

A CIÊNCIA COGNITIVA E A ROBOTIZAÇÃO DO HOMEM¹

Marcos Barbosa de Oliveira²

Universidade de São Paulo

Contra esta ideologia, insisto que não existe uma natureza humana imutável. Além e acima do nível animal, os seres humanos são maleáveis, em corpo e mente, até mesmo em sua própria estrutura instintiva. Homens e mulheres podem ser computadorizados transformando-se em robôs, sim – mas eles podem também se recusar a isso.

Marcuse³

Este ensaio tem a natureza de um esboço: pretende apenas delinear uma visão crítica sobre a ciência cognitiva — ou, mais precisamente, sobre a parte da ciência cognitiva que trata dos aspectos culturais da cognição humana —, e sobre o naturalismo — entendido como a posição que nega a existência de diferenças essenciais entre as ciências humanas e as naturais e, tomando estas como modelo, propõe que seus princípios epistemológicos e metodológicos sejam adotados nas ciências humanas. O desenvolvimento completo das idéias a serem apresentadas resultaria num texto muito mais longo e complexo, envolveria melhor explicitação e demonstração de inúmeras teses que serão a seguir pouco mais que enunciadas, a discussão de problemas a que as teses dão origem, etc. O trabalho seria sem dúvida algo de muito mais sólido, mas poderia, eventualmente deixar menos nítida a visada crítica que desejamos propor.

A crítica ao naturalismo apóia-se sobre a tese de que existem diferenças essenciais entre as ciências humanas e as naturais. Uma destas consiste na possibilidade, restrita às ciências humanas, de ocorrência daquilo que Popper denomina *efeito Édipo* – o fenômeno em que o anúncio de uma previsão afeta a ocorrência do evento previsto.⁴ O efeito Édipo pode ser visto como um caso particular de um fenômeno mais geral, a saber,

1. Uma versão ligeiramente diferente deste trabalho foi publicada como o cap. XI de *Da ciência cognitiva à dialética*.

2. Departamento de Filosofia da Educação e Ciências da Educação – USP – São Paulo.

3. 'Ecology and the critique of modern society', p.68. (Agradeço a Isabel Maria Loureiro pela indicação.)

4. Cf. Popper, *A miséria do historicismo*, cap.5; *A sociedade aberta e seus inimigos*, vol. I, pp.35-6; *Conjecturas e refutações*, p.68; *Autobiografia intelectual*, p.129-30. Em 'Popper e o efeito Édipo', cap. X de *Da ciência cognitiva à dialética*, mostra-se que há uma inconsistência entre o naturalismo defendido por Popper e sua admissão da possibilidade de ocorrência do efeito Édipo nas ciências humanas.

aquele em que o processo de investigação afeta seu próprio objeto, ou seja, a parte da realidade a que se refere. Vamos dar a este o nome de *transformação cognitiva*. Uma transformação cognitiva, convém enfatizar, não é uma transformação *no conhecimento*: é uma transformação *na realidade a que o conhecimento se refere*, provocada pelos processos envolvidos em seu desenvolvimento (isto é, no desenvolvimento do conhecimento).

O efeito Édipo é um caso particular de transformação cognitiva porque a emissão de previsões constitui apenas uma etapa do processo de investigação científica. Sem a pretensão de fazer uma lista exaustiva, e sem sugerir que a ordem dos itens corresponda à seqüência temporal real, podemos dizer que o processo de investigação científica envolve as seguintes etapas: 1) a escolha do tema, do objeto a ser estudado, ou, se se preferir, do problema a ser resolvido; 2) a realização de observações (observações “simples”, não observação dos resultados de experimentos), 3) a realização de experimentos (incluindo a observação dos resultados), 4) a formulação de teorias, e 5) a emissão de previsões.

Nossa tese é a de que em todas estas etapas pode haver, nas ciências humanas, e não nas ciências naturais, um impacto do processo de investigação sobre o objeto investigado. A cada etapa corresponde então um tipo de transformação cognitiva: a transformação cognitiva *temática* é aquela em que a escolha do tema por si só afeta a parte correlata da realidade; a transformação cognitiva *observacional*, aquela em que o impacto se deve à realização de observações, e assim por diante.

Cada um destes tipos merece um estudo especial. Vamos nos limitar a algumas observações.

A transformação cognitiva preditiva corresponde ao efeito Édipo, já mencionado.

A transformação cognitiva observacional – assim como vários outros tipos de transformação cognitiva – é discutida por Nagel em seu livro *A Estrutura da Ciência*, no capítulo dedicado aos problemas metodológicos das ciências sociais. Nagel discute, entre outras coisas, a influência que, nas pesquisas de opinião, a entrevista pode ter sobre as opiniões do sujeito entrevistado. Tendo em vista nossos objetivos, por outro lado, a transformação cognitiva observacional apresenta um problema. A tarefa em que estamos empenhados é a de apresentar a transformação cognitiva como uma diferença entre as ciências humanas e as naturais. No caso da transformação cognitiva observacional, muitas pessoas alegariam que ela não é exclusiva das ciências humanas, ocorrendo também nas ciências naturais. O fundamento desta alegação residiria na mecânica quântica, mais precisamente, nas idéias associadas ao Princípio da Incerteza de Heisenberg. Esta alegação entretanto só se sustenta se for adotada uma interpretação particular do Princípio da Incerteza, a saber, a interpretação que vincula a incerteza de

modo essencial ao processo de observação. Há outras interpretações em que a incerteza é apresentada como algo ontológico, não epistemológico; como algo presente na realidade, independente de existirem ou não observadores e observações. Este tópico faz parte da problemática dos fundamentos da mecânica quântica a qual, como se sabe, é das mais complexas. Sendo assim, vamos nos restringir a constatar que não há acordo entre os *experts*, e declarar nossa adesão a interpretações realistas da mecânica quântica, como a de Einstein. Embora se deva reconhecer que estas deixam algumas questões em aberto, ainda assim são preferíveis às interpretações idealistas, cujas dificuldades a nosso ver são ainda maiores. Sendo adotada uma interpretação realista, não sobra razão alguma para se defender a tese de que a transformação cognitiva observacional ocorre também nas ciências naturais.

O tipo de transformação cognitiva mais importante é com certeza o da transformação cognitiva *teórica*, o qual foi estudado com grande propriedade por Charles Taylor. A maneira usada por Taylor para expressar a tese correspondente consiste em dizer que nas ciências humanas, e apenas nestas, as teorias são *teorias transformadoras*, elas têm a capacidade de transformar seu próprio objeto. Em sua visão, são três as proposições que justificam esta tese.

1. Em qualquer sociedade, cada indivíduo tem certas idéias, certas concepções a respeito da sociedade a que pertence: a respeito de suas práticas características, de seus costumes, de suas instituições. A essas concepções Taylor dá o nome de *auto-entendimentos* (*self-understandings*). Estamos nos referindo aqui às pessoas comuns, aos leigos, não a cientistas sociais, ou filósofos. O auto-entendimento, portanto, é parte do senso comum.

2. Os auto-entendimentos não são um aspecto secundário, não são um epifenômeno da vida social: eles formam, como Taylor diz, um elemento *constitutivo*, pois, se se alteram, as práticas sociais também são alteradas.

3. As *teorias sociais* – ou seja, as teorias produzidas pelos sociólogos, economistas, e cientistas políticos – na medida em que se tornam conhecidas, afetam o auto-entendimento, isto é, afetam as concepções do senso comum sobre a sociedade, e, desta forma, afetam as próprias práticas sociais.

A seguinte passagem de Taylor resume o que a nosso ver é mais fundamental em sua abordagem, no que diz respeito às diferenças entre as ciências humanas e as ciências naturais. Taylor refere-se à ciência política, mas suas teses valem também, diríamos, para todas as ciências humanas.

Esta é a notável dessemelhança entre a ciência natural e as teorias políticas. Estas podem subverter, fortalecer ou moldar as práticas a que se referem. E isto porque (a) elas são teorias a respeito de práticas, as quais (b) são parcialmente constituídas por certos auto-entendimentos. Na medida em que (c) as teorias transformam

este auto-entendimento, elas subvertem, favorecem ou transformam as características constitutivas das práticas. Poderíamos expressar isto de outra maneira afirmando que as teorias políticas não dizem respeito a objetos independentes, como acontece na ciência natural. Nesta, a relação do conhecimento com a prática consiste na aplicação daquilo que se sabe sobre os poderes causais a casos particulares, porém as verdades acerca desses poderes causais nas quais se confia permanecem inalteradas. Esta é a razão para afirmar que a teoria, neste caso, refere-se a um objeto independente. *Na política, por outro lado, a aceitação da teoria pode ela própria transformar aquilo a que a teoria se refere.*

Uma terceira forma de expressar esta idéia consiste em dizer que enquanto na ciência natural a teoria também transforma a prática, a prática que ela transforma não é aquilo a que a teoria se refere. Ela é, neste sentido, externa à teoria. Pensamos nela como uma ‘aplicação’ da teoria. Mas na política, a prática é o objeto da teoria. *A teoria neste domínio transforma seu próprio objeto.* (‘Social theory as practice’, p.101, itálicos acrescentados)

Pensando agora em termos gerais, ou seja, deixando de lado as distinções entre os vários tipos, vamos dizer que a transformação cognitiva é uma transformação determinada, uma transformação que aponta numa certa direção, a qual depende do conteúdo da teoria, daquilo que a teoria afirma. Pode-se mostrar, por exemplo, através de estudos históricos, que as teorias econômicas clássicas que tomavam como básico o conceito de *Homo economicus*, contribuíram para que, ao longo do desenvolvimento do capitalismo, os seres humanos tenham se transformado numa direção que os fez se aproximar do ideal do *Homo economicus*. Vamos dizer então que nas ciências humanas cada produção teórica tem seu *vetor de transformação*, correspondente à direção em que ela tende a transformar a realidade a que se refere. O vetor de transformação depende, é claro, não apenas do conteúdo, mas também do contexto, das circunstâncias históricas em que a teoria está se encontra inserida.

Consideremos agora a pergunta: qual é a relação, em cada teoria, entre seu vetor de transformação e os desejos, intenções e valores daqueles que a defendem? Este é um problema de grandes proporções de que não vamos poder tratar aqui de forma completa. É possível, por exemplo, que o vetor de transformação de uma teoria seja contrário aos desejos e intenções de seu proponente. Apenas para ilustrar, consideremos um economista que esteja preocupado com a possibilidade de um aumento na taxa de inflação, e que procure alertar as autoridades e a população no sentido de que, se certas medidas não forem tomadas, a inflação vai subir. Em vez do efeito desejado, ou seja, o de fazer com que as medidas sejam tomadas, a divulgação das idéias do economista pode fazer com que os agentes formadores de preços procurem se defender aumentando

preventivamente seus preços, e assim fazendo com que a inflação venha a aumentar, em vez de diminuir, como era a intenção do economista.

Bem, em vez de examinar todas as possibilidades que podem existir neste domínio, vamos nos concentrar em um tipo particular de relação entre o vetor de transformação de uma teoria e os desejos, intenções e valores de seus proponentes. Para facilitar a exposição, vamos pensar simplesmente em termos de valores. Vamos dizer então que, nas ciências humanas, um tipo de relação entre o vetor de transformação de uma teoria e os valores de seus proponentes é aquele em que, 1) a relação é de concordância, ou seja, o vetor de transformação corresponde aos valores e, 2) esta correspondência não é reconhecida pelos proponentes da teoria.

Esta segunda característica aparece como algo necessário dentro das concepções naturalistas. O naturalismo adota uma concepção de ciência, supostamente válida tanto para as ciências naturais quanto para as ciências humanas, em que a transformação cognitiva não ocorre, ou seja, em que as teorias não são teorias transformadoras. Mas se as teorias não são transformadoras não há vetor de transformação, e não pode haver uma correspondência deste com os valores sustentados.

Os valores se situam na esfera da normatividade, e a correspondência entre o vetor de transformação e os valores sustentados justifica a afirmação de que, nas ciências humanas, toda teoria tenha um *conteúdo normativo*. No domínio das teorias propostas segundo as concepções naturalistas, tal conteúdo normativo não é reconhecido: os naturalistas propõem suas teorias como tendo a ver apenas com os fatos, como tendo por objetivo apenas *descrever* a realidade tal como ela é, independentemente de nossos desejos, intenções e valores. A esta normatividade não reconhecida vamos dar o nome de *criptonormatividade*. A criptonormatividade portanto é o fenômeno em que as proposições de uma teoria são apresentadas como puramente descritivas, quando na verdade são também normativas.

A forma mais simples de criptonormatividade corresponde ao que se pode chamar de *proposição criptonormativa predicativa*. Uma proposição criptonormativa predicativa é uma proposição da forma ‘A é B’ que deve ser interpretada como afirmando ‘A deve ser B’.

Na epistemologia naturalista, as evidências factuais desempenham o papel central na escolha das teorias. Como esta epistemologia não reconhece o efeito transformador das teorias sobre a realidade, ela não pode incluir o vetor de transformação como um outro fator a ser levado em conta na escolha das teorias. Sendo este efeito reconhecido, está claro que sua qualidade deve ser considerada. Ou seja, se uma teoria tende a produzir certa transformação na realidade, nada mais correto, nada mais racional, que, na avaliação da teoria, se leve em conta se esta transformação é boa ou má, se é algo desejável, ou algo que devemos evitar.

As idéias desta parte da exposição também já foram defendidas por Taylor. Veja-se, por exemplo, a seguinte passagem:

Começando com a ambição de ser teorias do mesmo tipo das teorias da Ciência Natural, elas⁵ na verdade acabam funcionando como teorias transformadoras, como acontece normalmente com as teorias políticas, porém inconscientemente e *malgré elles*. Elas evitam assim a interessante pergunta: ‘É esta a correta teoria transformadora?’, pois elas não a podem colocar; elas não podem ver que a pergunta deve ser colocada. (‘Social theory as practice’, p.103)

Essas considerações justificam então a afirmação de que, nas ciências humanas, encontram-se proposições da forma ‘A é B’ que devem ser interpretadas como ‘A deve ser B’. A vantagem desta interpretação é que ela traz à luz o conteúdo normativo da proposição, que é ignorado, ou negado por seus proponentes. Trazer à luz este conteúdo nos permite então refletir sobre ele, e avaliar racionalmente se a transformação que a teoria tende a produzir é desejável ou não.

Passemos agora à ciência cognitiva. Na ciência cognitiva os computadores desempenham um papel fundamental, não apenas como uma idéia abstrata, mas também como uma realidade concreta,. Sem os computadores, a ciência cognitiva, tal como a conhecemos, não existiria. A forma primordial em que se manifesta a centralidade dos computadores na ciência cognitiva corresponde à famosa *metáfora computacional* – a idéia de que a mente humana deve ser pensada como um computador digital, e explicada em termos computacionais. Expressa de maneira enfática, a metáfora computacional afirma: ‘A mente humana é um computador’.

Boa parte das críticas que se fazem à ciência cognitiva pode ser interpretada, em última análise, como contrapondo, à proposição cognitivista ‘a mente humana é um computador’, sua negação simples ‘a mente humana não é um computador’. Tais críticas, ou pelo menos algumas delas, não deixam de ter certa validade. Elas incorrem, todavia, num equívoco que também está presente na ciência cognitiva, e é solidário a seu naturalismo, a saber, o equívoco de considerar a essência do sistema cognitivo humano, e mais especificamente, o quanto esta essência se assemelha à de um computador, como um fato da natureza. Na concepção que defendemos, uma característica fundamental da mente humana é sua plasticidade.⁶ A plasticidade da

5. No contexto da citação, este ‘elas’ refere-se a certo tipo de teoria que procura explicar o comportamento político por meio de modelos econômicos. Nós diríamos que a afirmação – com certas restrições, por exemplo, a referente à economia, como Taylor aponta (*op. cit.*, pp.102-3) – vale para todas as teorias das ciências humanas, e vale em particular para a ciência cognitiva cultural — ou seja, a parte da ciência cognitiva que trata dos aspectos culturais da cognição.

6. Cf. ‘Natureza e cultura’, cap. 4 de *Da ciência cognitiva à dialética*.

mente humana não é total, e existem, em particular, aspectos da cognição humana que são naturais e, enquanto tal, dentro de certos limites, imutáveis. Estes podem, e estudados com os mesmos métodos que são utilizados nas ciências naturais. O problema da ciência cognitiva é que ela se propõe a estudar da mesma forma aspectos da cognição que são culturais, e desta forma variáveis ao longo da história das sociedades.

Entre tais aspectos culturais está o da semelhança ou não do sistema cognitivo humano com os computadores. *Pensar de uma maneira mais, ou menos análoga ao funcionamento dos computadores digitais não é uma característica natural da mente humana, é algo que varia de cultura para cultura, e ao longo da história de cada cultura.*

A questão fundamental a que somos levados então é a de se é ou não uma boa coisa, se é ou não desejável que a mente humana, a nossa mente, se torne semelhante aos computadores. O elemento criptonormativo da ciência cognitiva corresponde a uma resposta afirmativa a esta questão. Isto sugere que, para entender corretamente o cognitivismo, devemos interpretar a metáfora computacional como afirmando não ‘a mente humana é um computador digital’, mas sim ‘a mente humana deve ser um computador digital’.⁷

A inteligência artificial tradicional entretanto – e, junto com ela a metáfora computacional – já perderam boa parte da predominância que tiveram numa primeira etapa do desenvolvimento da ciência cognitiva. Como se sabe, numa sinopse muito rápida, de um lado os impasses da inteligência artificial tradicional, de outro a ascensão do connexionismo, e além disso uma percepção das limitações implícitas nas abordagens que separam a mente da matéria, ou a cognição do corpo, e abstraem das interações físicas dos sistemas cognitivos com o ambiente, todos estes fatores contribuíram para que hoje em dia predomine não a metáfora computacional, mas sim o que se pode chamar, se é que já não foi chamado, de *metáfora robótica*. A metáfora robótica é expressa com todas as letras como o título de uma entrevista concedida por Dennett à revista *Istoé*, e publicada em 1966: ‘*Somos todos robôs*’. Interpretada de acordo com o princípio da criptomatidade, ela significa: ‘*Devemos todos nos tornar robôs*’. A

7. Mencionamos há pouco o fato de que muitas das críticas à ciência cognitiva cometem o equívoco de defender uma essência não-computacional da mente humana como se esta fosse um fato de natureza. Um exemplo de uma crítica que não comete este equívoco encontra-se em Dreyfus, *computers can't do*: “A natureza do homem é na verdade tão maleável que pode estar a ponto de mudar mais uma vez. Se o paradigma computacional se tornar tão forte que as pessoas comecem a pensar a respeito de si próprias como aparelhos digitais segundo o modelo da inteligência artificial, então, uma vez que, pelas razões que estive expondo, as máquinas não podem ser como seres humanos, os seres humanos podem se tornar progressivamente como máquinas. [...] Nosso risco não reside no advento de computadores superinteligentes, mas de seres humanos sub-inteligentes.” (p.280)

questão portanto não é mais se computadorização da mente humana é desejável, mas sim se a robotização do homem é desejável.

Algumas das razões que podem fazer com que a robotização do homem seja considerada indesejável são as mesmas que podem nos fazer rejeitar a proposição de que *somos* robôs. Isto nos permite retornar por um momento à dimensão descritiva, para examinar a maneira como Dennett responde aos argumentos contrários à sua tese. Consideremos, por exemplo, a alegação de que robôs não têm sentimentos. Como nós certamente os temos, então não podemos ser robôs. (Ou então, no registro normativo: os robôs não têm sentimentos; como não queremos deixar de ter sentimentos, então não é uma boa idéia nos transformarmos em robôs.)

A resposta de Dennett consiste em rejeitar a alegação, afirmando que os robôs certamente podem ter sentimentos – e isto não apenas como uma possibilidade teórica, mas como um projeto cuja realização está bem próxima. Na entrevista mencionada ele declara:

Sim, certamente um robô poderá sentir felicidade, tristeza e dor. Cog, o robô humanóide que estamos construindo no Instituto de Tecnologia de Massachusetts, em Boston está sendo projetado para ter estas capacidades. (p.5)

Outras capacidades que Dennett atribui aos robôs incluem: ser consciente; sentir-se parte de uma comunidade; ter valores, e criar seus próprios códigos de ética; sentir respeito e admiração pelos seres humanos; chegar autonomamente ao *cogito*, ao “penso, logo existo”; cometer suicídio, e outras.

Esta liberalidade em relação às capacidades dos robôs é com certeza chocante para o senso comum, e é bem possível que as declarações de Dennett tenham em parte um intuito provocativo. Elas são úteis, entretanto, para colocar em questão o *conceito* de robô subjacente à tese robótica de Dennett. Com efeito, se os robôs podem em princípio ter todas estas capacidades, o que está sendo dito quando se afirma que os seres humanos são robôs? É evidente que, dependendo do conceito que se adote, a proposição ‘somos todos robôs’ pode não passar de uma tautologia.

O que gostaríamos de sugerir é que há uma característica dos robôs de que Dennett não pode abrir mão, sob pena de contrariar suas próprias premissas. Esta característica, se por um lado salva sua tese da trivialidade tautológica por outro – é o que pretendemos mostrar – a torna inaceitável. Para descrevê-la, entretanto, é necessário primeiro desenvolver uma nova linha de raciocínio, uma linha que consiste na aplicação do princípio de interpretação da criptonormatividade ao próprio naturalismo.

O naturalismo, segundo nossa definição, é a posição daqueles que negam a existência de diferenças essenciais entre as ciências humanas e as ciências naturais, e sustentam que aquelas devem adotar os mesmos métodos destas. Para expressar esta

posição numa forma adequada para a aplicação do tipo de análise que estamos propondo, podemos dizer que ela equivale à proposição: ‘o homem é um ser natural’. Se, conforme o nosso modelo, esta é uma proposição criptonormativa, então ela deve ser interpretada como: ‘o homem *deve se tornar* um ser natural’. Trata-se agora então não mais da computadorização, ou da robotização do homem, mas de sua *naturalização*.

Mas o que significa a naturalização do homem? Que sentido tem dizer que o homem deve se tornar um ser natural? Dizer que o homem é um ser natural é dizer que ele é parte da natureza, e num certo sentido esta afirmação é verdadeira. De acordo com a concepção mais aceita hoje em dia entre as pessoas educadas, a espécie humana provém de outras espécies de primatas não-humanos, num processo que se desenrolou segundo os princípios da seleção natural. Isto é o que diz a teoria da evolução de Darwin. Nossa origem, portanto, encontra-se na natureza, e neste sentido somos parte dela.

Mas há um outro sentido, evidentemente, em que não somos parte da natureza, e se não fosse assim, se não houvesse algo de não-natural em nós, os próprios termos ‘natural’ e ‘natureza’ perderiam o sentido. Sobre a linhagem das espécies que deu origem à nossa, podemos então dizer: era parte da natureza, e em certo período deixou de sê-lo.

Com isso podemos dar um primeiro passo em direção ao objetivo de explicitar o significado da naturalização do homem. A naturalização do homem significaria, segundo esta linha de pensamento, uma *volta à natureza*.

Parece então que a questão já está resolvida, uma vez que a volta à natureza é um tema bem conhecido, caracterizando na verdade toda uma vertente de críticos da civilização ocidental. É uma tema fortemente associado a Rousseau, por seu elogio do bom selvagem. Esteve muito em voga no movimento *hippie*, inspirando as tentativas (quase sempre malogradas) de formação de comunidades à margem do mundo moderno, isto é, do mundo da indústria, da tecnologia, da ciência, etc. E, de uma forma mais amena, o tema da volta à natureza manifesta-se também na obsessão do movimento ecológico com o *natural*.

Mas será esta a naturalização do homem que diríamos estar contida no *naturalismo*? Certamente não. O naturalismo, no sentido em que estamos empregando o termo, nasce de um entusiasmo com os sucessos do domínio das ciências naturais a partir da revolução científica dos séculos XVII e XVIII, e no contexto da ascensão do capitalismo. Historicamente, aparece como uma vertente do Iluminismo, uma vertente otimista, confiante dos poderes da racionalidade científica para impulsionar o progresso da humanidade. Em relação àquilo que os defensores da volta à natureza criticam, o naturalismo encontra-se certamente do outro lado da cerca. E, nos dias de hoje, em particular se pensarmos no caso da ciência cognitiva, não há dúvida de que a postura

básica dos naturalistas é favorável à ciência – isto se expressa já no nome com que foi batizada –, e à tecnologia.

Parece então não fazer sentido atribuir aos naturalistas em geral, e à ciência cognitiva em particular, um conteúdo normativo apontando na direção de uma volta à natureza.

Consideremos melhor a questão. Ao adotar os princípios epistemológicos das ciências naturais para o estudo do homem, o naturalista é levado a conceber o objeto de seu estudo como sendo a natureza do homem, sendo esta natureza pensada como algo fixo, imutável. Ora, aquilo que é fixo, que é imutável, nós não podemos mudar por um esforço de vontade, e conseqüentemente não podemos criticar. Pode-se criticar um costume, uma instituição, uma crença, etc., porque tais coisas podem ser mudadas. Mas não faz sentido criticar fatos da natureza: não faz sentido criticar a lei de gravidade, ou o princípio de conservação da energia, ou o fato de não termos um olho na nuca, para poder enxergar o que vem por trás. Todos estes são aspectos da natureza – da natureza física ou da nossa natureza – que só nos cabe aceitar.

Em conseqüência disso, tomar, como fazem os naturalistas, aspectos do homem que são culturais, e portanto modificáveis, como parte de uma natureza humana imutável, equivale a colocá-los além da possibilidade de contestação, equivale a torná-los *imunes à crítica*.

Neste ponto vamos introduzir um conceito-chave para a análise que estamos propondo. Trata-se do conceito de *segunda natureza*.

Todos sabem o que significa esta expressão. Um padrão de comportamento ou de pensamento, ou seja, um hábito, torna-se uma segunda natureza quando adquire algumas características daquilo que é natural, quando passa a se realizar como que por instinto. A maneira como os homens se cumprimentam corresponde obviamente a um padrão de cultural de comportamento. Mas, em nossa sociedade, por exemplo, quando alguém nos estende a mão reciprocamos o gesto instintivamente, num gesto tão quase tão automático quanto a reação de piscar quando algum objeto perigoso se aproxima de nossos olhos.

O aspecto que torna o conceito de segunda natureza particularmente adequado como instrumento para nossa análise reside no fato de que o processo de transformação de hábitos em segunda natureza, ou seja, por assim dizer, a *segunda-naturalização*, ou *naturalização de segunda ordem*, é ambígua quanto a seu valor: ele tem, por assim dizer, um lado bom e um lado mau.

Muitos processos de aprendizagem consistem na aquisição de certos hábitos, e se completam quando estes se transformam em segunda natureza. Aprender a dirigir é um caso paradigmático. Nas etapas iniciais, o aprendiz tem que pensar conscientemente pensamentos do tipo: ‘acho que já é hora de mudar para segunda’, ou ‘não posso me

esquecer de pisar na embreagem quando o carro estiver parando'. Já o motorista experiente executa todas as operações necessárias de forma automática, sem precisar pensar sobre o que está fazendo. A vantagem disso é evidente: em termos bem familiares aos cognitivistas, podemos dizer que a transformação dos hábitos que constituem a capacidade de dirigir em segunda natureza permite que o módulo do cérebro/mente responsável pelo pensamento verbal consciente fique liberado, podendo ser utilizado para outras ocupações, tais como ouvir rádio, conversar com outros passageiros, usar o telefone celular, etc. O lado positivo da naturalização de segunda ordem consiste assim da economia de recursos cognitivos que proporciona.

E qual é o lado negativo? O lado negativo consiste em que, ao se tornar segunda natureza, um hábito torna-se também difícil de ser mudado. É por isso que se insiste na importância de um processo de aprendizagem ser bem conduzido, tanto no caso da capacidade de dirigir, quanto, para variar o exemplo, no caso da prática de um esporte como o tênis. Se o aprendiz se acostuma a pegar na raquete de uma maneira errada – mas não tão errada que o impossibilite totalmente de rebater a bola –, então fica muito difícil corrigir este vício mais tarde, permanecendo o desempenho para sempre aquém do que poderia ter sido.

Trata-se de uma questão de dificuldade, naturalmente, uma dificuldade entretanto que no limite torna-se uma impossibilidade. Ou seja, a transformação completa de um hábito em segunda natureza significa que o hábito por assim dizer se petrifica, não podendo mais ser alterado.

Neste ponto a presente linha de raciocínio se entrelaça com a que vínhamos desenvolvendo antes. O resultado deste entrelaçamento consiste na conclusão de que, por se tornar imutável ao se transformar em segunda natureza, um hábito torna-se também imune à crítica.

O conceito de crítica, por outro lado, é inseparável do conceito de racionalidade. Nós exercemos nossa racionalidade quando submetemos nossos costumes, crenças e valores à crítica. Sendo assim, quanto mais nossos costumes, crenças e valores se transformam em segunda natureza, menor fica o domínio em que nossa racionalidade pode se exercer. No limite, quando toda a nossa cultura houver se transformado em segunda natureza, nossa racionalidade terá sido aniquilada.

Para o racionalista, a razão é aquilo que nos separa da natureza – sendo esta separação não necessariamente uma oposição; ela pode se dar com harmonia, desde que esta harmonia não seja confundida com unidade. É a razão que nos distingue dos animais. Porém esta não é uma afirmação válida para todo o sempre.

A respeito da linhagem de espécies que deu origem à nossa, dissemos que era parte da natureza, e depois deixou de sê-lo. Deixou de ser parte da natureza quando nos tornamos racionais. Mas assim como nos tornamos racionais, também podemos deixar

de sê-lo. Tal transformação tem pelo menos duas formas concebíveis: a primeira é um retorno a uma primeira natureza, à natureza primordial, de onde viemos – o qual é impossível, se aceitarmos a máxima de que não se pode girar para trás as rodas da história. A segunda consiste num avanço em direção a uma segunda natureza, e este sim é um caminho possível; é, como procuramos mostrar, o caminho para o qual aponta o naturalismo. Este então é o significado da naturalização que o naturalismo promove.

Para o racionalista, a conclusão é clara: a naturalização de segunda ordem, ou seja, o avanço em direção a uma segunda natureza, não menos que a volta à natureza primordial, significa uma queda no irracionalismo. O naturalismo, portanto, é uma forma de irracionalismo.

Com isso podemos voltar ao tópico da robotização. O que pretendemos sugerir agora – e esta é a etapa final do percurso – é que a robotização do homem é apenas uma das maneiras em que a naturalização pode se manifestar. A mediação é feita, de novo, pelo conceito de segunda natureza.

Num motorista experiente, como já observamos, os padrões de comportamento necessários para dirigir eficientemente um veículo tornam-se uma segunda natureza. De suas ações, por outro lado, podemos dizer que são realizadas *automaticamente*. É isto, naturalmente, que permite que pilotos de avião e navio – e talvez, no futuro, de carros também – sejam substituídos por *pilotos automáticos*. Mas a idéia de automação remete à de autômato, e ‘autômato’, salvo distinções mais refinadas, é sinônimo de ‘robô’.

Um caso extremo de naturalização de segunda ordem é o do operário em linha de montagem, e é isto, evidentemente, que permite sua substituição por robôs.

Mas a naturalização de segunda ordem, por outro lado, é a condição necessária para que qualquer padrão de comportamento ou aspecto dos seres humanos seja estudado naturalisticamente, ou seja, segundo os princípios em vigor nas ciências naturais.

As atribuições de capacidades aos robôs por parte de Dennett, de um lado, e as considerações que acabamos de expor, de outro, sugerem que o conceito de robô subjacente à sua afirmação de que somos todos robôs é a seguinte: *um robô é um ser que pode ser estudado de acordo com os princípios epistemológicos das ciências naturais*. Pode ser estudado como parte da natureza, e com auxílio das demais ciências naturais – das quais derivam, na verdade, todas as tecnologias utilizadas em sua construção.

Os motivos para rejeitar a robotização defendida por Dennett são desta forma os mesmos que se opõem à naturalização do homem: ambos os processos, como já mostramos para a naturalização, implicam numa queda no irracionalismo.

REFERÊNCIAS BIBLIOGRÁFICAS

- DENNETT, D. Somos todos robôs. Entrevista a *Istoé*. São Paulo, Editora Três, 16/11/1996, pp.5-7.
- DREYFUS, H.L. *What computers can't do: a critique of artificial reason*. Nova York, Harper and Row, 1972.
- MARCUSE, H. Ecology and the critique of modern society. Manuscrito inédito de uma palestra proferida na Califórnia em 1977. Tradução alemã publicada em Peter-Erwin JANSEN (org.), *Befreiung denken – ein politischer Imperativ*. 2ª ed. Offenbach/Main, Verlag 2000, 1990. (O número de página mencionado na nota 3 refere-se à publicação em alemão.)
- NAGEL, E. *The structure of science: problems in the logic of scientific explanation*. Nova York, Harcourt, 1961.
- OLIVEIRA, M. B. *Da ciência cognitiva à dialética*. São Paulo, Discurso Editorial, 1999.
- POPPER, K.R. *A miséria do historicismo*. Tradução de Octanny S. da Mota e Leonidas Hegenberg. São Paulo, Cultrix/EDUSP, 1980.
- POPPER, K.R. *A sociedade aberta e seus inimigos*. Tradução de Milton Amado. São Paulo, EDUSP, Belo Horizonte, Itatiaia, 1974.
- POPPER, K.R. *Autobiografia intelectual*. Tradução de Leonidas Hegenberg e Octanny S. da Mota. São Paulo, Cultrix/EDUSP, 1977.
- POPPER, K.R. *Conjecturas e refutações*. Tradução de Sérgio Bath, da quarta edição (revista), de 1972. Brasília, Editora Universidade de Brasília, s/d.
- TAYLOR, C. *Philosophical papers*. (Vol. 1, *Human agency and language*, vol. 2 *Philosophy and the human sciences*). Cambridge, Cambridge University Press, 1985.
- TAYLOR, C. Social theory as practice. Em *Philosophical papers*, vol. 2, pp.91-115.