmente encontram quando eles são partes envolvidas (aqui o exercício seria, com efeito, bastante semelhante ao exercício rawlsiano), eles poderiam ser considerados como portadores de perspectivas que são completamente diferentes daquelas geralmente aceitas na sociedade e na cultura local.

No exercício rawlsiano, enquanto existe uma exigência procedimental de ignorância dos interesses, objetivos e circunstâncias pessoais, nada é exigido acerca do conhecimento e da aprovação das crenças – e preconceitos – compartilhadas da sociedade na qual esses indivíduos vivem. Diferentemente, o que o modelo da arbitragem equitativa de Smith exige é que as pessoas, em qualquer sociedade, devam exercer um esforço para examinar como suas próprias práticas e convenções pareceriam aos outros, incluindo as pessoas que são informadas sobre elas, mas não completamente educadas nessa sociedade.<sup>26</sup>

## Perspectivas globais, interesses locais e paroquialismo<sup>27</sup>

As razões para ir além das partes contratantes em um Estado-nação incluem pelo menos três interesses distintos: (1) os interesses de outros povos podem ser afetados (por exemplo, por políticas públicas nacionais voltadas ao “aquecimento global” ou à assim chamada “guerra contra o terror”); (2) o paroquialismo, compartilhado por todas (ou quase todas) as pessoas dentro de uma dada sociedade, pode requisitar um desafio difícil com vistas à objetividade; e (3) o conhecimento adicional sobre o que é exequível, que pode ser adquirido a partir das experiências de outros países.<sup>28</sup>

---

26 - A formulação rawlsiana da “posição original” também sofre de um outro problema, ao especificar o conjunto supostamente fixado de negociadores, quando os resultados da negociação podem modificar o tamanho e a composição da população envolvida. Esse problema de “incoerência inclusiva” (como eu denomino), que a abordagem de A. Smith não possui, foi discutido, entre outros assuntos, em um artigo anterior nessa mesma revista “Open and Closed Impartiality”. Todavia, não buscarei prosseguir aqui com esse tipo diferente de dificuldade.

27 - N.T. O termo paroquialismo, na língua portuguesa, é definido como a “tendência a limitar interesses, atividades, pensamentos, opiniões, etc. a uma esfera puramente local, sem atenção a fatores externos ou mais amplos.” Cf. *Grande dicionário Houaiss da língua portuguesa*. Nos dicionários *Michaelis* e *Aulete*, há a referência a provincianismo, palavra que, por vezes, é usada no lugar de paroquialismo, como antônimo de cosmopolitismo.

28 - Sobre a última questão, vale ressaltar que o entendimento, em diferentes países, das viabilidades sociais (relevante, por exemplo, para a justiça de gênero) e viabilidades econômicas (relevante, por exemplo, para as políticas de combate à pobreza) é significativamente influenciado no mundo contemporâneo, a partir das experiências – sucessos e fracassos – de outros países; sobre isso, ver meu *Desenvolvimento como Liberdade*. Nova York: Knopf, 1999.

Sobre o primeiro item (ou seja, os interesses dos povos além das fronteiras que podem ser afetados de modo significativo), a abordagem da arbitragem permite a possibilidade de pesar alguns dos interesses mais amplos sobre a justiça global, incluindo os interesses distributivos, que nos levaram, nos anos recentes, a tentativas de considerar uma versão “cosmopolita” da posição original de modo que os interesses dos povos em outros países que podem ser influenciados por políticas públicas do país em questão, não são negligenciados.<sup>29</sup> Isto tem por objetivo ampliar o alcance da justiça.

Todavia, é importante ver claramente que buscar os espectadores imparciais em outro lugar não torna o exercício de arbitragem smithiano semelhante à versão “cosmopolita” da abordagem contratualista (que poderia configurar-se de modo a incluir todas as pessoas no mundo em uma posição original gigante, produzindo um enorme contrato social global para um mundo completamente justo). O rumo tomado pela arbitragem equitativa é fundamentalmente diferente daquele da negociação justa e do contrato social pela forma como a imparcialidade é interpretada.<sup>30</sup> As exigências institucionais do procedimento contratualista, se têm por objetivo a identificação da justiça transcendental, levam-nos imediatamente à necessidade de uma nação soberana, que foi o cerne da argumentação de Thomas Nagel. Esse é um problema que não advém de um modo similar, no caso de se invocar o esquema dos espectadores imparciais para auxiliarem na avaliação da justiça em uma abordagem comparativa.<sup>31</sup>

Contudo, existem algumas similaridades entre a versão cosmopolita do contrato social e a recorrência aos observadores imparciais como observadores. A exclusão rawlsiana de estrangeiros possuírem voz na avaliação de políticas públicas de um país que tem influência sobre o resto do mundo é restringida, tanto no contrato social cosmopolita quanto no exercício smithiano de convocar observadores de perto ou de áreas longínquas.

Segundo, a abordagem do espectador imparcial pode produzir, *inter alia*, perspectivas imparciais que são desvinculadas não apenas de interesses particulares conferidos aos cidadãos individuais, mas, também, de qualquer paroquialismo ou crenças locais que podem ser compartilhadas geralmente por todos os membros de um dado arranjo político ou comunidade. Assim, uma das vantagens possíveis da via da distribuição equitativa é a maior versatilidade em relação à outra perspectiva, que pode incorporar um desafio procedimental sistemático às distorções das convicções paroquiais.

29 - Ver, por exemplo, Thomas Pogge (ed.), *Global Justice*. Maiden, MA: Blackwell, 2001. Ver também Deen Chatterjee *The Ethics of Assistance: Morality and the Distant Needy*. New York: Cambridge, 2004.

30 - Por exemplo, pode ser mostrado que a versão cosmopolita da posição original de Rawls é também vulnerável à “incoerência inclusiva” de uma forma que a abordagem de Smith não é, como é discutido no meu artigo “Open and Closed Impartiality”.

31 - Aqui, novamente, deve-se notar que na versão do “contratualismo” apresentada por Scanlon (*What We Owe to Each Other*), que tem muitas afinidades com o exercício de imparcialidade smithiano, o foco não precisa ser somente na transcendência.

A ação de evitar o paroquialismo foi, de fato, uma das principais razões para a insistência de Smith para que os espectadores imparciais devessem representar *inter alia* as perspectivas de “uma certa distância” (como Smith coloca). Smith expõe o argumento assim: “não podemos fazer algo de outra forma que não seja pelo empenho para enxergá-lo com os olhos de outras pessoas, ou como outras pessoas provavelmente o veriam”.<sup>32</sup>

No capítulo de *A teoria dos sentimentos morais*, intitulado “Sobre a influência dos modos e costumes nos sentimentos de aprovação e desaprovação moral”, Smith argumenta que “as diferentes situações de diferentes épocas e países são capazes (...) de atribuir diferentes caracteres para a generalidade daqueles que vivem neles e aos seus sentimentos, no que tange aos graus particulares de cada qualidade, ou seja, culpável ou louvável, e variam de acordo com o grau que é usual em seus próprios países e nos seus próprios tempos”.<sup>33</sup> Uma das ilustrações de Smith de tais valores paroquiais era a tendência de todos os comentaristas políticos na Grécia antiga, incluindo os sofisticados atenienses, considerarem o infanticídio um comportamento social perfeitamente aceitável. Mesmo Platão e Aristóteles não deixavam de expressar aprovação, notou Smith, a essa prática extraordinária “de um costume ininterrupto existente nesse tempo (...) plenamente autorizado” na Grécia antiga.

O esquema rawlsiano de fazer perder as informações sobre as identidades pessoais em uma dada sociedade, o que contribui bastante para eliminar a influência de interesses individuais já fixados, não fornece qualquer modo sistemático de evitar os preconceitos que são amplamente compartilhados por todas as pessoas dentro de uma dada sociedade.

Na realidade, o localismo da abordagem contratualista é, de fato, reafirmado pela insistência de Rawls de que o exercício transcendental na posição original deva se concentrar na “estrutura básica” de “uma sociedade fechada: ou seja, devemos considerá-la como autocontrolada e sem possuir relações com outras sociedades”.<sup>34</sup> O argumento de Smith de que devemos ver nossos sentimentos *inter alia* a partir de “uma certa distância de nós mesmos” é motivado pela necessidade de se perguntar se alguma aparência de justiça é socialmente tendenciosa devido ao impacto da tradição recalcitrante e do costume local. O exemplo do infanticídio apresentado por Smith permanece relevante, de maneira preocupante, em algumas sociedades ainda hoje (embora não mais na Grécia), mas existem também muitas outras práticas para as quais o fato de a justiça ser percebida como realizada pode invocar eficientemente, como Smith coloca, “os olhos do resto da humanidade”.<sup>35</sup>

Enquanto um público norte-americano pode pensar ser fácil acreditar que perspectivas

---

32 - Smith, *The Theory of Moral Sentiments*, III, 1, 2, p. 110.

33 - Smith, *The Theory of Moral Sentiments*, Volume 2.7, p. 204.

34 - Rawls, *Political Liberalism*, p. 12.

35 - Smith, *Lectures on Jurisprudence*, R.L. Meek, D.D. Raphael, and P.G. Stein (Eds.). New York: Oxford, 1978. p. 104.

distantes podem ser importadas eficientemente, no caso de sociedades “retrógradas” tais como o Sudão ou o Afeganistão, nos quais, por exemplo, assassinatos pela honra ocorrem e mulheres adúlteras podem ser apedrejadas até a morte, pode não existir reconhecimento correspondente para a necessidade de fazer isso em países mais adiantados como os Estados Unidos. Todavia, práticas bem estabelecidas que recebem amplo suporte dentro das fronteiras podem ser vistas como inaceitáveis para pessoas em muitos outros países espalhados pelo mundo, desde a Europa até o Japão. Por exemplo, o uso abundante das punições capitais, com ou sem o acompanhamento da exaltação pública, pode precisar ser abordado não apenas ao se perguntar se tais punições parecem “cruéis e excepcionais” *dentro* da cultura local dos Estados Unidos, mas também “aos olhos do resto da humanidade” que, pensou Smith, deve ser invocado para entender se “uma punição parece justa” (*ibid.*, p. 104).

A relevância das perspectivas distantes tem um apoio claro em alguns debates atuais nos Estados Unidos, por exemplo, aquele que aconteceu na suprema corte há algum tempo, sobre a adequação do uso da pena capital para crimes cometidos por menores. As demandas da justiça, sendo vistas como realizadas, mesmo em um país com os Estados Unidos, não podem negligenciar inteiramente o entendimento que pode ser gerado pelo modo como o problema é avaliado em outros países do mundo. A maioria dos juízes da corte simplesmente não “diferiu dos estrangeiros com a mesma opinião” (como o juiz Scalia sugeriu), mas aceitaram que nas deliberações, para se chegar a julgamentos americanos bem fundamentados e não paroquiais, pode ser útil levar em consideração os esclarecimentos que perspectivas não locais propiciam, após submetê-las ao escrutínio crítico em uma abordagem local melhor informada.

De fato, a irrefutabilidade aparente dos valores coloquiais frequentemente depende da falta de conhecimento do que foi provado factível nas experiências de outros povos. A defesa do infanticídio na Grécia antiga poderia ter sido influenciada não apenas pelo conhecimento de sociedades nas quais o infanticídio é considerado como totalmente inaceitável, mas também devido ao fato de que essas sociedades não estão sucumbindo ao caos e à crise como um resultado da não permissão de tal assassinato. A despeito da importância indubitável do conhecimento local, o conhecimento global, também possui o seu valor e pode contribuir para os debates sobre os valores e as práticas paroquiais.

## A divergência e a exequibilidade da abordagem comparativa

Em um modelo de arbitragem, é provável que existam algumas vozes divergentes sobre vários problemas, especialmente quando as perspectivas buscadas advêm de todas as partes. Ainda que possa existir uma convergência considerável de valores entre a razão pública global, pode também existir diferenças residuais acerca do modo como os resultados de tais deliberações podem surgir em diferentes meios sociais e culturais. Para uma abordagem transcendental da justiça dependente da emergência de um acordo completo sobre a natureza e as demandas de uma “justiça global”, esse seria, evidentemente, um problema avassalador. Contudo, não é problemático, do mesmo modo, para o uso da abordagem comparativa da justiça. Aqui, o enfoque será se existem problemas significativos sobre quais acordos ou consensos podem surgir, especialmente depois da interação pública, com troca de conhecimento e entendimento. As demandas da justiça global podem não ir além dos modos acordados para fortalecer a justiça no mundo, no caminho comparativo dos juízos de justiça.

Ademais, a extensão do acordo sobre os rankings globais pode seguir se expandindo na medida em que o processo de interação continua. A abordagem comparativa não exige um “tudo ou nada” extremista e permite ao mundo enfrentar problemas intensos de injustiça global (tais como inanição, fome generalizada, analfabetismo crescente ou as mortes desnecessárias advindas de doenças evitáveis e curáveis), sobre os quais o consenso pode ser mais fácil de obter sem esperar por um acordo total sobre avaliações controversas.

Algo similar pode ser dito sobre a percepção de perspectivas globais ao se tomar decisões locais que não são, a princípio, questões de justiça global: por exemplo, ao se avaliar uma estrutura nacional de legislação punitiva, levando-se em consideração se “uma punição parece justa”, para usar a frase de Smith, sem negligenciar como isso é visto pelas pessoas educadas em um contexto diferente e com diferentes entendimentos informacionais. As demandas da objetividade não exigem apenas que se evite uma “inclinação pessoal” (como notou Rawls), mas também que se evite o paroquialismo nacional (como Smith enfatizou). Entretanto, isso não produz qualquer obrigação para se aceitar as visões de terceiros, apenas obriga que elas sejam levadas em consideração em um escrutínio geral (que leve à modificação das prioridades locais em alguns casos e à não revisão das prioridades em outros).

De fato, vozes distantes também estão submetidas à disciplina do escrutínio crítico, incluindo a convocação de espectadores imparciais de longe e de perto. A disciplina da arbitragem justa no contexto da justiça global pode ser vista como interativa de modo global. A *aceitabilidade* geral, que deve ser distinguida da aceitação universal preexistente, é um problema importante em qualquer avaliação social. Tentei discutir esse ponto em outro trabalho, no qual a razão pública interativa e aberta é crucialmente importante para o enten-

dimento das exigências feitas pelos direitos humanos, a despeito das diferenças em práticas manifestas *entre* países, e evidentemente também dentro de cada país.<sup>36</sup>

O ponto de vista comum que pode ser visto surgir com a base desse escrutínio associativo pode estar longe da totalidade e a forma de concordância necessária, em muitos casos, não vai além de notar que alguns arranjos sociais são seriamente injustos, de um modo que podem ser remediados, ainda que outras comparações possam ser difíceis de fazer sem uma ambiguidade substancial. Uma teoria da justiça não precisa torcer o nariz para o vasto alcance e a relevância das conclusões aceitáveis quanto aos direitos humanos ou à justiça social, produzidos com base na razão pública, mesmo quando eles não chegam a uma resolução completa de todos os problemas existentes sobre decisões na organização da sociedade.<sup>37</sup>

## Que diferença isso faz?

Não tentarei resumir o artigo, mas anotarei brevemente algum dos problemas discutidos aqui. Primeiro, argumentei pela rejeição da tradição de se focar na pergunta clássica – e muito invocada – “O que é uma sociedade justa?” e em favor de, ao invés disso, nos concentrarmos em questões comparativas de justiça (vinculadas aos questionamentos sobre a implementação da justiça ou a redução da injustiça). Apresentei algumas razões para se duvidar se a questão transcendental pode ser normalmente respondida, mas mesmo se pudesse ser (e uma teoria transcendental da justiça foi totalmente bem sucedida ao responder essa questão clássica), isso não produziria – direta ou indiretamente – uma estrutura comparativa que é necessária para uma avaliação real da justiça. Por outro lado, uma teoria comparativa da justiça pode ser totalmente viável e completamente utilizável sem conter – ou impor – qualquer resposta à grande questão “O que é uma sociedade justa?”.

Segundo, a especificação das demandas da justiça plena e da elaborada parafernália institucional que tem que ser posta em ordem para a busca do modelo rawlsiano de justiça deixa em aberto a questão sobre como devemos avaliar se algumas mudanças sociais implementam ou obstam a causa da justiça – questões que nós confrontamos constantemente no mundo dentro de cada país e no campo dos arranjos globais. Por exemplo, ter leis de paten-

---

36 - Ver meu artigo “Elements of a Theory of Human Rights”, *Philosophy and Public Affairs*, XXXII, 315-56, Outono de 2004. Ver também Joshua Cohen, “Procedure and Substance in Deliberative Democracy”, em Seyla Benhabib (Ed.), *Democracy and Difference: Contesting the Boundaries of the Political*. Princeton: University Press, 1996. pp. 95-119; e Charles Beitz, “Human Rights as a Common Concern”, *American Political Science Review*, XCV, 229-82, Junho de 2001.

37 - Eu me dedico a essas questões de forma mais completa em um livro a ser publicado pela Harvard University Press, chamado *Freedom and the Theory of Justice*.

tes decentes sobre a produção e distribuição de produtos farmacêuticos pode fazer muito para os pacientes miseráveis e desnecessariamente condenados à AIDS no mundo de hoje, e argumentos podem ser apresentados para sugerir que isso reduziria a injustiça evidente na sociedade global. Mas isso, por si só, não nos levará a qualquer lugar perto dos requisitos da justiça transcendental rawlsiana.

Terceiro, uma crença implícita que parece muito comum em uma parte substancial da filosofia política – a saber, a identificação de uma sociedade plenamente justa não é apenas crucial para a compreensão da natureza da justiça, mas é também essencial para um entendimento relacional da justiça bem fundamentado – causou o efeito de atribuir um papel ubíquo para a análise transcendental da justiça na filosofia contemporânea. Argumentei que essa crença subjacente pode estar completamente errada. Uma abordagem transcendental não é necessária nem suficiente para responder às questões sobre o avanço da justiça que exigem urgentemente nossa atenção e que pedem uma abordagem comparativa mais robusta. Essa abordagem é claramente diferente da investigação da transcendência.

Quarto, existe uma necessidade para permitir – e algumas vezes mesmo para afirmar – a incompletude das comparações relacionais da justiça. A incompletude pode advir de lacunas incomunicáveis nas informações, mas também de questões concernentes a decisões não resolvíveis que envolvem considerações díspares que podem resistir à gradação, mesmo com todas as informações. Todavia, a possível incompletude dos juízos que surgem da via relacional não é um constrangimento para a razão prática. De fato, ela está longe disso.<sup>38</sup> O desafio de avaliar o *avanço* ou de identificar o *retrocesso* não estará comprometido pela presença da incompletude substancial nos rankings de justiça com muita frequência.

Quinto, à parte da importância geral da incorporação de possíveis incompletudes das avaliações, a admissão da incompletude também torna mais fácil trazer as vozes distantes para a avaliação da justiça, e isso pode ser criticamente importante para o alcance e para a força da razão pública. Um requisito procedimental para considerar as perspectivas não locais pode ajudar a evitar o domínio indevido do interesse local e do paroquialismo da razão local moldada pela influência das convenções estabelecidas e das estruturas informacionais limitadas (sem que elas sejam intelectualmente desafiadas). Temos que passar adiante da confiança nas perspectivas ancoradas territorialmente nos “cidadãos livres e iguais nascidos na sociedade na qual levam suas vidas” (para usar a frase de Rawls).<sup>39</sup>

O mundo no qual vivemos não é apenas injusto, é extraordinariamente injusto. Não é algo

---

38 - Discuto esse assunto em “Consequential Evaluation and Practical Reason”, e alguns interesses analíticos subjacentes em “Maximization and the Act of Choice”.

39 - Rawls, *Political Liberalism*, p. 23.

frívolo buscar uma estrutura para uma teoria da justiça que se concentre no avanço, e não na transcendência, e também permita ser globalmente interativa, ao invés de ser intelectualmente isolada. Temos uma boa razão para deixarmos de nos concentrar exclusivamente em um programa que identifica demandas totalistas – e possivelmente paroquiais – da justiça transcendental e contratualista. Devemos retirar a teoria da justiça desse âmbito restrito.

# **Artigos Inéditos**



# Será que, como defendem os epicuristas, estar morto não é um mal para quem morre?

*As argued by the Epicureans,  
is being dead not bad to those who die?*

Domingos Faria\*  
Universidade de Lisboa

## Resumo

Os epicuristas argumentam que a morte não nos diz respeito. Porém, será esse argumento sólido? Neste artigo, vou examinar uma versão do argumento epicurista (a versão de Rosenbaum), que tenta mostrar que a morte não é má para nós. Vou apresentar alguns contraexemplos, tentando argumentar que o argumento epicurista não é sólido. Além disso, penso que existem boas razões para se aceitar a perspectiva da privação, de acordo com a qual a morte pode ser um mal para quem morre.

## Palavras-chave

Morte, Mal da Morte, Perspetiva Epicurista, Perspetiva da Privação

\* Agradeço ao professor Pedro Galvão as aulas de Filosofia da Morte e a indicação de referências bibliográficas que me motivaram e ajudaram a elaborar este artigo. Através de uma Bolsa de Doutoramento (SFRH/BD/85051/2012) me beneficieei, na realização deste trabalho, do apoio da Fundação para a Ciência e a Tecnologia.

## **Abstract**

Epicureans have argued that death is nothing to us. However, is this argument sound? In this paper, I will survey one version of the Epicurean's argument (Rosenbaum's version) which tries to show that death is not bad for us. I will put forward some counterexamples, trying to argue that Epicurean's argument is not sound. Moreover, I think there are good reasons for one to accept the deprivation account, according to which death might be a bad thing for the one who dies.

## **Keywords**

Death, Badness of Death, Epicurean Account, Deprivation Account

Seguramente, em um futuro breve ou mais distante, vamos estar mortos, mas será isso um mal para nós? Partindo da suposição que estar vivo é geralmente bom e que quando alguém morre deixa de existir, será que a abissal inexistência por estar morto constitui um mal para o próprio indivíduo que morre? Os epicuristas respondem ao problema do mal da morte alegando que estar morto não é um mal para quem morre. Com este ensaio pretendo apresentar a versão epicurista de Rosenbaum e defender que esta resposta não é plausível, pois parte de pelo menos um pressuposto errado. Além disso, essa objeção à resposta epicurista vai dar razões para se aceitar a perspectiva da privação como uma melhor resposta ao problema.

Antes de avançar para a discussão da perspectiva epicurista, é preciso salientar que o problema do mal da morte não pode ser confundido com a questão de a nossa própria morte ser má para as outras pessoas. É claro que a nossa morte poderá ser má para os outros, uma vez que os nossos amigos e familiares vão perder a oportunidade de continuar a interagir conosco. Ora, com a nossa morte, aquelas pessoas que gostam de nós ficarão impedidas de dialogar ou partilhar as suas vidas conosco, o que pode gerar imensa tristeza para elas. Porém, esse não é o problema que estamos a tratar aqui.

Do mesmo modo, não se pode confundir o problema do mal da morte com a questão de saber se o processo de morrer é mau ou não. Aliás, este processo de morrer pode ser bastante mau para uma pessoa se for acompanhado de um grande sofrimento. Ao contrário disso, o problema que se pretende discutir aqui é se estar morto é mau ou não para aqueles que estão mortos.

Epicuro, na *Carta a Meneceu*, defende que estar morto nunca é um mal para quem morre. Por isso, não nos devemos incomodar ou ter medo da nossa futura condição de estar

morto. Epicuro justifica que estar morto não é um mal para quem morre argumentando que “todo o mal e todo o bem assentam na sensação e a sensação acaba com a morte”. A ideia aqui presente parece ser que algo só pode ser mau ou bom para nós se pudermos sentir ou experienciar isso; ora, uma vez que com a morte finda com toda a sensação ou experiência, então esta não pode ser má para quem está morto. Desse modo, “a morte, o mais temido dos males, não nos diz respeito; pois enquanto existimos a morte não está presente, e quando a morte está presente nós já não existimos. Nada é portanto nem para os vivos nem para os mortos visto que não está presente nos vivos, e os mortos já não são” (Epicuro).

Apesar da passagem acima ser suscetível a múltiplas interpretações, pode-se fazer uma coerente reconstrução do argumento de Epicuro seguindo a proposta de Rosenbaum (1986: 121-122). Por isso, atendendo mais à consistência das ideias do que ao rigor histórico, pode-se representar canonicamente o argumento epicurista desta forma:

- (1) Um estado de coisas é mau para uma pessoa P só se P pode experienciar isso em algum momento.
- (2) Assim, P estar morto é mau para P apenas se isso é um estado de coisas que P pode experienciar em algum momento.
- (3) P pode experienciar um estado de coisas em algum momento só se começa a ocorrer antes da morte de P.
- (4) P estar morto não é um estado de coisas que começa a ocorrer antes da morte de P.
- (5) Desse modo, P estar morto não é um estado de coisas que P pode experienciar em algum momento.
- (6) Logo, P estar morto não é mau para P.

As premissas centrais desse argumento são (1), (3) e (4). Os outros passos (2) e (5) são apenas uma consequência lógica das premissas centrais, seguindo-se validamente a conclusão (6). Mas será que esse é um argumento sólido? Ou seja, serão as suas premissas centrais verdadeiras? A premissa (4) é verdadeira por definição, pois não se pode estar morto antes de haver morte. No entanto, (1) e (3) requerem justificação pormenorizada.

Um dos detalhes que é preciso esclarecer na premissa (1) é o conceito de experiência. Seguindo a fundamentação de Rosenbaum (1986: 124), uma pessoa experimenta um esta-

do de coisas apenas se isso pode afetá-la de alguma forma, o que implica um elemento causal na experiência. Portanto, uma condição necessária, embora não suficiente, para haver experiência é ser causalmente afetado por uma situação ou estado de coisas. Contudo, uma pessoa não experiencia algo meramente por acreditar que a situação ocorreu ou ocorrerá, ou por imaginar uma certa situação. Podemos nos imaginar numa viagem até à lua, mas isso não é uma situação que estamos experienciando, uma vez que aqui não está presente a conexão causal requerida entre a situação e a pessoa.

Alguns exemplos podem ajudar a compreender melhor esse conceito de experiência, bem como a premissa (1): vamos imaginar uma pessoa que é cega e que nunca conseguirá na sua vida adquirir a visão. Assim, um pôr-do-sol ou uma pintura esteticamente bela não podem causalmente afetar essa pessoa e, por conseguinte, ela não pode experimentar isso. Essa mesma pessoa não pode igualmente experimentar uma sequência de imagens chocantes que causam repulsa às pessoas que as veem. Portanto, essa sequência de imagens não pode ser chocante nem má para a pessoa cega porque ela não é capaz de ter essa experiência e de ser afetada por isso. Do mesmo modo, uma dor de cabeça só é má para nós se formos afetados por ela; isto é, caso experienciemos essa dor de cabeça. Se não experienciamos essa dor de cabeça, então isso não constituirá um mal para nós. Ora, esses exemplos parecem mostrar a plausibilidade do seguinte princípio de experiência:

(E) Se uma pessoa não pode experimentar um estado de coisas em algum momento, então o estado de coisas não é mau para a pessoa (cf. Rosenbaum 1986: 123).

Esse princípio é logicamente equivalente à premissa (1), o que constitui uma razão para se aceitar (1). E uma das consequências de (1) e do princípio (E) é que as pessoas que estão mortas não podem experimentar qualquer estado de coisas, pois ao estarem mortas não são capazes de ter sensações e experiências más. Assim, para quem está morto, nenhum estado de coisas pode ser mau.

Além disso, alguém pode pensar que (1) também leva ao seguinte princípio: o que um indivíduo não conhece não pode ser mau para ele próprio. Podemos imaginar, como Nagel (1979: 65), que um adulto inteligente sofreu de um grave acidente vascular cerebral que o reduz a capacidades mentais de uma criança de 3 meses. Parece que aquele sujeito com as capacidades mentais reduzidas não tem conhecimento ou consciência do que lhe ocorreu; desse modo, tal situação não seria má para ele. No entanto parece óbvio que esse sujeito

sofreu de um grande infortúnio e mal que está lhe prejudicando, pois ele perdeu imensas capacidades que limitaram completamente a sua vida, mesmo não tendo consciência disso. Assim, parece que não se pode aceitar (1). Será que com esse exemplo estamos a negar adequadamente a premissa (1)?

Para Rosenbaum (1986: 126) essa crítica falha no alvo, pois tudo o que requer a premissa (1) para um evento ou estado de coisas ser mau para a pessoa é que ela possa experienciar isso em algum momento, e não que a pessoa seja efetivamente consciente dos efeitos causais em algum momento. A ideia é que pode haver experiência de algo mau sem a pessoa ter consciência disso. Assim, aquele adulto inteligente que ficou com as capacidades de uma criança de três meses sofreu de um grande mal, pois foi causalmente afetado por esse estado de coisas, mesmo não tendo consciência disso.

Da mesma forma, não refuta (1), ao contrário do que Nagel pensava, se uma pessoa for traída pelos seus amigos sem o saber. Pois, é possível para essa pessoa descobrir que foi traída e, assim, experienciar esse mal em algum momento. Mas também é possível que essa experiência seja inconsciente, tal como é possível experienciar e ser diretamente afetado por graus nocivos de radioatividade sem o saber. Portanto, a ideia base da premissa (1) é a seguinte: o que um sujeito experencia (tanto de forma consciente como inconscientemente) pode fazer mal a esse sujeito, mas o que esse sujeito não pode experienciar não lhe pode fazer mal.

E que razões temos para aceitar a premissa (3)? Já vimos que uma condição necessária para uma pessoa experienciar um estado de coisas é ser causalmente afetada por esse estado de coisas. Ora, parece claro que uma pessoa que deixou de existir não pode ser causalmente afetada por eventos que ocorreram após a sua morte. Por exemplo, as pessoas que morreram em 1944 em Hiroshima não puderam experienciar os efeitos catastróficos e maléficos da bomba nuclear de 1945. Assim, uma pessoa só pode experienciar um estado de coisas se isso começar a ocorrer antes da morte da pessoa.

No entanto, pode acontecer um estado de coisas que causalmente afete uma pessoa antes de nascer, como uma mãe toxicod dependente que pretende ter um filho e que continua a injetar grandes quantidades de drogas até durante a gestação. Nesse caso, a futura criança que vai nascer já está sendo afetada pelo comportamento negligente da mãe mesmo antes de existir, pois tal comportamento materno vai prejudicá-la seriamente em seu desenvolvimento. Com esse exemplo, parece colocar-se ainda em causa o princípio de existência que fundamenta outros argumentos epicuristas, de acordo com o qual um estado de coisas pode ser mau para uma pessoa apenas se a pessoa existir no momento em que ocorre esse estado de coisas. Uma vez que pode haver estados de coisas que podem ser maus mesmo antes da pessoa existir, então o princípio de experiência (E) parece ser mais plausível do que

o princípio de existência. Assim, é compatível com (E), ao contrário do princípio de existência, afirmar que uma pessoa pode não existir quando a causa do mal que a afeta ocorre.

Porém, ao estar morto, por nada se pode ser causalmente afetado. O máximo que pode ser afetado depois da pessoa estar morta é o seu corpo inerte (o qual é suscetível de ser atingido pela decomposição). Mas, seguindo as principais teorias da identidade pessoal, uma pessoa não é a mesma que o seu corpo sem vida, nem existe qualquer tipo de preocupação prudencial para com um cadáver. Por isso, a pessoa não pode ser afetada ou experienciar qualquer evento depois da sua morte.

Em suma, com a justificação das premissas centrais (1) e (3), advoga-se que algo só é mau para uma pessoa se ela pode experienciar isso (isto é, se pode ser causalmente afetada por isso), consciente ou inconscientemente, em algum momento. Como nada se pode experienciar depois da morte, então estar morto não é um mal para quem morre.

Será que com isso se provou de forma cogente que estar morto não é um mal para quem morre? Penso que não, uma vez que considero que a premissa (1) e o princípio (E) são falsos. Negar o princípio (E) é dizer que «uma pessoa não pode experienciar um estado de coisas em algum momento, mas o estado de coisas é mau para a pessoa»; e negar a premissa (1) é defender que «um estado de coisas é mau para uma pessoa P, mas P pode não experienciar isso em algum momento». Portanto, nem todos os males para P são experienciados por P. Para mostrar que isso é plausível vou apresentar os seguintes contraexemplos a (1) e (E):

(I) Vamos imaginar que Afonso e Sofia são um casal romântico. Afonso é astronauta e foi chamado, juntamente com outros elementos, para uma missão de estudo em Marte, durante vários meses. No último mês de expedição, os aparelhos de comunicação com a Terra ficaram seriamente danificados e os astronautas foram obrigados a regressar imediatamente. Enquanto regressavam, Afonso envolveu-se intimamente com uma tripulante da nave espacial, traindo assim a sua esposa Sofia. No entanto, passado apenas alguns minutos a nave explodiu devido a uma grande falha nos motores e não sobreviveu ninguém, o que impede Sofia de experienciar a traição do seu marido.

(II) José é um multimilionário desconhecido, bastante idoso e sem família, que pretende doar toda a sua fortuna a uma instituição de solidariedade como a Unicef. Desse modo, assina um cheque com todo o dinheiro que conseguiu adquirir durante a sua vida e entrega-o ao seu único mordomo, Carlos, para o fazer chegar à Unicef.

Porém, em vez de entregar esse dinheiro à instituição, Carlos foge com o cheque para um país longínquo. Passado uma hora, José falece e a Unicef nunca poderá ter conhecimento ou qualquer experiência de que iria receber uma grande quantidade de dinheiro, que era o suficiente para acabar com a fome no mundo.

(III) Vera nasceu num país completamente isolado, fortemente armado, onde ninguém consegue sair ou entrar e no qual as mulheres não têm acesso à educação. Assim, ela não tem oportunidade de ler e escrever. Devido a esse sistema fundamentalista, a sua tendência natural para a poesia não se manifesta, nem para os outros nem para si mesma – apesar disso, ela tem uma vida confortável no seu país. Porém, se ela nascesse em outro país, livre e com igualdade de oportunidades, a sua propensão para a poesia ter-se-ia desenvolvido e manifestado, e, por causa disso, sentir-se-ia mais realizada. Mas nas condições em que se encontra no seu país, não pode experienciar isso.

No exemplo (I), Sofia não experienciou, nem pode experienciar o mal da traição, uma vez que não havia comunicações com a Terra, a nave foi reduzida a cinzas e não sobreviveu ninguém para contar o sucedido. Por isso, ela não pode ser causalmente afetada pela traição, até mesmo de forma inconsciente. Mas a traição de Afonso foi um mal para ela, pois podemos dizer que Afonso violou o acordo que livremente fizeram, bem como não retribuiu devidamente a lealdade de Sofia – e isso parece ser mau mesmo sem haver descoberta ou experiência. Portanto, a traição é má para Sofia, apesar de não ser possível que ela tenha qualquer experiência de tal ato. O mal que Sofia sofre é um mal por comparação: é pior uma situação em que é traída do que não o ser.

No exemplo (II), os membros da Unicef não podem ser causalmente afetados ou experienciar qualquer mal, pois José está morto e ninguém viu ou sabe do furto do cheque. Mas, mesmo assim, sofrem um grande mal, já que foram privados de acabar com a fome no mundo. Igualmente, as pessoas que a Unicef iria ajudar estão sofrendo um mal, uma vez que poderiam estar em melhores condições com aquele dinheiro. Portanto, a Unicef está a padecer de um mal sem o experienciar - ou seja, foi privada de fazer um grande bem. E o mal que a Unicef sofre é um mal por comparação: é pior uma situação em que esse donativo é roubado do que uma situação em que se recebe sem problemas esse dinheiro.

No exemplo (III), Vera não pode experienciar qualquer mal com o fato de possuir uma tendência natural para a poesia, que não foi desenvolvida nem descoberta, pois não tem

qualquer forma de saber se tem tal faculdade possível ou de sentir que está sendo prejudicada. Igualmente, parece que Vera não pode ser causalmente afetada ou experienciar um estado de coisas que não se manifesta e que não existe, nem existirá nas condições em que ela se encontra. Porém, apesar de Vera não experienciar isso, ela está a ser privada de um grande bem que a faria mais feliz (cf. Johansson 2005: 79). Aqui Vera sofre um mal que não é devido a alguma experiência, mas é um mal por comparação: é pior para ela a situação em que a sua propensão para a poesia não é desenvolvida e manifesta do que a situação oposta.

Em todos esses exemplos existem males para as pessoas que não são experienciados pelas próprias pessoas. E existe um mal para essas pessoas porque foram privadas de um acontecimento positivo, mesmo que não tenham qualquer experiência ou consciência disso. Portanto, parece legítimo dizer que nem todos os males são experienciáveis, o que refuta a premissa (I) e o princípio (E).

Além disso, é preciso sublinhar que o princípio (E) foi proposto a partir de exemplos que não permitem fazer uma generalização alargada, como ilustrar que o mau cheiro de um charuto aceso só é mau para quem pode experienciar isso; assim, para quem não tem olfato, esse cheiro não poderá ser mau. Com exemplos como esse, Rosenbaum generaliza que algo só é mau para nós se pudermos experienciá-lo. No entanto, presenciemos com isso a uma falácia de generalização apressada, pois apenas se consideraram exemplos que podem ser sensorialmente experienciados. E ao concebermos contraexemplos como (I), (II) e (III), vemos que não é adequado fazer tal generalização, uma vez que existem males que não são experienciados - e mesmo aqueles que são experienciados muitos não o são sensorialmente.

Em (E) constata-se uma categoria de males que é apenas uma apropriada subcategoria de todos os casos relevantes de mal (cf. Fischer 1997: 349). Porém, existem outros casos, como os contraexemplos acima, em que algo é mau para a pessoa mesmo que isso não possa afetar causalmente essa pessoa.

Seguindo uma proposta mais abrangente quanto à consideração dos males (cf. Kagan 2012: 210-212), podemos dizer que existem três grandes categorias nas quais algo pode ser mau para uma pessoa:

- (A) Intrinsecamente mau.
- (B) Instrumentalmente mau.
- (C) Comparativamente mau.

Algo é (A), intrinsecamente mau para nós, se pudermos experienciar isso diretamente. Por exemplo, ao experienciarmos qualquer sofrimento, isso é mau por sua própria natureza. Mas, quando se é reprovado em uma avaliação, isso não é intrinsecamente mau. Isso pode ser mau por levar à reprovação, que por sua vez pode originar a tristeza e outros males intrínsecos. Com isso vemos que algo pode não ser mau em si, mas mau em virtude do que causa e para onde nos leva - ou seja - é algo instrumentalmente mau (B). Além desses tipos de males, existem os males por comparação (C), no sentido em que o estado de coisas atual é pior para uma pessoa comparativamente com outro possível estado de coisas melhor. Por exemplo, dão-nos dois envelopes para escolhermos apenas um. Selecionamos o primeiro deles e dentro dele está um cheque de 5 euros. Contudo, no segundo envelope, sem termos conhecimento disso, estava um cheque com 50.000 euros. Assim, parece que foi mau escolher o primeiro envelope no sentido comparativo, pois teria sido melhor escolher o segundo envelope.

Enquanto (A) e (B) implicam experiência da pessoa para um estado de coisas ser mau para a pessoa, o tipo (C) não requer qualquer experiência como mostraram os contraexemplos (I), (II) e (III). Ao termos consciência de (C), da falácia patente em (E) e dos contraexemplos presentes acima, faz sentido substituir a premissa (1) do argumento epicurista pela seguinte premissa:

(1') Um estado de coisas é mau para uma pessoa P, mas P pode não experienciar isso em algum momento.

Porém, ao partirmos dessa premissa verdadeira, já não se pode validamente concluir (6) que estar morto não é um mal para a pessoa que está morta. A grande falha do argumento epicurista é que apenas considera a existência de males do tipo (A) e (B), desconsiderando completamente os males (C). Por isso, parte do pressuposto errado que todos os males para uma pessoa são experienciáveis (intrínseca ou extrinsecamente) por essa pessoa.

A objeção que apresentamos ao argumento epicurista pode igualmente dar-nos razões para se aceitar a perspectiva rival da privação, de acordo com a qual estar morto pode ser um mal para quem morre. Pois, ao admitirmos que existem males (C) que não são experienciáveis, como sugerem os contraexemplos, então parece razoável afirmar que estar morto poderá constituir um mal para a pessoa que morre no sentido comparativo. Ou seja, estar morto pode ser um estado de coisas mau para a pessoa que morre, uma vez que poderá privá-la de continuar a usufruir de coisas boas que, de outra forma, teria, caso estivesse viva.

De um modo mais formal o argumento é o seguinte:

- (7) Se uma pessoa P está privada de um bom estado de coisas, então P tem um mal comparativo.
- (8) Se P está morto, então P *pode* estar privado de um bom estado de coisas.
- (9) Logo, se P está morto, então P *pode* ter um mal comparativo. Ou seja, estar morto pode ser um mal para quem morre.

A premissa (7) parece devidamente fundamentada com o que já consideramos sobre os males tipo (C) e sobre os contraexemplos. Todavia, a justificação da premissa (8) requer mais atenção, pois a negação da ideia epicurista de que «nenhuma morte é um mal para quem morre» não é que «toda a morte é um mal para quem morre», mas sim que «alguma morte é um mal para quem morre». Então, o que se pretende transmitir com (8) é que estar morto pode ser mau para P caso prive P de um bom estado de coisas (ou seja, ao estar morto não se pode desfrutar das boas coisas da vida, como conviver com os amigos ou ler livros). Mas, também admite que estar morto pode não ser mau para P caso prive P de um mau estado de coisas (como ser poupado da fase final de uma doença degenerativa irreversível, grave e imensamente dolorosa). Por isso, é preciso questionar: o que é excluído ou privado com a morte? O que teria acontecido se a morte específica de P não tivesse ocorrido? A ideia é comparar a morte efetiva de P com a vida que ele teria caso não tivesse morrido naquela circunstância e naquele momento específico.

No entanto, seguindo a proposta de McMahan (2002: 117), essa comparação não deve ser feita apenas com a totalidade do que foi perdido, mas na medida em que P estaria ligado a si mesmo no futuro (caso não morresse) pelas relações de unidade prudencial e pela quantidade de bem que a sua vida conteria. Existem relações de unidade prudencial mínimas (isto é, relações que fundamentam a nossa preocupação com o que acontece conosco no futuro) se houver continuidade física e funcional de áreas do cérebro que sustentam a consciência. Porém, a força dessas relações de unidade prudencial varia de acordo com o grau de unidade psicológica entre nós agora e nós no futuro, o que depende: da proporção da vida mental que é sustentada ao longo desse período, da densidade dessa vida mental, e do grau de referência interna entre os vários estados mentais anteriores e posteriores (cf. McMahan 2002: 84).

À luz dessas considerações, será que, por exemplo, a morte de um embrião humano nas

primeiras semanas de gestação é um mal? Não, pois não há relações de unidade prudencial entre esse embrião e um indivíduo futuro. Do mesmo modo, penso que a objeção das “pessoas possíveis” (cf. Kagan 2012: 217-221) não funciona, pois se um embrião nas primeiras semanas não tem unidade prudencial com um indivíduo futuro, então também uma “pessoa possível” não tem tal unidade prudencial, aliás nem sequer está fecundada. Por isso, não se pode dizer que uma “pessoa possível” sofre de um mal. Pelo contrário, seria um grande mal a morte de uma criança de 7 anos para ela mesma; pois, além de ser privada de muitos bens, teria bastante unidade psicológica consigo mesmo no futuro.

A partir dessas premissas e fundamentações, parece justificar-se apropriadamente que estar morto *pode* ser um mal para quem morre, o que mostra que os epicuristas estão equivocados quando defendem, utilizando as palavras de Epicuro, que “a morte não nos diz respeito”.

## Bibliografia

Epicuro (s/d) “Carta a Meneceu”, in *Crítica: revista de filosofia*, trad. Desidério Murcho, <http://criticanarede.com/meneceu.html>.

Fischer, John Martin (1997) “Death, Badness, and the Impossibility of Experience”, in *The Journal of Ethics*, Vol. 1, No. 4, pp. 341-353.

Johansson, Jens (2005) “Epicurean Objections Against The Evil Of Death”, in *Mortal Beings: On the Metaphysics and Value of Death*. Stockholm, Sweden: Almqvist and Wiksell, pp. 68-99.

Kagan, Shelly (2012) “The Badness of Death”, in *Death*. New Haven: Yale University Press, pp. 205-233.

McMahan, Jeff (2002) *A Ética no Ato de Matar: problemas às margens da vida*, trad. Jônadas Techio. Porto Alegre: Artmed, 2011.

Nagel, Thomas (1979) “Death”, in *The Metaphysics of Death*, ed. John Martin Fisher. Stanford: Stanford University Press, 1993, pp. 61-69.

Rosenbaum, Stephen (1986) “How to Be Dead and Not Care: A Defense of Epicurus”, in *The Metaphysics of Death*, ed. John Martin Fisher. Stanford: Stanford University Press, 1993, pp. 119-134.



# **A moral sensualista de Helvétius pelo princípio do interesse**

## ***The sensualist moral of Helvetius by the principle of interest***

Camila Ferraz  
Universidade Federal do Paraná

### **Resumo**

O presente artigo pretende demonstrar como a tese sensualista, que considera que no homem tudo é fruto da sensibilidade física, está estritamente ligada à moralidade e à política propostas por Helvétius. Tal ligação é feita pela tomada do interesse, entendido, primeiramente, como busca ao prazer e fuga da dor, como princípio que rege tanto as paixões do homem como suas considerações sobre virtude ou vício. Dessa forma, o princípio do interesse se estende da sensibilidade física até a moralidade e a política, e postula a necessidade de que a legislação e a educação garantam a ligação do interesse pessoal ao interesse geral de uma nação, tornando-os compatíveis.

### **Palavras-chave**

Helvétius, sensualismo, interesse, moral, educação, legislação.

## **Abstract**

This article aims at demonstrating how the sensualist theory which considers all men as the result of physical sensualism is closely linked to morality and politics so considered by Helvétius. In such connection the concept interest is understood primarily as the search for pleasure and the avoidance of pain. Interest then becomes the principle that rules all passions of man as well as explains what is to be considered as a virtue or as a vice. Thus, the so called principle of interest starts from the physical sensualism and reaches morality and politics. It also states the need for legislation and education to ensure the connection of personal interest to the general interest of a nation, making them all compatible.

## **Keywords**

Helvétius, sensualism, interest, moral, education, law.

Helvétius faz parte de um corpo de autores do século XVIII que tinham por preocupação contrapor a ideia de que a moralidade e a política deveriam ser fundamentadas por concepções religiosas e dogmáticas. Tratava-se da busca de uma moral que não fosse fundamentada por dogmas religiosos, mas pela verdade sobre os próprios homens. Jacques Domenech nos fala da importância da fundamentação da moral no século XVIII, ressaltando alguns pontos cruciais da mudança de fundamento moral nesse século. Nesse sentido, Domenech aponta para a necessidade que muitos viam de laicizar a moral. Laicizar no sentido de “denunciar as regras morais erigidas em dogmas pelos padres da igreja” (DOMENECH, 1989, p.12). Tratava-se, então, de colocar motivos para a moral que fossem humanos e verdadeiros. Helvétius compartilhava desses pontos, afirmando que pretende tratar das “virtudes puramente humanas” e que, por isso, desconsideraria a virtude e a corrupção religiosas, pois estas não podem ser ligadas a questões morais e políticas amplas, como nos provam os costumes de várias nações que consideram como virtude religiosa o que na França, por exemplo, é tido como corrupção religiosa. A corrupção religiosa pode ser criminosa em determinado âmbito, mas ela não é incompatível com a felicidade dos membros de uma nação e, por isso, a moralidade religiosa não contribui para seus objetivos (HELVÉTIUS, 1973, p. 216). Essa negação de uma moral prioritariamente religiosa se deve a questões primordiais. Além de uma preocupação prática na forma como a igreja regia as normas de conduta, o principal nessa querela era trazer um fundamento universal que não dependia de uma concepção específica sobre um ser supremo, mas que correspondia a todo e qualquer

homem vivente. Não se trata, então, de uma negação da religiosidade, mas de garantir que a fundamentação da moral seja verdadeira e independente do que é apenas dogma. A moral, então, não se refere ao campo da religiosidade, mas, como veremos à frente, ao campo de uma ciência que tenha sua validade garantida pela experiência e que também seja útil aos homens e as nações em geral.

## O método

Tanto no prefácio de *Do espírito* (1758) quanto no de *Do homem* (1771), Helvétius deixa claro o objetivo de sua obra: trazer a verdade aos homens a fim de que possam estruturar a moral e, assim, tornarem-se felizes. A única forma de trazer a felicidade a eles é dar-lhes a instrução necessária e a legislação correta, que definem as regras morais. É indispensável que a verdade seja entregue aos homens para tirá-los da escuridão. Enquanto pautarmos as leis de conduta e a educação em um ideal longínquo e ilusório, não haverá avanços no campo da moral. E para que se possa retrair verdadeiramente a moral, devemos primeiramente retrair as próprias características do homem, por isso a investigação do *espírito*<sup>1</sup> do homem.

Devemos lembrar qual é o método destacado desde o prefácio de *Do Espírito*: “foi pelos fatos que cheguei às causas” (HELVÉTIUS, 1973, p. 179), ou seja, analisaremos os fatos aos quais temos acesso para, a partir deles, buscarmos suas causas, até o momento não explicitadas. Helvétius destacará especificamente que pretende tratar desse assunto assim como a física experimental trata de seu campo. É pela observação do que temos acesso e pela análise de seus desdobramentos que devemos alcançar as bases do problema, e, a partir disso, podemos fundar um conhecimento que nos auxiliará na melhora da sociedade. Helvétius pretende contemplar em seu trabalho tanto o interesse geral quanto a experiência; pretende que os princípios ali estabelecidos por ele estejam prioritariamente de acordo com esses expoentes. Dessa forma, sua obra pretende ser útil aos homens, totalmente ligada à realidade social e ainda pretende tratar da moral como as demais ciências, como exemplo, a física experimental. Ele pretende, então, analisar os fatos, como físicos analisam um experimento, e

---

1 - “O conhecimento do espírito, quando se toma esta palavra em toda a sua extensão, está tão estreitamente ligado ao conhecimento do coração e das paixões do homem, que seria impossível escrever sobre este tema, sem ter, pelo menos, que falar desta parte da moral comum aos homens de todas as nações, e que só pode ter, em todos os governos, apenas o bem público por objeto.” (HELVÉTIUS, C. Do Espírito, pág. 179.) O termo espírito será usado na obra de Helvétius em diferentes significações, ora mais abrangentes, ora mais específicas, dependendo do campo ao qual o autor se refere. Ao falar do espírito em relação ao indivíduo, ele é o conjunto de ideias e paixões descrito pela citação. Quando se trata da sociedade, o espírito é o conjunto de ideias novas e úteis à sociedade. Trataremos aqui, especificamente, do espírito em relação ao indivíduo.

daí retirar suas considerações para postular o princípio de conhecimento e ação do homem. Podemos lembrar os objetivos que Helvétius tinha em comum com outros filósofos da época, o de tornar a moral *sã* e afastá-la dos possíveis erros que o dogmatismo traz, por meio da ciência (DOMENECH, 1989, p. 11). Considerar a moral como uma ciência pretende ligar a moral à verdade acerca do homem. Por isso, a fundamentação da moral será colocada por um método científico capaz de nos trazer uma verdade demonstrável e condizente com “leis universais” obtidas por meio do que é comum no todo observado e que contempla todos os homens. Por isso ele inicia seu percurso na busca da formação das ideias do homem individualmente, ou seja, na investigação sobre o espírito.

Para que se investigue o espírito a fundo, devemos primeiramente pensar em como ele se forma. Como em Locke, é pela experiência que o homem é capaz de formar ideias. Ainda mais adiante, esta experiência será considerada especificamente a da sensibilidade física, como em Condillac. Helvétius tem como ponto inicial de sua definição do homem a tese sensualista de seu predecessor. Ela consiste em tomar como princípio de qualquer conhecimento, formação do entendimento do homem, e contato com o mundo, a sensibilidade física. Todas as características do entendimento que poderíamos tomar em separado serão vistas por este autor como formas da própria sensibilidade física.

## **A redução à sensibilidade**

Em *Do Espírito*, nosso autor demonstra como características tidas como individuais e constituintes do entendimento do homem (duvidar ou julgar, por exemplo) podem ser consideradas apenas formas de sensibilidade física. É interessante que vejamos que o percurso presente na obra *Do Espírito* é tomado, por Natalia Maruyama, como uma redução à sensibilidade. Essa redução tem a função de colocar um o ponto mais simples e originário da questão, o que condiz com o objetivo de fundamentar a moral de maneira precisa (MARUYAMA, 2005, p. 71). Para Helvétius, o homem é incapaz de não julgar, assim como é incapaz de não sentir. O julgamento não pode ser evitado na medida em que, ao abrir os olhos, já podemos diferenciar tamanhos e cores, por exemplo. Julgamentos sobre comparações entre objetos, sobre suas características físicas, como tamanho, são mais facilmente assimiladas à sensibilidade. Essa redução funciona como a desconstrução de teses de sua época, que supõem determinadas características (memória, razão, etc.) presentes no homem, como o entendimento, em separado do corpo e da sensibilidade. Veremos como isso aparece em sua obra.

Ele encontra primeiramente a sensibilidade e a memória no homem como o que daria a possibilidade de formação ao espírito e, conseqüentemente, à geração de nossas ideias. Ele recebe impressões de objetos exteriores através de sua sensibilidade física, que lhe causam prazer ou desprazer em relação a determinado objeto. Pela memória, ao deparar-se novamente com o mesmo objeto ou com algum objeto que remeta àquele, ele forma a ideia que tem desse objeto. Porém, mais à frente, ele assimila a memória a uma forma de sensação que estaria no presente, impulsionando o homem a sentir, novamente, de maneira mais enfraquecida, o que já sentiu antes. A própria memória seria sensibilidade física, na medida em que só ocorre por uma sensação. Mesmo lembranças relacionadas a ideias irão nos remeter à sensibilidade, visto que nos fazem sentir novamente o que sentimos quando formamos determinadas ideias. Por exemplo, quando nos lembramos de algo que comemos no passado é por causa de um aroma, sabor, ou outra sensação que estava presente no momento em que vivenciamos isso; ou quando nos lembramos de uma ideia abstrata é por sentirmos o mesmo que sentimos quando a obtivemos. A memória, entendida dessa forma, seria sensibilidade.

Depois de reduzir a memória à sensibilidade, nosso autor passa para as faculdades de nosso entendimento que, de maneira alguma, são vistas em separado. *“Ora, como o juízo é esta própria percepção, ou, pelo menos, o enunciado desta percepção, segue-se que todas as operações do espírito se reduzem a julgar.”* (Helvétius, 1973, p. 185). Operações como raciocinar, duvidar, fazer comparações e emitir juízos serão todas remetidas à última. Para este filósofo, o ato de julgar é o que ocorre em qualquer operação que estejamos exercendo e em relação a qualquer objeto, na medida em que definimos algo e delimitamos sua significação. Mesmo o julgamento é assimilado ao sentir. Todas as pretensas operações aqui citadas (duvidar, raciocinar, etc), seriam apenas formas de sentir. Elas, englobadas como julgamento, seriam reflexos do que é exterior e que atinge o homem de alguma forma. As operações da mente deixam de apenas servirem-se das sensações, mas são constituídas por elas. Nesse sentido, tanto Condillac como Helvétius radicalizam o empirismo de Locke. O homem permanece se desenvolvendo por meio de suas experiências, mas é especificamente da experiência sensível que esse desenvolvimento ocorre, e a sensibilidade física permanece, inclusive nas decisões sobre julgamentos morais. As operações que poderiam ser tomadas como pertencentes a algo como o entendimento se constituem da sensibilidade porque elas são o próprio desenvolvimento da sensibilidade.

Temos aí o primeiro vislumbre do que constitui o espírito: ele é totalmente pautado pela sensibilidade física. Em *Do Homem*, Helvétius passa de uma descrição do espírito para a descrição do próprio homem. O homem nasce ignorante e, como veremos à frente, é pela instrução que ele pode deixar esse estado (considerando a instrução advinda da própria sensibilidade física, já que é por ela que desenvolvemos a capacidade de julgamento). A

radicalidade desse modo de pensar nos faz chegar à sentença: “..no homem tudo é sentir.. A sociabilidade não é mais que uma consequência desta mesma sensibilidade” (HELVÉTIUS, 1771). Tal sentença traz desdobramentos que remontam da diferenciação entre o homem e o animal até a possibilidade da moral e da política.

## **A diferenciação entre homens e animais**

O homem parece nascer sem nenhuma característica que o diferencie do animal. A partir do que recebe pelos sentidos, pelo que lhe causa prazer ou dor, o espírito do homem vai se desenvolvendo e sendo capaz de formar ideias complexas, auxiliando sua capacidade de julgar e levando-o, então, à moralidade e à política.

Podemos nos perguntar se os animais são capazes de julgamentos complexos como os que são feitos pelos humanos, já que eles também possuem sensibilidade física, que é o que permite o desenvolvimento do espírito. Helvétius nos dirá, em *Do Espírito*, capítulo I, que a organização física de nossos corpos é o que define nossa diferenciação dos animais. Nesse sentido, por não possuírem polegar opositor, por exemplo, os animais não seriam capazes de desenvolver-se como os humanos. A única diferença trazida aqui pelo autor entre os homens e os animais é a diferença físico/anatômica, dividida em 4 tópicos: diferenças entre mãos e patas para manejar instrumentos; duração da vida; o número menor de necessidade que os animais possuem, pois a natureza já os deu força suficiente para se defender, e também características físicas que lhe trazem maior conforto; a inferioridade de domínio dos animais em relação às conquistas do homem não lhes permitem um maior desenvolvimento e, finalizando, o homem adapta-se melhor a vários ambientes, podendo estender seu habitat. Dessa forma, vemos que o que diferencia homens e animais não é, de forma alguma, essencial, mas é, na verdade, uma diferença de grau. O homem não é de uma natureza diferente do animal, ele é apenas mais desenvolvido fisicamente, o que possibilita uma formação de ideias mais desenvolvidas.

## **A alma e o espírito**

Um ponto importante na diferenciação de homens e animais relaciona-se com o método científico de Helvétius e, também, com suas considerações sobre o que se denomina geralmente como alma. No apêndice II de *Do Espírito*, ele nos fala ironicamente da ideia de

que há no homem algo outro presente em seu cérebro que não existe nos animais. Tal ideia parece a ele muito incoerente, pois precisar de um elemento a mais além da sensibilidade do homem para lhe conferir ideias não seria considerá-lo mais avançado que o animal, já que o animal consegue sobreviver, se alimentar, obter prazer e até formar ideias e paixões sem o auxílio de um “terceiro”. Esse terceiro é, para Helvétius, o que geralmente se considera a alma. Para ele, pressupor que uma inteligência superior tivesse que colocar em nós algo como a alma diminuiria essa inteligência, já que ela seria incapaz de por no próprio corpo dos homens diretamente a possibilidade de formar suas ideias. Propor a necessidade de algo que opera no homem, mas que é separado de seu corpo, faz com que essa inteligência superior precise, a cada concepção, criar pequenas almas e colocá-las no corpo do homem, e ainda dar algum destino a elas depois de sua morte. Tal engenho é desnecessário para Helvétius, e mais, é incompatível com uma aceção primordial para as ciências da época: a de que a natureza age pelas vias mais simples, curtas e rápidas para atingir a maior variabilidade, pois seria necessário um elemento a mais (a alma) para se atingir os mesmos resultados. Por isso, é demarcado que não há nada no homem em separado da sensibilidade física de seu corpo: nem a alma e nem mesmo o espírito. O espírito não está no homem como algo além de sua sensibilidade. Ele não existe previamente e é preenchido posteriormente por ideias. Ele só passa a existir se recebemos uma sensação que nos propicie prazer ou dor, e assim formamos a ideia. Ele é desdobramento da própria sensibilidade e inclui ideias julgamentos ou paixões.

## As paixões

Vemos nas paixões um ponto marcante para o tema tratado aqui. Como as ideias, as paixões constituem o espírito e são formadas a partir da sensibilidade física, mas, além disso, outro ponto é importante: *“As paixões são na moral o que o movimento é na física: cria, destrói, conserva, anima tudo; sem ele, só há morte. São elas também que vivificam o mundo moral”* (HELVÉTIUS, 1973, p. 262). É na força das paixões que podemos encontrar as inclinações e possibilidades do homem. A força das paixões seria proporcional à força de suas ações. O movimento do qual nos fala Helvétius é propriamente o que nos põe a agir e nos direciona segundo o que somos apaixonados. Ele acontece de forma dupla: a primeira, já colocada, em direção às paixões, e a outra, contra a direção do tédio, que nada mais é que ausência de estímulo, considerado como uma dor. Então, ainda lidamos com a dupla inclinação vinda da sensibilidade física, que nos leva ao prazer e nos afasta da dor, mas de forma mais complexa, por meio de paixões formadas em nosso espírito. *“O momento em que a paixão desperta mais fortemente em nós constitui aquilo que chamamos sentimento. Entende-se por paixão apenas*

*uma continuidade de sentimentos da mesma espécie.*” (Helvétius, 1973, p. 294) Dessa forma, enquanto o sentimento<sup>2</sup> é localizado e pontual, a paixão seria a reunião de sentimentos pelo mesmo objeto. O sentimento acontece imediatamente ao nos depararmos com uma sensação prazerosa e, na recorrência desse sentimento, nos apaixonamos por esse tipo de sensação e, a partir disso, somos impulsionados a procurar o que é objeto dessa paixão.

A impulsão ocorrerá a partir de dois tipos de paixões: as paixões naturais – da necessidade – e as paixões fictícias – supostas pela sociedade. As paixões naturais são as que atendem apenas à nossas necessidades físicas imediatamente. Tais paixões são compartilhadas por todos os homens e a data de sua origem é a mesma da sensibilidade física (respeitadas todas as diferenças moldadas pelos hábitos, todos possuem tais paixões). Já as paixões fictícias, são as que conhecemos como amor, ambição, ganância, etc. Elas têm como objetos de desejo os artificios, as pessoas, as coisas ou as ações. Mesmo que essas paixões sejam construídas socialmente, na verdade, sua raiz é comum às paixões naturais: a busca do prazer e a fuga da dor. É o que vemos nesta passagem de *Do Espírito*: “*Vou, portanto, seguindo a metamorfose das penas e dos prazeres, tais como a avareza, a ambição, o orgulho, a amizade, cujo objeto não parece fazer parte dos prazeres dos sentidos, é, no entanto, sempre a dor e o prazer físico que buscamos ou de que fugimos.*” (HELVÉTIUS, 1973, p. 269). Helvétius vê nas paixões ditas fictícias uma metamorfose de objeto. Ou seja, o homem passa a considerar o meio para saciar seus prazeres como próprio objeto de desejo e, dessa forma, como um prazer fictício.

Vemos, nessa divisão das paixões entre naturais e fictícias, a passagem clara entre o sensualismo e o âmbito social e, obviamente, também ao âmbito moral. Primeiramente, possuímos as paixões naturais, que são compartilhadas por todos e posteriormente, desenvolvemos as paixões fictícias, formadas socialmente e que variam entre os homens. As paixões que serão desenvolvidas são dependentes da sociedade em que se vive e da instrução que se recebe. É o que fica claro quando, no segundo capítulo de *Do Espírito*, Helvétius estabelece uma relação entre a forma de governo e a necessidade das artes. Os governos que reprimem determinados sentimentos ou paixões, que não os exaltam, fazem com que seus cidadãos busquem imagens de tais sentimentos pelas artes. “*Nos países em que os germes de certas paixões e sentimentos se acham abafados, o público só pode vir a conhecê-los nos quadros dos escritores célebres e sobretudo através dos poetas*” (HELVÉTIUS, 1973, p. 295).

Nesse momento, precisamos reforçar a importância das paixões no que tange a essa passagem da sensibilidade física ao âmbito moral. Como vimos, as paixões movimentam

---

2 - Os sentimentos são divididos entre os que se referem a necessidades físicas – chamados sensações – e os que se referem às paixões construídas. Dessa forma, vemos que diferente de Rousseau e Diderot, a diferença real entre sensação e sentimento é apenas classificatória, mas ambas são consideradas sentimentos. Não há diferenciação entre elas, a não ser no objeto a que se referem.

o mundo moral porque são elas que impulsionam as ações do homem e as direcionam a objetos determinados. Dependendo das paixões a que regem os homens de determinada sociedade, as ações que eles possam vir a cometer serão prejudiciais a esta sociedade ou favoráveis a ela. Aqui temos um ponto interessante que condiz com os objetivos práticos da filosofia de Helvétius: se são as paixões - tão ligadas ao sensualismo - que tornam as virtudes verdadeiramente fortes, vemos que não há contraposição entre a natureza humana - basicamente regida pela sensibilidade física - e a moralidade. Ao contrário, há continuidade entre elas: *“tratava-se de mostrar, mais precisamente, que as paixões e os prazeres não contradizem, por eles mesmos, o que podemos chamar de “bem”, “justiça, ou virtude, pelo simples fato de que são as bases de todas as ações, inclusive das boas ações.”* (MARUYAMA, 2005, p. 144.).

Isso que ocorre na obra de Helvétius está em conformidade com uma consequência primordial da obra de seu predecessor Condillac que, segundo Monzani, irá delinear um caminho para vários autores do fim do Séc. XVIII: *“O Traité des sensations mostra de forma inequívoca o primado da dimensão passional sobre a dimensão teórica (...) Nela, o teórico se subordina à dimensão prática, e é na camada mais originária, a das afecções elementares, prazer/dor, das necessidades e dos desejos, que brota um sentido original, primordial, mas que será determinante.”* (MONZANI, 1996, p. 208).

Trata-se de uma sobreposição do âmbito passional em relação ao âmbito teórico. Decorrente de aceções que remontam a Locke, mas que irão se cristalizar da obra de Condillac, o homem não se determina primeiramente por uma reflexão e posteriormente decide como agir. Essa sobreposição do prático acontece por um princípio bem definido, que já demonstramos anteriormente neste artigo: a busca pelo prazer e a fuga da dor.

Helvétius é herdeiro dessa importante consequência da obra de Condillac e toma a busca pelo prazer como princípio que determina e move o homem em toda e qualquer situação. Juntamente com La Mettrie, Helvétius dá um passo adiante de Condillac ao considerar esse princípio não apenas em relação a um indivíduo, mas busca as dimensões ético-políticas desse princípio.

## O interesse

Temos, então, uma busca pelo prazer e a fuga da dor como uma tônica que permanece na constituição de nossas ideias, paixões e ações. Essa tônica é reconhecida por ele como o *princípio do interesse*. Esse interesse primordial do homem remete, automaticamente e novamente, à sensibilidade física. Seja na formação do espírito, nos julgamentos, na tomada de ação ou

nas considerações sobre virtude, é a sensibilidade física que constitui o homem. Como vimos na metamorfose das paixões, o interesse não é sempre visto da mesma forma: *“o interesse é na terra o poderoso sedutor, que modifica aos olhos de todas as criaturas a forma de todos os objetos”* (HELVÉTIUS, 1973, p. 206). O interesse, assim como modifica nossa atenção ao olhar os objetos e dá aos homens ideias diferentes, também é motor de nossas ações. Mesmo que haja a modificação desse interesse, e a diferenciação dos modos de atingir a felicidade em um ou outro indivíduo, o interesse dos indivíduos aponta sempre para a própria felicidade.

Podemos fazer um breve retorno à descrição do sensualismo para, agora, clarificar a tese “julgar é sentir”, com o auxílio do princípio do interesse. Mesmo os julgamentos morais são considerados advindos da sensibilidade física, porque o homem é regido por este interesse primordial que é buscar o prazer e evitar a dor. Todos os nossos julgamentos pautam-se exclusivamente sobre esse princípio e, visto dessa forma, não podemos desatrelar o julgamento da sensibilidade.

Assim, o princípio do interesse mantém-se, apesar de todas as metamorfoses que pode sofrer por conta da variabilidade das paixões. Ele é o pano de fundo que todos os homens recebem igualmente da sensibilidade: *“o leitor esclarecido sentirá que tomo esta palavra [interesse] em um sentido mais extenso e que a aplico geralmente a tudo o que nos pode proporcionar prazeres ou nos evitar dores”* (HELVÉTIUS, 1973, p. 202).

## **O interesse como base da moralidade**

Como dissemos anteriormente, a obra de Helvétius é direcionada às preocupações de cunho prático e moral, bem como às consequências ético-políticas, ressaltadas por Monzani. Dessa forma, o mesmo princípio que impera na formação de ideias e paixões impera também na moralidade, já que o que parte da sensibilidade física permanece na moralidade: o interesse pessoal é o que define a conduta e, conseqüentemente, torna o homem vicioso ou não. Helvétius é, segundo Domenech, *“o primeiro que fundou a moral sobre a base inabalável do interesse pessoal”* (DOMENECH, 1989, p. 35). Porém, ao fundar a moralidade pelo princípio do interesse pessoal, temos a colocação de algumas consequências à moralidade que podem se mostrar problemáticas.

Tais consequências são vistas na necessidade primordial na fundamentação da moral a partir do interesse: coadunar os interesses pessoais para que se possa viver em sociedade e para que essa sociedade progrida. *“Em matéria de moral como em matéria de espírito, é o interesse pessoal que dita o juízo dos indivíduos e o interesse geral que dita o das nações”* (HEL-

VÉTIUS, 1973, p. 203)

O interesse geral e o interesse pessoal devem estar totalmente articulados e não podem ter discordâncias. Para que a moral fundamentada no interesse pessoal possa ser posta em prática, o interesse geral não pode estar em confronto com o interesse pessoal, mas, sim, ter estreita ligação com o mesmo. Tal ligação entre o interesse pessoal e o interesse geral é o que Domenech chama de “interesse bem compreendido”. Trata-se, então, de que se compreenda o princípio do interesse de forma mais ampla, e se ultrapasse a noção de interesse pessoal e atinja o âmbito público. A moral só pode então ser constituída desse interesse bem compreendido, não desligando, assim, a ideia de interesse geral à de interesse pessoal.

Nessa vinculação do interesse pessoal ao interesse geral, a moralidade ganha uma variabilidade que depende da sensibilidade e dos interesses pessoais dos indivíduos que constituem esta sociedade. Essa variabilidade é vista nas considerações do autor sobre a virtude.

## Virtude e probidade

Veremos, em breve, que o mesmo princípio que rege as paixões e as ações do homem irá definir, em primeira instância, o que deve ser considerado virtuoso ou não. Assim como nas paixões, temos virtudes verdadeiras e virtudes de preconceito. Além disso, Helvétius dirá que é das *fortes paixões* que surge a verdadeira virtude (ou probidade), aquela que é esclarecida e ativa (HELVÉTIUS, 1973, p. 274) e consiste em atos úteis à felicidade do homem e, consequentemente, de seu estado. Esse tipo de virtude depende necessariamente das *fortes paixões*, enquanto paixões não tão vívidas, paixões que Helvétius atribui aos “homens de bem”, podem levar apenas a uma virtude passiva, que de forma alguma é a verdadeira: “...a força de seus vícios ou de suas virtudes será sempre proporcional à vivacidade de suas paixões, cuja força se mede pelo grau de prazer que ele encontra em satisfazê-las” (HELVÉTIUS, 1973, p. 276).

Se a virtude é dependente das paixões e, consequentemente, do interesse pessoal, parece não haver espaço para um conceito fixo de virtude que julgue atos específicos sempre da mesma forma, como virtuosos ou viciosos. Em *Do Espírito*, depois de definir o espírito como o conjunto de ideias do homem, Helvétius parte para as considerações do espírito em relação à sociedade. Ele constata que, devido a toda ação estar embasada no interesse, é por ele que também devemos julgar um grande espírito. O que se julga não são as intenções de um homem, mas suas ações. Helvétius trata, então, do conceito de probidade, que seria a virtude aplicada à prática. Dessa forma, “de dois homens, o mais honesto em suas ações é, muitas vezes o menos apaixonado pela virtude” (HELVÉTIUS, 1973, p. 204). O homem que possui

probidade a possui por suas ações úteis ou inúteis aos seus interesses, primeiramente, mas a verdadeira probidade se dá em relação ao interesse geral. A probidade é o hábito de ações particularmente úteis à sua sociedade. Ora, para que essa probidade seja atingida, ela deve estar em conformidade com a probidade em relação a cada indivíduo, ou seja, o interesse individual deve estar associado ao interesse geral e nunca ser contrário a ele.

Devemos lembrar o quão variável pode ser a utilidade para cada nação, e para cada homem. Parece estar posto aí um relativismo moral, do qual Helvétius será acusado várias vezes (DOMENECH, 1989, p. 36). Desse ponto de vista, parece arbitrário definirmos a virtude. Nesse sentido, Helvétius parece primeiramente pensar na impossibilidade de pensarmos a moral universalmente, postulando um fundamento válido para qualquer território. Porém, como já vimos, o interesse pessoal é colocado como princípio primeiro da moralidade e ele delinea a forma como Helvétius considera a virtude. Sua definição de virtude será anterior a qualquer norma de conduta, mas também não as excluirá: “... *por esse termo virtude, apenas se pode entender o desejo da felicidade geral...*” (HELVÉTIUS, 1973, p. 211). Dessa forma, Helvétius, ao mesmo tempo em que adequa a definição de virtude ao moto inicial do homem, pode adequá-la também ao fato das grandes diferenças entre os costumes das nações. A virtude se torna moldável em relação ao contexto histórico analisado.

Mas isso não colocará qualquer ação como virtuosa? Não. Helvétius dá como exemplos vários costumes de outros povos que parecem bárbaros. Mas considera que haver um motivo útil para que aquele costume tenha se instituído não é, de maneira alguma, aceitá-lo incondicionalmente. Por isso, considerar as paixões fonte da virtude, o interesse pessoal base do que é considerado virtuoso e da própria moral, acarreta em uma grande possibilidade de variabilidade sem que se perca a moral. Assim, ao constatar a grande variabilidade do que é considerado virtude, o conceito não é abandonado, e nem acarreta normas de conduta positivas. Ele se mantém, no fundo, comum a todos os homens e nações. A virtude é um uníssono que deve se manter, independente das formas que ela irá adquirir.

## **A moral**

Tomar o interesse como única permanência em uma constituição moral dá a ela uma maleabilidade tão abrangente que não podemos calcular quais são as consequências de tal tese. Mesmo que ela se baseie primeiramente na sensibilidade física, o que poderia fazer com que nenhum sistema moral fosse tão complexo, a possibilidade de desenvolvimento do julgamento – como próprio desenvolvimento da sensação – e a metamorfose que ocorre

em nossas paixões fazem com que haja uma variabilidade no próprio interesse e na própria utilidade que tornam a moralidade e a política complexas.

Nesta altura do percurso, vemos que, para que os homens sejam virtuosos, há a necessidade de uma passagem do interesse pessoal ao interesse geral, que não ocorra de forma abrupta. Deve-se fazer o que é útil a todos, inclusive a você mesmo, e, para isso, não se pode perder de vista que o interesse geral deve partir dos interesses pessoais de quem o compõe.

Mas, então, como formular a moralidade se não parece haver nenhuma imposição positiva para o que consideramos bom, justo, ou virtuoso? Como a sensibilidade física pode garantir a probidade dos indivíduos? Como é possível uma conciliação branda entre interesse pessoal e geral, se a sociedade é constituída de indivíduos movidos apenas pelo interesse pessoal? Podemos analisar os problemas que a junção do sensualismo e da moral trazem em dois níveis.

O primeiro, em relação ao indivíduo na formação do interesse pessoal: como fazer com que um indivíduo que recebe tantas sensações diferentes torne-se um indivíduo capaz de moralidade? Trata-se de questionarmos a forma como a sensibilidade física recepciona as sensações que recebe, para daí formar uma noção de interesse pessoal que propicie a junção dela à noção de interesse geral.

O segundo, publicamente em relação à formação do interesse geral: como unir todos os interesses pessoais em prol de um interesse geral que não seja castrador e que leve ao progresso de um Estado? Como serão preservados os interesses pessoais em um Estado? Aqui questionamos como ocorre dentro de um Estado a noção de interesse geral que consiga abarcar os interesses pessoais dos que ali residem.

Não há mesmo nenhuma garantia natural da conciliação entre os dois tipos de interesse. Mas também não há nada que impeça auxiliares artificiais para que ela ocorra: a legislação e a educação.

A primeira forma de problematizar a ligação entre o sensualismo e a moral é respondida por Helvétius pelo conceito de educação. Para nosso autor, como já foi dito aqui anteriormente, a instrução é indispensável para que os homens tornem-se felizes. Ela parece ser o primeiro ponto que pode levar a sensibilidade a constituir um espírito e torná-lo político. O primeiro ponto por que a educação, tal como Helvétius trata, não é apenas a educação institucional. Amplamente entendida, a educação acontece desde a primeira sensação do homem. Ela começa com o primeiro movimento da criança e só termina com a morte. Dessa forma, nossos sentidos são os professores, e, por isso, ninguém pode receber a mesma educação (HELVÉTIUS, 1771). Com isso, o que entendemos como educação institucio-

nal é posterior à educação de nossos sentidos e não deve ser contrária a ela, mas, sim, sua continuação. O que se recebeu dos sentidos não pode nunca ser apagado pela educação institucional, já que o papel dela é trazer os interesses pessoais à tona, a fim de regulá-los ao interesse geral. Parece-nos, então, que a educação institucional também não pode ser castradora. Se ela não pode anular o que recebemos da educação que temos desde o primeiro movimento, ela deve adaptar o indivíduo e auxiliá-lo em sua interpretação do que é interessante. Essa educação é capaz de desenvolver o espírito do homem e torná-lo sociável. É notável que essa noção de educação também atrela o sensualismo à moral, já que se inicia com a primeira sensação. Por isso, a educação nunca pode contrariar o princípio do interesse pessoal e tudo que não traga prazer aos indivíduos deve ser descartado de seu domínio. Ela faz com que quem é educado aprenda a obter seu prazer físico pelas mesmas vias em que se conquista o interesse geral.

A segunda forma de problematizarmos a obra de Helvétius é respondida pela perfectibilidade das leis. A legislação faz com que o interesse geral descrito nela não seja incoerente em relação aos interesses pessoais dos participantes dessa nação. Trata-se não de impor algo que deva ser seguido pelos cidadãos, independentemente de seus interesses pessoais, mas de coadunar os interesses pessoais e fazer com que eles propiciem o melhor para toda a nação. Isso só pode ser feito com uma profunda análise dos costumes da nação que será legislada, o que fica a encargo do legislador, que deve ter a habilidade de adequar o interesse geral a fim de fazer com que os homens não vejam o Estado como inimigo, mas, sim, que seja impossível não agir de acordo com o interesse geral aos buscar seus interesses pessoais.

Ao dizer que o homem não é essencialmente mau, Helvétius nos mostra que, se os homens fazem ações más, é pelo trabalho mal feito dos legisladores, que, ao escreverem as leis, pensam primeiramente no interesse pessoal, sem ligá-lo ao interesse geral, que é o interesse ao qual as leis devem se reportar. Justamente pela incompatibilidade entre o que lhes diz a natureza e o que lhes diz a legislação, os homens devem seguir. É com essa pretensa oposição que surgem as más ações humanas. Então, ele coloca exemplos da relatividade de julgamentos de determinadas ações. Um homem que ajuda um familiar seu a defender-se da acusação de assassino, um ministro que dá a seus amigos e parentes bons cargos, são julgados por sua família, que foi favorecida pela ação, como bons homens, enquanto que, pelo interesse geral, são desonestos e prejudiciais. Ora, tais ações só são possíveis pela má legislação e pela oposição entre o interesse pessoal e o interesse geral. Assim, torna-se raro o verdadeiro amor por sua pátria.

É papel do legislador formular leis que correspondam ao interesse geral da nação, fazendo com que os homens possam usufruir de suas mais fortes paixões para serem virtuosos

e agirem com probidade, pelo mesmo caminho pelo qual buscam o prazer. Isso também inclui sempre reavaliar as leis de seu Estado para que elas estejam em conformidade com as mudanças que ocorrem através do tempo. A variabilidade da moral também coloca a necessidade de uma constante revisão das normas de conduta, pois o que foi útil um dia pode deixar de ser em outro.

A legislação e a educação parecem trabalhar conjuntamente em direções diferentes na mesma via para promover a ligação entre os interesses pessoais e o interesse geral: a educação vem auxiliar o indivíduo em seu desenvolvimento, adequando o interesse pessoal ao interesse geral, enquanto a legislação vem no sentido inverso: adequar o interesse geral ao interesse pessoal, tornando o interesse público interessante também aos indivíduos isoladamente. Ressaltamos aqui que o uso desses dois auxiliares acontece durante todo o processo de desenvolvimento do homem: da formação das ideias, paixões, atos e considerações sobre a virtude, por exemplo. Elas proporcionam o acabamento final que visa garantir a possibilidade de uma moralidade que não seja contrária à definição primária do homem como sensível.

Ao tomar a verdade sobre o homem como o que pode nos garantir uma vida em sociedade que leve os homens à felicidade, Helvétius formula uma ideia de moral composta por conceitos que, sempre remontando à sensibilidade física, são capazes de ser adequados a formas variáveis de sistemas políticos. Isso ocorre, prioritariamente, pela busca de um princípio que, de tão inegável e mínimo, pode ser reconhecido em toda formação do indivíduo, mesmo que muito profundamente. Essa característica é passada para a consideração da moralidade, que toma, da mesma forma, um princípio mínimo e inegável, para daí construir a sociabilidade que leve à felicidade. Pelo sensualismo se manter em todos os pontos da moralidade por meio do princípio do interesse, a universalidade da moral é atingida e coloca a moral como pano de fundo de um edifício a ser construído de modo dependente das sensações de seus habitantes.

## Referências bibliográficas

CHAUSSINAND-NOGARET, G. 2009 *Lês Lumières ao péril du bûcher: Helvétius et d’Holbach*. Paris: Fayard.

DOMENECH, J. 1989. *L’Éthique des Lumières: Les fondements de la morale dans la philosophie française du XVIII<sup>e</sup> siècle*. Paris: Librairie Philosophique J. VRIN.

HELVÉTIUS, C-A. 1989. *De l'Homme*. Paris: Fayard.

\_\_\_\_\_. 1973 *Do Espírito*. Coleção Os Pensadores. São Paulo: Abril Cultural.

LADD JR, E. C. 1962. Helvétius And D'Holbach: La Moralisation De La Politique. *Journal of the History of Ideas*, Philadelphia: University of Pennsylvania Press, 23, ap. - je, p. 221 – 238.

MARUYAMA, N. 2005. *A moral e a filosofia política de Helvétius: uma discussão com J.-J. Rousseau*, São Paulo: Humanitas.

MONZANI, L. R. 1995. *Desejo e Prazer na Idade Moderna*. Campinas, SP: Editora da Unicamp.

\_\_\_\_\_. 1996. Origens do Discurso Libertino. In: NOVAES, A. (Org.) *Libertinos Libertários*. São Paulo, SP: Companhia das letras.

PUIGGRÓS, R. 1945. *Los Enciclopedistas – Diderot – Holbach – Helvetius*. Buenos Aires: Futuro.

ROSTAND, J. 1951. La conception de l'homme selon Helvétius et selon Diderot In: *Revue d'histoire des sciences et de leurs applications*. Paris, Presses Universitaires de France. Tome 4 n°3-4. pp. 213-222.

SILVA, G. F. 2006. *Moral do interesse e a crítica rousseauiana*. In: SANTOS, A. C. (Org.). *História, pensamento e ação*. Aracaju/SE: Editora da Universidade Federal de Sergipe, v. I, p. 261-273.

# **Robôs podem ser moralmente responsabilizados? Reflexões à luz do pragmatismo de John Dewey**

## **Can robots be morally responsible? Reflections from John Dewey's pragmatism**

*Fernando Cesar Pilan  
Universidade Estadual do Norte do Paraná*

### **Resumo**

Os dois principais objetivos do presente trabalho são: (1) brevemente apresentar os avanços da Inteligência Artificial e (2) problematizar o “dilema moral” que esses avanços geram, buscando contribuições na Teoria Moral de John Dewey. Na esteira dos modelos éticos tradicionais, esse tipo de debate parece não encontrar muitos avanços, pois se assenta na noção de dever moral e imputação de culpa. Nosso objetivo é mostrar que, no viés filosófico pragmatista de Dewey, seria possível uma reflexão sobre a conduta no plano da ação não mais detida no domínio representacionista, no plano do dever. Para Dewey, se um ser é capaz de atualizar seus hábitos, este pode, em algum sentido, ser considerado responsável por suas ações, pois os procedimentos morais não se distinguiriam dos procedimentos cognitivos. Nessa perspectiva, precisaríamos de uma redefinição da noção de responsabilidade, não mais pautada na imputação de culpa, mas na correção de hábitos. Acreditamos que a sugestão de Dewey, de que a conduta seria um processo dinâmico e intimamente ligado à experiência e à cognição, nos oferece ferramentas para pensar alternativamente a questão da responsabilidade no contexto filosófico contemporâneo.

## **Palavras-chave**

Inteligência Artificial; cognição incorporada e situada; responsabilidade; moral; pragmatismo.

## **Abstract**

The two main goals of this work are: (1) briefly to explain how the advances in the Artificial Intelligence has allowed the development of machines with increasingly performances similar to those of living beings, and (2) before these advances, present the moral dilemma generated by them seeking contributions in Moral Theory of John Dewey. In the approach of traditional ethical models this kind of debate does not seem to find many advances, because the moral duty is taken as the starting point of responsible in dualists terms. To Dewey's pragmatism any robots are able to update our habits, they may in some sense be held responsible for their actions: moral procedures are not distinguished from cognitive procedures. We believe that Dewey's suggestion in which the conduct is a dynamic process and deeply linked to the experience offers us tolls to think the question of responsibility alternatively in the contemporary philosophical context.

## **Keywords**

Artificial intelligence; embodied and situated cognition; responsibility; moral; pragmatism.

## **Introdução**

Um dos principais objetivos da chamada “Inteligência Artificial Forte” (IA) nas últimas décadas tem sido desenvolver máquinas capazes de imitar ações e condutas tipicamente de seres vivos. Esse objetivo tem levado a IA a conseguir avanços significativos, principalmente no que se refere ao desenvolvimento de máquinas capazes de agir no mundo de forma cada vez mais sofisticada e autônoma. Um dos problemas que emergem desse panorama atual da IA diz respeito ao âmbito moral destes seres artificialmente desenvolvidos. A partir desse pano de fundo, nossos principais objetivos serão: (1) apresentar uma perspectiva recente em IA – a abordagem da Cognição Incorporada e Situada (C.I.S.) –, considerada um marco

no desenvolvimento de máquinas com condutas e ações cada vez mais próximas às de seres vivos; e (2) mostrar a necessidade de um debate ético não mais pautado nos modelos filosóficos tradicionais, sugerindo possíveis contribuições da teoria da ação contemporânea do pragmatista John Dewey.

## **A abordagem da Cognição Incorporada e Situada como alternativa à Ciência Cognitiva Tradicional**

Por um lado, podemos compreender que a Ciência Cognitiva e a Inteligência Artificial como projetos de pesquisa datam da década de 1950. Nesse sentido, seriam áreas muito recentes que teriam em comum o interesse em compreender a natureza e o funcionamento dos processos cognitivos, com o intuito de construir modelos capazes de imitar nosso comportamento. Por outro lado, podemos também compreender esses dois campos como possuindo uma origem muito mais longínqua. Segundo Georges Vignaux (1991, p. 15),

*O desejo de realizar máquinas inteligentes data da antiguidade. E, sem dúvida, uma das primeiras menções do gênero poderia encontrar-se no canto XVIII da Ilíada, no qual Hefesto, deus do fogo, tinha construído mesas com três pés munidas de rodízios, que se deslocavam sozinhas nos palácios dos deuses, uma espécie de robots antes do tempo!*

Como exemplos de que a preocupação com a construção de máquinas inteligentes é um anseio humano de longa data, o autor também cita a construção de um autômato chamado *Golem*, feito de madeira e argila, do qual encontramos registro na tradição judaica. Também podemos mencionar a ideia de um “animal-máquina”, proposta pelo mecanicismo cartesiano do século XVII. Enfim, com essas elucidações, Vignaux pretende mostrar que, diferentemente do que muitos afirmam, as preocupações da IA e da Ciência Cognitiva não são produtos do século XX, mas constituem um corpo de problemas relacionados à natureza dos processos cognitivos que inquietam a humanidade há séculos.

Apesar da inquietação com os problemas relacionados à mente poder ter sua origem identificada há séculos, foi somente a partir dos avanços lógico-matemáticos do início do século XX que tivemos o marco de surgimento da IA e da Ciência Cognitiva enquanto áreas de pesquisa. Em primeiro lugar, destacamos que o surgimento da Ciência Cognitiva, conforme definido por Dupuy (1996, p. 9), pode ser datado em 1940, oriundo principalmente

das “Conferências Macy”. Estas foram uma série de conferências levadas a cabo por um grupo de pesquisa formado por lógicos, matemáticos, engenheiros, fisiologistas, psicólogos, antropólogos e economistas que buscavam empreender uma *ciência geral do funcionamento da mente*. Em sua origem enquanto área de pesquisa, a Ciência Cognitiva recebeu o nome de Cibernética, sendo desenvolvida, segundo Dupuy, a partir do choque provocado pelas descobertas lógicas da década de 1930, possibilitando avanço nas áreas de investigação sobre o conhecimento. Destacamos que as referidas descobertas lógicas foram devidas, em parte, às abordagens alternativas surgidas na filosofia contemporânea, que reconheceram a relevância dos processos e fenômenos contingentes no âmbito da investigação filosófica. Um marco que pode ser sublinhado foi a “Virada Naturalista”<sup>1</sup>, proveniente do darwinismo e sistematizada pelo pragmatismo de C. S. Peirce, W. James e John Dewey, quando propõe a mudança de um padrão lógico meramente formal, fixo e de valores de verdade estáticos, para um padrão lógico natural, contingente, movido pelo acaso, que permearia todos os processos evolutivos naturais.

Destacamos que, mesmo tendo influenciado o surgimento das lógicas não clássicas, a filosofia naturalista, que pretendia aproximar lógica e filosofia dos processos evolutivos, relacionando-as à vida, acabou sendo deixada de lado pelos avanços que se seguiriam no âmbito da Ciência Cognitiva. Acreditamos que isso tenha ocorrido principalmente devido ao que Dupuy chama de “fascínio pelo modelo”. Esse encanto pela modelagem algorítmica fez com que, após, a originária cibernética fosse superada, principalmente, a partir dos trabalhos de Herbert Simon, quando tivemos o início da Ciência Cognitiva conforme a concebemos hoje. Passou-se, então, a pressupor os processos cognitivos como reproduzíveis a partir de programação de funções normativas, sem a necessidade de se levar em consideração o plano biológico e corpóreo.

*Em clara oposição às ideias relacionais da cibernética a abordagem simbólica, foca-se num produto final (resolução de um problema), ignorando por completo o processo. O processo cognitivo resume-se, portanto, ao processamento sequencial da informação de entrada e a consequente transformação desta informação numa saída, que é a*

---

1 - A origem da “virada naturalista”, segundo Gonzalez e Broens (2011), pode ser datada na metade do séc. XIX, principalmente com a publicação da obra *A origem das espécies*, de Darwin. Conforme ressalta Dewey em seu artigo *The influence of Darwinism in Philosophy* (1909), Darwin teve grande influência no pensamento filosófico e científico, o que propiciou o surgimento de uma *nova lógica* de investigação, não mais pautada em pressupostos fixos, imutáveis, sobrenaturais e transcendentais. A virada naturalista trouxe novidades em filosofia como, por exemplo, a aproximação da reflexão filosófica dos problemas cotidianos (*os life problems*) e o diálogo com as ciências como a biologia, a química e a física em prol de uma concepção de mundo dinâmica, evolutiva e contingente.

*resposta do sistema ou resolução do problema.* (PINTO, 2007, p. 14)

Como vemos, a partir dos trabalhos de Marvin Minsky, Allen Newell e Herbert Simon na década de 1950, o principal intuito da Ciência Cognitiva passou a ser a simulação dos processos cognitivos em termos de meios e fins, manipulando tão somente regras e símbolos digitais. O chamado *General Problem Solver* (1957) foi um dos primeiros programas de computador concebidos para simular os processos inteligentes do pensamento humano em suas atividades de resolução de problemas. O modelo do computador digital seria, para Simon, a melhor imitação dos processos cognitivos e desse tipo de investigação, segundo ele, resultariam os futuros avanços na Ciência Cognitiva. A partir daí, inaugurou-se o que podemos chamar de paradigma simbólico/computacional da Ciência Cognitiva.

Esses avanços marcaram, conseqüentemente, o reconhecimento da Inteligência Artificial enquanto área de pesquisa. A inauguração da IA teria sido na reunião ocorrida no colégio de Darmouth, EUA, em 1956, na qual Newell e Simon apresentam o seu *Logic Theorist* e propõem a denominação de Inteligência Artificial (IA) para essa nova fase de trabalhos. Seguidamente, J. McCarthy e P. Hayes aperfeiçoam os trabalhos iniciais da promissora IA desenvolvendo o programa *Lisp*, em 1960, com pretensões de que esse fosse um sistema processador capaz de agir inteligentemente no mundo. (VIGNAUX, 1991)

*Foi a partir do desenvolvimento da IA, nas últimas décadas, que toda a ideia de uma ciência da mente se desenvolveu. A IA proporcionou o passo fundamental para se tentar relacionar mentes e computadores e estabelecer o que passamos a chamar de “modelo computacional da mente”. Não fossem os desenvolvimentos e realizações da IA nas últimas décadas suas máquinas de jogar xadrez, demonstrar teoremas matemáticos, realizar diagnósticos médicos toda uma polêmica sobre a natureza da mente e da inteligência não teria surgido.* (TEIXEIRA, 1998, p. 13)

Ainda segundo Teixeira (1998, P. 104), “a Inteligência Artificial simbólica herdou uma metafísica cartesiana ao estabelecer uma dualidade entre *software* e *hardware* como metáfora para conceber as relações mente/cérebro.” Portanto, a Ciência Cognitiva e a IA tomaram como ponto de partida metodológico o pressuposto da separação entre a mente e matéria, ou seja, tinham como princípio a ideia de que as capacidades cognitivas seriam instanciáveis, independentemente de sua base física/biológica. Assim, a mente seria uma funcionalidade passível de ser simulada por um processador central que manipulasse corretamente

algoritmos. Esse paradigma inspirou-se principalmente nas contribuições advindas de Alan Turing (1950) – Funcionalismo –, que permitiu teoricamente a concepção de uma máquina abstrata (Máquina de Turing) capaz de realizar qualquer tipo de operação manipulando apenas símbolos. Desse paradigma resultou o estabelecimento de duas posturas na IA e a consequente distinção entre “IA Forte” e “IA Fraca”: a versão “Forte” da IA postularia justamente a certeza de que a mente poderia ser integralmente reproduzida em um computador, incluindo estados conscientes, intencionais e comportamentos autônomos, pois o mental não passaria de um processo sequencial de procedimentos passíveis de transcrição algorítmica. A “IA Fraca”, por sua vez, defende uma postura menos radical. Resumidamente, podemos dizer que a versão “Fraca” da IA teria como principal pressuposto que as modelagens artificiais servem apenas para aprofundar e melhorar nosso conhecimento sobre a mente, mas sem ter o objetivo de criar inteligência autônoma equivalente à humana.

Após essa etapa inicial de trabalhos, um segundo momento marcou o desenvolvimento da Ciência Cognitiva e da IA. Foi o surgimento da abordagem chamada *conexionista*. Pretendendo-se colocar como alternativa à postura simbólico/funcionalista, o conexionismo baseou-se nos estudos de redes neuronais naturais para modelar os sistemas artificiais, criando, com isso, as *Redes Neurais Artificiais*. Em linhas gerais, podemos dizer que esse paradigma se opõe, em certa medida, à abordagem meramente sintático/simbólica das capacidades mentais, indicando para a necessidade de se considerar a emergência semântica que ocorre nos processos cognitivos. Enquanto que para a I.A. simbólica, pela simulação algorítmica (sem plausibilidade biológica) se chegaria ao conhecimento da mente, para as Redes Neurais Artificiais seria pela simulação do cérebro (também por meio de algoritmos, mas considerando a plausibilidade biológica) que teríamos uma adequada interpretação dos processos cognitivos.

A cognição, no viés conexionista, consistiria na emergência de estados globais em uma rede de componentes materiais simples. Esses elementos simples que podemos chamar de nódulos conectar-se-iam entre si sem a necessidade de uma central de gerenciamento responsável pelo processamento da informação. Assim, a mente seria resultado do funcionamento de toda essa rede constituída por esses nódulos.

*A abordagem conexionista é uma tentativa de construir um modelo de mente mais próximo de sua realidade biológica. Embora estes sistemas não sejam um modelo completo do cérebro e de seu funcionamento, pode-se pelo menos dizer que eles são inspirados na estrutura do cérebro. (TEIXEIRA, 1998, p. 85).*

Ao que parece, com o conexionismo tivemos uma aproximação entre a mente e a estrutura biológica, na medida em que se entende o processo de produção cognitiva como resultado estrito das conexões neuronais e da estrutura cerebral, rechaçando a postura da IA *simbólica* que preservava a múltipla instanciação da mente. Nesse sentido, para o conexionismo, o significado, enquanto resultado dos processos cognitivos, estaria intimamente ligado às ocorrências sinápticas do cérebro e não ao jogo simbólico, como pretendia a I.A. *simbólica*. No processo cognitivo, ao atingir o nível chamado de *subsímbolo*, o significado se descolaria do plano biológico e da estrutura cerebral, emergindo daí o conhecimento, sem, no entanto, deixar de estar relacionado com alguma parte específica na rede neural do cérebro. Apesar de o conexionismo ter se colocado como alternativa à I.A. *simbólica*, apontando para a importância cognitiva dos sistemas biológicos, seu pressuposto básico continua sendo o de que a formalização seria o melhor meio para se reproduzir a inteligência humana, sem considerar a corporeidade. Por isso, convém destacar que no conexionismo, mesmo considerando-se a plausibilidade biológica, ainda não temos a preocupação com as capacidades cognitivas típicas das interações ambientais e perceptuais ligadas ao plano da ação.

Em resumo, as abordagens tradicionais da Ciência Cognitiva e IA – *simbólica* e *conexionista* – não consideravam a relevância da dinâmica do plano da ação, da corporeidade e da constituição biológica. Nesse sentido, apresentaremos a proposta da *Cognição Incorporada e Situada (CIS)*, que surge na década de 1980, como alternativa à mencionada tradição, desenvolvendo uma teoria dinâmica dos processos cognitivos que admite interações entre corpo, cognição e mundo. (SMITH, 1999, p. 769).

Podemos dizer que a teoria da *Cognição Incorporada e Situada (CIS)* desenvolveu-se tendo como objetivo evitar as dificuldades enfrentadas pelo paradigma computacional/simbólico da IA, conforme destacaremos nos trabalhos de Andy Clark (1998, 2001, 2008), Thelen & Smith (1994), Tony Chemero (2007), Haselager & Gonzales (2006) e Port & Van Gelder (1995). Em linhas gerais, a proposta da CIS é de mostrar o lado dinâmico e contínuo dos processos cognitivos. A CIS, ao absorver contribuições de áreas como da filosofia, da psicologia do desenvolvimento, dos estudos da percepção, das neurociências e da robótica, constitui-se então como um programa de pesquisa interdisciplinar com o objetivo de desenvolver modelos inteligentes corpóreos, capazes de interagir com o ambiente e de resolver situações problema. Assim, o viés da *Cognição Incorporada e Situada* tem como principal pressuposto que os sistemas orgânicos são constituídos por capacidades interacionais com o ambiente, que não podem ser dissociadas de sua história evolutiva, de sua corporeidade e de seu contexto.

Andy Clark pode ser apontado como um dos principais proponentes do tipo de abor-

dagem situada e incorporada. O autor desenvolveu vários temas importantes sobre o papel epistêmico do plano da ação nas performances inteligentes. Dentre esses, destacamos a criação do conceito de *ação epistêmica* enquanto importante avanço no sentido de diluir as separações e divisões entre pensamento e ação. Ações epistêmicas seriam certos tipos de ações que se dirigem ao aperfeiçoamento de nossa *performance* mental: em outras palavras, são alterações que fazemos no mundo visando não adaptar-nos à necessidades impostas fisicamente pelo ambiente, mas visando facilitar o desempenho de nossas capacidades cognitivas. Como vemos, Clark assume uma postura continuísta, segundo a qual não haveria uma ruptura entre as capacidades cognitivas do plano da ação e as capacidades cognitivas do plano racional/teórico/formal. Para o autor as ações epistêmicas:

*Alteram o mundo para ajudar e auxiliar os processos cognitivos enquanto reconhecimento e pesquisa. Meras ações pragmáticas, ao contrário, alteram o mundo porque alguma mudança física é desejável por si própria (por exemplo, a ação de tapar o buraco de uma represa). Ação epistêmica, sugerimos, exige a propagação de crédito epistêmico. Se, diante de uma tarefa, uma parte do mundo funciona como se fosse um processo feito na cabeça, não hesitaríamos em reconhecer como parte dos processos cognitivos, de forma que esta parte do mundo é (nós defendemos) parte dos processos cognitivos. Processos cognitivos não estão (todos) na cabeça! (CHALMERS & CLARK, 1998, p. 3-4)*

Como vemos, Clark defende a ideia, segundo a qual os processos cognitivos não estão sediados em um órgão ou instância superior ou primaz. Tal ideia se torna fundamental para sua proposta de mente estendida, pois, para ele, haveria uma extensão inteligente da mente que extrapola os limites meramente cerebrais, racionais, *conscientes*, espalhando-se por todo o conjunto corpóreo de que somos constituídos. Apoiando-se nas investigações de Kirsh e Maglio (1994) sobre os padrões de comportamento de seres humanos ao jogar *Tetris*, Clark conclui que, nesse tipo de situação, temos presente uma coparticipação direta da ação do indivíduo em seu processo cognitivo.

Segundo Clark, a pesquisa elaborada por Kirsh e Maglio (CHALMERS & CLARK, 1998) nos remete ao comportamento de seres humanos em uma situação na qual estão jogando *Tetris*. Nessa pesquisa, tomou-se o pressuposto de que haveria dois tipos de conduta para solução do encaixe das peças. Uma primeira opção seria realizar a rotação mentalmente para, em seguida, executar a rotação física por meio de um botão de comando ligado à

tela de um computador. A outra opção de resolução seria realizar primeiro a rotação física para ver qual encaixe se adequa melhor ao contexto que se apresenta na tela. A conclusão à que Kirsh e Maglio chegam, e da qual Chalmers e Clark se apropriam, é a de que as rotações feitas com a utilização do botão de comando, ou seja, na prática, tem uma maior eficácia do que as realizadas mentalmente. Enquanto a rotação “prática” demora 300 milissegundos, a rotação “mental” demora 1000 milissegundos.

De acordo com o autor, a partir desse experimento, é possível percebermos a importância do plano da ação e do conhecimento prático para o aperfeiçoamento de performances cognitivas. Nesse sentido, parece ser adequada a defesa de que a mente não se reduz a um órgão central como o cérebro, por exemplo, mas que é própria da natureza da mente ser dinâmica e estendida ao corpo e ao mundo, ou seja, a mente seria algo encarnado, incorporado e situado que se desenvolveria sempre a partir da relação com o ambiente.

Também em favor da relevância cognitiva do plano da ação, destacamos o pensamento de Thelen & Smith (1994). As autoras se posicionam criticamente em relação às abordagens lineares e dualistas dos processos cognitivos, propondo uma perspectiva *sistêmica* no âmbito da psicologia do desenvolvimento, superando os paradigmas tradicionais ao considerar a relevância cognitiva da ação. Os estudos de Thelen & Smith debruçam-se principalmente sobre o comportamento humano, em especial, no desenvolvimento cognitivo de bebês. A abordagem sistêmica de Thelen & Smith busca compreender os processos cognitivos considerando o todo como um único sistema, constituído de sistemas, dentro de outros sistemas, sempre em relação dinâmica. Nesse sentido, tentam explicar o desenvolvimento enquanto um processo complexo, abrindo mão de pressupostos dualistas como, por exemplo, mente/corpo, função/estrutura, cérebro/comportamento, etc.

Na concepção das autoras, a atividade auto-organizada teria um papel fundamental no desenvolvimento das capacidades comportamentais motoras, perceptivas, cognitivas e sociais, pois estas enfatizam o caráter auto-organizado dos sistemas inteligentes justamente porque defendem uma postura segundo a qual a mente seria processo complexo, resultado da dinâmica histórica, o contextual, cultural, social, etc. Consequentemente, as capacidades orgânicas básicas são consideradas como tendo um papel fundamental na cognição, pois são os ajustes cognitivos e auto-organizativos primitivos que revelam a inteligência de nossa corporeidade.

*Por exemplo, as nossas rótulas limitam os graus de movimento possível com as pernas, fazendo o equilíbrio e locomoção muito mais fácil. É apenas um pequeno exagero dizer que aprender a andar é fácil para os seres humanos, já que as pernas já*

*sabem como fazer.* (THELEN & SMITH, 1994. Apud Chemero, 2007, p. 27)

Nessa abordagem, entende-se que a herança ontogenética está estreitamente ligada às capacidades cognitivas do conhecimento do senso comum, manifestadas no desenvolvimento de sistemas biológicos. Nesse sentido, ao analisarem o desenvolvimento de crianças, as autoras concluem, por exemplo, que a formação de categorias (a formação da capacidade de classificar coisas diferentes, mas que guardam certas características comuns) é uma capacidade cognitiva fundamental para operações mentais de qualquer tipo, forjadas ontogeneticamente, sem precedentes inatos.

*Sem esta habilidade [de formação de categorias], não há base para operações mentais de qualquer espécie, tanto de pensamento, quanto de ação. A formação de categorias é o início primitivo da vida mental; a ontogenia da formação de categorias é a base para o desenvolvimento do comportamento.* (THELEN & SMITH, 1994, p. 162)

Assim, vemos que a CIS tende a recuperar os processos biológicos nas investigações dos processos cognitivos e da mente. Thelen & Smith consideram os processos físicos os responsáveis por estimular e impulsionar o desenvolvimento, argumentando a favor de uma relação direta do sistema inteligente com o mundo, sem a necessidade estrita de representações que façam a mediação dos processos cognitivos. “Mentes não representam necessariamente o mundo, elas vivem nele e fazem parte da realidade física, uma realidade do eu incorporado” (THELEN & SMITH, 1994, p. 164). Com essa afirmação, vemos o quanto a abordagem sistêmica pode contribuir para uma reavaliação do papel do conhecimento comum nos processos cognitivos, pois parece ser plausível considerarmos estes como os primeiros tipos de manifestação cognitiva constatada em seres humanos. Como citamos no exemplo da formação de categorias em crianças, esse desenvolvimento se dá a partir da experiência com o mundo, sem a necessidade prévia de teorização ou de alguma faculdade inata, evidenciando o caráter incorporado e evolutivo da inteligência.

Em suas críticas às abordagens tradicionais e lineares dos processos de desenvolvimento, Thelen & Smith ainda argumentam que a cognição se trata de um processo dinâmico no qual não há sistemas que se sobressaem a outros ou que sejam centrais. Nesse viés, todos os componentes de um sistema teriam igual importância dentro do processo cognitivo, cada qual desempenhando seu papel adequadamente e de forma contínua e dinâmica. Segundo Thelen (2000, p. 8 – tradução nossa),

*É precisamente essencial a continuidade dos sistemas dinâmicos incorporados e acoplados no tempo para fluir comportamento adaptativo que dá sentido à noção de cognição incorporada. Não há momento em que estes processos parem e outra coisa assuma. Assim, há boas razões para acreditar não somente nas origens sensorio-motoras da cognição, mas também na relação íntima e intrincada entre agir e pensar ao longo da vida.*

Por meio da citação, podemos ver o quão relevante se torna o fator *tempo* no âmbito da perspectiva incorporada e situada. Segundo Port & Van Gelder (1995, p. 3 - tradução nossa),

*O sistema cognitivo não é um computador, ele é um sistema dinâmico. Não se trata de algo no cérebro, isolado e encapsulado; ao contrário, ele é um todo composto por sistema nervoso, corpo e ambiente. O sistema cognitivo não é um manipular sequencial discreto de estruturas representacionais estáticas; ao contrário, é uma estrutura contingente de influência mútua e simultânea. O seu processo não figura no tempo arbitrário e discreto dos passos de um computador; ao contrário, ele manifesta-se no tempo real de mudança contínua no ambiente, no corpo e no sistema nervoso. O sistema cognitivo não interage com os outros aspectos do mundo pela passagem de mensagens ou comandos; ao contrário, está continuamente evoluindo com eles.*

Podemos dizer que a proposta da CIS herdando contribuições da abordagem dinâmica assume a “mudança” como palavra-chave. A mudança se torna fundamental na medida em que os processos cognitivos passam a ser vistos como incorporados no tempo e em um contexto a partir dos quais se torna possível a evolução de um sistema. Em suma, a CIS aplicada à robótica pretende desenvolver agentes artificiais que evoluam e aprendam com o ambiente, adquirindo padrões de conduta ao lidarem de forma apropriada com desafios e imprevistos apresentados em situações não controladas.

## **A relevância cognitiva do corpo**

De acordo com o exposto até o momento, a corporeidade passa a ser entendida como tendo um papel fundamental dentro dos processos cognitivos. Esse pressuposto se contrasta com toda a tradição dualista filosófica e o paradigma computacional/funcionalista da mente. Ao compreender uma continuidade entre corpo, cognição e ambiente, teríamos a vantagem de dissolver alguns problemas como, por exemplo, o da relação entre mente e corpo, tão caro à filosofia tradicional. No paradigma incorporado e situado, a cognição surgiria como *emergência* dos processos biológicos corporais, de forma contínua e natural. Quando dizemos que há uma relação entre corpo e cognição não queremos estabelecer uma identidade de propriedades (como, por exemplo, a tipo-tipo<sup>2</sup>). No viés incorporado e situado, corpo e cognição não seriam a mesma coisa, mas também não seriam coisas distintas e excludentes. Trata-se de compreender a cognição como um *continuum* auto-organizado, sempre em transformação pela dinâmica que a constitui como um todo.

De acordo com os trabalhos de Clark (1998), por exemplo, o corpo possuiria uma dinâmica própria, reveladora de sua flexibilidade e plasticidade, o que nos permite estabelecer distinções, por exemplo, entre o corpo de um robô e o corpo de um ser vivo. Os movimentos do corpo de um ser vivo possuem uma flexibilidade e plasticidade espontâneas que se adaptam dinamicamente em relação ao ambiente. Já os movimentos de um corpo robótico não teriam tal flexibilidade espontânea, pois necessitariam de uma série de regras e parâmetros para o reconhecimento do local, para poderem tentar adaptar-se. Tal flexibilidade e plasticidade espontânea do corpo de um ser vivo são expressas por Clark como sendo um comportamento *soft assembly*. Essa expressão pode ser definida como movimentos produzidos por nosso corpo, nos quais há uma mistura de sutileza e força ao mesmo tempo, que se dão de forma espontânea, sem previsão, digamos, consciente: trata-se de um tipo de conhecimento tácito do ambiente.

No plano da ação, cognição e corpo fundem-se em uma continuidade dinâmica da qual resultam performances inteligentes de alta eficiência em situações do senso comum que realizamos. Segundo Haselager & Gonzalez (2006, p. 9)

*Se olharmos para o exemplo da ‘quebra do ovo’, o problema de como quebrar um ovo não se resume a uma consulta na base de dados de nosso conhecimento proposicional para formular um planejamento simples. Ao invés disso, o meio ambiente e suas pro-*

---

2 - A Teoria da Identidade Tipo-tipo, criada por UT. Place e J.J Smart na década de 1950, afirma que cada tipo de estado mental corresponde a um tipo de estado cerebral. Esse viés materialista/fisicalista da mente propõe, portanto, uma redução dos estados mentais aos estados físicos, o que não ocorre na abordagem da Cognição Incorporada e Situada.

*priedades possibilitam, selecionam ou explicitam certos comportamentos. A borda da vasilha permite a quebra da casca do ovo. O movimento que fazemos com nossos braços e mãos não é planejado e calculado, mas emerge da dinâmica intrínseca de nosso corpo, do peso do ovo, da influencia da gravidade e de muitos outros fatores. Nesse sentido, os aspectos gerais do organismo na situação do senso comum se auto-organizam em uma sequência apropriada de comportamentos emergentes.*

Ressaltamos que as contribuições da CIS permitiram à IA e à robótica avanços significativos nos últimos anos, justamente por destacar o relevante papel do corpo nos processos cognitivos em nível de senso comum. Por isso, esse tipo de abordagem tem sido acolhida nas ultimas décadas, principalmente no âmbito da IA e da robótica, e já deixa de ser apenas uma abordagem alternativa, ganhando o *status* de teoria científica corrente. (GUDWIN, 2005, p.14).

Como vemos, a partir da abordagem da CIS, diminui-se a distância entre o comportamento de seres artificiais e o de seres vivos. Uma das principais questões que pretendemos colocar é: seria cabível atribuir moralidade e responsabilidade a seres artificiais? Uma das primeiras propostas para esse debate foi apresentada por Isaac Asimov, em sua obra de ficção científica, “*Eu, Robô*”, quando ainda não tínhamos tipos tão avançados de “humanoides” como os de hoje. Segundo as leis de Asimov: 1) um robô não pode ferir um ser humano ou, por omissão, permitir que um ser humano sofra algum mal; 2) um robô deve obedecer às ordens que lhe sejam dadas por seres humanos, exceto nos casos em que tais ordens contrariem a Primeira Lei; 3) um robô deve proteger sua própria existência, desde que tal proteção não entre em conflito com a Primeira e Segunda Leis. Atualmente, com os avanços na área da IA, as leis de Asimov parecem não dar conta deste que podemos chamar de dilema moral dos robôs, principalmente porque possuem um valor prescritivo ainda aliado a concepções tradicionais de moral e ética.

Com os atuais avanços da IA, é razoável pressupormos que robôs podem ser relativamente considerados autônomos de diversas maneiras, desde aqueles que executam tarefas que exigem alta precisão ou repetitividade como os utilizados em linha de produção, até aqueles que oferecem certo grau de risco à vida do homem, pois lidam com os seres humanos atuando como recepcionistas (robô *Reem*, 2011, da *Palm Robotics*), como interlocutores (robô *Zeno*, 2010, da *Hanson Robotics*) e como atores (robô *Geminoid F*, 2012, Universidade de Osaka), capazes de interagir com seres humanos. Nesse sentido, temos um problema que parece extrapolar os limites dos modelos éticos prescritivos, do *agir por dever*, conforme expressos pelo código criado por Asimov e pelas filosofias tradicionais.

Sugerimos que algumas filosofias contemporâneas podem trazer contribuições para o referido debate. Tentamos mostrar que, num viés filosófico pragmático, não dualista, é possível uma reflexão sobre a conduta no plano da ação não mais detida no domínio *racionalista/representacionista/internalista*. Entre essas filosofias contemporâneas, destacamos o pensamento de John Dewey e sua proposta pragmatista, que enfatiza a relevância do plano da experiência e da ação para a superação do dualismo tipo “mente *versus* corpo”, “eu *versus* ação”, na compreensão das questões morais e éticas.

## **O pragmatismo de John Dewey: uma teoria da ação *inusitada***

Ao tratar da moral e da ética, Dewey toma sempre como objeto de investigação a conduta dos indivíduos, enfatizando que ações e condutas não se devem unicamente a um *eu* interno do qual partem as coordenadas. Inicialmente, podemos dizer que um dos principais pressupostos de John Dewey é o de que a formação e a constituição da conduta dar-se-ia de forma contínua; além disso, o ambiente seria decisivo para direção e formação dos hábitos daquele que está atuando ou agindo. Assim, os estímulos externos (que incitam o agente a ter ações e reações) contribuem no sentido de auxiliar aquele que age a adquirir hábitos dentro do processo de aprendizado.

Segundo Dewey, as vertentes de tendência metafísica/internalista (como a kantiana, por exemplo, e as *éticas do dever*, em geral) possuem uma concepção moral fortemente dualista, pois não compreendem a formação moral em sua totalidade e continuidade entre agente e ambiente. O “eu” é fundamentalmente um “eu” corpóreo que age, um agente como um todo com o ambiente que o circunda. A proposta deweyana é de que, entre o “eu” e a ação, não há uma distinção do tipo meio e fim. Diferentemente, “há um arranjo circular. O ‘eu’ não é simples meio para produzir consequências, porque estas, quando de espécie moral, entram na formação do ‘eu’ e este na formação delas.” (Dewey, 1964, p. 134). Em suma, a constituição de uma identidade dependeria do processo de aprendizagem que envolve tanto o plano individual (“eu”), quanto o ambiental. Como vemos, a ambiência ganha um papel fundamental na discussão da moral e é apontada por Dewey como sendo também relevante para a constituição da *consciência*.

*Está a região moral isolada do restante da atividade humana? Somente uma classe especial de objetivos e relações humanos tem valor moral? Essa conclusão é o resultado necessário da teoria [tradicional] segundo a qual nossa consciência e nosso*

*conhecimento morais são únicos na espécie. Mas se a consciência moral não é separada, então não se pode traçar qualquer linha rígida e firme na conduta isolando o reino da moral do reino não-moral. (DEWEY, 1964, p. 110)*

A consciência, para Dewey, antes de ter um caráter subjetivo, interno, racional, possui um caráter objetivo que a constitui. Esse caráter objetivo da consciência é fundamental para que compreendamos o genuíno significado de um julgamento moral, pois, nas perspectivas filosóficas tradicionais, tal julgamento se dava em uma perspectiva *internalista*. Para Dewey, a consciência deveria ser compreendida, então, de forma objetiva, experiencial, o que permitiria à reflexão sobre a moralidade expandir seu alcance, não limitando o julgamento moral apenas a um procedimento punitivo.

*Um ser humano é considerado responsável a fim de que possa aprender; a fim de que possa aprender não teórica e academicamente, mas de forma a modificar e – até certo ponto – refazer seu “eu” anterior [...] o que interessa é se é capaz de agir diferentemente na próxima vez, a importância prática e modificações no caráter humano é que torna importante a responsabilidade. (DEWEY, 1964, p. 153)*

Em suma, podemos dizer que o autor pretende que todo julgamento moral considere aspectos como o contexto, a sociedade, a cultura, as redes relacionais, etc., que são cruciais para a constituição de uma determinada identidade. “A consciência genuína tem perspectiva objetiva; é atenção e cuidado inteligentes para com a qualidade de um ato, em vista de suas consequências para a felicidade geral; não é solicitude ansiosa para com o estado virtuoso do próprio indivíduo” (DEWEY, 1964, p. 120).

## **O dilema moral dos robôs à luz do pragmatismo deweyano**

De acordo com o que expusemos até aqui, pretendemos agora mostrar que a proposta de John Dewey parece plausível de ser relacionada à conduta não somente humana. Nesse sentido, tentando compreender o âmbito da conduta de seres artificiais à luz das considerações deweyanas, pode-se dizer que, a partir do momento em que um robô, em dado momento, desempenha um padrão de ação como, por exemplo, o de andar de bicicleta, de

tocar um violino ou de recepcionar pessoas em um evento, ele o faz imitando um comportamento típico, que expressa determinado objetivo. Ressaltamos que tal objetivo provém estritamente de todo o processo interacional com contexto do qual o robô participa.

De acordo com o projeto da robótica atual, a saber, de criar robôs capazes de aprender e de alterar hábitos conforme a dinâmica ambiental, torna-se possível que os mesmos tenham uma conduta *responsável* e sejam *responsabilizados* por suas ações. Não estamos debatendo se robôs teriam ou não consciência, estamos apenas indicando que o julgamento moral não se deve exclusivamente a aspectos internos e subjetivos. Em outras palavras, da mesma forma que um homem pode ser caracterizado como responsável, um robô também poderia vir a sê-lo, se considerado um agente que assume determinada postura, em um determinado contexto, pois sua conduta implicaria consequências práticas. Nossa interpretação é a de que a filosofia de John Dewey parece não encontrar problemas nessa afirmação, pois a responsabilidade, tanto no âmbito humano, quanto no âmbito de um agente artificial, sempre seria dividida entre ambiente e agente.

O julgamento moral, nesse caso, não deveria servir meramente como instrumento para culpar aquele que errou: responsabilizar um indivíduo, como vimos, implicaria, antes de qualquer punição, a expectativa de mudança de conduta. Assim, quando um robô erra, responsabilizá-lo implicaria não meramente uma acusação ou condenação; indo além, a responsabilidade deve levar esse robô à necessidade de mudança de padrões de hábito. Nesse sentido, na medida em que robôs forem capazes de tomadas de decisão, de mudança de hábitos nesses moldes, podem ser considerados seres *responsáveis* por suas ações, mas concebendo sempre a *responsabilidade* diferentemente dos pressupostos tradicionais.

O que deve ser redefinida, ou talvez até banida, é a noção de imputação de culpa e penalização enquanto algo que recai sobre uma consciência prévia que planeja e executa uma ação. A concepção de uma mente interna que coordene as ações tidas como corpóreas é recusada por Dewey, em prol de uma definição de conhecimento e ação oriunda de um processo complexo que se inicia pelo fluxo contínuo da experiência. Em outras palavras, a sociedade e o contexto são também *responsáveis* pela *má ação cometida* por um indivíduo. Nesse sentido, a sociedade e o criador do robô também teriam um papel em sua atuação, por isso, uma análise moral em termos tradicionais não pode dar conta do complexo problema que envolve o dilema moral dos robôs. Trata de um processo complexo que envolve uma dinâmica relacional pautada no plano da ação, e não em instâncias e faculdades subjetivas responsáveis pela *agência* do robô. Nesse sentido, responsabilizar um robô implica, antes de tudo, compreender a totalidade da situação em que se deu o erro implicando que, ao invés de possuir um caráter punitivo, essa responsabilização possua um caráter educativo.

## Considerações finais

Ao que parece, abordar o chamado “dilema moral dos robôs” implica uma renúncia a éticas dualistas, o que traz implicações morais no sentido de que conceitos como de responsabilidade e liberdade sejam tratados filosoficamente de maneira alternativa. Como vimos na filosofia de John Dewey, os procedimentos morais são um *continuum* dos procedimentos cognitivos, e a responsabilidade de um agente estaria ligada à sua capacidade de prever as consequências de seus próprios atos. O autor, em seu ideal democrático, sempre se mostrou avesso aos sistemas prisionais e punitivos, pois acreditava que um julgamento moral devesse ser compreendido como algo educativo e não meramente punitivo.

Se, por exemplo, um robô recepcionista como o *Reem* (2011) causasse mal a algum visitante desse evento, o que diríamos dessa situação? De alguma maneira, esse robô, um agente interacional que cometeu um erro, não teve previsão suficiente de sua ação. Talvez isso nos permita interpretá-lo como sendo, em algum sentido, *responsável* por sua ação, mas não com o intuito primordial de penalizá-lo, como é corrente em nossos códigos morais. Nesse sentido, a *responsabilidade moral* redefinida parece ser aplicável a robôs.

No entanto, é preciso que façamos uma ressalva em nossa análise, pois para Dewey a predisposição de natureza biológica é fundamental para o desenvolvimento de um padrão de conduta e o estabelecimento de uma *identidade pessoal*. Por isso, antes que cometamos alguma injustiça com o pragmatismo de Dewey, sublinhamos que, como ainda não temos *efetivamente* modelos biologicamente desenvolvidos, conceber os robôs como possuindo uma identidade pessoal, um agir consciente, ainda não é possível. Restringimo-nos apenas a relacionar Dewey e o dilema moral dos robôs nos aspectos mencionados aqui, segundo os quais a ação não é algo de origem estritamente interna, mas um processo dinâmico entre agente e ambiente.

## Bibliografia

ASIMOV, I. *Eu, Robô*. Trad. Luiz Horácio da Mata. São Paulo: Circulo do Livro, 1976.

BROOKS, R. *Inteligência sem representação*. Tradução: Pedro Rocha de Oliveira e João de Fernandes Teixeira. Em: [www.filosofiadamente.org/images/stories/textos/representacao.doc](http://www.filosofiadamente.org/images/stories/textos/representacao.doc). Acesso em: 28/11/2011.

\_\_\_\_\_. Intelligence without representation. In: *Artificial Intelligence* 47 (1991), pp. 139-159.

CHEMERO, A. *Radical Embodied Cognitive Science*. Massachusetts: MIT/Bradford, 2009.

CLARK, A. *Being there: putting brain, body, and world together again*. Massachusetts: MIT/Bradford, 1998.

\_\_\_\_\_. *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press, 2001.

\_\_\_\_\_. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. New York: Oxford University Press, 2008.

CLARK; A.; CHALMERS, D. The Extended Mind. In: *Analysis*, 58 (1), p.10-23, 1998.

DEWEY, J. *Democracia e educação: introdução à filosofia da educação*. 3. ed. Tradução Godofredo Rangel e Anísio Teixeira. São Paulo: Companhia Editora Nacional, 1959.

\_\_\_\_\_. *Experiência e natureza*. São Paulo: Abril Cultural, 1980.

\_\_\_\_\_. *Reconstrução em Filosofia*. São Paulo: Companhia Editora Nacional, 1959.

\_\_\_\_\_. *Human Nature and Conduct: an introduction to social psychology*. New York: Modern Library, 1950.

\_\_\_\_\_. *Teoria da vida moral*. São Paulo: IBRASA, 1964.

DUPUY, J.-P. *Nas origens das Ciências Cognitivas*. Tradução: Roberto Leal Ferreira. São Paulo: Editora Unesp, 1996.

GONZALEZ, M. E. Q.; BROENS, M. C. Darwin e a virada naturalista na filosofia. In: MORAES, J. Q. (Org.). *Darwin e a origem do homem*. Campinas: UNICAMP, 2011.

GUDWIN, R.R. *Novas Fronteiras na Inteligência Artificial e na Robótica*. In: *Dincom'2005*. Campinas. 2005. pp. 01-18. Disponível em: [http://www.dca.fee.unicamp.br/~gudwin/ftp/publications/Dincom05\\_Gudwin.pdf](http://www.dca.fee.unicamp.br/~gudwin/ftp/publications/Dincom05_Gudwin.pdf). Acesso em: 20/09/2012

HASELAGER, W.F.G.; GONZALEZ, M.E.Q. Conhecimento comum e auto-organização. Disponível em: <http://www.nici.kun.nl/~haselag/publications/ConhecimentoComumAutoorg.pdf>. Acesso em: 10/07/2011

HASELAGER, W.F.G. Auto-organização e comportamento comum: opções e problemas. In: GONZALES, M.E.Q; D'OTTAVIANO, I.M.L; SOUZA, G.M. (orgs). *Auto-organização: estudos interdisciplinares*. Vol. 38. Campinas, Brasil: Coleção CLE, 2004. pp. 213-235.

HASELAGER, W.F.G. O mal estar do representacionismo: sete dores de cabeça da Ciência Cognitiva. In A. Ferreira, M.E.Q. Gonzalez & J.G. Coelho (Eds.). *Encontros com as Ciên-*

cias Cognitivas, Vol. 4. São Paulo: Coleção Estudos Cognitivos, 2004. pp. 105-120.

KIRSH, D.; MAGLIO, P. On distinguishing epistemic from pragmatic action. In: *Cognitive Science*, 18, pp. 513-549, University of California, San Diego, 1994. Disponível em <http://adrenaline.ucsd.edu/kirsh/publications.html>. Acesso em: 23/06/2012.

PINTO, S.M.C.R. A Natureza Histórica da Cognição: Debates filosóficos na Teoria dos Sistemas Dinâmicos na Ciência Cognitiva. Dissertação de Mestrado. Universidade do Porto. Portugal, 2007.

PORT, R.; VAN GELDER, T. *Mind as Motion: Explorations in the dynamics of cognition*. Massachusetts: MIT/Bradford, 1995.

SMITH, B.C. Situatedness/Embeddedness. In: WILSON, R.A.; FRANK, C.K. (orgs). *The MIT Encyclopedia of the Cognitive Sciences*. Massachusetts: MIT, 1999. pp. 769-770.

TEIXEIRA, A. A pedagogia de Dewey. In: *Vida e Educação*. São Paulo: Companhia Editora Nacional, 1959.

\_\_\_\_\_. Bases da teoria lógica de Dewey. *Revista Brasileira de Estudos Pedagógicos*. Rio de Janeiro, v. 23, n. 57, jan/mar, 1955 p. 3-27.

THELEN, E. Grounded in the world: developmental origins of the embodied mind. In: *Infancy*, 1 (1), 3-28, 2000.

THELEN, E.; SMITH, L. *A dynamic systems approach to the development of cognition and action*. Cambridge: MIT Press, 1994.

VIGNAUX, G. *As Ciências Cognitivas: uma introdução*. Coleção Epistemologia e Sociedade. Tradução: Maria Manuela Guimarães. Lisboa: Instituto Piaget, 1991.



# **Sistema Político Comunal, Liberdade, Igualdade e Religião entre os Norte-Americanos segundo Tocqueville**

## ***Communal Political System, Freedom, Equality and Religion in North America according to Tocqueville***

*Cícero J. Silva Oliveira*  
*PUC-Rio*

### **Resumo**

O ineditismo que Tocqueville acredita encontrar no fenômeno político americano está radicado no ponto de partida que constituiu aquela nação. De um lado, entre as experiências determinantes para a formação do que compõe o “caráter nacional” dos estadunidenses está a formação política dos imigrantes constitutivos daquele povo em uma “escola rude”, isto é, expostos às turbulências políticas da Europa. De outro, ao fato de a chegada ao Novo Mundo ter sido acompanhada das instituições de relações sociais e políticas que exigiam e favoreciam o desprendimento das relações aristocráticas, símbolos das relações sociais e políticas na Europa, mas cujos princípios, sobretudo políticos, se radicavam na austeridade dos costumes da religião puritana. Para nosso propósito, importa expor o modo como da combinação do espírito de liberdade ao de religião resulta o aparecimento das instituições políticas livres, a saber, as comunas.

### **Palavras-chave**

Igualdade; liberdade, política; religião, sistema comunal, política norte-americana

## **Abstract**

Tocqueville finds that a novelty in the North-American political phenomenon is rooted at the starting point which will constitute one nation. Firstly, he highlights one point in the formation of a “national character” in United States citizens, namely, the fact that immigrants were trained in a so-called “hard school” which exposed and dealt with the political turmoil in Europe. Secondly, he sustains that the arrival in the new world was accompanied by a range of social relations and the establishment of policies that demanded a disengagement from the aristocratic European relations but whose principles, especially political, were linked to the austerity of Puritan morals and religion. For our purpose, it is important to expose how the combination of the spirit of the freedom of religion results in the appearance of free political institutions, namely the communes.

## **Keywords**

Equality; freedom; politics; religion; communal system; North-American politics

A então incipiente nação dos Estados Unidos da América apresentava, para Tocqueville, o privilégio de permitir ver, de modo claro e preciso, o aspecto fundamental do ponto de partida que perpassa todo o seu desenvolvimento: a despeito da ocupação da América ter sido feita por homens que guardavam inúmeras diferenças entre si, as características que esses diferentes homens comungavam concorreram, de modo determinante, para a composição daquilo que Tocqueville chama de *causas primeiras* da América.

Do ponto de vista político, o traço elementar do ponto de partida dos Estados Unidos da América reside no fato de, na época das primeiras emigrações, o governo comunal, esse germe fecundo das instituições livres, já ter impregnado profundamente os hábitos dos ingleses, e, com ele, a ideia da soberania do povo foi introduzida no próprio seio da monarquia dos Tudor (TOCQUEVILLE, 2005, tomo I, p. 39). Em grande parte, tal se deu como implicações políticas dos dogmas puritanos em favor da liberdade e, assim, a relevância política da religião na obra de Tocqueville é o assunto que também precisamos ressaltar aqui.

Contudo, qualquer divisão entre as esferas política e social que se pretenda ser estanque foge à perspectiva de Tocqueville porque o estado político de liberdade é também um reflexo das condições sociais de igualdade. Não há qualquer relação de necessidade ou causalidade que faça com que o estado social de igualdade resulte inevitavelmente na expressão política da liberdade. A expressão “marcha da democracia”, empregada pelo autor em *A Democracia*

na América, diante do que ele chama de a vitória do princípio democrático sobre o aristocrático, restringe-se à noção do vir a ser da *igualdade de condições*, apresentada já na introdução como um projeto providencial, como algo para cujo desenvolvimento “todos os homens ajudaram com seus esforços: os que tinham em vista concorrer para seus sucessos e os que não pensavam em servi-la; todos foram levados de roldão pelo mesmo caminho, e todos trabalharam em comum, uns contra a própria vontade, outros sem o saber, *cegos instrumentos nas mãos de Deus*” (TOCQUEVILLE, 2005, tomo I, p. 10-1). Como justificativa de semelhante argumento, Tocqueville apresenta-nos três aspectos que o faz reconhecer a dimensão providencial desse fenômeno, a saber, a universalidade, a durabilidade (que para o autor perfaz um processo com cerca de setecentos anos), e o fato de escapar ao homem a possibilidade de deter esse mesmo processo.

O autor nos adverte, entretanto, que o movimento irresistível dessa força não tem um percurso determinado. A *revolução social* cuja tradução é o estabelecimento inelutável e gradual da *igualdade de condições*, depende da ação humana para, no plano político, transfigurar-se em *liberdade*. Embora Tocqueville não nos apresente uma definição precisa do que seja *democracia*, seu pensamento parece nos autorizar a entendê-la de um modo bastante sumário como a confluência entre a *igualdade social e a liberdade política*. O fato de que Tocqueville tenha pensado a chamada marcha da democracia como a realização irresistível da igualdade tomada isoladamente, não parece refutar essa hipótese, pois para o autor, ainda que inerente ao conceito de democracia, a liberdade não é o elemento distintivo da democracia moderna, já que ela pode ser verificada em uma dimensão bastante restrita mesmo nos Estados aristocráticos, constituindo um privilégio de poucos. Segundo Tocqueville, o elemento inteiramente novo da democracia que ele tem diante dos olhos é a “incomensurabilidade da igualdade moderna” (JASMIM, 1997, p. 196) que estende a liberdade à maioria. Essa igualdade incomensurável é, para o autor, o princípio elementar ou a condição *sine qua non* do sistema social e político chamado democracia, enquanto fenômeno moderno.

Embora a liberdade política que se verifica na análise das comunas dependa, em certa medida, das condições sociais dos anglo-americanos, em seu germe mais originário ela se encontra no que Tocqueville chama de *caráter nacional* do povo anglo-americano, já formado quando os ingleses aportaram no novo mundo.

*Nascidos num país agitado desde séculos pela luta dos partidos em que as facções tinham sido obrigadas, sucessivamente, a se colocar sob a proteção das leis, sua educação política fizera-se nessa rude escola, e viam-se difundidas entre eles mais*

*noções de direitos, mais princípios de verdadeira liberdade do que entre a maioria dos povos da Europa. Na época das primeiras emigrações, o governo comunal, esse germe fecundo das instituições livres, já havia profundamente impregnado os hábitos dos ingleses (TOCQUEVILLE, 2005, tomo I, p. 37).*

Cabe aqui estabelecer a razão do caráter restritivo do sistema político comunal, expresso no título desse artigo como uma alusão apenas ao caso dos anglo-americanos. Embora os migrantes de todas as colônias europeias tivessem em comum o fato de não poderem levar para o Novo Mundo qualquer princípio válido de superioridade de uns sobre os outros, (dado que a situação de pobreza expunha a uma condição de igualdade os que buscavam o exílio em virtude dessa situação, tanto quanto o infortúnio põe no mesmo nível os poderosos que, por motivo de desavenças políticas ou religiosas, viam-se forçados ao exílio), as colônias inglesas do norte, cujo contraexemplo eram as situadas ao sul, foram as únicas a reunir as condições propícias para o desenvolvimento das comunas, isto é, as instituições provinciais, livres por excelência. Entre suas singularidades figura o estabelecimento de classes abastadas que não guardavam qualquer sorte de diferenças extremas, seja quanto às posses ou quanto às luzes. Por isso, embora o nosso foco de análise recaia sobre a o papel político da religião, não se deve deixar passar ao largo a importância da questão social para o surgimento do sistema político comunal entre os anglo-americanos.

Entre os acontecimentos que concorreram para o desenvolvimento e a formação do espírito instruído e contestador dos anglo-americanos que estabeleceram colônias no Norte dos Estados Unidos, Tocqueville assinala as querelas religiosas nas quais a Inglaterra se envolveu. As disputas foram determinantes para a difusão geral das Luzes e das várias noções de direito e liberdade subjacentes a essa difusão. Ademais, Tocqueville considera que mesmo o simples envolvimento dos ingleses nos assuntos relativos à religião forjou um conjunto de costumes mais puros sobre os quais se calca o sistema político que os ingleses fundariam no novo mundo.

Dos três modos de colonização postos em prática pela coroa inglesa, apenas à região conhecida pelo o nome de Nova Inglaterra foi concedido (sob a condição de concessão do governo inglês e os limites por ele estipulados de não contrariar qualquer um de seus princípios) o direito de governar-se. Quanto às colônias inglesas do sul, o governo britânico determinou ou um governo para administrá-las segundo suas ordens, ou adotou a prática de confiar a exploração e a administração aos grandes senhores e à companhias, por meio de contratos.

Para Tocqueville, é sintomático que, para as colônias inglesas estabelecidas ao sul, tenha se dirigido, em virtude da exploração do ouro e da prata, “gente sem recursos e sem modos”, porque o “espírito inquieto e turbulento” dos que ali se estabeleceram sob o signo comum da busca por fortuna, “perturbou a infância da colônia” (TOCQUEVILLE, 2005, tomo I, p. 39).

Exilavam-se também no sul homens que se viram impelidos a deixar a mãe-pátria, quer por motivo de pobreza, quer em razão de má conduta. Além disso, os empreendimentos industriais que essa região atraía trouxeram consigo a cristalização da desigualdade nas relações sociais, com a introdução da escravidão em 1620. Por tudo isso, não se desenvolveram no sul as qualidades sociais verificadas na Nova Inglaterra, como eram conhecidas as colônias inglesas do norte dos Estados Unidos.<sup>1</sup> Desde o nascimento, essas colônias foram acompanhadas das condições sociais de igualdade material e intelectual, aliadas ao espírito de ordem, à moralidade que inspirou seus habitantes e ao fato do empreendimento geral dos colonos ser motivado não pelo desejo de fortuna, mas pela implementação de um ideal religioso de estabelecimento da fé cristã no Novo Mundo, com implicações democrático-republicanas.

De acordo com Tocqueville, foram os princípios políticos subjacentes ao espírito religioso que tornou a Inglaterra inóspita aos seguidores do puritanismo que “[...] era quase tanto uma teoria política quanto uma doutrina religiosa” (TOCQUEVILLE, 2005, tomo I, p. 43). Esse acontecimento estimulou – e mesmo forçou – a emigração dos que não podiam viver junto à mãe-pátria sem ter os princípios morais e religiosos feridos pelos seus adversários políticos, aqueles que determinavam os rumos da sociedade inglesa.

Os ingleses que aportaram no norte dos Estados Unidos viam o Novo Mundo, ou, mais especificamente, a Nova Inglaterra como o cenário no qual se estabeleceriam em definitivo, a fim de fazer prosperar uma comunidade cuja glória e consolidação do espírito piedoso seria o triunfo da obra de Deus, a quem essa comunidade religiosa tomava como dever assegurar a memória e perpetuar o louvor.

Para Tocqueville, as implicações políticas da orientação religiosa puritana mostram-se desde o início da chegada dos seus seguidores ao novo continente, precisamente na atitude ousada de fundar, em um plano prático, aquilo que até então estivera relegado ao âmbito da *abstração* da literatura de filosofia política moderna: um corpo político via um acordo mútuo. Tal foi um dos primeiros atos empreendidos pelo grupo de cento e cinquenta pessoas que partiram da Inglaterra, com o propósito de estabelecer a primeira colônia puritana no novo continente. Segundo as palavras de Nathaniel Morton (um dos primeiros historiadores da Nova Inglaterra, que ocupava a condição de sectário), de quem Tocqueville cita

---

1 - Ao todo, eram seis os Estados que constituíam a Nova Inglaterra: Connecticut, Rhode Island, Massachusetts, Vermont, New Hampshire e Mayne.

trechos extensos, os fundadores ingleses estavam tão conscientes deste ato de dar início a uma comunidade política que formalizaram em documento escrito o acordo voluntário ou o pacto fundamental que institui tal comunidade.<sup>2</sup> O pacto forjou o princípio elementar de acordo com o qual as emigrações religiosas seriam fundadas.<sup>3</sup>

Como o expediente de desterrar os que sustentavam crenças indesejáveis aos interesses do governo de Carlos I tornou-se uma prática comum na Inglaterra – fato que fez da região onde se estabeleceu a primeira colônia puritana uma válvula de escape para problema que o governo inglês tinha com o puritanismo –, é provável que a prática do autogoverno tenha sido em seu nascimento algo a que o governo colonial foi indiferente, tendo apenas posteriormente assumido certo controle.

O fato de que havia colônias regidas pelo autogoverno sem a concessão da coroa, já no período em que a Inglaterra buscava assumir esse controle, mais do que um claro sinal da ingerência inglesa,<sup>4</sup> é um indício patente de que essa forma de governo surgiu não de um programa de colonização previsto pela coroa, mas da própria iniciativa dos novos habitantes, que, sem se remeterem a uma autoridade política de ultramar, fundaram no Novo Mundo o poder político que, nas palavras de Hannah Arendt, baseia-se na capacidade humana de agir em concerto, e tem origem apenas nas faculdades humanas de fazer e cumprir pactos, contratos e promessas.<sup>5</sup>

A inclinação de Abraão para fazer pactos, a importância que a promessa tinha já no sistema jurídico romano e a ênfase na influência que esse tipo de associação política exerceu sobre

---

2 - O trecho do documento citado por Nathaniel Morton e reproduzido por Tocqueville é o seguinte: “Nós, cujos nomes seguem e que, para a glória de Deus, o desenvolvimento da fé cristã e honra de nossa pátria, empreendemos estabelecer a primeira colônia nestas terras longínquas, acordamos pelo presente ato, por consentimento mútuo e solene, e diante de Deus, formar-nos em corpo e sociedade política, com o fim de nos governar e de trabalhar para a consumação de nossos propósitos; e, em virtude desse contrato, acordamos promulgar leis, atos, decretos, e instituir, conforme as necessidades, magistrados a quem prometemos submissão e obediência” (TOCQUEVILLE, 2005, tomo I, p. 44). Essa formalização que, para Tocqueville, não revela senão a flagrante influência das teorias contratualistas da modernidade política sobre os colonos, é posta em dúvida por Hannah Arendt, para quem os colonos “certamente não tinham qualquer noção de nenhuma teoria”, o que a faz pensar que Locke “foi mais influenciado pelos fatos e acontecimentos na América, e talvez de uma maneira mais decisiva, do que os fundadores o foram por seus *Tratados do governo civil*” (ARENDR, 1988, p. 135).

3 - É digno de observação – como nota o próprio Tocqueville –, que até este momento histórico o puritanismo tenha encontrado na classe média inglesa sua principal fonte de adesão. O que resulta desse fato é de grande relevância política, já que os sectários exilados ou banidos pelo governo inglês mantinham equilibradas, na Nova Inglaterra (em função da condição social que de um modo geral desfrutavam), as condições de igualdade, imprescindíveis para o desenvolvimento da democracia. Daí porque, expulsando dos seus domínios aqueles cuja orientação religiosa era danosa à prática política monarquista, a Inglaterra contribuía para a ampliação de uma sociedade no seio da qual a igualdade se consolidava, sem o saber, o sistema feudal inglês criou as condições para o vir a ser da democracia.

4 - “Plymouth, New Haven, o Estado de Connecticut e o de Rhode Island foram fundados sem o concurso e, de certa forma, sem o conhecimento da mãe-pátria” (TOCQUEVILLE, 2005, tomo I, p. 45).

5 - “At any rate, the great variety of contract theories since Romans attests to the fact that the power of making promises has occupied the center of political thought over the centuries” (ARENDR, 1998, p. 244).

os teóricos da modernidade, são exemplos do itinerário histórico que Arendt perfaz a fim de aduzir a centralidade que o poder humano de prometer assume em matéria de política.

Para Hannah Arendt, a possibilidade de iniciar algo novo que está radicada na natalidade de cada homem que vem ao mundo tem como subsequente potencial a arregimentação de mais agentes. Isso acontece quando uma multiplicidade de homens se associa para levar a cabo algo cuja iniciativa se deve apenas a um deles. É dessa “ação em concerto” que irrompe o espaço público sem o qual não faz sentido falar de política. O princípio elementar que gera e mantém vivo um organismo político é esse tipo de relação que torna indivíduos esparsos em agentes políticos.

Arendt, tanto quanto Tocqueville, estava completamente convencida de que os corpos autenticamente políticos só podem nascer dessa forma de interação discursiva que levam os homens a estabelecer, por meio de pactos e promessas, os compromissos que os vinculam também no futuro. Entretanto, para Tocqueville, havia algo mais que presidia o elemento fundamental que, de acordo com Arendt, era a confiança que se tinha em “[...] como estabelecer organismos políticos” (ARENDDT, 1988, p. 141).

Embora possamos dizer que esses autores concordassem quanto ao fato de que essa confiança nas capacidades humanas de prometer e fundar corpos políticos estivesse animada pelo espírito de liberdade, a convicção (e mesmo o entusiasmo diante da crença)<sup>6</sup> de que esse espírito de liberdade estava aliado ao de religião não era compartilhado por Arendt.

Nesse sentido, o autor de *A Democracia na América* acredita encontrar na religião um recurso pedagógico que permite levar os homens individualistas dos tempos democráticos a servir ao interesse público. Essa compreensão revela o modo como Tocqueville espera tirar proveito político dos hábitos e das virtudes religiosas, pois, na concepção do autor, o estabelecimento do sistema político comunal naquelas circunstâncias históricas reais foi possível porque havia um conjunto de preceitos com conteúdos políticos encerrados no interior de uma orientação religiosa chamada de puritanismo.

De uma perspectiva política, o puritanismo representava a promessa de realização de uma democracia cujo arranjo no plano social foi, por assim dizer, um evento simultâneo ao desembarque dos colonos ingleses no norte do Novo Mundo. Foi graças aos princípios políticos de um republicanismo democrático, que “na época das primeiras emigrações, o governo comunal, esse germe fecundo das instituições livres, já havia profundamente im-

---

6 - O estado de espírito de Tocqueville, diante do que ele acredita ser a união entre liberdade e religião, pode ser sugerido pelas palavras que, numa carta de 1843, ele endereça a um amigo: “Um dos meus sonhos, o principal ao entrar na vida política, era o de trabalhar por conciliar o espírito liberal e o espírito religioso, a nova sociedade e a Igreja”. (Citado por Jean-Jacques Chevallier, em *As Grandes Obras Políticas de Maquiavel a Nossos Dias*. Rio de Janeiro – RJ, Agir Editorial, 2002).

pregnado os hábitos dos ingleses” (TOCQUEVILLE, 2005, tomo I, p. 37).

*A maior parte da América inglesa foi povoada por homens que, depois de terem se furtado à autoridade do papa, não se haviam submetido a nenhuma supremacia religiosa; eles levavam ao novo mundo um cristianismo que eu não poderia pintar melhor do que chamando-o democrático e republicano. Isso favorecera singularmente o estabelecimento da república e da democracia nos negócios. Desde o princípio, a política e a religião estavam de acordo, e desde então não deixaram de estar. (TOCQUEVILLE, 2005, tomo I, p. 338).*

A independência do sistema de comunas, que o espírito político do puritanismo fez nascer, oferecia a particularidade de haver dado, ao curso do encadeamento das instâncias políticas na América, uma ordem inversa à do quadro geral da Europa<sup>7</sup>. O processo de formação da autoridade política no Novo Mundo seguiu um movimento ascendente para o qual a comuna foi o ponto de partida ou o primeiro estágio; não por ser cronologicamente o primeiro a se ter constituído, mas antes em virtude de ser a unidade simples de associação política da qual nascem o poder e a soberania das comunidades locais politicamente livres. Dessas instituições comunais, que nas palavras de Thomas Jefferson são “as repúblicas elementares”,<sup>8</sup> emanavam a autoridade que é o único meio legítimo (e consoante ao espírito de liberdade que presidiu a fundação da comuna) para a integração dessas unidades em instituições político-administrativas superiores, como os condados, os Estados, e, por fim, a União.

Internamente, a consequência direta da independência comunal, associada à não admissão da lei de representação para as questões políticas locais, foi a liberdade, isto é, a existência de uma vida política autêntica e ativa. Porque, com uma população média estimada entre 2000 e 3000 habitantes e um total de 19 cargos político-administrativos com duração de um ano, tinha-se uma grande rotatividade no exercício desses cargos políticos<sup>9</sup>, somada às demandas participativas das reuniões comunais a que todos os cidadãos eram

---

7 - Com o inconveniente de terem sido conduzidas por um conjunto de princípios morais de extrema rigidez (no qual o puritanismo se baseava), as primeiras legislações penais da Nova Inglaterra apresentavam o problema jurídico de compreender toda a sorte de pecados como crimes que, em alguns casos, deviam ser punidos com a morte – como o adultério em Massachusetts e Hutchinson. O estreito espírito de seita foi, para Tocqueville, o aspecto indesejável que se desenvolveu ao lado das virtudes políticas do puritanismo.

8 - Citado por Hannah Arendt em *Da Revolução*, p. 181. São Paulo: Ática, 1988.

9 - “Contam-se dezenove funções principais na comuna. Cada habitante é obrigado, sob pena de multa, a aceitar essas diferentes funções; mas também a maioria delas é retribuída, de modo que os cidadãos pobres possam consagrar-se a elas seu tempo sem ter prejuízo” (TOCQUEVILLE, 2005, tomo I, p. 75).

convocados e àquelas que eles mesmos exigiam junto aos magistrados eletivos (chamados de *select-men*<sup>10</sup>), os quais nada eram senão “executantes das vontades populares”<sup>11</sup> (TOCQUEVILLE, 2005, tomo I, p. 73).

Por essa razão, o poder político que surgiu na Nova Inglaterra havia se baseado no fato consumado de que sua fonte reside no povo. Nisso consiste a soberania que Tocqueville afirma pairar sobre no sistema político anglo-americano que permitia aos cidadãos tratar “na praça pública” ou “no seio da assembleia geral dos cidadãos, como em Atenas, os assuntos públicos concernentes ao interesse de todos” (TOCQUEVILLE, 2005, tomo I, p. 49).

Um dos mais claros sinais de que a liberdade política é um projeto contido nas próprias entranhas do puritanismo se mostra, para Tocqueville, nas atribuições que essa orientação religiosa conferiu à *educação pública*. A atitude política do puritanismo, no que diz respeito à educação de todos os cidadãos, representa para o autor o encerramento do quadro geral que ele deseja pintar para expor o ponto de partida e o caráter singular da civilização americana, em cujo seio se deu a combinação dos espíritos de religião e liberdade.

A legislação comunal, que no plano político recusou a representatividade e fez da participação direta um aspecto indispensável à vida política, entendendo a educação pública como algo que é de sumo interesse para o Estado – por ajudar a preparar os homens para a liberdade com o auxílio das luzes –<sup>12</sup>, tornou-a universal e obrigatória. À formação escolar destinada a preparar os homens para vida social e política, seguia-se o processo político-pedagógico prático, pressuposto nas relações instituídas pelo sistema de comunas, que nas palavras de Tocqueville:

---

10 - O número dos *select-men* variava entre três e seis, segundo as dimensões das comunas.

11 - “Somente os *select-men* têm o direito de convocar a reunião comunal (*town-meeting*), mas pode-se leva-los a fazer isso. Se dez proprietários concebem um novo projeto e querem submetê-lo ao assentimento da comuna, eles reclamam uma convocação geral dos habitantes, os *select-men* são obrigados a inscrevê-la e conservam tão só o direito de presidir a assembleia” (TOCQUEVILLE, 2005, tomo I, p. 74).

12 - A lei educacional citada por Tocqueville, da qual se faz seguir as disposições que autorizam as taxações para a construção e manutenção das escolas públicas, é a seguinte: “Considerando que satanás, o inimigo do gênero humano, encontra na ignorância dos homens suas mais poderosas armas e que é importante que as luzes que nossos pais trouxeram não fiquem sepultadas em seus túmulos; considerando que a educação das crianças é um dos primeiros interesses do Estado, com assistência do Senhor...” (Citado por Alexis de Tocqueville em *A Democracia na América*, tomo I, p. 49. São Paulo: Martins Fontes, 2005 – minha ênfase). Note-se que as luzes não são mais do que um auxílio na tarefa de preparar os cidadãos para a vida social como um todo e para o exercício da participação política. A parte mais importante dessa tarefa cabe à religião, porque a ela cabe o papel de forjar virtudes e costumes sobre os quais a democracia se alicerça. Essa é tanto a posição de Tocqueville quanto a convicção geral dos anglo-americanos, que o autor trata de expor em passagens como a que se segue: “Com os que não creem escondendo sua incredulidade e os que creem mostrando sua fé, cria-se uma opinião pública favorável à religião: amam-na, sustentam-na, honram-na [...]” (TOCQUEVILLE, 2005, tomo I, p. 353), seja porque acreditam nela ou porque reconhecem os benefícios que acompanha os costumes que elas faz nascer.

“[...] estão para a liberdade assim como as escolas primárias estão para a ciência: elas colocam ao alcance do povo, fazem-no provar seu uso tranquilo e habituam-no a empregá-la. Sem instituições comunais uma nação pode se dotar de o governo livre, mas não possui o espírito de liberdade” (TOCQUEVILLE, 2005, tomo I, p. 71).

O benefício geral que resultou da conjugação anglo-americana do assim chamado espírito de liberdade ao espírito de religião consistiu em fazer cooperar duas forças (reunidas em um só povo) que, embora não fossem contrárias, eram, por assim dizer, distintas e indiferentes, uma em relação à outra, isto é, a força conservadora<sup>13</sup> dos *ardentes sectários* e a força revolucionária dos *inovadores exaltados*. Desta (a força revolucionária dos inovadores exaltados), adveio a maleabilidade inovadora com que, tratando os assuntos humanos, os habitantes das colônias inglesas estabelecidas ao norte dos Estados Unidos puderam apresentar ao mundo um espetáculo político sem precedentes. Ao passo que daquela (a força conservadora dos *ardentes sectários*), surgiram os princípios morais e os costumes que fizeram da religião “a primeira instituição política”, embora ela “nunca se envolvesse no governo da sociedade”<sup>14</sup> (TOCQUEVILLE, 2005, tomo I, p. 344).

Os riscos da liberdade e da flexibilidade com que os habitantes da Nova Inglaterra tratavam as leis e as instituições políticas eram, em certa medida, redimidos pelo espírito de religião. O temor religioso inspirava os limites morais de que a ação livre precisava. “Democracia é movimento perpétuo, agitação permanente do mundo político. Religião é imutabilidade, imobilidade do mundo moral” (CHEVALLIER, 2002, p. 276). Para Tocqueville, “longe de se prejudicarem, estas duas tendências, em aparência tão opostas, caminham em concórdia e parecem prestar-se um apoio mútuo” (TOCQUEVILLE, 2005, tomo I, p. 52). A liberdade encontra no espírito de religião os limites morais que o seu bom exercício político requer.

Quando Tocqueville fala do espírito religioso e do modo como ele concorre para a realização do Estado democrático, o espaço destinado à relevância política da religião (inicialmente limitado àquela religião que presidiu a instituição do sistema comunal por conter em si princípios políticos democráticos e republicanos) se alarga para dar lugar a todas as

---

13 - Não me refiro à aceção política do termo, mas a um conjunto de princípios morais fixos, como é próprio de qualquer religião.

14 - “A liberdade vê na religião a companheira de suas lutas e de seus triunfos, o berço de sua infância, a fonte divina de seus direitos. Ela considera a religião como a salvaguarda dos costumes; os costumes como a garantia das leis e penhor de sua própria duração” (TOCQUEVILLE, 2005, tomo I, p. 52).

religiões que, embora não encerrem preceitos políticos consoantes à liberdade, facilitam singularmente o seu uso, na medida em que modelam os costumes e dão à ação os limites morais que refreiam a possibilidade sempre aberta de tudo poder fazer.

A moderação advinda do auxílio mútuo que os espíritos de religião e liberdade se prestaram, aliada às condições de igualdade que a Nova Inglaterra conheceu em virtude de causas acidentais, foram as condições fundamentais para o surgimento do sistema político comunal entre os anglo-americanos, de acordo com Tocqueville.

Por compreender a religião como um instrumento no desenvolvimento de costumes que importam à vida política, transparece nas ideias de Tocqueville a compreensão de Montesquieu, segundo a qual a democracia tem como móvel a virtude. Contudo, para o autor de *A Democracia na América*, apesar de servir à política, a religião exerce um papel exclusivamente moral de vincular um homem ao outro, munindo-os de valores comuns, sem realizar qualquer intervenção direta na política. Seu papel político se restringe ao de um *instrumental pedagógico*, que prepara e fomenta a vida pública.

## Referências Bibliográficas

ARENDDT, Hannah. *The Human Condition*. Chicago, The University of Chicago Press, 1998.

\_\_\_\_\_. *Da Revolução*. São Paulo: Ática, 1988.

BOESCHE, Rober. *The Strange Liberalism of Alexis de Tocqueville*. Ithaca and London, Cornell University Press, 1987.

CHEVALLIER, Jean-Jaques. *As Grandes Obras Políticas de Maquiavel aos Nossos Dias*. Rio de Janeiro, Editora Agir, 2002. 8ª edição e 4ª impressão.

JASMIM, Marcelo Gantus. *Alexis de Tocqueville: A historiografia como ciência da política*. Rio de Janeiro, ACESS, 1997.

TOQUEVILLE, Alexis. *A Democracia na América*. Tomos I e II. Tradução: Eduardo Brandão. Prefácio, bibliografia e cronologia: François Furet. São Paulo, Martins Fontes, 2004.



# Como tornar a bivalência compatível com a intuição de que o futuro é aberto?

## *How to make bivalence compatible with the intuition that the future is open?*

Matheus Silva  
Universidade Federal do Rio Grande do Sul

### Resumo

Suponha que um almirante diga que haverá uma batalha naval amanhã. Essa proposição, considerada no momento em que é asserida, é verdadeira ou falsa? Essa pergunta gera perplexidade, pois, por um lado, temos a intuição de que qualquer proposição é verdadeira ou falsa, ainda que não saibamos qual é o seu valor de verdade. Essa é a intuição que sustenta o princípio da bivalência, segundo o qual toda proposição é verdadeira ou falsa. Por outro lado, também temos a intuição de que essa proposição não pode ser verdadeira ou falsa agora, no momento em que é asserida, pois o que a tornará verdadeira ou falsa ocorrerá apenas amanhã. Essa é a intuição de que o futuro é aberto. Contudo, como eu posso aceitar a intuição de que o almirante tem escolhas livres se os valores de verdade das proposições acerca de suas escolhas futuras já estão determinados? Esse é o problema dos futuros contingentes. Neste artigo, apresento soluções para esse problema que nos permitem manter o princípio da bivalência sem abdicar da intuição de que o futuro é aberto. Embora alguns argumentos apresentados sejam de minha autoria, este artigo deve ser entendido como um texto introdutório sobre o assunto e não como um artigo de investigação: o meu objetivo foi é recolher e apresentar de maneira acessível os diferentes argumentos da literatura e não fornecer uma contribuição inovadora ao assunto.

## **Palavras-chave**

Bivalência, futuro aberto, determinismo, valores de verdade.

## **Abstract**

Suppose an admiral says that there will be a sea-battle tomorrow. Is the relevant proposition, considered in the moment of its assertion, true or false? That question is perplexing, since on the one hand we have the intuition that any proposition is true or false, even if we do not know its truth-value. This is the intuition supporting the principle of bivalence, according to which every proposition is either true or false. On the other hand, we also have the intuition that this proposition is not true or false now - in the moment of its assertion - since what makes it true or false will obtain only tomorrow. This is the intuition that the future is open. However, how can I accept the intuition that the admiral has free choices if the truth-values of the propositions about his futures choices are already determined? This is the problem of the future contingents. In this paper, I present solutions to this problem that uphold the principle of bivalence without abdicating the intuition that the future is open. Although some arguments presented are my own, this paper should be understood as an introductory text on the subject and not as a research paper: my aim was not to present a contribution to the subject but rather to collect and present in an accessible manner different arguments from the extant literature.

## **Keywords**

bivalence, open future, determinism, truth-values.

## **1. O problema dos futuros contingentes**

Suponha que um almirante diga que haverá uma batalha naval amanhã. Essa proposição, considerada no momento em que é asserida, é verdadeira ou falsa? Essa pergunta gera perplexidade, pois, por um lado, temos a intuição de que qualquer proposição é verdadeira ou falsa, ainda que não saibamos qual é o seu valor de verdade. Essa é a intuição que sustenta o princípio da bivalência, segundo o qual toda proposição é verdadeira ou falsa. Por outro lado,

também temos a intuição de que essa proposição não pode ser verdadeira ou falsa *agora*, no momento em que é asserida, pois o que a tornará verdadeira ou falsa ocorrerá apenas amanhã. Essa é a intuição de que o futuro é aberto. Mas como posso conciliar a intuição de que o futuro é aberto se agora é verdadeiro ou falso que haverá uma batalha naval amanhã, i.e., como posso conciliar a intuição que o almirante tem escolhas livres se o valor de verdade da sua escolha futura já está determinado? Esse é o problema dos futuros contingentes<sup>1</sup>.

É importante observar que o que torna atraente a intuição de que o futuro é aberto não é o simples fato de a proposição em causa referir um evento futuro, pois há eventos futuros que são perfeitamente determinados – posso afirmar seguramente que daqui a cinco anos “se P então P” ainda será uma verdade lógica. O ponto do exemplo da batalha naval é que se trata de uma proposição que refere um evento futuro que envolve ações humanas. Como posso afirmar que o futuro é aberto e que o almirante tem escolhas livres se posso determinar com antecedência qual será sua escolha futura? Isso nos forçaria a admitir o determinismo, a hipótese de que todos os acontecimentos e estados de coisas do mundo estão determinados por acontecimentos e estados de coisa que lhes são anteriores, ou o fatalismo, a hipótese de todas as ações humanas e todos os acontecimentos estão predeterminados e nada do que fizermos poderá alterar o rumo dos acontecimentos. Essas hipóteses representam uma ameaça à nossa intuição de que as pessoas têm livre-arbítrio e de que as suas ações futuras não podem ser determinadas pelo presente. De fato, essa intuição motiva outra intuição: a ideia de que o futuro é aberto e não está determinado. Se as pessoas podem decidir que haverá uma batalha naval amanhã ou não, até amanhã o futuro permanecerá aberto e indeterminado, pois ambas as opções estão sobre a mesa.

## 2. Soluções para o problema

Há pelo menos três tipos de soluções para o problema: (1) Manter as intuições de que somos livres e o futuro é aberto, e admitir que as proposições acerca dos futuros contingentes não têm valores de verdade no presente, abandonando a bivalência; (2) Sustentar que essas proposições possuem valores de verdade determinados mesmo no presente, assim mantendo assim a bivalência, e descartar a intuição de que o futuro é aberto – essa é a saída do determinista e do fatalista; (3) Defender que as proposições acerca de futuros contingentes são bivalentes e manter a intuição de que o futuro é aberto. Os argumentos que apresento em seguida são do terceiro tipo.

---

1 - O problema foi apresentado originalmente por Aristóteles, no capítulo 9 de *De interpretatione*.

*O que será o caso não tem de ser o caso*

Se a proposição de que haverá uma batalha naval amanhã é verdadeira hoje, a aparente inevitabilidade da ocorrência da batalha naval amanhã implicaria que o futuro é determinado. O erro desse raciocínio é confundir o que será o caso com o que não pode deixar de ser o caso. O fato de que não haverá uma batalha naval amanhã não implica que não poderia haver uma batalha naval amanhã. De fato, tudo poderia ser diferente, pois o que será o caso não precisa ser o caso. O que irá acontecer não é inevitável: algo que irá acontecer não é algo que tem de acontecer (TORRE, 2011, p. 365).

Eu não posso alterar o futuro em relação ao que ele será, mas posso alterar o futuro em relação ao que ele poderia ter sido. Se eu optar por comer uma maçã, ao invés de um pão francês no café da manhã, altero o futuro em relação ao que ele poderia ter sido – eu poderia ter comido um pão francês ou outra coisa qualquer. Mas eu não tenho o poder de alterá-lo em relação ao que ele será. Similarmente, eu não posso alterar o presente em relação ao que ele é, mas apenas em relação ao que ele poderia ter sido, e não posso alterar o passado em relação ao que ele foi, mas no passado eu poderia alterá-lo em relação ao que ele poderia ter sido (SWARTZ, 2001, p. 227).

Na realidade, os meus próprios atos determinam em parte o que o futuro será e os fatos contingentes sobre como as coisas serão dependem do que fazemos agora (TORRE, 2011, p. 370). Eu poderia agir de uma maneira na qual o café da manhã que eu terei amanhã seja diferente do que ele de fato será – eu poderia deixar de comprar pães hoje, por exemplo. Há aqui uma dependência contrafactual do futuro em relação ao presente: o que eu terei no café amanhã depende contrafactualmente do que eu fizer agora. A abertura do futuro é determinada por sua dependência contrafactual e não por uma suposta indeterminação de valores de verdade das proposições acerca dos futuros contingentes (LEWIS, 1987, p. 38).

A habilidade de agir livremente influencia o futuro apenas no sentido de determinar aquilo que será o caso, mas não altera o fato de que algo será o caso. O futuro é aberto no sentido de que poderia ser diferente do que de fato será. Da mesma forma que há apenas um passado que poderia ser diferente se o agente livre escolhesse outra opção, também há apenas um futuro, mas que poderia ser diferente se o agente livre escolhesse outra opção. Os eventos presentes estarão relacionados com um único futuro, que será a única continuação dos eventos atuais. Há vários futuros possíveis, mas apenas um será efetivo.

O fato de ser verdade que amanhã haverá uma batalha naval não implica que é necessário ou inevitável que amanhã ocorrerá uma batalha naval. Uma vez que o erro desse raciocínio tenha sido exposto, a fragilidade da ameaça determinista torna-se óbvia. A propo-

sição de que haverá uma batalha naval é interpretada pelo determinista como “É inevitável que haja uma batalha naval amanhã” ou “É necessário que haja uma batalha naval amanhã”. Porém, se aceitássemos essa interpretação, não deveríamos interpretar essas proposições como indeterminadas, mas sim como falsas. Essa interpretação parece ser pressuposta por autores que mantêm a bivalência, admitindo ao mesmo tempo a posição equívoca de que todas as afirmações acerca de futuros contingentes são falsas. A ideia é que, se interpretarmos afirmações do gênero “Haverá uma batalha naval amanhã” como “É inevitável que haja uma batalha naval amanhã”, podemos seguramente concluir que qualquer afirmação desse gênero é falsa, pois se algo é resultado de um agente livre não será o caso até que o agente decida o que é. Para mantermos ainda a bivalência, basta sustentar que a sua negação, i.e., “Não haverá uma batalha naval amanhã”, é verdadeira, pois o que ela realmente significa é “Não é inevitável que haja uma batalha naval amanhã” (PRIOR, 1967 apud RICE, 2010). Porém, uma explicação tão convoluta é dispensável, uma vez que o que será o caso não tem de ser o caso.

### *O argumento da rigidez modal*

Toda proposição é logicamente impossível ou logicamente possível. Proposições logicamente possíveis, por sua vez, dividem-se entre necessariamente verdadeiras e contingentes. Proposições logicamente impossíveis são necessariamente falsas. O estatuto modal de uma proposição diz respeito ao fato de ela ser contingente, necessariamente verdadeira ou necessariamente falsa. O estatuto modal de uma proposição é rígido: uma proposição contingente nunca pode se tornar necessária ou impossível, assim como uma proposição impossível nunca pode se tornar necessária ou contingente e uma proposição necessária nunca pode se tornar impossível ou contingente.

Considerando agora o problema dos futuros contingentes, podemos dizer que uma proposição acerca de um futuro contingente permanece contingente ainda que ela seja verdadeira. Certamente, a proposição não se tornará necessária por corresponder ao curso que a história irá tomar. Assim, se a proposição de que haverá uma batalha naval amanhã é verdadeira agora, haverá uma batalha naval amanhã. Porém, esse evento não é inevitável, pois a proposição é contingente e poderia ser igualmente falsa (SWARTZ, 2004, seção 5).

*A intuição de que o futuro não é aberto resulta de uma falácia modal*

O argumento anterior é necessário para compreender a acusação de que o problema dos futuros contingentes resulta de uma falácia modal. Podemos representar o problema dos futuros contingentes da seguinte forma: suponha que dois almirantes, A e B, estão se preparando para a batalha naval que ocorrerá amanhã. Sabemos que toda proposição é verdadeira ou falsa (princípio do terceiro excluído) e que nenhuma proposição é ambas verdadeira e falsa (princípio da não-contradição). Assim, considerando as proposições “A vence” e “é falso que A vence” devemos admitir que uma deve ser verdadeira, enquanto a outra é falsa. Se a proposição “A vence” é verdadeira hoje, então, independentemente de todo o planejamento dos almirantes e esforço dos soldados, A vencerá, i.e., a proposição “A vence” será uma verdade necessária. Inversamente, se a proposição “A vence” for falsa hoje, então, independentemente do que aconteça durante a batalha, A perderá, ou seja, a proposição de que A vence será necessariamente falsa. Se interpretarmos a proposição “O almirante A vence” como “A” e a proposição “O almirante B vence” como “B”, podemos apresentar o raciocínio que leva a essa conclusão determinista da seguinte maneira:

**(Premissa)** Ou A ou B [uma das duas proposições deve ser verdadeira - admissão do princípio de legítima dicotomia]

**(Premissa)**  $\neg\Diamond(A \wedge B)$  [não é possível que ambas as proposições sejam verdadeiras - admissão do princípio da não-contradição]

**Logo,**

$A \supset \Box A$  (Se A é verdadeira, então A é necessariamente verdadeira)

$A \supset \neg\Diamond\neg A$  (Se A é verdadeira, então não é possível que A seja falsa)

$A \supset \Box\neg B$  (Se A é verdadeira, então B é necessariamente falsa)

$A \supset \neg\Diamond B$  (Se A é verdadeira, então não é possível que B seja verdadeira)

$B \supset \Box B$  (Se B é verdadeira, então B é necessariamente verdadeira)

$B \supset \neg\Diamond\neg B$  (Se B é verdadeira, então não é possível que B seja falsa)

$B \supset \Box \neg A$  (Se B é verdadeira, então A é necessariamente falsa)

$B \supset \neg \Diamond A$  (Se B é verdadeira, então não é possível que A seja verdadeira)

Cada uma das conclusões afirma que da verdade de A ou B se segue que A e B são ou necessariamente verdadeiras ou necessariamente falsas, mas nenhuma dessas conclusões se segue das premissas. Na realidade, todas as conclusões são falsas, pois A é contingente e B é contingente. Pensar diferente seria violar o princípio de rigidez modal. Na realidade, o argumento é um tipo de falácia modal, pois atribui impossibilidade à verdade ou falsidade de A ou B quando deveria atribuir a impossibilidade à relação entre a verdade de A e a falsidade de B. Se A é verdadeira, não é impossível que B seja verdadeira, pois B é contingente. O que é impossível é que B seja verdadeira enquanto A é verdadeira, o que é bem diferente. As únicas conclusões que podemos inferir validamente das premissas são “ $\Box(A \supset \neg B)$ ” e “ $\Box(B \supset \neg A)$ ” ou, para utilizar outras expressões que são logicamente equivalentes, “ $\neg \Diamond(A \wedge B)$ ” e “ $\neg \Diamond(B \wedge A)$ ”. Mas como “ $\neg \Diamond(A \wedge B)$ ” é equivalente a “ $\Box(A \supset \neg B)$ ” e não a “ $A \supset \Box \neg B$ ”, o argumento é inválido (SWARTZ, 2004, seção 6.a).

### *Uma teoria da correspondência mal formulada*

Outra maneira de abordar o problema é identificar a sua fonte na admissão de uma teoria da verdade como correspondência (IACONA, 2007, p. 55) (RICE, 2010). Essa teoria pode ser resumida na afirmação de que uma proposição é verdadeira se, e somente se, existe um fato correspondente a ela. Uma consequência dessa teoria é que se uma proposição P é verdadeira em certo momento, então o estado do mundo naquele momento deve ser tal que torne P verdadeira. Assim, se uma afirmação acerca de um futuro contingente for verdadeira agora, algo deve ocorrer agora que determine o que ocorrerá amanhã. Se quisermos ainda manter a intuição de que o futuro é aberto sem abdicar da teoria da correspondência, parece não haver alternativa que não seja o abandono da bivalência. Há, contudo, um modo de resolver esse conflito de intuições sem que tenhamos que abdicar da bivalência ou da teoria da correspondência. A solução consiste em apresentar uma versão do princípio da correspondência que pode ser considerada plausível por razões independentes e que não irá gerar os mesmos problemas.

Começemos com a suposição de que a aceitação da teoria da correspondência im-

plica que deveríamos abdicar da bivalência nesses casos. A proposição de que haverá uma batalha naval amanhã ainda não possuiria valores de verdade no momento em que fosse asserida, pois não haveria nada no mundo para determinar que ela fosse verdadeira. Essa suposição parece indefensável se considerarmos as atribuições de verdade às proposições acerca de eventos passados. Suponha que alguém diga, apontando para um local, “Um dinossauro estava situado neste ponto a um milhão de anos atrás”. Para que essa proposição seja verdadeira, é necessário que o estado presente do mundo seja tal que determine a sua verdade? Isso não parece ser o caso (RICE, 2010). Na realidade, afirmações acerca do passado geralmente são verdadeiras quando dizem respeito a algo que existiu, mas não existe mais. A inexistência do objeto referido pela proposição pode até mesmo ser o que torna a proposição verdadeira. Ao observar o galinheiro, eu digo “Havia uma raposa no galinheiro”. Essa proposição somente é verdadeira quando não há mais uma raposa no galinheiro (WESTPHAL, 2006, p. 138). Outro exemplo é a proposição “Brutus assassinou César”, asserida hoje. É plausível supor que, se a proposição for verdadeira, ela é tornada verdadeira por algo, mas não parece razoável afirmar que esse algo deve existir hoje. Como a proposição é acerca de um evento que ocorreu em 44 a. C., é razoável pensar que o que a torna verdadeira é o próprio evento do passado.

O mesmo raciocínio aplica-se a afirmações contingentes acerca do futuro. Embora elas sejam verdadeiras em virtude de algo, é natural assumir que esse algo existirá no futuro. A suposição de que a verdade é correspondência com a realidade é perfeitamente compatível com a suposição de que o que torna uma proposição asserida em certo tempo verdadeira é algo que existe em um tempo diferente (IACONA, 2007, p. 55-56). Isso mostra que a teoria da correspondência precisa ser ligeiramente modificada. Ao invés de afirmarmos que uma proposição é verdadeira num certo momento em virtude do estado do mundo naquele momento, devemos dizer que uma proposição é verdadeira somente se há algo ou havia algo em virtude do qual ela é verdadeira, ou se haverá algo em virtude do qual ela é verdadeira (WESTPHAL, 1996, p. 4).

Assim, a proposição “Haverá uma batalha naval amanhã” corresponderá à realidade somente se a batalha naval ocorrer amanhã. Correspondência à realidade nesse caso significa corresponder ao fato de que a história ocorrerá desse modo, i.e., de que este futuro, e não outros futuros possíveis, será atualizado. O que será o caso ainda não está atualizado. Eventos que ainda serão presentes e tornam afirmações acerca do futuro verdadeiras são exatamente como eventos presentes, quando o seu tempo chega. Se eu digo hoje “O dia será ensolarado amanhã” essa afirmação não é verdadeira hoje porque no futuro o dia está ensolarado, mas sim porque o dia será ensolarado amanhã, quando o seu tempo chegar (WESTPHAL, 2006, p. 2).

*Uma disjunção não pode ser verdadeira se suas disjuntas não têm valores de verdade*

Se alguém deseja sustentar que o futuro é aberto e que a bivalência não se aplica às afirmações acerca do futuro, terá de admitir que uma proposição como “Ou haverá uma batalha naval amanhã ou não haverá uma batalha naval amanhã” é verdadeira, ainda que suas disjuntas não tenham valores de verdade no presente. Essa resposta é difícil de aceitar, pois temos a forte intuição de que uma disjunção não pode ser verdadeira se suas disjuntas não possuem valores de verdade (SWARTZ, 2001, p. 226) (RICE, 2010). Como está estabelecido que haverá ou não haverá uma batalha naval amanhã (essas são as únicas possibilidades dada a natureza do mundo), também está estabelecido que ou é verdade que haverá uma batalha naval amanhã ou é falso que haverá uma batalha naval amanhã.

Um exemplo torna esse ponto mais claro. Suponha que no segundo turno das próximas eleições presidenciais haverá um candidato do PT e um candidato do PSDB. A propriedade de ser petista é possuída por um candidato, mas não pelo outro, e, no dia anterior às eleições, eu não estou em posição de saber quem vencerá. Ambos os candidatos partilham a propriedade de ser petista ou não petista. Contudo, mesmo se no dia anterior às eleições eu não estiver em posição de saber quem vencerá, eu estarei em posição de saber que o vencedor será ou petista ou não petista e isso é tudo o que é preciso para estabelecer a verdade de uma afirmação acerca de um futuro contingente. Eu sei que a disjunção “O futuro presidente será ou petista ou não petista” é verdadeira e, conseqüentemente, também posso afirmar que a disjunção “Ou ‘O futuro presidente será petista’ é verdadeira ou é falsa” é verdadeira. Assim, eu posso dizer algo verdadeiro sobre o candidato vitorioso antes de saber quem será o vencedor (IACONA, 2007, p. 58-59).

Uma maneira de contornar essas dificuldades seria manter que, embora as disjuntas não tenham valores de verdade determinados agora, - pois o futuro é metafisicamente indeterminado no presente -, elas são determinadamente ou verdadeiras ou falsas. O valor de verdade que elas poderão adquirir será determinado com o decorrer do futuro (BARNES & CAMERON, 2011, p. 4-5). O problema dessa solução é que ela introduz uma distinção artificiosa e *ad hoc* entre a atribuição de valores de verdade à disjunção e a atribuição de valores de verdade às suas disjuntas. Não parece haver qualquer razão para admitir essa distinção que não seja o desejo de evitar uma conclusão filosófica rival.

Diante dessas observações parece mais plausível simplesmente manter a bivalência. Para utilizar o vocabulário proposto na objeção anterior, podemos dizer que é presentemente determinado se haverá uma batalha naval amanhã, ainda que o futuro seja contingente e sejamos livres. Afirmações acerca de futuros contingentes também possuem valores de verdade determinados. Os valores de verdade que elas possuem podem ser identificados agora ou na

medida em que os eventos se desdobram. O desdobramento do futuro dependerá dos eventos e escolhas realizadas agora. Não há uma ausência de valor de verdade ou um suposto terceiro valor de verdade (o indeterminado) que possamos atribuir às disjuntas no momento em que a disjunção é asserida – há apenas a nossa ignorância acerca do valor de verdade das disjuntas. O futuro é aberto, mas com o decorrer inevitável do tempo, haverá apenas um único futuro determinado e assim podemos dizer que as afirmações contingentes acerca do futuro são verdadeiras ou falsas agora, pois têm como referência o futuro determinado.

*Em retrospectiva as predições parecem ter valores de verdade no momento em que são asseridas*

Um dos melhores argumentos contra a recusa da bivalência consiste em observar que, uma vez que algo tenha ocorrido de acordo com as nossas previsões, nós podemos olhar para trás e determinar quais de nossas predições passadas eram verdadeiras ou falsas (BARNES & CAMERON, 2011, p. 20). Se ontem eu afirmei que choveria hoje, mas isso não aconteceu, não poderei defender que essa predição era indeterminada. Pelo contrário, parece que a predição já era falsa no momento em que foi asserida, ainda que eu não soubesse disso. Esse fato parece indicar precisamente que a nossa incapacidade de atribuir um valor de verdade a algumas afirmações acerca de eventos futuros se deve a razões meramente epistêmicas.

É claro que o opositor da bivalência poderia replicar que este é precisamente o ponto: a proposição possuirá um valor de verdade apenas quando o evento que a tornar verdadeira ocorrer. Suponha que eu tenha afirmado ontem que choveria hoje e tenha acertado em minha previsão. De fato choveu. Estou justificado em dizer que a minha afirmação é verdadeira, mas apenas agora, em função de um evento que ocorreu hoje. Seria um equívoco dizer que a proposição era verdadeira no momento em que foi asserida, pois o evento que a tornaria verdadeira não tinha ocorrido ainda.

O problema dessa objeção é que a nossa intuição é de que a proposição já era verdadeira no passado e não que ela tenha se tornado verdadeira agora. A única diferença entre o momento passado em que ela foi asserida e o momento presente em que constatamos o fato é de que antes não sabíamos qual era o seu valor de verdade.

*Em nosso cotidiano temos conhecimento de futuros contingentes*

Em nosso cotidiano, muitas vezes atribuímos valores de verdade a afirmações acerca de futuros contingentes. Todos os meses as pessoas compram bilhetes da loteria federal com a esperança de ganhar milhões de reais. Elas certamente admitirão que uma proposição como “Algum bilhete será o vencedor na data marcada para o sorteio dos números” era verdadeira muito antes da data do sorteio (SWARTZ, 2004, seção 4.a). Não é preciso fazer muito esforço para encontrar inúmeros exemplos semelhantes.

Uma consequência natural desses exemplos é que em muitos casos também alegamos ter conhecimento de eventos futuros. De fato, a nossa própria sobrevivência depende desse conhecimento. Antes de atravessar uma avenida, eu espero o sinal fechar, pois sei que os motoristas não imaginam que qualquer pedestre seja maluco o bastante para atravessar a avenida com o sinal aberto. Eu também sei que, se atravessar a rua quando o sinal estiver aberto, serei atropelado violentamente em poucos segundos. Exemplos similares abundam e representam um forte indício de que sabemos em inúmeros casos como serão os eventos futuros envolvendo ações humanas (SWARTZ, 2004, seção 4.b).

De fato, parece que temos o conhecimento acerca de eventos futuros, ainda que eles ocorram num período muito distante. Suponha que um fiscal sanitário municipal apareça religiosamente todo primeiro dia do mês em uma loja para realizar uma inspeção. Há dez anos que o fiscal nunca falhou ao serviço. Eu posso assegurar, portanto, que o fiscal virá no primeiro dia do próximo mês de maio, julho, etc, nos próximos vinte anos, enquanto não estiver aposentado ou morto. Eu tenho o conhecimento de um evento envolvendo uma ação humana que ocorrerá daqui a vinte anos, mas o futuro ainda permanecerá aberto.

Alguém poderia objetar aqui que não temos conhecimento do futuro nesses casos. Afinal de contas, podemos errar flagrantemente em nossas previsões (SWARTZ, 2004, seção 4.b). Para manter um dos exemplos anteriores, se eu atravessasse a avenida enquanto o sinal estivesse aberto, o motorista que viesse em minha direção poderia estar atento a ponto de desviar no último segundo. Como eu não tenho garantias de que isso não poderia ocorrer, eu não posso alegar ter conhecimento de eventos futuros nesses casos.

Essa objeção, contudo, peca por estabelecer critérios exigentes demais. Não precisamos ter garantias de que nunca cometeremos um erro para assegurar que temos conhecimento de algo. É claro, se cometermos um erro em uma previsão não teremos conhecimento nesse caso, mas “ter conhecimento” não equivale a “ter conhecimento infalível”. As expressões “x sabe” e “x sabe com certeza” não são mais equivalentes do que “x vê a montanha distante” e “x vê a montanha distante perfeitamente, até mesmo as veias das folhas nas árvores”. Somos

falíveis e podemos cometer erros. Além disso, se o único conhecimento genuíno for infalível, muito provavelmente não teremos conhecimento de boa parte do passado (SWARTZ, 2004, seção 4.b). O passado aqui pode incluir até mesmo as nossas vidas pessoais. Se eu não tenho garantias infalíveis de que a minha memória é precisa, não poderei dizer que tenho conhecimento de quem são meus pais, em qual curso estou matriculado, etc. Essa consequência, no entanto, é implausível.

Um argumento que é uma consequência natural dos argumentos anteriores é o seguinte: se o futuro é reconhecidamente aberto quando posso identificar qual é o valor de verdade da proposição acerca do futuro, também será aberto quando não posso identificar o seu valor de verdade. Se o futuro que envolve as minhas ações no café de amanhã é aberto, também será aberto o futuro que envolve a ocorrência ou ausência da batalha naval. Qualquer razão para admitir que o futuro é aberto no primeiro caso será também uma razão para admitir que o futuro será aberto no segundo caso. Além da nossa manifesta incapacidade de determinar os valores de verdade da segunda proposição, não há nada que o primeiro caso possua que o segundo caso não possua. Essa constatação leva o opositor da bivalência a um dilema: ou ele admite que a proposição acerca do futuro contingente no primeiro caso não possui valores de verdade no presente ou assume que apenas essas proposições têm valores de verdade no presente. A primeira saída é implausível demais para ser levada a sério e a segunda é simplesmente incoerente.

### *Proposições são entidades abstratas que não mudam*

Se proposições contingentes acerca do futuro são indeterminadas no presente e somente terão valores de verdade futuramente, teremos de admitir que as suas propriedades mudam com o passar do tempo. Porém, uma proposição é uma entidade abstrata que não está localizada no espaço-tempo e parece absurdo manter que as propriedades de uma entidade dessa natureza possam mudar com o passar do tempo (SWARTZ, 2001, p. 226).

Alguém poderia objetar que o fato de uma proposição ser atemporal apenas implica que a sua posse de valores de verdade ou a sua ausência nunca irão mudar (KASPAR, 2002, p. 279-280). Essa objeção, contudo, não vai muito longe. O indeterminista irá recusar a bivalência, motivado pela intuição de que algumas proposições têm valores de verdade variáveis devido a mudanças temporais. Ele acredita que a proposição da batalha naval não tem valores de verdade no momento em que é enunciada, mas será verdadeira ou falsa depois que os fatos relevantes ocorrerem no mundo. Reconhecer que a ausência de valores de verdade

dessa proposição só pode ser mantida se for mantida sempre é pagar um preço alto demais para recusar a bivalência. A suposição da atemporalidade por si só não irá determinar quantos valores de verdade uma proposição pode ter, mas fornecerá boas razões para manter a bivalência em detrimento das demais alternativas teóricas.

Outra saída seria sustentar que entidades abstratas de fato possuem propriedades que mudam com o tempo, a saber, propriedades relacionais. Ao referir o número 10, eu não modifico o número 10, mas altero apenas as suas propriedades relacionais (RICE, 2010). Poderíamos estender essa explicação para as proposições acerca de futuros contingentes se os valores de verdade forem interpretados como propriedades relacionais. Porém, muito trabalho precisa ser desenvolvido nessa direção.

#### *Argumentos dedutivamente válidos com proposições acerca de futuros contingentes*

A posição de que proposições acerca de futuros contingentes não são bivalentes também se revela implausível quando consideramos as suas consequências para avaliarmos argumentos válidos que as utilizam como premissas (DOWDEN, 2011)<sup>2</sup>. Se proposições acerca de futuros contingentes não possuem valores de verdade agora, não podemos determinar, pelos métodos tradicionais, a validade dedutiva do seguinte argumento:

Se haverá uma batalha naval amanhã, devemos acordar o almirante.

Haverá uma batalha naval amanhã.

Logo, devemos acordar o almirante.

Se as premissas desse argumento não têm valores de verdade no presente não podemos avaliar adequadamente o argumento utilizando os padrões de validade dedutiva, pois estes padrões dizem respeito ao relacionamento entre os valores de verdade das proposições componentes – um argumento é dedutivamente válido quando é impossível que ele tenha premissas verdadeiras e conclusão falsa.

Uma objeção imediata é de que quem adotar uma concepção distinta de validade dedutiva e negar que proposições futuras possuam valores de verdade não será convencido pelo argumento. Alguém poderia adotar a concepção de validade entendida como preservação de

---

2 - A autoria da objeção é atribuída a Quine.

informação, por exemplo. Essa objeção tem alguma força, mas também é muito limitada. Se o argumento não convence quem nega essa concepção de validade, então não temos muito com o que nos preocupar, pois a generalidade dos filósofos parece aceitar essa concepção.

Uma objeção mais difícil de ser respondida é a acusação de que ele depende de uma confusão entre a validade e a solidez do argumento. Podemos determinar a validade do argumento, pois podemos conceber facilmente que nas circunstâncias em que as premissas são verdadeiras a sua conclusão também será verdadeira. A controvérsia envolvendo o problema dos futuros contingentes é saber em que circunstâncias as premissas poderão ser verdadeiras e o argumento poderá ser sólido. Se negarmos que proposições acerca de futuros contingentes possuem valores de verdade no momento em que são asseridas, o argumento não poderá ser sólido no momento em que as premissas forem asseridas, pois os eventos que as tornam verdadeiras ainda não ocorreram. De fato, a própria concepção de validade dedutiva associada a valores de verdade não descarta essa hipótese. Tudo o que ela afirma é que um argumento dedutivamente válido é aquele que não pode ter premissas verdadeiras e conclusão falsa, mas ela não determina em que circunstâncias as premissas podem ter valores de verdade.

Essa objeção é plausível, mas podemos reformular o argumento de modo a contorná-la. O argumento pode ser reformulado da seguinte maneira: se proposições acerca de futuros contingentes não possuem valores de verdade agora, argumentos que utilizam premissas acerca de futuros contingentes não podem ser sólidos. Mas sabemos que há inúmeros argumentos com premissas acerca de futuros contingentes que são sólidos. Logo, as proposições acerca de futuros contingentes possuem valores de verdade agora.

Para tornar essa formulação do argumento mais convincente, basta considerar os casos de afirmações acerca de futuros contingentes que indubitavelmente têm valores de verdade no presente. Por exemplo:

Se a Fernanda irá me visitar no próximo mês, terei de limpar o meu apartamento.

A Fernanda irá me visitar no próximo mês.

Logo, terei de limpar o meu apartamento.

Como a Fernanda sempre me visita no início do mês, posso assegurar que ela irá me visitar no mês seguinte. O argumento, no presente, não é apenas válido, mas perfeitamente sólido. De fato, apenas uma intuição corrompida pela discussão dos futuros contingentes declararia que o argumento apresentado não pode ser sólido no presente.

*Relações semânticas não são relações causais*

A verdade de uma proposição não faz com que os eventos ocorram. A direção é inversa: é o evento que torna a proposição verdadeira. Se eu tomo café enquanto estudo às 6 da manhã de hoje, 9 de janeiro, isso torna verdadeira a proposição de que “Matheus toma café enquanto estuda às 6 da manhã no dia 9 de janeiro”. A ameaça determinista inverte a ordem semântica e cria a ilusão de que a verdade de uma proposição acerca do futuro causa a ocorrência do evento. Pelo contrário, é o fato de que o evento irá ocorrer no futuro que torna a proposição verdadeira hoje, mas essa é uma relação semântica e não causal. Que se trata de uma relação semântica e não causal nota-se pelo fato de que seria inconsistente manter que o evento previsto ocorreu sem admitir a verdade da proposição acerca desse evento. De qualquer modo, o evento futuro não causa de maneira retrocedente a verdade de uma proposição asserida no presente (SWARTZ, 2001, p. 227).

Isso se torna mais óbvio se considerarmos proposições acerca do passado. Suponha que um grupo de dez pessoas esteja debatendo sobre o ano da morte de John Lennon, que foi assassinado a tiros em 1980. Do grupo, apenas uma pessoa, Fernanda, acerta a data. No entanto, não diríamos que o fato de Fernanda ter feito uma afirmação verdadeira hoje acerca do assassinato de Lennon tenha de algum modo causado o assassinato. Suponha agora que o ano da discussão fosse 1975 e o mesmo grupo de pessoas estivesse debatendo a data em que John Lennon será assassinado. Apenas uma pessoa, Lucas, acerta a previsão. Contudo, não iremos dizer que Lucas é de alguma maneira causalmente responsável pelo assassinato de Lennon – isso seria absurdo (SWARTZ, 2001, p. 227).

*Ligações de valores de verdade*

Se uma afirmação asserida no tempo presente com a forma “Um evento do tipo K está ocorrendo” é verdadeira, então a afirmação correspondente no tempo passado, “Um evento do tipo K ocorreu um ano atrás”, asserida daqui a um ano, também deve ser considerada como verdadeira (DUMMETT, 2004, p. 75). Suponha agora que alguém assira que “Um evento do tipo K irá ocorrer daqui a um ano” e o evento de fato ocorre no período de um ano. Em um período de dois anos alguém também poderia dizer “Um evento do tipo K ocorreu um ano atrás”. Ambas as afirmações, sendo a primeira uma previsão e a segunda uma retrodição, estão separadas por um período de dois anos e ambas estão ligadas por seus valores de verdade.

Esse padrão de ligações de verdade é tudo o que precisamos para demonstrar que afir-

mações acerca de futuros contingentes têm valores de verdade no momento em que são asseridas. Se uma afirmação no tempo presente, asserida daqui a um ano, com a forma “Um evento do tipo K está ocorrendo”, for verdadeira daqui a um ano, então a afirmação correspondente no tempo futuro, “Um evento do tipo K irá ocorrer daqui a um ano”, asserida agora, deve também ser considerada verdadeira (WESTPHAL, 2006, p. 6-8). Desse modo, demonstramos, a partir das ligações de verdade, que afirmações acerca de futuros contingentes devem possuir valores de verdade no tempo em que são asseridas.

*Não podemos demarcar de maneira muito precisa quando uma proposição se torna verdadeira*

Uma objeção à afirmação de que proposições acerca de eventos futuros não possuem valores de verdade no presente é de que é muito problemático colocar em termos precisos a ocorrência de uma proposição tornando-se verdadeira. Quando se tornou verdade que o Bush pai venceu a eleição de 1988? Quando os votos foram contados? Quando a Alemanha perdeu a Segunda Guerra Mundial? Quando a força de invasão dos Aliados desembarcaram nas praias de Normandia? Quando Alan Turing e seu time decifraram o código secreto alemão? (SWARTZ, 2001, p. 226).

### **3. Conclusões**

A principal virtude das soluções apresentadas é o seu caráter econômico e pouco revisionista: podemos manter a bivalência sem fazer maiores revisões em nossa lógica e manter a intuição de que as proposições acerca de ações futuras podem ter valores de verdade no presente sem ter de adotar teorias contra-intuitivas, como o determinismo. É claro que alguém pode estar disposto a negar essas suposições por razões independentes, mas o problema dos futuros contingentes certamente não será uma dessas razões. De fato, o problema só parece surgir se pensarmos que o que será o caso tem de ser o caso, violarmos o princípio de rigidez modal, cometermos uma falácia modal, adotarmos uma concepção infalível de conhecimento e negarmos nossas atribuições de conhecimento do dia a dia, entre muitos outros erros. Se alguém pretende negar a bivalência ou sustentar o determinismo a partir desse problema e os erros que o motivam, ele precisará procurar razões em outro lugar.

## Referências bibliográficas

- ARISTOTLE. De interpretatione. In: MCKEON, R. (ed.), *The Basic Works of Aristotle*. Trad. E. M. Edghill. New York: Random House, 1941.
- BARNES, E; CAMERON, R. Back to the open future. *Philosophical Perspectives*, Vol. 25, Issue 1, 2011, p. 1–26.
- DOWDEN, B. Time, In: *Internet Encyclopedia of Philosophy - A Peer-Reviewed Academic Resource*, 2011. Disponível em: <http://www.iep.utm.edu/time/>. Acesso em: 10 de agosto de 2011.
- DOWDEN, B; SWARTZ, N. Truth, In: *Internet Encyclopedia of Philosophy - A Peer-Reviewed Academic Resource*, 2004. Disponível em: <http://www.iep.utm.edu/truth/>. Acesso em: 17 de dezembro de 2012.
- DUMMETT, M. *Truth and the Past*. New York: Columbia University Press, 2004.
- IACONA, A. Future Contingents and Aristotle's Fantasy, *CRÍTICA - Revista Hispanoamericana de Filosofía*. Vol. 39, No. 117, 2007, p. 45–60.
- KASPAR, D. The end of the sea battle story, *Philosophia*, 29, (1-4), p. 277-286, 2002.
- LEWIS, D. Counterfactual Dependence and Time's Arrow, In: *Philosophical Papers*, 2. Oxford: Oxford University Press, 1987.
- PRIOR, A. *Past, Present and Future*. Oxford: Oxford University Press, 1967.
- RICE, H. Fatalism, In: *The Stanford Encyclopedia of Philosophy*, 2010. Disponível em: <http://plato.stanford.edu/archives/win2010/entries/fatalism/>. Acesso em: 10 de dezembro de 2012.
- SWARTZ, B. *Beyond Experience*, University of Toronto Press, 2° ed., 2001.
- \_\_\_\_\_. Foreknowledge and Free Will, In: *Internet Encyclopedia of Philosophy - A Peer-Reviewed Academic Resource*, 2004. Disponível em: <http://www.iep.utm.edu/foreknow/>. Acesso em: 09 de dezembro de 2012.
- TORRE, S. The Open Future, *Philosophy Compass*, vol. 6, nº 5, p. 360–373, 2011.
- WESTPHAL, J. The Future and the Truth-Value Links: A Common Sense View, *Analysis*, Vol. 66, No. 1, p. 1-9, 2006.



# Quem Sabe Outra Hora, Outro Mundo

## *Maybe another time, another world*

Fernando Furtado

Universidade Federal de Minas Gerais

### **Resumo**

Neste artigo, procuro trabalhar nossas intuições modais e temporais com base em alguns exemplos que penso serem interessantes. Vou procurar tratar dois problemas metafísicos: identidade pessoal e identidade transmundial. Além disso, vamos ver neste artigo como a semântica modal pretende traduzir a nossa conversa modal e até onde ela é capaz de chegar. Este artigo, portanto, tem a finalidade de desenvolver as intuições modais e temporais do leitor forçando essas intuições, em alguma medida, contra a tradução da conversa modal para semântica modal ou semântica dos mundos possíveis.

### **Palavras-chave**

mundo possível, identidade pessoal, identidade transmundial, modalidade, tempo, existência necessária

### **Abstract**

In this paper I try to work our modal and temporal intuitions based on some examples that I

think are interesting. I will seek to address two metaphysical problems: personal identity and transworld identity. Moreover, I will see in this paper how the modal semantic want to translate our modal conversation and how far it can reach. This paper therefore aims to develop modal and temporal intuitions forcing the reader insights against the translation of our modal conversation for modal semantics or semantics of possible worlds.

## Keywords

possible world, personal identity, transworld identity, modality, time, necessary existence.

## I

Neste artigo, procuro trabalhar nossas intuições modais por meio de alguns exemplos particularmente interessantes. Pretendo, com isso, tornar mais clara a diferença entre intuições modais e intuições temporais ou, mais especificamente, clarificar a diferença entre a *existência necessária* e a *existência eterna*. Dois problemas distintos em metafísica estão, pelo menos a princípio, diretamente relacionados com essa distinção; o *problema da identidade pessoal* e o *problema da identidade transmundial*. No primeiro caso, procuram-se encontrar critérios que sirvam para identificar um determinado particular ao longo do tempo. No segundo, a procura é por critérios para identificar um particular ao longo dos mundos possíveis ou em situações contrafactuais.

Pense no seguinte caso: uma semente de laranjeira que é dada a uma criança para que ela possa plantar no quintal de casa. Essa semente, caso germine, será, depois de alguns anos, uma árvore muito grande e talvez dê origem a várias outras sementes. A semente, ao longo de sua existência, sofrerá várias modificações em nível material; a pequenina semente poderá dar origem a uma bela e frondosa árvore que periodicamente produzirá laranjas que poderão resultar em outras laranjeiras. Nesse caso, quase automaticamente, respondemos positivamente para seguinte pergunta: a grande laranjeira é o mesmo particular que a semente que foi dada para a criança muitos anos atrás? Esse é um caso típico de *identidade pessoal*. Queremos saber o que faz com que a pequena semente seja o mesmo particular depois de muito tempo, sabendo que a semente se tornou uma grande laranjeira. Precisamos de critérios de identidade ao longo do tempo para dar conta desse caso.

Um caso bastante diferente é perguntar o que faz com que a laranjeira seja a mesma

laranjeira em situações contrafactuais. Antes de tratar de casos como esse vejamos, pois, o que são, afinal, situações contrafactuais. Tomemos os seguintes dois exemplos, que são particularmente esclarecedores: 1) ‘se eu não estivesse escrevendo este artigo, estaria na sala vendo TV.’ 2) ‘se eu estivesse na seleção brasileira de 1950, o Brasil não teria perdido a copa.’ Tanto num caso como no outro, o antecedente das condicionais é falso: é por isso que se chamam *contrafactuais*, pois são contrárias aos fatos. Mas 1 é intuitivamente verdadeira, ao passo que 2 é intuitivamente falsa. Como explicar isso (sobretudo tendo em conta que, na lógica clássica, qualquer condicional com antecedente falsa é verdadeira)? Saber como determinar o valor de verdade de uma condicional contrafactual é uma questão filosófica em disputa, mas uma análise muito comum nos termos da linguagem dos mundos possíveis é a seguinte: uma contrafactual é verdadeira se, e só se, em todos os mundos possíveis *mais próximos* do mundo atual onde o antecedente é verdadeiro, o consequente também é verdadeiro. Por outras palavras, uma contrafactual é verdadeira se, e somente se, uma modificação mínima no mundo atual, de modo a admitir a verdade do antecedente, produz a verdade do consequente. Pensando nos exemplos anteriores, em 1 nada de especial é necessário para que, não estando a escrever este artigo, estivesse a ver TV. Já em 2 muitas outras coisas seriam necessárias, além de eu estar na seleção, eu teria de ser um jogador excelente, tão excelente que influenciaria o resultado de todo o time. Ora, é perfeitamente razoável pensar que há outro mundo possível no qual eu jogo na seleção e mesmo assim o Brasil não ganha a copa. Daí que pensemos que essa contrafactual é falsa. Pense agora: o que cara que estaria, em uma situação contrafactual, na copa 50, ainda seria eu? Esse é o segundo caso que nos interessa aqui. Vejamos, pois, a seguir, em maior detalhe, um exemplo desse tipo.

Pense agora na criança do caso da laranjeira. O que faz com que a criança – suponhamos: ‘Carlos’ – seja a mesma criança ao longo de situações contrafactuais? Certamente, assim como a laranjeira, Carlos perde e ganha matéria ao longo do tempo, perde e ganha propriedades ao longo do tempo. Sendo assim, temos também acerca de Carlos o problema da identidade pessoal. Porém, no caso da identidade transmundial, a pergunta é sobre as propriedades que Carlos poderia não apresentar (ou exemplificar) em situações contrafactuais e, ainda assim, não perderia a sua identidade e seria o mesmo Carlos. Certamente diríamos que Carlos ainda seria Carlos, mesmo que ele não tivesse a cabeleira que atualmente possui, mas talvez não disséssemos que Carlos ainda seria Carlos se, numa situação contrafactual, ele tivesse pais diferentes daqueles que atualmente possui. Por outras palavras, queremos saber quais são as propriedades essenciais de Carlos; queremos saber quais propriedades Carlos, caso não as instanciasse (exemplificasse) em outros mundos possíveis, deixaria de ser Carlos. Assim, queremos critérios de *identidade transmundial*.

Certamente o leitor mais atento já percebeu que as nossas intuições oscilam acerca des-

ses dois problemas, porém, são dois problemas, pelo menos à partida, claramente diferentes. Temos então – e este é o objetivo fundamental deste artigo – que treinar nossas intuições para evitar essas oscilações. Certamente, os problemas que, do ponto de vista técnico, são fundamentalmente diferentes, podem, na realidade, ser meras ilusões teóricas. Porém, para tanto, argumentos que tenham clareza com relação a esses casos devem ser avançados na direção de mostrar que os dois problemas – identidade pessoal e identidade transmundial – são realmente o mesmo problema.

## II

Nesta seção, desenvolvo, de modo bastante simplificado, a parte mais técnica deste artigo; vários exemplos serão apresentados visando tornar a discussão tão intuitiva quanto possível. Primeiro, apresento algumas noções básicas acerca da tradução dos termos modais para a linguagem dos mundos possíveis. Depois, apresento duas consequências diretas dessa tradução e uma explicação técnica, que aparece como pano de fundo nessa tradução. E, por fim, apresento algumas complicações que podem dificultar a aceitação dessa tradução.

### *Linguagem dos mundos possíveis e conversa modal*

A semântica dos mundos possível foi desenvolvida principalmente por Kripke e é capaz de integrar, em uma mesma semântica formal, os vários sistemas de lógica modal incompatíveis formalmente entre si e que não tinham uma semântica formal capaz de captar de forma clara a diferença entre eles. A semântica dos mundos possíveis é uma reconstrução técnica da linguagem dos mundos possíveis tornada conhecida entre os filósofos por Leibniz (1646-1716). Com a semântica dos mundos possíveis (ou semântica modal) que é capaz de integrar todos os sistemas em uma única semântica formal, tornou-se muito mais fácil ver a diferença entre esses sistemas. E recorrendo à semântica dos mundos possíveis, a lógica modal tornou-se ferramenta imprescindível em filosofia e é usada em várias áreas para clarificar nossas intuições acerca dos termos modais.<sup>1</sup> Veremos agora, brevemente, como a *semântica dos mundos possíveis* pretende traduzir a nossa *conversa modal alética*.<sup>2</sup>

---

1 - Há vários tipos de modalidades: modalidades epistêmicas, temporais e deônticas, além de várias outras. Entre elas estão as modalidades aléticas; tema específico deste artigo.

2 - Formalmente, a semântica dos mundos possíveis, ou melhor, um modelo  $\mathbf{M}$  para a linguagem modal é tipicamente definido como um triplo ordenado  $\langle \mathbf{W}, \mathbf{R}, \mathbf{V} \rangle$  onde  $\mathbf{W}$  é um conjunto não vazio de *mundos possíveis*,  $\mathbf{R}$  é uma *relação de acessibilidade*, e  $\mathbf{V}$  uma função que atribui valor de verdade para cada sentença atômica  $p$  da linguagem  $\mathbf{V}(p)$  em cada mundo. Para ver detalhes dessa formalização da semântica modal, ver (FITTING, 1998) ou (HUGHES, 1996) ou ainda (KRIPKE, 1963).

Com a frase “Dilma é, necessariamente, a presidente do Brasil”, o que queremos dizer, traduzindo para a linguagem dos mundos possíveis, é que Dilma é a presidente do Brasil em todos os mundos possíveis.<sup>3</sup> Com a frase “Dilma poderia ter sido uma partidária do PSDB”, o que queremos dizer é que há pelo menos um mundo possível em que Dilma é uma partidária PSDB. A frase “Dilma é, essencialmente, uma partidária do PT”, quer dizer que, em todos os mundos possíveis onde o particular Dilma existe, ela é uma partidária do PT.<sup>4</sup> Quando dizemos que “Dilma foi, contingentemente, uma guerrilheira”, o que estamos a dizer é que Dilma foi uma guerrilheira em alguns mundos possíveis e noutros mundos possíveis não foi. Com a frase “Dilma é, acidentalmente, a atual presidente do Brasil”, o que estamos a dizer é que, no mundo atual, Dilma é a atual presidente do Brasil, mas há outros mundos possíveis em que Dilma existe e não é a atual presidente do Brasil.

A frase “Dilma é Dilma” exprime uma proposição necessariamente verdadeira; na linguagem dos mundos possíveis, dizemos que a proposição em causa é verdadeira em todos os mundos possíveis. A frase “Dilma não é Dilma” exprime uma proposição necessariamente falsa; na linguagem dos mundos possíveis, dizemos que a proposição em causa é falsa em todos os mundos possíveis. A frase “Dilma é a atual presidente do Brasil” exprime uma proposição contingentemente verdadeira; na linguagem dos mundos possíveis, dizemos que essa proposição é verdadeira no mundo atual, mas é falsa em algum mundo possível. A frase “Dilma é a mulher mais importante do mundo” exprime uma proposição contingentemente falsa; na linguagem dos mundos possíveis, dizemos que a proposição em causa é falsa no mundo atual, mas é verdadeira em alguns mundos possíveis.

### *Semântica modal: duas consequências diretas*

A lógica modal alética e a semântica dos mundos possíveis, como tivemos a oportunidade de ver acima, pretendem, dentre outras coisas, dar conta do nosso discurso modal. Agora apresento duas consequências diretas da aceitação dessa análise, ou tradução, dos termos modais – necessário e possível – que talvez sejam indesejadas.

3 - O nosso objetivo neste artigo não é ver os detalhes técnicos da semântica dos mundos possíveis, mas, antes disso, ver se a tradução do que consideramos intuitivo acerca dos termos modais é captado de modo adequado pela linguagem formal. Portanto, como veremos, o conceito de *acessibilidade* não é mencionado ao longo do texto a fim de simplificar os trabalhos e por não resultar em problema algum para as finalidades deste artigo. **Se fossemos levar em consideração o conceito de acessibilidade** (o que pode ser feito pelo leitor), a tradução anterior seria algo mais complexo como o seguinte: com a frase “Dilma é, necessariamente, a presidente do Brasil”, no mundo atual, o que queremos dizer é que Dilma é a presidente do Brasil em todos os mundos possíveis acessíveis a partir do mundo atual. Essa frase seria, portanto, verdadeira, mesmo que houvesse mundos possíveis onde Dilma não é a presidente do Brasil, desde que esses mundos não fossem acessíveis ao mundo atual.

4 - Para um ataque direto a essa tradução ver (FINE, 1994). Nesse artigo, Kit Fine ataca a ideia de identificar propriedades necessárias (propriedade que o particular exemplifica em todos os mundos possíveis em que existe) com propriedades essenciais apresentando o seu famoso contraexemplo do “*singleton* de Sócrates”.

### ***Existente eterno contingente***

A leitura dos termos modais tem que permitir que haja um existente eterno contingente. O que isso quer dizer? Quer dizer que pode haver um particular que exista em um mundo possível ao longo de toda a linha do tempo e que, ainda assim, não seja um existente necessário. Pense no exemplo de Deus: Deus pode existir durante toda a eternidade – ao longo de toda a linha do tempo – no mundo atual e não ser um existente necessário. Pode haver um mundo possível com ligeiras modificações (onde, por exemplo, o leitor nunca leu este artigo) em que, nesse mundo possível, Deus não exista, mesmo que Ele seja um existente eterno no mundo atual. Se, por outro lado, tivermos em mente as leis da natureza, o mesmo acontece: uma determinada lei da natureza pode existir, ocorrer ou ser verdadeira ao longo de toda a linha do tempo do mundo atual e não existir em mundo possível com ligeiras modificações. Esses exemplos mostram que defender a eternidade de Deus ou das leis da natureza é algo fundamentalmente diferente de defender a necessidade dessas entidades.

### ***Existente necessário que existe por um intervalo infinitamente pequeno de tempo***

Por outro lado, se a leitura dos termos modais permitir os exemplos apresentados acima, também permite, por razões semelhantes, exemplos de existentes necessários que existam apenas em um intervalo da linha do tempo de um determinado mundo possível. O intervalo de tempo pode ser, certamente, infinitamente pequeno, portanto, temos que admitir que possa haver um existente que existiu por um intervalo de tempo infinitamente pequeno. Sendo assim, talvez tenha havido, no mundo atual, um particular que existiu por um período infinitamente pequeno de tempo e com o qual nunca podemos contatar-nos, porém, existe em todos os mundos possíveis e, portanto, é um existente necessário.

Ambos os casos apresentados não são novidades para aqueles que estão familiarizados com a lógica modal e semântica dos mundos possíveis, portanto, não pretendo mostrar que essas consequências são indesejadas e que devemos abandonar a semântica dos mundos possíveis e a lógica modal. Meu objetivo com esses exemplos é apenas clarificar o que estamos dizendo quando usamos a semântica dos mundos possíveis para traduzir a nossa conversa modal.

### ***Mundos possíveis e linha temporal***

Em vários momentos neste artigo, vimos, de modo informal e intuitivo, como os proble-

mas temporais e os problemas modais<sup>5</sup> devem ser estudados separadamente, mesmo que nossas intuições acerca deles sejam confusas e estejam misturadas. Aqui, vou caracterizar de forma mais técnica essa ideia e argumentar, de forma mais direta, em favor dela.

A ideia técnica que está por trás da distinção é que os mundos possíveis podem ser encarados de duas formas diferentes: na primeira, como entidades cuja dimensão temporal é irrelevante, por estarem todos situados no mesmo instante de tempo. Portanto, não há um mundo possível onde a linha do tempo está atrasada ou adiantada com relação à linha do tempo de outro mundo possível. Na segunda, entidades cuja dimensão temporal é relevante, por poderem estar situados em distintos instantes de tempo, defasados uns em relação aos outros. Nesse caso, há mundos possíveis deslocados na linha do tempo com relação à linha do tempo de outros mundos.

Vejam alguns exemplos para clarificar essas ideias: se encararmos os mundos possíveis como entidades cuja dimensão temporal é irrelevante, teremos que dizer que há um mundo possível onde eu já (agora) tenha terminado de escrever este artigo, ou seja, há um mundo possível com algumas horas adiantadas do mundo atual – tempo que eu ainda, no mundo atual, vou gastar para terminar este artigo. Por outro lado, se encararmos os mundos possíveis como entidades cuja dimensão temporal é irrelevante, ou seja, se não admitirmos mundos possíveis que estejam deslocados temporalmente do mundo atual, teremos que dizer que não há um mundo possível onde eu (agora) já tenha terminado de escrever este artigo. A forma mais precisa de demonstrar essa ideia é como segue no gráfico abaixo:



Figura 1.

5 - Os problemas temporais são, estritamente, problemas modais, mas para simplificar a discussão, aqui reservo o termo “modal” para dizer coisas acerca das modalidades aléticas e chamo de “temporais” as modalidades temporais.

O gráfico acima pressupõe que os mundos possíveis são entidades cuja dimensão temporal é irrelevante, ou seja, os mundos possíveis estão no mesmo momento do tempo e não há um mundo possível deslocado no tempo. Todas as cinco situações – a linha do *mundo* O representa o mundo atual e as demais os outros mundos possíveis – estão no mesmo momento do tempo que o mundo atual e não admitem situações contrafactuais no que diz respeito ao tempo. Portanto, em nenhum dos mundos possíveis eu terminei (agora) este artigo, pelo simples fato do mundo estar adiantado na linha do tempo.<sup>6</sup>

A forma de representar a posição contrária – os mundos possíveis são entidades cuja dimensão temporal é relevante – é, conforme gráfico abaixo, a seguinte:

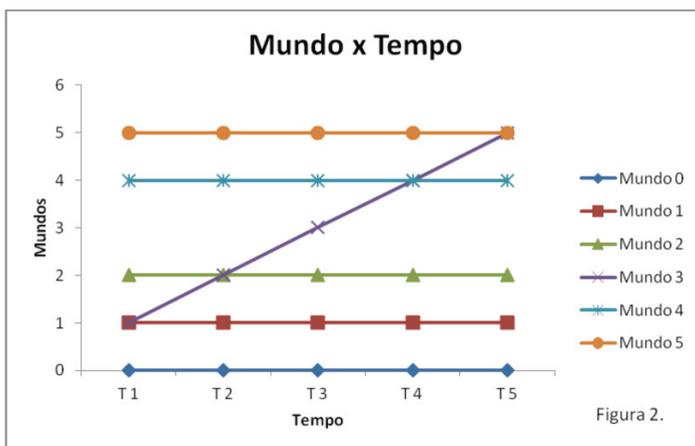


Figura 2.

Se encararmos os mundos possíveis como entidades cuja dimensão temporal é relevante, vamos permitir que o *mundo* 3, como parece no gráfico acima, seja considerado um mundo possível. O *mundo* 3, ao contrário dos outros mundos, não está em uma linha do tempo paralela ao mundo atual, portanto, para esse modelo, os mundos possíveis são entidades cuja dimensão temporal é relevante. Em outras palavras, o comportamento modal e temporal é relevante para sabermos a extensão total dos mundos possíveis.

Dois exemplos podem tornar essa diferença mais clara. Primeiro, pense na seguinte afirmação: “Eu posso vir a ser rico”. Essa afirmação é ambígua para duas interpretações, à luz das distinções que fizemos: na primeira encaramos a afirmação como uma potencialidade

6 - Obviamente, se mudarmos as condições, por exemplo, se eu tivesse me empenhado mais que me empenhei no mundo atual, poderia agora já ter terminado este artigo, porém, não é esse o caso. O que interessa é haver ou não um mundo possível onde eu já terminei o artigo apenas porque este mundo está adiantado na linha do tempo em relação ao mundo atual e todo o restante está mantido.

– análoga à semente se tornar uma laranjeira; nesse caso estamos afirmando que, em algum lugar no futuro, seguindo a linha do tempo no mundo atual, vou me tornar uma pessoa rica. Essa é a interpretação não modal da afirmação. Na segunda, estamos afirmando, supondo que estamos no momento  $T_2$  no mundo atual, que no futuro, momento  $T_3$ , em algum mundo possível, *mundo 4*, por exemplo, eu sou uma pessoa rica neste mundo. Essa é a interpretação modal da afirmação. Se pensarmos novamente acerca das leis da natureza, veremos que as coisas são semelhantes: podemos querer dizer apenas que as leis da natureza são eternas, ou seja, que as leis da natureza são verdadeiras ao longo de toda a linha do tempo do mundo atual; nesse caso, não estamos nos comprometendo com o comportamento modal das leis da natureza. Ou podemos dizer que as leis da natureza são verdadeiras ao longo de toda a linha do tempo do mundo atual e também ocorrem em todos os mundos possíveis; nesse caso, estamos nos comprometendo com o comportamento modal das leis da natureza.

Até aqui tentei apenas clarificar as coisas para algo que estamos completamente familiarizados. A partir de agora, podemos ver o que talvez seja o mais estranho que pretendo apresentar.

### *Argumento de Zenão*

Para tentar tornar implausível a ideia de que os mundos possíveis são entidades temporais, lanço mão de um argumento baseado no famoso *Paradoxo de Zenão*. Imaginemos a seguinte situação: enquanto eu escrevia este artigo, ouvi várias vezes a canção *Gymnopédies 1*. de Erik Satie. Eu poderia não tê-la escutado enquanto escrevia este artigo? Certamente, diríamos, sim! Talvez eu não tivesse me lembrado da canção ou não tivesse a canção no computador em que escrevi este artigo ou talvez eu simplesmente não gostasse de música clássica. São vários os motivos que poderiam me levar a não ter ouvido à *Gymnopédies 1*. de Erik Satie enquanto escrevia este artigo. Suponhamos que, por algum desses motivos, eu, de fato, não ouvi essa canção enquanto escrevia este artigo. Nesse caso, estamos falando de uma situação contrafactual, um mundo possível alternativo ao atual. Essa situação contrafactual é possível? Parece que não temos problemas em admitir que sim. Agora perguntemos: eu poderia ter ouvido pela primeira vez, enquanto escrevia este artigo, à *Gymnopédies 1*., três minutos antes do que efetivamente eu ouvi? Há um mundo possível idêntico em todos os aspectos (não temporais) no qual a primeira vez que eu ouvi à *Gymnopédies 1*. enquanto escrevia este artigo ocorreu três minutos antes do que ocorreu no mundo atual? Se a resposta para essas perguntas for não, então os mundos possíveis são entidades cuja dimensão do tempo é irrelevante e ficamos com a resposta contraintuitiva de que não há um mundo possível idêntico em todos os aspectos (não temporais) onde eu ouvi pela primeira vez à

*Gymnopédies 1*. enquanto escrevia este artigo, três minutos antes do que eu ouvi no mundo atual, e aceitamos o primeiro gráfico como modelo adequado.<sup>7</sup> Se a resposta às perguntas for sim, então os mundos possíveis são entidades cuja dimensão do tempo é relevante e o gráfico que melhor representa a relação entre mundos possíveis e tempo é o segundo gráfico. Se encararmos os mundos possíveis como entidades cuja dimensão temporal é relevante e admitirmos que há um mundo possível deslocado três minutos no tempo, temos que admitir, só por isso, que há infinitos mundos possíveis. Pense no seguinte: três minutos é um recorte temporal convencional que pode ser novamente repartido infinitamente, portanto, se há um mundo possível deslocado três minutos no tempo, há uma infinidade de mundos possíveis deslocados partes de três minutos no tempo. Assim, temos, quase trivialmente, uma infinidade de mundos possíveis, apenas por admitir um deslocamento temporal mínimo de mundos possíveis.

Depois desses esclarecimentos, pense novamente acerca do segundo gráfico: encarar a linha roxa no gráfico como um mundo não é tecnicamente correto, cada deslocamento no tempo, mesmo que seja mínimo – e há infinitos em apenas um segundo –, geraria um mundo possível alternativo. Portanto, a linha roxa, a linha do *mundo 3*, deve ser encarada como a representação técnica de uma infinidade de mundos possíveis.

Por esses motivos, penso que encarar os mundos possíveis como entidades cuja temporal seja relevante pode ser enganoso. Logo, parece que temos que olhar para os mundos possíveis com alguma precaução, evitando encará-los, sem que nos percebamos, como entidades cuja dimensão temporal é relevante procurando evitar que essa pressuposição ‘contamine’ nossas intuições modais e temporais.

### **Conclusão: Soluções para o problema**

Frente aos argumentos e exemplos apresentado até agora, temos três vias de solução para o problema: primeira, rejeitar a semântica dos mundos possíveis como capaz de traduzir adequadamente a nossa conversa modal; a nossa conversa modal apresenta intrinsecamente intuições temporais (não estamos dispostos a admitir existentes necessários que existam em intervalo infinitamente pequeno de tempo), a semântica modal parece não ser

---

7 - Mais uma vez, lembre-se de que, por outros motivos que não apenas por uma questão de deslocamento temporal de um mundo possível, eu poderia ter ouvido à *Gymnopédies 1*, pela primeira vez três minutos, antes do que efetivamente eu ouvi, por exemplo, num mundo possível onde me deu vontade de ouvir a canção três minutos antes do que me deu vontade no mundo, eu de fato ouvi a canção três minutos antes e, certamente, este é um mundo possível, porém, não se trata apenas de um deslocamento temporal de mundos possíveis.

capaz de captá-las, portanto, devemos rejeitar a tradução dos termos modais à semântica dos mundos possíveis. Segunda, aceitar que a semântica modal alética não tem que dar conta de questões temporais e explicar porque nossas intuições modais muitas vezes se confundem com intuições temporais. Terceira, aceitar que a semântica dos mundos possíveis dá conta das nossas intuições modais e elas estão intrinsecamente ligadas a intuições temporais. Nesse caso, seria necessário fornecer um tratamento técnico diferente aos exemplos e problemas apresentados. Assim, questões temporais e modais ficam no mesmo pacote e podem ser tratadas pela semântica dos mundos possíveis.<sup>8</sup> A mim (o autor deste artigo), a melhor solução parece ser a segunda. E quanto ao leitor, o que pensa?

## Referências

- CANIELLI, . e PIZZI, C. 2009. *Modalities and Multimodalities*. Springer.
- HUGHES, G. e CRESSWELL, M. 1996. *A New Introduction to Modal Logic*. London: Routledge.
- CRESSWELL, M. e RINI, A. 2012. *The World-Time Parallel*. Cambridge: Cambridge University Press.
- FINE, K. 1994. Essence and Modality. *Philosophical Perspectives* 8:1-16.
- FITTING, M. e MENDELSON, R. 1998. *First-order Modal Logic*. Norwell: Kluwer Academic Publishers.
- FURTADO, F. F. R. 2010. Determinismo Modal: Um Estudo Conceitual. *XIII Encontro de Pesquisa na Graduação em Filosofia da USP. Caderno de resumos*. São Paulo – SP.
- KRIPKE, S. 1963. Semantical Considerations on Modal Logic. *Acta Philosophica Fennica*, 16, pp. 83–94.
- \_\_\_\_\_ (1980) *Naming and Necessity*. Oxford: Blackwell.
- MURCHO, D. 2002. *Essencialismo Naturalizado*. Coimbra: Angelus Novus.
- PLANTINGA, A. 1974. *The Nature of Necessity*. Oxford: Clarendon Press.
- WILLIAMSON, T. 2002. Necessary Existents. Em O’Hear, org., *Logic, Thought and Language*. Cambridge: Cambridge University Press, pp. 233-251.

---

8 - Há muitos trabalhos técnicos que fornecem tratamentos técnicos para dois ou mais tipos de modalidades e pode, talvez, tratar dos problemas apresentados aqui. São os sistemas híbridos ou multimodais. Para ver em detalhes esses trabalhos sugiro (CANIELLI, W. e PIZZI, C. 2009)



# Tradução



# Notas sobre o preconceito\*

## *Notes on Prejudice*

Isaiah Berlin

Tradução: Mário Nogueira de Oliveira \*\*  
UFOP

Poucas coisas causaram maior dano que a crença por parte de indivíduos ou grupos (ou tribos, Estados, nações, igrejas) de que são os únicos possuidores da verdade. Especialmente a verdade sobre como viver, o que ser e o que fazer – e que aqueles que diferem dos que se consideram possuidores de tal verdade não estão meramente enganados, mas são perversos ou loucos, e precisam de repreensão ou desaparecimento. É uma terrível e perigosa arrogância acreditar que somente você está correto: acreditar que você possui um olho mágico que vê a verdade e que os outros não podem estar corretos se discordarem de você.

Isso faz um indivíduo ter a certeza de que existe um único objetivo para a sua nação, sua igreja ou para toda a humanidade e que qualquer quantidade de sofrimento vale a pena

---

\* **N.T.:** Este texto foi escrito por Isaiah Berlin em 1981 em forma de anotações rápidas destinadas a um amigo que iria debater sobre o tema e pediu ajuda a Berlin. Por constituírem anotações rápidas, o texto não apresenta o rigor da escrita encontrado nos demais textos do autor. Este texto é mantido pelo responsável pela administração da obra de Berlin, Henry Hardy. Pesquisas sobre textos de Berlin podem ser feitas em <http://berlin.wolf.ox.ac.uk/lists/nachlass/index.html> e <http://www.bodley.ox.ac.uk/dept/scwmss/wmss/online/modern/berlin/berlin.html>.

\*\* **N.T.:** O tradutor gostaria de agradecer a Livia Reis por sua leitura da tradução e revisão da Língua Portuguesa. Na tradução optou-se por usar a pontuação tal como costumamos fazer no Português praticado no Brasil, evitando a repetição do uso do sinal de pontuação dois-pontos (colon) existente no texto em inglês.

(particularmente por parte das outras pessoas) se o único objetivo somente é alcançado “através de um oceano de sangue para o Reino do Amor” (ou algo parecido), como disse Robespierre.<sup>1</sup> Hitler, Lenin, Stalin, e eu ousaria dizer que os líderes nas guerras religiosas de cristãos contra muçulmanos ou católicos contra protestantes, sinceramente acreditaram nisso: a crença de que existe uma, e somente uma, resposta verdadeira para as perguntas centrais que afligem e afligiram a humanidade e que somente um indivíduo possui essa verdade – ou o seu líder a possui – foi responsável por oceanos de sangue. Mas nenhum “Reino do Amor” emergiu deles, nem poderia: existem muitos modos de viver, de acreditar e de se comportar. O mero conhecimento fornecido pela História, a Antropologia, a Literatura, as Artes e o Direito deixa claro que as diferenças de culturas e de indivíduos são tão profundas como são as semelhanças (o que torna os mais diferentes homens, igualmente humanos), e que nós não somos os mais pobres nessa rica variedade: esse conhecimento abre as janelas da mente e da alma e torna as pessoas mais sábias, mais gentis e mais civilizadas. A ausência desse conhecimento gera o preconceito irracional, os ódios, a exterminação odiosa de hereges e daqueles que são diferentes: se as duas grandes guerras e os genocídios de Hitler não nos ensinou isso, nós somos incuráveis.

O mais valioso elemento na tradição britânica, ou um dos mais valiosos, é precisamente a liberdade relativa à política, às raças, ao fanatismo religioso e à monomania<sup>2</sup>. Ficar em paz com as pessoas que você não simpatiza ou que não entende é indispensável para qualquer sociedade decente. Nada é mais destrutivo do que um feliz sentimento da sua própria infalibilidade – ou da infalibilidade da sua própria nação - que lhe permite destruir outros com a consciência tranquila, uma vez que você está fazendo o trabalho de Deus (por exemplo, atos da Inquisição Espanhola ou dos Aiatolás) ou o trabalho da raça superior (por exemplo, Hitler) ou o trabalho da História (por exemplo, Lenin-Stalin).

A única cura é entender como outras sociedades vivem no espaço e no tempo e entender que é possível viver vidas diferentes da sua própria vida, e ainda assim, ser completamente humano, digno de amor, de respeito ou pelo menos de curiosidade. Jesus, Sócrates, João Huss da Boêmia<sup>3</sup>, o grande químico Lavoisier, os socialistas e os liberais (assim como os

1 - N.T.: Cf. 'Rapport sur les principes de morale politique qui doivent guider la Convention nationale dans l'administration intérieure de la République, fait au nom du Comité de salut public, le 18 pluviôse, l'an 2e de la République, par Maximilien Robespierre; imprimé par ordre de la Convention nationale (18 pluviôse an II - 5 février 1794' Em *Discours par Maximilien Robespierre* — 17 Avril 1792-27 Juillet 1794. The Project Gutenberg EBook of *Discours par Maximilien Robespierre* - 17 Avril 1792-27 Juillet 1794, by Maximilien Robespierre. Produced by Daniel Fromont. Texte en français moderne établi par Charles Vellay. <http://www.gutenberg.org/files/29887/29887-h/29887-h.htm>

2 - N. T.: Monomania: forma de insanidade em que um único pensamento ou ideia absorve a mente do indivíduo. Cf. *Grande Dicionário Houaiss da Língua Portuguesa*.

3 - N.T.: Jan Hus (ou John Huss) costumava denunciar abusos da igreja nos seus sermões. Ele criticava a disciplina e alguns costumes dentro da Igreja. Defendia a ideia de que membros da Igreja deveriam exercer seus poderes apenas no âmbito espiritual e nunca como governantes no mundo secular. Foi excomungado em 1412 por insubordinação. Hus também combatia a

conservadores) na Rússia, os judeus na Alemanha, todos pereceram nas mãos de ideologias “infalíveis”: uma certeza intuitiva não é o substituto para o conhecimento empírico cuidadosamente testado baseado na observação, em experimentos e na livre discussão entre os homens. As primeiras pessoas que os totalitaristas destroem ou silenciam são os homens de ideias ou de mentes livres.

2. Os estereótipos são outra fonte de conflitos evitáveis. As tribos odeiam as tribos vizinhas pelas quais se sentem ameaçadas e, assim, com seus medos, representam-nas como perversas ou inferiores, irracionais ou desprezíveis. Ainda assim, esses estereótipos modificam-se rapidamente em muitas ocasiões.

Tomemos como exemplo o século XIX: por volta de 1840, os franceses eram considerados fanfarrões, galantes, imorais, militaristas, homens com bigodes curvados, perigosos para mulheres e prontos para invadir a Inglaterra como vingança por Waterloo. Os alemães eram tidos como bebedores de cervejas, provincianos, musicais, cheios de metafísicas nebulosas, inofensivos e, de algum modo, tolos. Por volta de 1871, os alemães já são vistos como a cavalaria que atemoriza a França, comandados pelo terrível Bismarck – ou vistos como assustadores militares prussianos cheios de orgulho nacional. Já a França passa a uma pobre terra civilizada e sufocada, necessitando a proteção de todos os homens de bem, para que sua arte e sua literatura não sejam esmagadas pelos calcanhares dos terríveis invasores.

Os russos do século XIX são servos subjugados pelos semirreligiosos eslavos, místicos que escrevem romances profundos e vistos também como uma horda enorme de cossacos leais ao Czar que canta lindamente. Nos nossos tempos, tudo foi dramaticamente modificado: a população foi subjugada, sim, mas há também a tecnologia, os tanques, o materialismo ateu, a cruzada contra o capitalismo, etc.

Os ingleses eram imperialistas cruéis que se impunham sobre os soldados estrangeiros, olhando com seus narizes empinados para o resto do mundo. Depois passam a empobrecidos, liberais, beneficiários do Estado de Bem-estar Social e necessitados de aliados. Todos esses estereótipos são substitutos do conhecimento real que nunca é algo tão simples ou permanente como uma imagem particular generalizada de estrangeiros, os quais são estímulos para a autossatisfação nacional e para desdém de outras nações. Isso é um sustentáculo ao nacionalismo.

---

ideia da venda de indulgências para arrecadar dinheiro para a igreja. Foi acusado de divulgar doutrinas heréticas e condenado à morte em 1415. Cf. [www.justus.anglican.org/resources/bio/7.html](http://www.justus.anglican.org/resources/bio/7.html)

3. O Nacionalismo, que todos no século XIX pensavam que estava decaindo, tornou-se a mais forte e mais perigosa força de hoje. Ele é geralmente uma ferida causada por uma nação no orgulho ou no território de outra. Se Luís XIV (o Rei Sol cujo Estado deu as leis a todos na política, na arte da guerra, nas artes, na filosofia e na ciência) não houvesse atacado, devastado e humilhado os alemães por anos, então, estes talvez não houvessem se tornado tão agressivos no início do século XIX, quando foram fortemente nacionalistas contra Napoleão. De modo similar, se os russos não tivessem sido tratados como uma massa de bárbaros pelo Ocidente no século XIX, ou se os chineses não tivessem sido humilhados pelas guerras do ópio ou pela exploração em geral, eles não teriam caído tão facilmente em uma doutrina que prometia que eles herdariam a terra depois que tivessem, com ajuda de “forças históricas” que ninguém poderia parar, esmagado todos os descrentes capitalistas. Se os indianos não tivessem sido tratados como inferiores, etc., etc.

As conquistas, a escravidão de povos, o imperialismo, etc., não são alimentados apenas pela ganância ou desejo de glória, mas têm de justificar a si mesmos por alguma ideia central, a saber, os franceses como os detentores da única cultura verdadeira; o “fardo” do homem branco; o comunismo; os estereótipos dos outros como inferiores ou perversos. Apenas o conhecimento, cuidadosamente adquirido, e não adquirido por atalhos, pode dissipar tudo isso. Mesmo que não venha a dissipar a agressividade humana em si mesma nem a antipatia pelos diferentes (na pele, na cultura, na religião). Ainda assim, a educação em História, Antropologia e Direito ajuda (especialmente se são comparadas, e não apenas consideradas isoladamente, como geralmente o são).

# Resenha



## ***Obra: Philosophical Devices, Probabilities, Possibilities, and Sets***

Autor: David Papineau

Resenha por: Matheus Martins Silva - UFRGS

Contemporary philosophy has a proper language, a dense and technical vocabulary that has been created along the years by philosophers themselves or imported from other areas. ‘Analytic/synthetic’, ‘a priori/a posteriori’, ‘contingent/necessary’, ‘possible worlds’, ‘rigid designators’, ‘alethic modalities’, ‘conditional probability’, ‘conditionalization’, ‘denumerability’ and ‘incompleteness theorem’ are some of the terms in the philosophical agenda. So much technicism could lead one to believe that contemporary philosophy is a world unto itself. The semantic of possible worlds, to cite just one example, is an idiom as fundamental to understand contemporary philosophy as the evolutionary biology is fundamental to understand contemporary biology. Thus, it is not possible to understand how philosophy is currently made without mastering this technical vocabulary. ‘Philosophical Devices’ is a book that aims to present this vocabulary to philosophy undergraduates.

Papineau selects some of the most important and fundamental concepts used in contemporary philosophy and present them in an accessible way, but without lack of rigor. The book is divided into four parts; each composed by three chapters with corresponding exercises. The first part, ‘Sets and Numbers’, presents the basic vocabulary of set theory, Russell’s paradox (the set of all sets that does not include themselves), transfinite sets, de-

numerability, and the continuum hypothesis.

The second part, 'Analyticity, A Prioricity, and Necessity', presents the main concepts involving possible worlds and modalities, such as semantic modalities (analytic and synthetic), epistemic modalities (a priori and a posteriori), alethic modalities (necessary and contingent), as well as discussions about the necessary a posteriori, rigid designators, the contingent a priori, the nature of necessity, and the causal theory of reference.

The third part, 'The Nature and Uses of Probability', presents basic notions of probability and their application in philosophy. Some of the notions presented are objective and subjective probability, Kolmogorov's axioms, conditional probability, Bayes' theorem, probabilistic independence, expected utility, Dutch books, correlations, causality, and Simpson's paradox.

The fourth part, 'Logics and Theories', presents some fundamental distinctions such as logic and metalogic, syntactic and semantic consequence, soundness and completeness, and Gödel's incompleteness theorem. Apart from the indispensable references, the book contains a list of bibliographical suggestions, solutions to exercises and a good index.

Papineau's choice of concepts and topics included in the book is reasonable given the wide range of contemporary philosophy. One of the main virtues of the book is that Papineau presents the chosen topics and concepts in a philosophically interesting manner, informing the context, and telling the reader of some disagreements about what he just presented. Papineau uses plain language and dispenses formalisms or technical details that are not strictly necessary to an elementary understanding of what is presented. This option for skipping some details could be seen as a slip or lack of rigor by some specialists used to more formal investigations, mainly those who also have a mathematical training. But any reaction of that type would be misplaced. We should have in mind that the main target readers of the book are freshmen making their first contact with these notions, and not specialists who are willing for more substantial materials.

Of course, the book is not perfect. One flaw is its size. Its short dimension forces the author to present some concepts way too fast, sometimes in a manner that is superficial even to the standards of a textbook. In chapter seven, for instance, Papineau does not explain trivialities such as the manner to determine the value of the probability of  $p \& q$  when  $p$  and  $q$  are propositions about independent events. Of course that this is a truism for one who already mastered the basics about the subject, but the whole purpose of the book is precisely to explain these truisms in an accessible manner to students who still do not know the basics. The explanation of the Gödel's theorem presented in the last chapter is also very

brief and can let some readers disappointed<sup>1</sup>. I am not saying that the book should be longer in order to explore more complicated topics, but that the book should be longer in order to explain some basic topics in a clearer and accessible way.

The book could also have been more informative in some points. In chapter 8, for instance, the author introduces some concepts about conditionals (conditional probability, material conditional, the distinction between indicative and subjunctive conditionals), but mentions without going into further details that the conditional probability of a conditional cannot be its probability of being true. What happens is that in the beginning of the seventies some philosophers started to conjecture that the probability of a conditional being true should be measured by its conditional probability – this conjecture is usually known as ‘Stalnaker’s hypothesis’<sup>2</sup>. At first sight, this seems true. The probability of ‘the match will light if you struck it’ is the same probability than ‘if you strike the match, it will light’.

Against our intuitions, David Lewis (1976) (1986) showed that this is not possible: there is no proposition that can fulfill this requirement. To be more precise, he showed that a proposition to which its truth probability is correspondent to its conditional probability would have so trivial probabilities distributions that it would be too simplistic or unrealistic, for representing human beliefs<sup>3</sup>. There are still some controversies about this proof, but most philosophers nowadays accepted it. Since the author uses three sections of the chapter eight explaining only notions involving conditionals he could have easily created an entire chapter about conditionals, thus having the necessary space to explain in an accessible way not just Lewis’ proof but other important topics.

In any case, these problems do not make the book less worthy to philosophy undergrads or less recommendable. Rather, it is a very important textbook either for doing an introductory exposition of important technical notions that normally would require different textbooks to comprise them, as for providing the most recent developments in philosophy. It is a book that should be a mandatory reference to aspirant philosophers who want to do their first steps in the field. It is a reading not to be missed.

---

1 - A clearer explanation can be found in Ernest Nagel and James Newman’s book, ‘Gödel’s Proof’.

2 - See Stalnaker (1970).

3 - The reader can find an accessible presentation of Lewis’ proof in Bennett (2003).

## References

- Bennett, J. (2003). *A Philosophical Guide to Conditionals*, Oxford University Press.
- Lewis, D. K. (1976). "Probabilities of Conditionals and Conditional Probabilities," *Philosophical Review*, 85: 297–315
- Lewis, D. K. (1986). "Probabilities of Conditionals and Conditional Probabilities II," *Philosophical Review*, 95: 581–589.
- Stalnaker, R. C. (1970). "Probability and Conditionals," *Philosophy of Science*, 37, 64-80. Reprinted in W. L. Harper, R. Stalnaker, and G. Pearce (eds). *Ifs*. Dordrecht: D. Reidel, 107-28.



