

# ENCONTRO COM AS CIÊNCIAS COGNITIVAS

ORGANIZADORES:

Ademar Ferreira

Maria Eunice Quilici Gonzalez

Jonas Gonçalves Coelho

Volume 4



# **Encontro com as Ciências Cognitivas**

**Volume 4**





*Aos nossos queridos amigos Michel Debrun,  
Michael B. Wrigley e Renato Schaeffer, in memoriam,  
cujos trabalhos servem de inspiração a todos nós*



# **Encontro com as Ciências Cognitivas**

**Volume 4**

Ademar Ferreira  
Maria Eunice Quilici Gonzalez  
Jonas Gonçalves Coelho

Organizadores

**Coleção Estudos Cognitivos**

2004

**Universidade Estadual Paulista  
Faculdade de Filosofia e Ciências**

**Conselho Científico**

Ana Maria Pellegrini  
Francisco Cock Fontanella  
Franklin Leopoldo e Silva  
Itala Maria Loffredo D'Ottaviano  
Lauro Frederico Barbosa da Silveira  
Osvaldo Frota Pessoa Junior  
Paula Moucinho Martins

**Editoração**

Mariana Claudia Broens  
Edevaldo D. Santos

**Assessoria Técnica**

Maria Luzinete Euclides (Bibliotecária)  
Vânia Maria Silveira Reis Fantin (Bibliotecária)

**Capa**

Edevaldo D. Santos  
Enrique Raul Broens

Encontro com as Ciências Cognitivas / organização de  
Ademar Ferreira, Maria Eunice Quilici Gonzalez e  
Jonas Gonçalves Coelho. São Paulo, Cultura  
Acadêmica, 2004.

p. 236; 21 cm. (Encontro com as Ciências Cognitivas; v. 4)

**ISBN 85 — 7139520 — 9**

1. Naturalismo. 2. Informação e conhecimento. 3.  
Perspectivismo *versus* reducionismo. I. Ferreira, Ademar,  
org. II. Gonzalez, Maria Eunice Quilici, org. III. Coelho,  
Jonas Gonçalves, org.

DOI: <https://doi.org/10.3611/2004.85-7139520-9>

CDD — 153.4

Índice para catálogo sistemático

1. Cognição — conhecimento — 153.4
2. Epistemologia — 121

## Índice

|                                                                                                                                                                                |            |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>Introdução .....</b>                                                                                                                                                        | <b>i</b>   |
| <b>Parte I – Contribuições da Filosofia às Ciências Cognitivas.....</b>                                                                                                        | <b>03</b>  |
| Naturalismo em Filosofia da Mente                                                                                                                                              |            |
| <i>Paulo Abrantes .....</i>                                                                                                                                                    | <i>05</i>  |
| O reducionismo impiedoso de John Bickle e seus problemas                                                                                                                       |            |
| <i>Paulo Roberto Margutti Pinto .....</i>                                                                                                                                      | <i>39</i>  |
| Considerações sobre o perspectivismo e a lógica da<br>Ciência Cognitiva                                                                                                        |            |
| <i>Elias Humberto Alves .....</i>                                                                                                                                              | <i>75</i>  |
| O naturalismo biológico de John Searle e as ambigüidades<br>de uma concepção internalista da mente                                                                             |            |
| <i>Mariana Claudia Broens e Carmen Beatriz Milidoni .....</i>                                                                                                                  | <i>81</i>  |
| Bergson: consciência, corpo e ação                                                                                                                                             |            |
| <i>Jonas Gonçalves Coelho.....</i>                                                                                                                                             | <i>89</i>  |
| <b>Parte II – Auto-organização, informação e o problema das<br/>Representações nas Ciências Cognitivas 103</b>                                                                 |            |
| O mal estar do representacionismo: sete dores de cabeça da<br>Ciência Cognitiva                                                                                                |            |
| <i>Willem F. G. Haselager.....</i>                                                                                                                                             | <i>105</i> |
| Informação e auto-organização: tramas conceituais revisitadas                                                                                                                  |            |
| <i>Maria Luísa Bissoto .....</i>                                                                                                                                               | <i>121</i> |
| Atividade cerebral em múltiplas escalas e a experiência consciente                                                                                                             |            |
| <i>Alfredo Pereira Junior.....</i>                                                                                                                                             | <i>135</i> |
| A postura da mão influencia a compatibilidade estímulo-resposta<br>numa tarefa de reconhecimento da lateralidade                                                               |            |
| <i>Lameira, A. P.; Gawryszewski, L. G.; Carvalho, A. A.; Soeiro-Santos, R.;<br/>Guimarães-Silva, S.; Ferreira, F.M.; Pereira Jr, A.; Vargas, C. D.<br/>e Umiltà, C.A. ....</i> | <i>153</i> |
| O amplo conceito peirceano da mente: sua relevância para a biologia,<br>inteligência artificial e a cognição                                                                   |            |
| <i>Lúcia Santaella .....</i>                                                                                                                                                   | <i>167</i> |

*Encontros com as Ciências Cognitivas*

|                                                                                                                                                                                |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Notas sobre a lei da mente peirceana<br><i>Max Rogério Vicentini</i> .....                                                                                                     | 181 |
| Informação e conhecimento: notas para uma taxonomia da informação<br><i>Maria Eunice Quilici Gonzalez, Thiago Carreira Alves Nascimento e<br/>Willem F. G. Haselager</i> ..... | 195 |
| Aspectos da teoria de controle na cibernética<br><i>Ademar Ferreira</i> .....                                                                                                  | 221 |
| <b>Índice remissivo</b> .....                                                                                                                                                  | 233 |

## Introdução

O presente volume de *Encontro com as Ciências Cognitivas* reúne trabalhos apresentados nos V e VI “Encontro Brasileiro-Internacional de Ciências Cognitivas” (EBICC) realizados em Agosto de 2003 e Setembro de 2004, respectivamente, na UNESP, USP e UFMA, além de textos apresentados no Centro de Lógica, Epistemologia e História da Ciência (CLE) da UNICAMP.

A coletânea traz a seus leitores reflexões inéditas sobre os processos cognitivos e informacionais desenvolvidos pelos sistemas naturais e artificiais, elaboradas a partir de uma perspectiva interdisciplinar que envolve a filosofia, a neurociência, a computação, a engenharia, a biologia, a semiótica e a psicologia. A análise apresentada neste volume focaliza temas polêmicos, como a natureza da relação mente/corpo, reducionismo versus anti-reducionismo, internalismo versus externalismo, o papel da informação e das tecnologias computacionais no estudo dos estados mentais, a natureza das representações mentais, entre outros. O caráter polêmico desses temas reflete o estado atual das pesquisas em desenvolvimento nas ciências cognitivas no Brasil e no exterior. Em processo de consolidação, o campo das ciências cognitivas no Brasil se encontra disperso em departamentos de engenharia, filosofia, biologia e computação, sendo a sua característica interdisciplinar ainda pouco explorada. A série de publicações de *Encontro com as ciências cognitivas* representa uma tentativa de superação dessa deficiência, caracterizando-se como uma contribuição à expansão da concepção interdisciplinar do conhecimento manifesta nas ciências cognitivas.

O presente volume está dividido em duas partes, as quais dão especial ênfase aos tópicos do naturalismo e do reducionismo na Filosofia da Mente, bem como à abordagem informacional do conhecimento em suas várias facetas.

A *Parte I*, denominada *Contribuições da Filosofia às Ciências Cognitivas*, compreende os trabalhos de filósofos cujas investigações trazem análises sobre a natureza da mente e do conhecimento, a partir de diferentes perspectivas metodológicas, ontológicas e epistemológicas.

No capítulo 1, *Naturalismo e Filosofia da Mente*, Paulo Abrantes transpõe para a Filosofia da Mente a classificação de Goldman das diferentes modalidades de naturalismo em Epistemologia. De forma extremamente



didática, o autor propõe uma classificação do naturalismo em quatro grupos principais: o ontológico, o conceitual-lingüístico, o metodológico e o substantivo, caracterizando-os de modo a dar especial atenção a propriedades como intencionalidade e consciência, além de explorar as implicações de diferentes soluções para o problema da relação mente-corpo.

No capítulo 2, *O reducionismo impiedoso de John Bickle e seus problemas*, Paulo Roberto Margutti Pinto analisa criticamente o livro *Philosophy and neuroscience – A ruthlessly reductive account* de John Bickle. Apresenta inicialmente a teoria de Bickle mostrando que ela se filia à teoria da identidade mente-cérebro e se inspira no materialismo eliminativista, rejeitando a noção de propriedade emergente em sentido forte e a tese da múltipla realização. Com muita elegância e clareza, Margutti critica a viabilidade da redução proposta por Bickle e o seu conceito de metaciência ascendente, argumentando que considerações na direção inversa (descendente) são necessárias para um bom entendimento da problemática em questão. O autor conclui pela necessidade das duas descrições, a neurocientífica e a psicológica, no estudo dos estados mentais.

No capítulo 3, *Considerações sobre o Perspectivismo e a Lógica das Ciências Cognitivas*, Elias Humberto Alves analisa a posição do perspectivismo, tal como formulado pelo filósofo Donald Peterson, argumentando que, por meio da noção de “forma de representação”, o Perspectivismo indica uma metodologia para as ciências cognitivas. Como qualquer sistema lógico é uma forma de representação, Alves explicita a questão sobre a existência de uma lógica adequada para as investigações das ciências cognitivas. A sua sugestão é a de que essa lógica poderia ser a Lógica Illocucionária, a qual tem origem nas reflexões de Frege, passando depois por Austin, Montague e Searle, dentre outros.

No capítulo 4, *O Naturalismo biológico de John Searle*, Mariana C. Broens e Carmen Beatriz Milidoni discutem a concepção internalista dos estados mentais defendida por John Searle na perspectiva da filosofia da mente e das ciências cognitivas. Apesar de apreciarem a clareza expositiva do autor, elas argumentam que tal concepção é problemática na medida em que se situa simultaneamente no contexto biológico das neurociências e da experiência subjetiva de primeira pessoa. Sugerem as autoras que o projeto de pesquisa cognitivista, no seu viés da cognição situada e incorporada, pode, senão resolver

o problema da interface objetivo/subjetivo, pelo menos apontar caminhos mais profícuos relacionados a uma subjetividade corpórea.

A **parte I** se completa com o capítulo 5, *Bergson: consciência, corpo e ação*, elaborado por Jonas Gonçalves Coelho, que procura mostrar a forma original pela qual Bergson relaciona o surgimento e os vários modos de consciência à estrutura e ao funcionamento do sistema nervoso em seus respectivos corpos, ao longo da evolução dos seres vivos. O autor argumenta que a originalidade de Bergson está na importância concedida à ação na produção dos processos conscientes, ou seja, que a constituição da consciência está intimamente associada à inserção pragmática dos corpos no mundo.

A **Parte II**, que denominamos *Auto-organização, informação e o problema das representações nas ciências cognitivas*, está endereçada, principalmente, às questões relativas ao papel das representações mentais nas explicações de atividades relacionadas ao conhecimento comum, à organização cerebral e informacional na caracterização da atividade perceptual, aos processos de auto-organização e à causação final, quando se trata de estabelecer as bases conceituais para a ciência cognitiva. Elaborada por filósofos e cientistas com formações diversas, esta parte do livro apresenta resultados de uma cooperação interdisciplinar no estudo da cognição

No capítulo 6, *O mal estar do representacionismo: sete dores de cabeça da ciência cognitiva*, Willem F. G. Haselager aponta as dificuldades que os cientistas cognitivos enfrentam diante da abordagem representacionista do comportamento inteligente. Argumenta que a crítica ao representacionismo é uma tarefa difícil, pois a própria ciência cognitiva se tornou dependente de representações. Haselager apresenta alguns sintomas dessa dependência, propondo não a eliminação definitiva das representações, mas o seu uso em explanações e modelagens apenas quando for necessário, ou seja, quando houver um bom motivo, e não como ponto de partida de toda e qualquer explicação da atividade cognitiva.

Em *Informação e auto-organização: tramas conceituais revisitadas*, Maria Luísa Bissoto estabelece, no capítulo 7, correlações entre os conceitos de auto-organização e de informação. A autora procura mostrar que o modo como o conceito de informação é concebido no contexto dos estudos de auto-organização e das formas de apreensão da informação determina diferentes significados tanto para o conceito de auto-organização como para o de

informação. Um questionamento sobre o alcance e a possibilidade de aplicação desses conceitos para a explicação dos fenômenos aos quais estão ligados constitui um dos tópicos centrais do capítulo.

Em *Atividade cerebral em múltiplas escalas e a experiência consciente*, Alfredo Pereira Júnior critica, no capítulo 8, os modelos reducionistas que explicam a atividade do cérebro a partir de um patamar privilegiado de descrição, argumentando que o cérebro atua em diferentes escalas organizacionais, em complexa interação com o corpo e o ambiente. Partindo da hipótese segundo a qual os processos mentais correspondem à auto-organização do sistema cerebral em suas interações com o corpo e o ambiente, o autor argumenta também que, no tocante ao cérebro, o estudo das diversas escalas de organização – e respectivos modos de processamento de informação – é relevante para se entender os intrincados aspectos da cognição. Considera ainda que a experiência consciente se correlaciona com atividades de todas as escalas da atividade cerebral, encontrando um suporte na informação e computação quânticas em populações de íons.

No capítulo 9, *A postura de mão influencia a compatibilidade estímulo-resposta numa tarefa de reconhecimento da lateralidade*, Lameira A. C. et al. procuram mostrar que a discriminação de imagens especulares da lateralidade da figura de uma mão apresentada em tela de computador envolve não só o sistema visual mas também os sistemas somestésico e motor. Os autores analisam alguns dos fatores que podem afetar o tempo de reação necessário para determinar a lateralidade da figura de uma mão e apontam a interação entre fenômenos de compatibilidade espacial e a postura da mão. Com esse estudo, eles esperam contribuir para a compreensão dos mecanismos de reconhecimento de padrões realizados pelos seres humanos e máquinas, bem como para os estudos sobre a interação entre mecanismos sensoriais e motores em humanos e robôs.

Em *O amplo conceito peirceano de mente: sua relevância para a biologia, inteligência artificial e cognição*, Lúcia Santaella interpreta, no capítulo 10, a concepção peirceana de mente, mostrando que ela é mais ampla do que a concepção convencional que a equipara à consciência. A autora apresenta o conceito peirceano de mente como coextensivo de causação final. Esta caracterização da mente abrange tanto suas formas mais desenvolvidas, como, por exemplo, o ato humano intencional de raciocínio, até suas formas

mais rudimentares de atividades direcionadas para um fim ainda não consciente. Para Santaella, a equiparação entre mente e causação final explica a atribuição de mente a outros seres além dos humanos e mesmo fora do reino biológico. Além disso, ela defende a hipótese de que o modelo lógico básico da semiose, que se expressa na definição de signo, não é apenas um modelo para a descrição da mente, pensamento, inteligência, continuidade e crescimento; ele é também um modelo para o entendimento dos processos evolutivos que, para Peirce, são manifestações da mente entendida em sentido lato.

No capítulo 11, *Notas sobre a lei da mente peirceana*, Max Rogério Vicentini apresenta e analisa a idéia peirceana de “lei da mente” revelando elementos que permitam compreender o pensamento de Peirce e contribuir para o debate contemporâneo sobre a noção de auto-organização. Vicentini aponta similaridades e convergências entre o princípio que rege a lei da mente e as principais características hoje atribuídas à classe de fenômenos descrita pela noção de auto-organização. Duas sugestões são propostas pelo autor para o desenvolvimento dos estudos de auto-organização: avaliação e análise cuidadosa da noção de *continuum* e consideração da causalidade final como constituinte primário dos fenômenos de crescimento que ocorrem no universo, seja na própria mente ou na matéria considerada como um tipo especial de pensamento.

Em *Informação e conhecimento: notas para uma taxonomia da informação*, Maria Eunice Q. Gonzalez et al. propõem, no capítulo 12, uma geografia conceitual das principais abordagens contemporâneas do conceito de informação. A partir dessa classificação, os autores analisam problemas contemporâneos concernentes ao estatuto ontológico da informação e sua relação com o conhecimento. Argumentam que, em consonância com os instrumentos tecnológicos de nossa época, as reflexões sobre a natureza do conhecimento, no passado pertencentes essencialmente ao domínio clássico da filosofia, têm sido objeto de estudo interdisciplinar, em termos informacionais, pelos cientistas cognitivos. O sujeito cognitivo deixa de ser visto, nessa perspectiva, como um sujeito absoluto ou transcendental, situando-se no plano contingente dos sistemas informacionais que se auto-organizam no embate entre elementos físicos, biológicos e sociais. Um dos objetivos propostos consiste em elucidar a natureza do conhecimento comum, erroneamente considerado pela vertente tradicional das ciências cognitivas como resultante do processamento mecânico da informação num âmbito representacional.

Finalmente, no capítulo 13, *Aspectos da teoria de controle na cibernética*, Ademar Ferreira explora a trajetória das disciplinas de controle e sistemas dinâmicos na construção da primeira cibernética, procurando verificar as condições em que exerceram sua influência, ou deixaram de fazê-lo. São analisados alguns dos desdobramentos mais recentes da cibernética, os quais, segundo o autor, estão alterando nossas concepções do sistema nervoso central, da mente e do próprio fenômeno da vida. Ferreira analisa também algumas possibilidades que se apresentaram à teoria de controle de representar um papel mais significativo no projeto de edificar uma ciência da mente, tal como idealizado pelos primeiros cibernéticos. Finalmente, o autor apresenta algumas idéias que contribuíram para o desenvolvimento da neurociência e da inteligência artificial.

Ao trilhar os diversos caminhos de leitura propostos nesta coletânea, o leitor encontrará opiniões e argumentos nem sempre convergentes a respeito de temas comuns. Entendemos que esta característica se deve ao fato de que a presente coletânea constitui uma amostra de pesquisas realizadas em uma área de investigação ainda pouco consolidada não só no Brasil, como no mundo. Nos seus pouco mais de 50 anos de existência, as ciências cognitivas se encontram ainda em processo de construção, em um terreno árduo que envolve o trabalho da mente se auto-investigando com pretensões de cientificidade. Nesta progressão, o presente volume constitui apenas mais uma etapa na trilha do caminhante, que constrói o seu caminho ao caminhar. A expectativa dos organizadores é que a sua leitura possibilite debates e reflexões profícuas sobre a natureza e a dinâmica constitutiva das mentes.

*Ademar Ferreira  
Maria Eunice Quilici Gonzalez  
Jonas Gonçalves Coelho*

**PARTE I — CONTRIBUIÇÕES DA FILOSOFIA ÀS  
CIÊNCIAS COGNITIVAS**





# Naturalismo em Filosofia da Mente

Paulo Abrantes<sup>1</sup>

## Naturalismo: considerações preliminares

Uma das dificuldades em avaliarmos o naturalismo — mesmo se nos restringimos às suas versões mais recentes em epistemologia — é a variedade de teses filosóficas englobadas sob essa denominação. Não é óbvio que haja um núcleo comum de compromissos aceitos por todas as variedades de naturalismo em epistemologia. Apesar disso, é corrente o uso do termo 'naturalismo' para designar uma orientação bastante robusta em epistemologia.

O naturalismo em epistemologia corresponde a uma combinação de todas, ou de parte, das seguintes teses, configurando programas mais ou menos radicais:

- a) Rejeição da possibilidade de uma justificação *a priori* para crenças e, de modo particular, a contestação do pretenso *status a priori* da epistemologia;
- b) Rejeição do fundacionalismo;
- c) Externalismo (em teoria da justificação);
- c) Psicologismo;
- d) Fisicalismo;
- e) Monismo metodológico;
- f) Cientificismo.<sup>2</sup>

Como a temática deste artigo não se situa primordialmente no domínio da epistemologia, limito-me aqui a nomear essas teses, sem enunciá-las ou analisá-las (várias delas serão, de todo modo, discutidas em diferentes momentos neste artigo, particularmente o fisicalismo).<sup>3</sup>

---

<sup>1</sup> Universidade de Brasília – UnB. Endereço eletrônico: [abrantes@unb.br](mailto:abrantes@unb.br)

<sup>2</sup> Há quem inclua o 'darwinismo' nessa lista, embora isso não seja usual. Ver, por exemplo, Rosenberg (1996). Para Papineau (1993), o 'naturalismo' é entendido pelos filósofos de várias maneiras, e pode estar associado às seguintes teses: a) continuidade entre filosofia e ciência; b) rejeição do dualismo; c) rejeição do internalismo em epistemologia; d) fisicalismo.

<sup>3</sup> O leitor interessado no naturalismo epistemológico poderá consultar Abrantes (1995, 1998) e Abrantes e Bensusan (2003).



O termo 'naturalismo' não é empregado, usualmente, para distinguir programas em filosofia da mente, como é o caso em epistemologia. Neste artigo, proponho uma classificação de diferentes posições que poderiam ser consideradas 'naturalistas' em filosofia da mente.

Com esse fim, adoto estrategicamente o "guia" de Goldman (1998) para o naturalismo em epistemologia. Vou distinguir, com base nessa tentativa de classificação, modalidades análogas de naturalismo em filosofia da mente, dividindo-as em três grupos: o naturalismo enquanto uma postura ontológica, conceitual-lingüística ou metodológica. Essa transposição da epistemologia para a filosofia da mente não deve surpreender, na medida em que os problemas fundamentais de ambas as áreas estão intimamente conectados.

Como as propriedades de 'intencionalidade' e de 'consciência' são especialmente críticas para um programa naturalista em filosofia da mente, sirvo-me, também, da classificação que propõe Chalmers (1996) de distintas posturas frente ao *status* ontológico da consciência, bem como da que discutem Stich et al. (1994), mais voltada para as atitudes proposicionais (e, conseqüentemente, para os problemas colocados pela intencionalidade de alguns tipos de estados mentais).

Na medida do possível, tento ilustrar, com referências a diferentes filósofos contemporâneos, cada uma das modalidades identificadas de naturalismo e as correspondentes modalidades de antinaturalismo. Dou uma atenção especial, mas não exclusiva, às posições de Chalmers, de Searle e de Dennett.

Embora não seja este o objetivo principal do artigo, acredito que ele possa ser utilizado como uma introdução ao problema mente-corpo, explorando as implicações de diferentes soluções propostas para este que é o problema central da filosofia da mente.

### **Naturalismo ontológico (NATO)**

Há controvérsias sobre em que medida o naturalismo implica uma ontologia particular como, por exemplo, o fisicalismo redutivo — posição que rejeita objetos, estados, propriedades e processos além daqueles postulados pela física. Mostrarei que o naturalismo também é compatível com um fisicalismo não-redutivo, por exemplo.

De modo geral, podemos considerar naturalistas ontológicas aquelas orientações que rejeitam o dualismo de substância e que postulam que os estados e propriedades mentais são idênticos, reduzem-se ou supervêm a estados neurais.<sup>4</sup> Cada uma dessas postulações pode ser considerada 'fiscalista'.

NATO tem diversas variantes, correspondendo a diferentes concepções de 'redução' e de 'superveniência'. Combinações entre os tipos de redução e os tipos de superveniência geram várias sub-modalidades de naturalismo ontológico. Algumas delas serão exemplificadas a seguir.

### **1. Fiscalismo<sup>5</sup> redutivo**

O reducionismo é frequentemente apresentado como uma posição que admite a possibilidade de se reduzir as ciências de nível alto, ou 'especiais' (como a psicologia, por exemplo) às ciências 'fundamentais', como a biologia ou a física.<sup>6</sup> No contexto da presente seção, o reducionismo relevante é, propriamente, o ontológico.

Numa ontologia fiscalista redutiva, não há nada além das propriedades e leis descritos pela física. No naturalismo fiscalista redutivo, postula-se uma redução do mental (do fenomênico, do intencional ou de estados com conteúdo) ao físico. Essa redução pode tomar a forma de uma identidade entre propriedades mentais, de um lado, e propriedades físicas, de outro.

A teoria da identidade (*brain state theory*) exemplifica esta última posição, ao postular uma identidade entre (tipos de) estados mentais e (tipos de) estados físicos. Por isso, essa solução para o problema mente-corpo também é conhecida como um 'fiscalismo de tipo', que definirei de forma precisa na próxima seção. Na avaliação de J. Kim, o grande legado da teoria da identidade foi de tornar praticamente hegemônica, entre os filósofos da mente contemporâneos, uma perspectiva fiscalista no tratamento do problema mente-

---

<sup>4</sup> Mostrarei que posturas não-naturalistas correspondem a negar que estados mentais com conteúdo, por exemplo, ou que possuem *qualia*, possam ser idênticos, redutíveis ou mesmo que supervenham a estados físicos (e.g. neurais).

<sup>5</sup> 'Fiscalismo' e 'materialismo' estarão sendo usados, no que se segue, como sinônimos, salvo observação em contrário. Por vezes distingue-se, entretanto, tais termos. Ver os verbetes correspondentes em Guttenplan (1995); Kim e Sosa (1995).

<sup>6</sup> Isso remete ao que Searle chama de 'redução teórica'. Ele distingue, como apresentarei adiante, diferentes tipos de redução.

corpo. Kim caracteriza tal perspectiva, basicamente, pela rejeição do dualismo de substância proposto por Descartes.

Proponho que se inclua também posições eliminativistas nessa categoria de fisicalismo redutivo, embora não se deva confundir 'eliminar' (e.g. conceitos ou termos associados ao mental) com 'reduzir'. O eliminativismo é a tese de que determinados estados ou propriedades mentais simplesmente não existem, embora tendamos a vê-los como a referência de conceitos utilizados na linguagem mentalista ordinária. O eliminativismo pode ser associado, portanto, ao não-realismo com respeito aos conceitos ou termos mentalistas, como eles são compreendidos, por exemplo, no âmbito da psicologia de senso comum. Se não existem os estados e propriedades a que se referem os conceitos associados ao mental, não faz sentido tentar reduzi-los. Podemos ser não-realistas com respeito às chamadas atitudes proposicionais ('crenças', 'desejos', etc. e, de modo geral, com respeito a estados mentais com conteúdo), mas também com respeito às chamadas propriedades fenomênicas associadas à consciência.<sup>7</sup> O chamado 'nilismo de *qualia*', defendido por Dennett entre outros, exemplifica esta última posição.

É problemático tentar enquadrar na modalidade de naturalismo ontológico o chamado "materialismo eliminativo" de Paul Churchland.<sup>8</sup> Apesar do nome, essa corrente não apresenta contornos ontológicos claros no que diz respeito, propriamente, ao problema mente-corpo. A tese central defendida por Churchland é que a linguagem mentalista (a linguagem da psicologia de senso comum) deve ser eliminada e, não reduzida, por exemplo, à linguagem da neurofisiologia. Como o que se pretende eliminar aqui não são eventos ou propriedades, mas sim elementos lingüísticos, essa orientação talvez localize-se

---

<sup>7</sup> O termo "consciência" refere-se, ordinariamente, a diferentes capacidades. Mas há um crescente consenso, entre filósofos da mente, em considerar a "experiência qualitativa", em especial as qualidades (*qualia*) associadas às nossas sensações, como o "problema difícil" (CHALMERS, 1996, 1997) no domínio do mental. Thomas Nagel popularizou a expressão "what it is like to be X" - que poderíamos traduzir por "como é ser X" - para referir-se à experiência de um indivíduo X, possuidor de uma estrutura cognitiva particular e, portanto, de uma "consciência" particular do mundo. Nessa expressão, a variável X pode ser substituída por um ente — um homem, um morcego (como prefere Nagel) ou um indivíduo de outra espécie biológica — capaz de ter estados mentais conscientes, e possuidor de um ponto de vista particular diante do mundo. Ver também nota 40.

<sup>8</sup> Embora, por comodidade, eu vá, a seguir, omitir o prenome, não deve haver confusão com uma outra filósofa, a Patricia Churchland, cujas posições não são objeto deste artigo.

melhor na modalidade de naturalismo conceitual-lingüístico. Voltarei, portanto, ao materialismo eliminativo mais adiante.

O fisicalismo, entretanto, não implica necessariamente um reducionismo ou um eliminativismo. Kornblith (1998), um destacado naturalista, embora considere que o progresso científico forneça evidências a favor do fisicalismo (ou seja, a favor da “constituição física” de todas as coisas, como ele define essa doutrina), sustenta que o estágio atual das ciências não apóia o reducionismo e, muito menos, o eliminativismo. Kornblith é realista com respeito às espécies naturais postuladas por ciências como a psicologia (p.ex. em psicologia cognitiva, tipos de estados mentais) e considera que tais espécies têm poder causal genuíno. Ele rejeita, conseqüentemente, o epifenomenalismo<sup>9</sup> em filosofia da mente. Há, portanto, compatibilidade entre uma posição naturalista e a autonomia das ciências especiais.<sup>10</sup>

Pettit (1994) também percebe uma ambigüidade ontológica no naturalismo: há aqueles que defendem um reducionismo, ou mesmo um eliminativismo, restringindo portanto a sua ontologia; e os que adotam uma ontologia mais rica, na qual também têm lugar, por exemplo, propriedades que supervêm às propriedades físicas.

## **2 Fisicalismo não-redutivo**

Acredito que, hoje em dia, seja este o âmbito no qual a maioria dos filósofos da mente se deslocam. O fisicalismo não-redutivo é uma perspectiva que rejeita o dualismo de substância — levando a sério as evidências empíricas de que há covariância entre o mental e o físico — e pressupõe a dependência do mental em relação ao físico sem, no entanto, admitir uma redução do mental ao físico.

### **2.1 A relação de superveniência**

A superveniência é normalmente empregada em tentativas de articulação de um fisicalismo não-redutivo, enquanto pano de fundo ontológico no tratamento do problema mente-corpo. Pretende-se, com base nessa relação, assegurar a autonomia das ciências especiais, como a psicologia, com respeito às

---

<sup>9</sup> O epifenomenalismo é a doutrina que defende que estados mentais são causados por estados cerebrais, mas que aqueles não têm poder causal, sendo meros ‘epifenômenos’.

<sup>10</sup> Para uma exposição mais detalhada do naturalismo de Kornblith, ver Abrantes (2004).



fundamentais, como a física. A relação de superveniência pode ser formulada nos seguintes termos: “Indiscernibilidade física implica indiscernibilidade psicológica; ou, equivalentemente, nenhuma diferença mental sem diferença física” (KIM, 1996, p.10).

A partir dessa definição básica, a relação de superveniência pode ser formulada de forma mais fraca ou mais forte.<sup>11</sup>

O princípio de superveniência do mental ao físico é usado para comparar as posições conhecidas como 'fiscalismo de particulares' (*token physicalism*) e 'fiscalismo de tipo' (*type physicalism*). O 'fiscalismo de particulares' é definido por Kim nos seguintes termos:

Todo evento que é uma ocorrência de um tipo de evento mental também é uma ocorrência de um tipo de evento físico (ou, em outras palavras, todo evento que possui uma propriedade mental possui também alguma propriedade física)” (KIM, 1996, p. 59).

Este princípio, por si só, não impõe uma correlação entre propriedades mentais e físicas e, portanto, pode valer mesmo se não há superveniência do mental ao físico. Dois eventos físicos particulares, ambos de um mesmo tipo, podem estar, cada um deles, correlacionado a um evento mental de um tipo diferente. Nesse sentido, o fiscalismo de particulares não é propriamente uma doutrina fiscalista (isso se aceitarmos que o princípio de superveniência define um fiscalismo mínimo, tese que analisarei na próxima seção)<sup>12</sup>.

<sup>11</sup> Assim, a indiscernibilidade física pode dizer respeito exclusivamente a indivíduos tomados num mesmo mundo, ou então envolver comparações de indivíduos tomados em diferentes mundos possíveis (superveniência local); num outro registro, a superveniência pode unicamente comparar as distribuições globais de propriedades físicas, de um lado, e mentais, de outro, em diferentes mundos possíveis, não se comprometendo com comparações de como essas propriedades são instanciadas e estão relacionadas em indivíduos, comparados dois a dois, nesses mundos (superveniência global). A superveniência local implica a global, mas não o contrário. Por outro lado, o modo como especificamos a noção de ‘possibilidade’, se lógica ou nomológica pode, por sua vez, gerar variantes das relações anteriores. Não caberia, para os fins deste artigo, discutir essas diversas noções de superveniência, um tópico por si só bastante complexo.

<sup>12</sup> O fiscalismo de particulares é uma doutrina fraca demais, pois não pressupõe que a cada vez que um evento seja a ocorrência de um certo tipo de estado mental — por exemplo, uma dor — ele também seja a ocorrência do mesmo tipo de evento físico — digamos, a ativação da fibra C. Num determinado momento, um evento pode instanciar esses dois tipos e, em outro momento, instanciar o mesmo tipo mental (uma dor) e um tipo físico totalmente diferente (por exemplo, a ativação da fibra D, e não mais a fibra C, como no evento anterior). Na verdade, o fiscalismo de particulares é compatível até com a existência de um outro mundo, idêntico ao nosso em todos os seus aspectos físicos, mas no qual não haja sequer mentalidade (ver KIM, 1996, p. 61).

A teoria da identidade mente-corpo implica um fisicalismo de tipo: “Tipos de eventos mentais são tipos de eventos físicos; em outras palavras, propriedades mentais são propriedades físicas” (KIM, 1996, p. 59).

O fisicalismo de tipo é, portanto, uma posição reducionista: não há propriedades mentais distintas de propriedades físicas, e além destas últimas. Propriedades mentais e físicas são, na verdade, identificadas. A superveniência do mental ao físico permite definir, portanto, uma posição que é mais forte do que o fisicalismo de particulares, porém mais fraca do que o fisicalismo de tipo.

Filósofos com tendências naturalistas ou fisicalistas mais radicais consideram, entretanto, o conceito de ‘superveniência’ bastante suspeito, por parecer ressuscitar, com uma nova roupagem, posições consideradas problemáticas como o epifenomenalismo, o emergentismo, etc. Uma das críticas à noção de superveniência é que ela conduz a aceitar relações, por exemplo, entre o mental e o físico, enquanto fatos brutos, não passíveis de serem, por sua vez, explicados.<sup>13</sup> Veremos que esta crítica pode também ser feita ao dualismo que propõe Chalmers.

## 2.2 O fisicalismo mínimo de J. Kim

Para Kim, o fisicalismo adota, no mínimo, as três teses seguintes:

- i) *Superveniência mente-corpo*;
- ii) *Princípio anticartesiano*: não podem existir puros seres mentais (ou seja, propriedades mentais sem contrapartida física);<sup>14</sup>
- iii) *Dependência mente-corpo*: propriedades mentais são determinadas por propriedades físicas (KIM, 1996, p. 11)<sup>15</sup>.

Qualquer doutrina mais fraca é considerada, por Kim, não-fisicalista. O reducionismo, por sua vez, corresponde a uma espécie mais forte de fisicalismo.

Kim não acredita que a relação de superveniência configure uma solução para o problema mente-corpo: diferentes soluções incompatíveis para esse problema implicam a superveniência (KIM, 1998, p. 12). A superveniência

---

<sup>13</sup> Para uma defesa da emergência na explicação das relações ontológicas entre níveis, em lugar da relação de superveniência, que não proporcionaria essa explicação, ver Humphreys (1997a, 1997b).

<sup>14</sup> Pode-se mostrar que este segundo princípio não é implicado, estritamente, pela superveniência.

<sup>15</sup> O princípio da dependência mente-corpo é mais forte que o da superveniência, que não pressupõe determinismo, mas só covariância. A relação de determinação é assimétrica, mas não a relação de superveniência.

não pretende, portanto, oferecer uma explicação para as correlações entre estados mentais e estados físicos; postular essa relação simplesmente exclui formas extremas de dualismo, como o cartesiano (KIM, 1998, p. 15).<sup>16</sup>

Kim (1998, p. 101), de forma um tanto surpreendente, considera o funcionalismo como uma concepção reducionista (embora, normalmente, se tenha considerado o funcionalismo como uma alternativa à teoria da identidade mente-corpo).<sup>17</sup> Ele critica a análise da redução teórica proposta por E. Nagel<sup>18</sup> e defende que a redução de um estado ou propriedade (e.g. mental) pressupõe a 'funcionalização' desse estado ou propriedade.<sup>19</sup> O problema, contudo, é que se fizermos esse tipo de redução, os estados mentais perdem poder causal autônomo: toda causalidade dar-se-á no nível dos realizadores físicos dos estados funcionalizados.

Chalmers concorda com Kim: a múltipla realizabilidade de um evento (e.g. mental) não é um empecilho para a sua explicação reduitiva. Desde que a ocorrência do evento possa ser explicada em termos de uma ocorrência (*token*) de nível mais baixo, há redução. Isso quer dizer que o funcionalismo *tout court* em filosofia da mente é, no final das contas, reducionista (CHALMERS, 1996, p. 43).

Num balanço final das conseqüências desse tipo de redução via funcionalização, Kim enfatiza a (provável) irredutibilidade da consciência:

[...] a notícia realmente ruim é que algumas propriedades mentais, notadamente as propriedades fenomenológicas das experiências

---

<sup>16</sup> Kim sentença: "[...] A superveniência não é, portanto, uma relação metafísica 'profunda' sobre padrões de covariância de propriedades, padrões que possivelmente são manifestações de relações de dependência mais profundas. Se isso é correto, a superveniência mente-corpo [só] coloca o problema mente-corpo, não sendo uma solução para ele" (KIM, 1998, p. 14).

<sup>17</sup> De modo a possibilitar a múltipla realização de estados mentais — incompatível com a teoria da identidade mente-corpo, uma posição que, como vimos, é reducionista — o funcionalismo caracteriza os estados mentais de forma abstrata, em termos de sua função na psicologia de um sistema cognitivo. Tal função envolve as relações de um tipo de estado mental com outros estados mentais, com as entradas (estímulos) e saídas (respostas) do sistema.

<sup>18</sup> Sobre E. Nagel, ver nota 18.

<sup>19</sup> Ao reduzir uma propriedade nas ciências físicas, em primeiro lugar fazemos, segundo Kim, uma reconstrução funcional dessa propriedade. Deixamos de ver essa propriedade como uma propriedade intrínseca e passamos a vê-la como uma propriedade extrínseca, por exemplo, caracterizada em termos de relações causais (KIM, 1998, p. 24-25; 98 et seq.). Esse é o procedimento-padrão para se funcionalizar uma propriedade. São exemplos de reconstruções desse tipo as definições de propriedades como 'temperatura', 'ser transparente' e aquelas associadas ao 'gene'. A distinção entre propriedades de primeira-ordem e de segunda-ordem também é relevante nesse contexto (Ver ABRANTES e AMARAL, 2002).

conscientes, parecem resistir à funcionalização e isso significa que não existe maneira de responder por sua eficiência causal dentro de um esquema fisicalista" (1998, p. 118-119).

E conclui com um tom pessimista: "[...] todos os caminhos que se bifurcam a partir do fisicalismo parecem, ao final, convergir possivelmente para o mesmo ponto, a irreabilidade do mental" (1998, p. 118-119).

Apresentarei, numa próxima seção, o argumento em que Kim, partindo dos pressupostos básicos do fisicalismo, demonstra que estados mentais têm poder causal somente na medida em que os indentifiquemos a estados físicos. Esse argumento, se válido, compromete a esperança de se poder articular um fisicalismo de tipo não-redutivo, no qual estados mentais possam ter um poder causal autônomo (condição *sine qua non* de um realismo quanto ao mental).

### 2.3 O naturalismo não-materialista de Searle

Os "materialistas", como os caracteriza Searle, pretendem "naturalizar" a intencionalidade e a consciência, no sentido de reduzir os "fenômenos mentais" aos "fenômenos físicos" (SEARLE, 1992, p. 2).

Searle tem o cuidado de distinguir, entretanto, vários tipos de redução:

a) redução ontológica: objetos/entidades de certos tipos não são considerados nada mais que objetos de outros tipos;

b) redução ontológica de propriedades: um caso especial de (a), em que se reduz não um objeto/entidade, mas sim propriedades;

c) redução teórica: este é o tipo de redução usualmente estudada pela filosofia da ciência de cepa empirista-lógica<sup>20</sup>;

d) redução lógica ou definicional;

e) redução causal: os poderes causais da coisa reduzida são "inteiramente explicáveis" em termos dos poderes causais dos fenômenos redutores;

Searle é partidário de um naturalismo comprometido com uma redução do último tipo (1992, p. 115). Passo a caracterizá-la.

<sup>20</sup> O modelo de redução inter-teórica mais conhecido foi proposto por E. Nagel. Mencionei, acima, este influente filósofo da ciência quando estava apresentando as posições de Kim. Não confundi-lo com o filósofo da mente T. Nagel, mencionado em nota anterior.



*Reduccionismo causal*

Searle só usa duas vezes a expressão “naturalismo biológico” em seu livro *A Redescoberta da Mente*, e a faz corresponder à tese segundo a qual

“... eventos e processos mentais são causados por processos neurofisiológicos no cérebro e são eles próprios características do cérebro [...] Eventos e processos mentais são parte de nossa história natural biológica tanto quanto a digestão, a mitose, a meiose ou a secreção de enzimas” (SEARLE, 1992, p. 1, 106)<sup>21</sup>.

Em *O Mistério da Consciência*, a expressão “naturalismo biológico” ocorre duas vezes (SEARLE, 1997, p. xiv, 210), sempre no contexto de uma discussão sobre a consciência. Ele refere-se à consciência do seguinte modo:

“A consciência é um fenômeno biológico [...] Ela é causada por microprocessos no cérebro, num nível mais baixo [*lower-level*] e ela é uma característica do cérebro nos níveis de ordem mais alta [*higher-order*]” (SEARLE, 1997, p. xiv).

Searle enfatiza que o seu naturalismo biológico não é uma forma de “materialismo”, já que ele vincula o materialismo ao reduccionismo. Ao contrário, para ele “a consciência é uma parte real do mundo real e não pode ser eliminada a favor de alguma outra coisa, ou reduzida a ela” (SEARLE, 1997, p. 210).

Segundo Searle, toda redução, incluindo a causal, objetiva, em princípio, uma redução ontológica. A consciência, contudo, possuiria um caráter especial, e sua redução causal não implicaria uma redução ontológica. Searle apresenta dois argumentos para sustentar essa tese:

A) A consciência é uma propriedade causalmente emergente do cérebro.

Uma propriedade de um sistema é causalmente emergente quando ela não pode ser “deduzida, concebida [*figured out*] ou calculada” a partir simplesmente da composição e arranjo dos elementos do sistema<sup>22</sup>. Uma propriedade emergente, para ser “explicada”, tem que se levar em consideração também “as interações causais entre os elementos” do sistema (SEARLE, 1992, p. 112).

<sup>21</sup> Todas as traduções do livro de Searle de 1992 são minhas, feitas livremente a partir da edição original, em inglês. O mesmo vale para os textos de Chalmers e de Dennett.

<sup>22</sup> Searle distingue dois tipos de emergência. Aqui trata-se da emergência do primeiro tipo. Searle não acredita que existam propriedades emergentes de um segundo tipo. Para essa distinção, ver Searle (1992, p. 112).

A consciência, embora emergente, é considerada por Searle uma propriedade física: para ele não existiriam propriedades fenomênicas ao lado (ou além) de propriedades físicas. Aquelas são causadas por estas últimas<sup>23</sup>.

B) Mesmo no caso de propriedades físicas (como calor, cor, etc.) a redução ontológica só é conseguida fazendo-se abstração das características mentais (fenomênicas) associadas a tais propriedades.

No caso da 'cor', do 'calor', etc. a redução não constitui uma nova descoberta, mas resulta de uma redefinição desses conceitos de modo a se excluir a "parte subjetiva" dessas qualidades<sup>24</sup>. Os aspectos subjetivos ou "epistêmicos" (sic.) que são deixados de lado para se fazer a redução não deixam, contudo, de existir, de ter "realidade"; esse é o padrão tanto na redução de qualidades primárias (e.g. solidez, liquidez, etc.), quanto na de qualidades secundárias (e.g. calor, cores, etc):

Em geral, o padrão de nossas reduções apóia-se na rejeição da base epistêmica subjetiva para a presença de uma propriedade, como uma parte da constituição última daquela propriedade. Nós descobrimos [fatos, coisas] a respeito do calor ou da luz por meio da sensação/tato [*feeling*] e da visão, mas nós em seguida definimos o fenômeno de um modo que é independente da epistemologia (SEARLE, 1992, p. 122).

Tomemos como exemplo a redução ontológica da qualidade (ou propriedade) 'vermelho'. A partir do momento que se sabe que a experiência do vermelho é causada pela emissão de luz com uma certa frequência  $\lambda$ , propõe-se a seguinte redução ontológica: o vermelho é a luz de frequência  $\lambda$ . Para fazer-se essa redução, para afirmar-se essa identidade, deixa-se de lado os aspectos fenomênicos, subjetivos, ligados à experiência do vermelho (o *quale* do vermelho).

No caso da consciência, Searle defende que ela não é passível desse tipo de redução (ou seja, através de uma redefinição) pois seus aspectos essenciais

---

<sup>23</sup> Nesse sentido, Searle talvez pudesse ser considerado um fisicalista não-redutivo. Normalmente, fisicalismo e materialismo são considerados termos sinônimos (ver nota 4), mas Searle, como vimos, não se considera materialista, pois associa essa posição ao reducionismo. Ver também, abaixo, a terceira tese defendida por Chalmers.

<sup>24</sup> As reduções de entidades/propriedades físicas envolvem, portanto, redefinições, nas quais uma parte do (significado do) conceito inicial é deixada de lado em função de nossos interesses (dimensão pragmática da redução ontológica). Nas ciências físicas, o interesse é o de controle, daí a importância dada, nas redefinições, aos nexos causais/funcionais, deixando-se de lado os aspectos subjetivos/experienciais (cf. SEARLE, 1992, p.123).

[meu termo] são subjetivos (experenciais/aparenciais). No caso da consciência, diz Searle, "a aparência é a realidade" (1991, p. 122); o que nos interessa é a própria experiência subjetiva<sup>25</sup>.

Na perspectiva defendida por Searle, eventos de alto-nível (ou num macronível) têm poderes causais. Em particular, no caso de um evento mental como uma dor, ele causaria outros eventos mentais e, eventualmente, de modo descendente, eventos no nível físico. Searle rejeita, portanto, o epifenomenalismo e, como veremos, ao admitir a causação descendente, compromete a tese, central para o fisicalismo, do fechamento causal do mundo físico<sup>26</sup>.

Com respeito à superveniência, a posição de Searle é bastante heterodoxa. Ele define uma "superveniência causal" do mental ao neurofisiológico nos seguintes termos: "causas neurofisiológicas tipo-idênticas [*type-identical*] teriam efeitos mentais tipo-idênticos" (SEARLE, 1992, p. 124). Essa superveniência causal implica que estados neurofisiológicos são causalmente suficientes, embora não necessários, para estados mentais<sup>27</sup>.

Dennett (1993) interpreta o naturalismo biológico de Searle como a negação da existência de um nível intermediário de processamento de informação, entre o nível neurofisiológico e o nível dos processos mentais. Searle é, de fato, um crítico do funcionalismo em filosofia da mente e do recente flerte desta última área com as chamadas ciências cognitivas, em especial a inteligência artificial. Ele ficou famoso com a controvertida experiência de

---

<sup>25</sup> Para Searle, isso não afetaria o nosso "quadro científico de mundo", pois não teria implicações metafísicas "profundas". Seria unicamente uma decorrência "trivial" de nossas "práticas definicionais" ou padrões de redução: "O contraste entre a redutibilidade do calor, da cor, da solidez, etc., de um lado, e a irredutibilidade dos estados conscientes, de outro, não reflete nenhuma distinção na estrutura da realidade, mas uma distinção nas nossas práticas definicionais. Nós podemos dizer o mesmo, adotando o ponto de vista do dualista de propriedade: o contraste aparente entre a irredutibilidade da consciência e a redutibilidade da cor, do calor, da solidez, etc., na verdade é só aparente. Nós realmente não eliminamos a subjetividade da cor, por exemplo, quando nós reduzimos o vermelho a reflexos de luz; nós simplesmente paramos de chamar de 'vermelho' a essa parte subjetiva. Nós não eliminamos quaisquer fenômenos subjetivos com essas reduções; nós simplesmente paramos de chamá-los pelos nomes antigos. Seja que tratemos a irredutibilidade de um ponto de vista materialista, seja dualista, nós ainda ficamos com um universo que contém um componente físico [*sic!*] subjetivo irredutível, enquanto um componente da realidade física" (Searle, 1992, p. 123).

<sup>26</sup> Trata-se da tese de que qualquer evento físico só pode ter como causa um outro evento físico.

<sup>27</sup> Nos termos de Searle, "identidade [*sameness*] neurofisiológica garante a identidade na mentalidade, mas identidade na mentalidade não garante a identidade neurofisiológica" (SEARLE, 1992, p. 125).

pensamento (*Gedankenexperiment*) do "quarto chinês", que usou repetidas vezes na tentativa de refutar um dos elementos do credo cognitivista: o modelo computacional de mente.

#### 2.4 J. Kim e o problema da causação mental

O problema da causação mental representa, para Kim, um desafio para o fisicalismo, constituindo, ironicamente, uma revanche do cartesianismo.

Para Kim, o desafio é defender um fisicalismo robusto — que pressupõe o fechamento causal do domínio físico — sem, no entanto, cair no reducionismo. Há, porém, vários problemas com a causação mental, sendo que o da exclusão causal tem implicações diretas para o fisicalismo (e, portanto, para o naturalismo).

Kim discute diversos modelos de causação mental, sendo um deles o de causação superveniente. O modelo superveniente admite, por um lado, que haja causação num macronível, mas que essa causação seja superveniente à causação no micronível, isto é, no nível físico. Ou seja, eventos no macronível mantêm, nesse modelo, poder causal, não desembocando no epifenomenalismo.

O modelo de causação superveniente mantém-se fisicalista em sua perspectiva por fazer, assim mesmo, depender a causação no macronível da causação no micronível. Kim faz-nos ver que esse é um resultado que o emergentista, por exemplo, dificilmente aceitaria. Para o emergentista, a partir do momento em que as propriedades no macronível 'emergem', os estados nesse nível passam a ter um poder causal autônomo, não-derivado dos poderes causais dos eventos no micronível. No modelo de causação superveniente, o poder causal no macronível é derivado do poder causal no micronível.

Kim mostra, entretanto, que a causação superveniente é ameaçada por um dilema:

a) se a superveniência não vale, o mental torna-se um domínio ontologicamente autônomo, não se ancorando no físico, havendo a possibilidade do mental ter efeitos no mundo físico sem uma correspondente causa física. Perdemos, desse modo, o fechamento causal do mundo físico, e o fisicalismo não se sustenta.

b) se a superveniência vale, então temos sobredeterminação causal (um estado mental sendo causado tanto por um outro estado mental quanto por um

estado cerebral, portanto, físico). Para se evitar a sobredeterminação, o papel causal do estado mental torna-se dispensável e/ou temos novamente uma violação do fechamento causal do mundo físico. As relações de estados mentais entre si, ou entre estados mentais e estados físicos, não são genuinamente causais (KIM, 1998, p. 45).

Conclusão do dilema: se a superveniência mente-corpo não vale, a causação mental é ininteligível; se ela vale, a causação mental é, de novo, ininteligível. Logo, a causação mental é ininteligível.<sup>28</sup>

É claro que esse dilema não se coloca para posições reducionistas, como a teoria da identidade. Kim assinala que Searle, com a sua tese de que estados mentais são causados por estados cerebrais, não escapa a esse dilema, também caindo numa sobredeterminação causal (ver KIM, 1998, p. 48). A posição de Searle (de que não há sobredeterminação, mas simplesmente descrições feitas em diferentes níveis, que seriam compatíveis) colapsaria numa redução (do mental ao cerebral), o que é inconsistente com outras posições defendidas por ele.

## 2.5 O monismo não-redutivo de Davidson

O "monismo anômalo" de Davidson pode ser considerado como um fisicalismo de particulares (*token physicalism*). Davidson é monista por defender que se considerarmos os nexos causais e adotarmos a descrição adequada, isto é, a física, todo evento mental torna-se idêntico a um evento físico.<sup>29</sup> Ele rejeita, nesse sentido, qualquer tipo de dualismo.<sup>30</sup>

Davidson nega, entretanto, que se possa estabelecer relações nomológicas entre tipos mentais, ou ainda entre tipos mentais e tipos físicos. Nesse sentido, o mental seria anômalo:

O monismo anômalo assemelha-se ao materialismo ao afirmar que todos os eventos são físicos, mas rejeita a tese, usualmente considerada essencial ao materialismo, de que podem ser dadas explicações

---

<sup>28</sup> Kim apresenta esse impactante argumento no seu livro de 1998. Para uma discussão detalhada do mesmo, no contexto do funcionalismo em filosofia da mente, ver Abrantes e Amaral (2002).

<sup>29</sup> Davidson defende uma ontologia em que eventos são indivíduos e, não, universais, como em Smart e outros.

<sup>30</sup> Davidson rejeita tanto o dualismo de substância cartesiano, que ele classifica como um dualismo anômalo, quanto outras formas de "dualismo nomológico", como paralelismos, o interacionismo, o epifenomenalismo, etc. Provavelmente Davidson também rejeitaria o dualismo nomológico de Chalmers, que discutirei abaixo, embora este filósofo se apresente como um naturalista.



puramente físicas dos fenômenos mentais" (DAVIDSON, 1991, p. 250).

O "monismo nomológico", ao contrário do monismo anômalo, é materialista por supor a existência de leis psicofísicas. Davidson nega essa possibilidade: "Não há leis psicofísicas estritas dado o caráter díspar dos compromissos [*commitments*] dos esquemas mental e físico" (1991, p. 253). As atribuições de atitudes proposicionais a agentes não somente é holista — pressupondo uma ampla teoria a respeito dos estados mentais do agente — como os ajustes a serem feitos nessa teoria comprometem-se com um "ideal constitutivo de racionalidade". Necessariamente, "concebemos o homem como um animal racional" (1991, p. 254) ao fazer ajustes nessa teoria a respeito de indivíduos, que tomamos como pessoas<sup>31</sup>.

O monismo anômalo é compatível com a superveniência do mental ao físico. Para Davidson, essa relação de superveniência asseguraria o caráter não reducionista da sua posição.<sup>32</sup>

A teoria física constitui um sistema fechado, no sentido de um fechamento causal. O mental, ao contrário, não constituiria um sistema fechado já que Davison admite o princípio de que há dependência causal de (pelo menos alguns) eventos mentais com respeito a eventos físicos. Portanto, há dependência causal, mas independência nomológica, do mental com respeito ao físico:

[...] eventos mentais como uma classe [tipo] não podem ser explicados pela ciência física; eventos mentais particulares podem [ser explicados] quando conhecemos as identidades particulares (1991: p. 255).

Mas não é esse o tipo de explicação que buscamos, por exemplo, para a ação de uma pessoa, e sim uma que relacione eventos mentais com outros

---

<sup>31</sup> Poder-se-ia explorar as conexões entre essa noção davidsoniana de "anomalia" do mental e a "postura intencional" de Dennett, que se recusa a adotar o realismo com respeito às nossas atribuições de estados intencionais a outros agentes. A psicologia de senso comum (*folk psychology*) não teria, para Dennett, o *status* de uma teoria à qual se possa atribuir valores veritativos (ou seja, ela não seria verdadeira nem tampouco falsa), mas estaria envolvida nas nossas práticas interpretativas cotidianas, visando a compreender e a prever o comportamento dos nossos semelhantes. Como Dennett é um naturalista, isso indica que o naturalismo não se compromete, necessariamente, com um realismo a respeito do mental (pelo menos como ele é entendido com base nas categorias da psicologia de senso comum). Cf. o "materialismo eliminativo" de Churchland.

<sup>32</sup> Vimos, na última seção, que o argumento de Kim contesta isso.



eventos mentais (e a ação<sup>33</sup>). Adotando um explícito viés kantiano, Davidson admite que "o anomalismo do mental é, portanto, uma condição necessária para que se veja a ação como [nomologicamente] autônoma" (1991, p. 255).

Não tenho a pretensão de avaliar em que medida a proposta de Davidson é consistente. O meu objetivo neste artigo é mais modesto: mostrar que o fisicalismo não-redutivo, além de não se apresentar como uma posição unívoca, ainda enfrenta problemas sérios para afirmar-se como uma posição naturalista aceitável e consistente em filosofia da mente<sup>34</sup>.

### 3 Posições naturalistas não-fisicalistas

#### 3.1 O dualismo naturalista de Chalmers

Existe uma variante exótica de naturalismo que é assumidamente não-fisicalista: o "dualismo naturalista" de Chalmers. Esta variante do naturalismo, embora rejeite o dualismo de substância, compromete-se, no entanto, com um dualismo de propriedades juntamente com a existência de leis psicofísicas irreduzíveis (o que, como vimos, Davison rejeita).

Chalmers resume a sua posição em quatro teses:

**Tese 1.** "A experiência consciente existe" (CHALMERS, 1996, p. 161)

**Tese 2.** "A experiência consciente *não* é logicamente superveniente ao físico" (Ibid. id.). Logo, como vimos acima, essa experiência não seria redutível ao físico.<sup>35</sup> Em princípio, zumbis<sup>36</sup> são concebíveis.

<sup>33</sup> O termo 'ação' é usualmente empregado, em lugar de 'comportamento', quando há causas mentais envolvidas, como no caso aqui considerado.

<sup>34</sup> Godfrey-Smith (comunicação pessoal) propõe que se veja o trabalho filosófico-naturalista como o de coordenar dois tipos de "fatos" (sic.) tomados, em princípio, como irreduzíveis: os relativos às nossas práticas interpretativas, de um lado, e os relativos à nossa constituição neurofisiológica e seus vínculos com o mundo físico, de outro lado. O primeiro tipo de fato é particularmente enfocado pela tradição davidsoniana, sellarsiana e também, de certa forma, pela dennettiana. A proposta de Godfrey-Smith parte do reconhecimento de que falharam, até agora, as ousadas tentativas de reduzir propriedades semânticas a propriedades físicas, como as de Millikan e de Dretske.

<sup>35</sup> Sobre a relação de superveniência lógica, ver notas 9 e 10.

<sup>36</sup> 'Zumbis' são personagens que povoam os mundos imaginários dos filósofos da mente. Os zumbis têm, por definição, comportamentos idênticos a pessoas como nós (ou seja, não podem ser distinguidos de nós com base no seu comportamento unicamente, naquilo a que se pode ter acesso de um ponto de vista de terceira pessoa), mas ao mesmo tempo admite-se que não tenham estados de consciência, como nós. Em outras palavras, os zumbis não têm experiência fenomênica, embora possam ter estados internos, 'mentais', de certo tipo (sem *qualia*), controlando o seu comportamento.

**Tese 3.** “Se existem fenômenos que não são logicamente supervenientes aos fatos físicos, então o materialismo é falso” (CHALMERS, 1996, p. 161)

O ‘dualismo naturalista’ de Chalmers pode ser resumido então nos seguintes termos: a consciência supervém naturalmente (ou seja, nomologicamente) ao físico, mas não supervém nem lógica nem metafisicamente ao físico (CHALMERS, 1996, p. 71).

Um aspecto fundamental do naturalismo (não-materialista) de Chalmers é o pressuposto de que existem leis que vinculam propriedades físicas a propriedades fenomênicas: “a experiência consciente surge do [*arises from*] físico de acordo com algumas leis da natureza, mas não é, ela própria, física” (CHALMERS, 1996, p. 161).

Essas leis psicofísicas<sup>37</sup> assegurariam, para Chalmers, o caráter naturalista da sua posição, por fazê-la compatível, desse modo, com o quadro de mundo que nos traça a ciência contemporânea:

Na visão que eu advogo, a consciência é governada pela lei natural e, eventualmente, pode haver uma teoria científica razoável dela. Não há nenhum princípio *a priori* que afirma que todas as leis naturais serão leis físicas; negar o materialismo não é negar o naturalismo. Um dualismo naturalista expande a nossa visão de mundo, mas ele não invoca as forças do obscurantismo” (CHALMERS, 1996, p. 170, grifo nosso).

É importante notar como se amplia a concepção de ‘natureza’, com a distinção entre lei natural e lei física. Chalmers joga também com a ambiguidade do termo ‘naturalismo’: dependendo da concepção de natureza que se tenha, teremos diferentes ‘naturalismos’.

Chalmers, em seu debate com Searle, contesta o ponto de vista deste último que, como vimos, considera suficiente, para explicar o mental, oferecer uma descrição neurofisiológica completa: “Para explicar por que e como cérebros dão suporte [*support*] à consciência, não seria suficiente contar uma história sobre o cérebro; para vencer o fosso, precisamos adicionar leis-ponte independentes” (SEARLE, 1997, p. 165). [Chalmers classifica a posição de Searle como a de um “materialismo não-redutivo”: “... embora não possa haver

<sup>37</sup> Faço notar que a posição de Chalmers parece ser essencialista, no sentido definido por Popper (1972), porque não admite que leis fundamentais (físicas ou psicofísicas) possam vir a ser explicadas apelando-se para mecanismos ou níveis ontológicos ‘inferiores’. Tais leis são, além disso, atemporais para Chalmers (elas não surgiram ou evoluíram). Ver Chalmers, 1996, p. 170.

uma implicação [entailment] lógica dos fatos físicos para os fatos acerca da consciência e, portanto, nenhuma explicação redutiva da consciência, esta última é somente [just] física..." (1996, p.162; cf. p. 164).

É importante ressaltar que surge uma nova ambigüidade no contexto desse confronto, gora no sentido do termo 'materialismo'. Searle, como vimos, não se considera materialista porque associa o materialismo ao reducionismo. Mas Searle é, para Chalmers, um materialista por rejeitar tanto o dualismo de substância quanto o dualismo de propriedades. Vimos, de fato, que para Searle não existem propriedades fenomênicas ao lado (ou além) de propriedades físicas. A posição de Searle apresenta, para Chalmers, problemas "internos" e colapsa numa das outras alternativas disponíveis para o problema mente-corpo, provavelmente no dualismo de propriedades. Vimos que Kim também aponta para outras inconsistências no naturalismo biológico de Searle.

**Tese 4.** "O domínio físico é fechado causalmente" (CHALMERS, 1996, p.161).

Chalmers vê como um problema da sua posição que haja uma "irrelevância das propriedades fenomênicas para a explicação do comportamento", o que, nesse aspecto, o conduziria a um certo tipo de epifenomenalismo (Ibid. p. 165). Apesar dessa consequência, sua posição teria a vantagem de ser "naturalista" em seu caráter, retirando qualquer "mistério" da consciência, enquanto fenômeno do mundo.

Searle certamente rejeitaria essa proposta de Chalmers, defendendo que propriedades de alto nível, como a consciência, são eficazes causalmente (ver SEARLE, 1997, p. 161).

### 3.2 Sellars e as duas imagens de homem

A posição de Sellars é, em vários aspectos, análoga à de Davidson. Contudo, embora seja possível ver em Sellars um compromisso com um naturalismo metodológico (ver ABRANTES e BENSUSAN, 2003), ele certamente não é um naturalista ontológico. Sellars é cético quanto à possibilidade de se poder integrar as imagens "manifesta" e "científica" de homem (como a proposta pela biologia, pela neurofisiologia, pela física, etc.): "... para completar a imagem científica nós precisamos enriquecê-la, não com



mais modalidades de se dizer o que é o caso, mas com a linguagem da comunidade e das intenções individuais..." (SELLARS, 1963, p. 40).

Com essa proposta, Sellars pretende evitar, de um lado, o dualismo (de tipo cartesiano), de outro uma postura não-realista com respeito às entidades postuladas pela imagem científica de homem.

Nesse trecho, Sellars está se detendo, em particular, no tema do livre arbítrio, nas dimensões ética, dos direitos e deveres, etc. Ele, de fato, acredita que "... a irredutibilidade do que é pessoal é a irredutibilidade do 'dever ser' ao 'ser' " (1963, p. 39). Essa objeção à famigerada falácia naturalista é bastante comum, mas tem sido enfrentada pelos naturalistas. Antes disso, Sellars também apresenta objeções a tentativas de se identificar as sensações a processos neurofisiológicos, antecipando a discussão atual em filosofia da mente em torno da irredutibilidade das propriedades fenomênicas (*qualia*).

### **Naturalismo conceitual-lingüístico (NACON)**

Esta modalidade de naturalismo não tem compromissos ontológicos, como o anterior (NATO), mas envolve teses a respeito dos nossos conceitos (morais, epistêmicos, mentais, etc.) e dos termos da linguagem que utilizamos nesses vários domínios.

Especificamente com respeito à filosofia da mente, NACON corresponde ao tratamento que damos à linguagem mentalista ou intencional, e aos conceitos do mental referidos por essa linguagem.

Stich et al. (1994) propõem uma categorização de estratégias naturalistas em filosofia da mente, especificamente com respeito ao problema da intencionalidade, que me sugeriu algumas idéias para desenvolver a modalidade NACON de naturalismo. Eles distinguem duas estratégias de naturalização de conceitos ou predicados intencionais.

A. Naturalizar predicados ou conceitos intencionais (como 'representa', 'crê', 'quer', etc.) corresponde a fazer uma análise conceitual (ou análise de significado). Ou seja, naturalizar é mostrar que o predicado intencional é coextensivo, em todos os mundos possíveis, a uma conjunção de predicados não-intencionais (ou seja, predicados usados em teorias nas ciências físicas)<sup>38</sup>.

<sup>38</sup> Se I é um predicado intencional (e.g. uma relação como 'representa'), então naturalizar I é propor uma análise do seguinte tipo: "I(a,b) se e somente se N ", onde a condição N está descrita

Essa primeira estratégia de naturalização — explicitada por Fodor, entre outros, nas idas épocas em que grande parte do que se fazia em filosofia era análise conceitual — não atende, contudo, ao tipo de orientação metodológica adotada pelos filósofos naturalistas, como veremos na seção que dedico à modalidade metodológica de naturalismo. É certo que aqui não está em foco a metodologia, mas o naturalismo entendido como uma estratégia para lidar com os conceitos e predicados intencionais, que permitiria, caso tenha sucesso, substituí-los por conceitos e predicados físicos de boa estirpe. Isso corresponde bastante bem ao que Searle, ao elencar os vários tipos de redução, chama de "redução lógica ou definicional" (SEARLE, 1992, p. 115; ver seção II-3 acima)<sup>39</sup>. Ele sugere que há uma expectativa de que o sucesso de tais reduções abriria caminho para reduções propriamente ontológicas, mas grande parte dos filósofos ditos 'analíticos' pretendiam, justamente, evitar comprometimentos ontológicos, restringindo com segurança o trabalho filosófico à análise da linguagem, tendo como referencial as nossas intuições. De toda forma, indo de encontro a Stich et al. (1994), é duvidoso que tal estratégia possa ser considerada naturalista, em sentido próprio.

B. Naturalizar um predicado ou conceito intencional é descobrir as propriedades essenciais a que ele se refere. O conceito intencional é entendido, neste caso, como um termo de espécie natural [*natural kind term*]. Do mesmo modo como descobrimos que 'água  $\equiv$  H<sub>2</sub>O', naturalizar um conceito intencional seria descobrir a sua essência (ou as condições de sua aplicação a algo no mundo)<sup>40</sup>.

---

num vocabulário não-intencional. Notar que o bicondicional é de verdade necessário, já que verdadeiro em todos os mundos possíveis. Stich et al. (1994) apresentam a análise conceitual (também conhecida como análise de significado) como sendo *a priori*. Trata-se, portanto, de um bicondicional necessário *a priori*. Esse caráter *a priori* da análise filosófica contradiz a tese (a), explicitada na p. 1 do presente artigo, onde tento caracterizar o naturalismo em epistemologia, uma indicação de que essa estratégia não seria, a rigor, aceita pelos naturalistas.

<sup>39</sup> Stich et al. (1994) também indicam, pertinentemente, que esse tipo de análise pressupõe uma teoria clássica de conceitos. A respeito de teorias alternativas de conceitos, ver Oliveira (1995).

<sup>40</sup> Isso corresponde a escrever o bicondicional "I(a,b) se e somente se N", onde a condição N está, como no caso anterior, escrita num vocabulário não-intencional. Porém, este bicondicional possui agora um caráter distinto: afirma uma descoberta científica, *a posteriori* mas também necessária, se aceitamos os argumentos de Kripke. Portanto, uma alternativa de naturalização de conceitos (da epistemologia, da filosofia da mente, etc.) seria vê-los como termos de espécies naturais, referindo-se a propriedades essenciais, físicas ou biológicas. Caberia às ciências descobrir a que propriedades tais termos se referem. Em outras palavras, a tarefa de chegar a tais bicondicionais

Nesta segunda estratégia, a tarefa de naturalizar um predicado não pode ser realizada *a priori*, como no caso anterior, apelando-se para nossas intuições de senso comum, mas seria uma descoberta *a posteriori*, realizada propriamente no âmbito da atividade científica. Os autores levantam dúvidas a respeito do sucesso de qualquer uma dessas estratégias de naturalização, mas não é minha intenção neste artigo fazer uma avaliação das tentativas de levá-las a cabo.

Stich et al. (1994) também mencionam, após as duas estratégias acima apresentadas, uma terceira segundo a qual naturalizar é mostrar que as propriedades intencionais são idênticas ou supervêm a propriedades não-intencionais. Essa estratégia naturalista enquadra-se, perfeitamente, na modalidade de naturalismo ontológico (NATO) que discuti na última seção<sup>41</sup>. Portanto, o que me interessa para caracterizar modalidades de NACON são somente as duas primeiras estratégias.

### **Materialismo eliminativo**

Referi-me, anteriormente, ao materialismo eliminativo, uma posição em filosofia da psicologia (e não propriamente em filosofia da mente) pelo fato de discutir o *status* da psicologia intencional, da psicologia de senso comum (*folk psychology*). Churchland, o principal articulador dessa posição, defende que essa psicologia é uma teoria falsa e que, portanto, não faria sentido tentar reduzi-la a uma teoria neurofisiológica (uma redução teórica). A situação seria análoga a de certas teorias da física — como a teoria do flogisto, a teoria do calórico, as teorias do éter, etc. — que são falsas e, por isso, seus termos teóricos não se referem a nada no mundo. Não faria sentido, portanto, reduzir tais teorias, mas sim eliminá-las. Com respeito à psicologia de senso comum, Churchland defende que devemos, do mesmo modo, eliminá-la e, claro, com ela todos os seus termos teóricos, como os de atitudes proposicionais que, numa perspectiva realista, são equivocadamente associadas a estados nas mentes (ou nos cérebros) dos agentes aos quais as imputamos.

---

envolveria de forma crucial a pesquisa científica, não se limitando a um trabalho filosófico-analítico, como na estratégia anterior.

<sup>41</sup> Notar que as estratégias (A) e (B) tratam de predicados e conceitos, enquanto a terceira estratégia trata de propriedades, configurando um claro comprometimento ontológico.



Podemos entender o materialismo eliminativo a partir das duas estratégias descritas por Stich et al. (1994). O fracasso de ambas as estratégias, aplicadas ao caso da linguagem intencional (a linguagem teórica da psicologia de senso comum), levaria à sua eliminação. Não caberia fazer uma análise conceitual (estratégia A) de conceitos associados às atitudes proposicionais, particularmente o de intencionalidade. Outra maneira de se entender o materialismo eliminativo seria a de ver os termos e conceitos da psicologia de senso comum como não se referindo a espécies naturais. Uma das implicações disso é a irrealidade do mental. Churchland pressupõe, claro, que os conceitos e predicados da neurofisiologia referem-se a espécies naturais e, conseqüentemente, podem figurar em leis, suas ocorrências tendo poder causal.

### **Naturalismo metodológico (NAME)**

O naturalismo metodológico afirma a continuidade entre a investigação filosófica e a científica; em particular, defende que não há uma distinção nítida entre os métodos empregados em cada um desses domínios de investigação.

Em epistemologia, NAME é a versão quineana do naturalismo. Segundo essa variante do naturalismo epistemológico, “a epistemologia deve ser uma ciência empírica ou deve, pelo menos, ser informada e devedora [*beholden*] dos resultados de disciplinas científicas” (GOLDMAN, 1998, p. 305). Para Quine (1987a, 1987b), essas disciplinas incluem, pelo menos, a psicologia e a biologia evolutiva.

De modo análogo, o naturalismo metodológico em filosofia da mente consideraria esta área como situando-se propriamente no domínio das ciências (e.g. as ciências cognitivas) adotando, como consequência, os métodos científicos empregados nas mesmas, e apoiando-se no conhecimento produzido por elas. Esta seria a maneira correta de enfrentar problemas fundamentais como o problema mente-corpo, o problema de outras mentes, etc.

Encontramos uma outra formulação do naturalismo em Guttenplan, que pode ser traduzida em termos metodológicos: “O naturalismo com respeito a algum domínio é a visão de que tudo o que existe naquele domínio, e todos aqueles eventos que ocorrem nele, são características empiricamente acessíveis do mundo” (1995, p. 449).

Se algo (uma entidade, um processo) é “empiricamente acessível”, então pode ser descrito de um ponto de vista de terceira-pessoa. Vimos que Searle, certamente, não é naturalista nesse sentido (1992, p. 71, 73), defendendo, ao contrário, que, com respeito à consciência, o único ponto de vista admissível é o de primeira-pessoa — embora com respeito ao comportamento, obviamente, possamos adotar um ponto de vista de terceira-pessoa. Tais posições epistemológicas têm diversas implicações metodológicas.

Um exemplo de naturalismo metodológico com respeito ao mental é o “método heterofenomenológico” proposto por Dennett, que pressupõe explicitamente o ponto de vista de terceira-pessoa, em vez do ponto de vista de “primeira-pessoa plural”, comumente adotado nas discussões sobre a fenomenologia do mental (DENNETT, 1991, p. 70, 96).

No método heterofenomenológico, os relatos de vários sujeitos são tomados como ficções, como descrições de mundos heterofenomenológicos (mundos povoados de ‘objetos intencionais’). Com base nessas ficções, nesses fatos, que são categorizados pelo teórico, este se pergunta se os objetos intencionais mencionados nesses relatos correspondem a objetos, eventos, processos, etc. reais no cérebro (ou na mente). Esta é, para Dennett, uma “investigação empírica” (1991, p. 98).

Ele descreve o método heterofenomenológico nos seguintes termos:

[...] aqui temos um caminho *neutro* indo da ciência física objetiva, e de sua insistência no ponto de vista de terceira-pessoa, a um método de descrição fenomenológica que pode (em princípio) fazer justiça às experiências subjetivas mais privadas e inefáveis, ao mesmo tempo que nunca abandona os escrúpulos metodológicos da ciência” (DENNETT, 1991, p. 72)<sup>42</sup>.

Mais recentemente, Dennett reforça esse naturalismo que caracteriza a sua investigação filosófica:

[...] desde o início, eu trabalhei a partir do ‘ponto de vista de terceira-pessoa’ adotado pela ciência, e considerei minha tarefa a de construir uma estrutura física — ou melhor, a de esboçar as suas linhas gerais —

---

<sup>42</sup> Em “What it is like to be a bat”, título do famoso artigo de T. Nagel (1991) que poderíamos traduzir por “Como é ser um morcego”, esse renomado filósofo da mente, embora radicalmente contrário ao fisicalismo, considera possível avançarmos em direção a uma “fenomenologia objetiva”, com o desenvolvimento de novos conceitos com essa finalidade, embora esse empreendimento não possa esgotar o que está envolvido na experiência, que teria uma irredutível componente subjetiva. Cf. Dennett (1991, p. 71).

que poderia ser vista como realizando o enigmático *legerdemain* da mente" (DENNETT, 1995, p. 236-237).

Dennett adota a postura de 'engenheiro', vendo a ciência cognitiva comprometida com algo como a engenharia reversa (DENNETT, 1995, p. 242; ver também Dennett, 1991). Como Quine, Dennett defende um intercâmbio entre a filosofia e as ciências:

Minha insistência sobre a necessidade dos filósofos, antes de qualquer desenvolvimento [*holding forth*], abastecerem-se [*to stoke up*] na ciência relevante, e a recusa de conduzir minhas investigações pelo método tradicional de definição e argumento formal, tornou-me um filósofo da mente especialmente impuro (1995, p. 242).

Nessa passagem, Dennett é explícito em recusar a análise conceitual (que descrevi na parte referente à modalidade NACON de naturalismo) como sendo o método característico da filosofia.

### **Posições contrárias ao naturalismo metodológico**

Em vários momentos deste artigo, indiquei que posições em ontologia, em epistemologia e em metodologia restringem-se mutuamente. Pode-se esperar, por exemplo, que os filósofos que adotam uma postura claramente contrária ao fisicalismo em filosofia da mente também abracem posições metodológicas antinaturalistas. Um exemplo disso é o modo como Searle e Dennett se confrontam no plano metodológico.

Searle, como vimos, considera-se um 'naturalista biológico', mas ele seguramente não é naturalista no plano metodológico, apontando como uma das causas dos erros filosóficos a "aplicação dos métodos da ciência em áreas para as quais eles não são apropriados" (SEARLE, 2000, p. 18). Uma dessas áreas é, justamente, a dos fenômenos mentais.

Na sua crítica ao materialismo, Searle aponta as limitações de uma metodologia que adota um ponto de vista de terceira-pessoa quando aplicada ao domínio do mental. Um exemplo disso seria a ciência cognitiva, que estaria segundo ele cometendo o mesmo erro do behaviorismo, qual seja, o de "estudar somente fenômenos objetivamente observáveis" (SEARLE, 1992, p. xii; cf. Guttenplan, acima). Searle enfatiza, ao contrário, a interdependência entre ontologia e metodologia: "Já que é um erro supor que a ontologia do mental é

objetiva, é um erro supor que a metodologia da ciência da mente deve preocupar-se somente com o comportamento objetivamente observável" (SEARLE, 1992, p. 20).

Além de afirmar a interdependência entre a epistemologia, a metodologia e a ontologia, Searle defende que esta última possui precedência com respeito às demais dimensões:

"A epistemologia para o estudo do mental não determina mais a sua ontologia do que a epistemologia de qualquer outra disciplina determina a sua ontologia. Ao contrário, no estudo do mental, como em qualquer outro domínio, a questão central da epistemologia é a de capturar *[to get at]* a ontologia preexistente" (SEARLE, 1992, p. 23).

Percebe-se uma tensão entre esta postura e, ao mesmo tempo, a defesa de uma interrelação estreita da filosofia com as ciências. Searle afirma, efetivamente, que os objetivos da ciência e da filosofia são os mesmos: construir uma teoria "ao mesmo tempo verdadeira, explicativa e geral" dos fenômenos do mundo. A ênfase na generalidade não é gratuita, pois vê nisso uma das características distintivas do trabalho filosófico. Essa teoria deve ser, além disso, adequada aos fatos. Os "fatos brutos" revelados pelos avanços no conhecimento científico, como por exemplo pela neurofisiologia, podem exigir o abandono de "compromissos filosóficos", o que constituiu o ponto de partida do seu "naturalismo biológico".

Essa posição certamente não é simpática para os que reivindicam uma total autonomia, ou mesmo uma precedência, da filosofia com respeito às ciências. Entretanto, os naturalistas mais convictos tampouco ficam satisfeitos com a tentativa de Searle de restringir a tarefa da filosofia, que seria a de mostrar, quando muito, como, por exemplo, uma explicação neurofisiológica da consciência é "possível" (SEARLE, 1992, 55, 92), aguardando que a ciência nos explique como estados neurofisiológicos causam "realmente" estados mentais conscientes. Embora Searle defenda uma mútua fertilização da filosofia pelas ciências, fica-se com a impressão de que ele simplesmente coloca os problemas, em vez de fazer um esforço sério de solucioná-los, deixando para os cientistas, digamos, o trabalho pesado.

Para Searle, a investigação filosófica caracteriza-se, sobretudo, por fazer uma análise crítica dos pressupostos, do "pano de fundo" (*background*) composto por "posições-padrão" (*default positions*) que precedem a reflexão,



distinguindo as falsas das verdadeiras e, eventualmente, rejeitando as categorias mesmas nas quais estão formuladas. As soluções tradicionais para o problema mente-corpo, por exemplo, estariam marcadas por erros conceituais tão flagrantes que Searle surpreende-se que os seus pares, mesmo os mais ilustres, não o percebam.

É curioso, nesse sentido, o lugar que ocupa o senso comum nas estratégias argumentativas searleanas. Apesar da ênfase que dá à tarefa crítica da filosofia, Searle defende, surpreendentemente, as "posições-padrão" associadas ao senso comum (como a posição realista, a que reconhece a existência da consciência, etc.) — que considera em sua maioria verdadeiras — contra as posições defendidas por muitos filósofos, o que avalia como "um fato triste" da sua "profissão".

McGinn ocupa, nesse registro antinaturalista, uma posição *sui generis*. Segundo ele, tanto reducionistas e eliminativistas, como Churchland, quanto antireducionistas como Searle, evidenciam a nossa incapacidade para lidar com o problema mente-corpo, para oferecer uma visão unificadora do lugar que mentes ocupam no mundo físico. Há, para ele, mistérios insondáveis nesse campo. McGinn admite que essa visão unificadora possa existir numa esfera platônica de idéias, mas ela permanecerá, contudo, inacessível a seres com as nossas limitações cognitivas:

"[...] Não deveria nos surpreender a descoberta de que nem todo aspecto do mundo natural seja acessível aos nossos poderes de compreensão. Nós não esperamos que outras espécies, que resultaram da evolução, sejam oniscientes; portanto, por que assumir que a nossa inteligência tenha evoluído com a capacidade de resolver qualquer problema que possa ser colocado a respeito do universo, do qual somos uma parte tão pequena e contingente? [...] Uma teoria do vínculo psicofísico pode existir em algum lugar no céu Platônico; só ocorre que as nossas mentes estão a milhas de distância de capturar em que consiste essa teoria. Portanto, devemos estar preparados a nos mover erratically [*to flail around*] na ignorância, indo de um extremo implausível a outro" (MCGINN, 1999).

Esse pessimismo epistemológico é, sem dúvida, contrário ao otimismo que usualmente caracteriza os naturalistas. O fato é que nunca teremos certeza se a nossa incompreensão do problema mente-corpo (ou de qualquer outro problema) revela uma limitação cognitiva insuperável, ou se simplesmente as dificuldades estão a exigir mais esforço, persistência na investigação ou mesmo

mudanças radicais em nossas concepções. Posições 'misteristas' como a de McGinn parecem-me obscurantistas, por nos fazerem parar prematuramente, se é que algum pressuposto epistemológico ou ontológico deva nos fazer parar...

### **Outras modalidades de naturalismo**

Goldman menciona uma outra modalidade de naturalismo em epistemologia, além das que tentei aqui adaptar para distinguir posições em filosofia da mente: o "naturalismo substantivo" (GOLDMAN, 1998, p. 113-117).

Ele inclui, nessa modalidade de naturalismo, a busca de explicações causais, informacionais, contrafactuais ou confiabilistas (*reliabilist*) do conhecimento ou da justificação. Por exemplo, seriam naturalistas aquelas epistemologias que vêem o sujeito epistêmico em interação com o meio ambiente físico, e o conhecimento como efeito dessa interação<sup>43</sup>. Tais posições estão usualmente associadas ao externalismo em epistemologia, e talvez possam ser aproximadas do externalismo em filosofia da mente. As epistemologias evolutivas poderiam, também, enquadrar-se nessa variante de naturalismo substantivo, já que propõem explicações do conhecimento com base no processo darwinista de seleção natural (ver ABRANTES, no prelo).

Acredito que possam ser incluídas nessa modalidade de naturalismo substantivo tentativas de aproximar a filosofia da mente das ciências cognitivas, em particular das neurociências, na busca de explicações causais para estados e processos cognitivos. Também poderiam ser enquadradas como 'naturalistas substantivas' as tentativas de se adotar categorias emprestadas à biologia para abordar problemas tradicionais em filosofia da linguagem e filosofia da mente, relativos às noções de significado, referência e intencionalidade. Tenho em mente, aqui, o trabalho de Edelman, embora este autor não se considere um filósofo, mas acredite que a biologia possa dar uma "nova direção" à filosofia (1992, p. 159). De toda forma, as fronteiras entre ciência e filosofia são, nesta modalidade de naturalismo, ainda menos nítidas. O trabalho de Millikan (1995) situa-se, indiscutivelmente, na confluência de vários programas em filosofia, e

---

<sup>43</sup> Exemplos incluem Paul Churchland e a sua tentativa de nos ver como "máquinas epistêmicas" (1979, p. 125), bem como a teoria informacional do conhecimento de Dretske (1981). Traduzo aqui 'epistemic engine' por 'máquina epistêmica'.



na apresentação ao livro desta filósofa, Dennett é particularmente enfático a respeito da sua orientação naturalista:

[...] Já que nós, seres humanos, somos somente uma parte da natureza — supremamente complicados, mas porções não privilegiadas da biosfera — explicações filosóficas de nossas mentes, conhecimento e linguagem devem, ao fim e ao cabo, estar em continuidade com as ciências naturais e em harmonia com estas” (apud MILLIKAN, 1995. p. ix).

Também poderiam ser considerados 'naturalistas substantivos' os programas que se apóiam na biologia evolutiva para reconstruir a filogenia de diferentes tipos de sistemas cognitivos (ou de tipos de mentes)<sup>44</sup>. Por razão de espaço, e também porque essas abordagens 'substantivas' distanciam-se do que é usual se fazer em filosofia da mente, limito-me, neste artigo, a essas poucas referências a programas nessa modalidade de naturalismo.

## **Conclusão**

Com base nas discussões anteriores proponho, então, a seguinte classificação de modalidades de naturalismo em filosofia da mente, acompanhados de alguns exemplos de propostas que se enquadrariam em cada uma delas (embora em algumas modalidades não tenha conseguido encontrar bons exemplos na literatura pertinente):

### Naturalismo ontológico

- Teoria da identidade
- Eliminativismo
- Fisicalismo mínimo
- Naturalismo biológico<sup>45</sup>
- Naturalismo evolutivo<sup>46</sup>

<sup>44</sup> Estou engajado, atualmente, num projeto de pesquisa voltado para o tema da evolução de diferentes tipos de sistemas cognitivos, incluindo sistemas intencionais como nós. Esse projeto pode ser visto como tendo uma orientação 'naturalista substantiva', inspirada nos trabalhos de Godfrey-Smith (1998, 2002) e de Sterelny (2003), entre outros. Eles classificam a sua orientação como um "naturalismo evolutivo".

<sup>45</sup> Refiro-me ao uso que faz Searle desta expressão (ver seção II-3). Em que medida Searle é, de fato, um naturalista ontológico conseqüente, é algo passível de discussão, como mostrei naquela seção. Ver nota 47.

<sup>46</sup> Uso essa expressão no sentido em que a empregam Godfrey-Smith e Sterelny (ver nota 44).

- Dualismo naturalista

Posições (ontológicas) não-naturalistas incluiriam o dualismo de substância de Descartes e, talvez, o emergentismo<sup>47</sup>, entre outras.

Naturalismo conceitual-lingüístico (NACON)

Naturalismo metodológico (NAME)

- O método heterofenomenológico

Em metodologia, encontramos atitudes não-naturalistas em autores como Searle, McGinn e, talvez, T. Nagel.

Naturalismo substantivo

- Aplicações das ciências cognitivas e da biologia ao problema do surgimento e da inserção de mentes no mundo físico.

Algumas orientações em filosofia da mente enquadram-se dificilmente nessas modalidades, como o eliminativismo de Churchland, o monismo anômalo de Davidson e, mesmo, o naturalismo biológico de Searle, para dar alguns exemplos. Isso não deve surpreender em qualquer tentativa de classificação, sobretudo de programas em filosofia.

A despeito da grande diversidade de tradições que poderiam ser classificadas de naturalistas em filosofia da mente, acredito que essa classificação possa servir, nessa área, como um "guia", à semelhança do que fez Goldman (1998) em epistemologia. Esse guia de 'naturalismos' em filosofia da mente pode também contribuir, por contraste, para uma melhor delimitação de posturas antinaturalistas nessa área<sup>48</sup>.

---

<sup>47</sup> Penso aqui na tese de Searle de que a consciência é um fenômeno emergente. Como esta tese está associada à possibilidade de uma causalidade descendente — o que entra em conflito com o fisicalismo — o naturalismo biológico de Searle também se enquadra mal na categoria de 'naturalismo ontológico'. Não excluo a possibilidade, contudo, de se articular um emergentismo que possa ser, consistentemente, incluído nessa categoria de naturalismo.

<sup>48</sup> Agradeço ao CNPq por uma bolsa que possibilitou o desenvolvimento desta pesquisa. Também sou grato a um parecerista anônimo pelos comentários, que me permitiram identificar problemas numa versão anterior deste texto.

## Referências

- ABRANTES, P. Naturalizando a epistemologia. In: \_\_\_\_\_ (Org.) *Epistemologia e cognição*. Brasília: Editora da UnB, 1995, p. 171-218.
- \_\_\_\_\_. Resenha de: KIM, J. *Philosophy of Mind, Principia*, Florianópolis, v.1, n.2, 1997.
- \_\_\_\_\_. Naturalismo epistemológico: apresentação. *Cadernos de História e Filosofia da Ciência*, Campinas, v. 8, n. 2, 1998, p. 7-26.
- \_\_\_\_\_. Resenha de: SEARLE, J. Mente, linguagem e sociedade: filosofia no mundo real'. *Folha de São Paulo*, 13 jan. 2001, Caderno de Resenhas, p. 8.
- \_\_\_\_\_. Resenha de: NAGEL, T. A última palavra. *Folha de São Paulo*, 10 nov. 2001. Caderno de Resenhas.
- \_\_\_\_\_. Metafísica e ciência: o caso da filosofia da mente. In: CHEDIAK, K. e VIDEIRA, A. A. P. (Org.) *Temas de Filosofia da Natureza*. Rio de Janeiro: UERJ, 2004, p. 210-239.
- \_\_\_\_\_. O programa de uma epistemologia evolutiva. *Revista de Filosofia da PUCPR*, Curitiba, no prelo.
- ABRANTES, P.; AMARAL, F. Funcionalismo e causação mental. *Manuscrito*, Campinas, v. 25, p. 13- 45, 2002. Número especial: "Mental Causation".
- ABRANTES, P.; BENSUSAN, H. Conhecimento, ciência e natureza: cartas sobre o naturalismo. In: SIMON-RODRIGUES, S. (Ed.) *Filosofia e conhecimento: das formas platônicas ao naturalismo*. Brasília: Editora da UnB, 2003, p. 273-333.
- CHALMERS, D. *The conscious mind: in search of a fundamental theory*. Oxford: Oxford UP, 1996.
- \_\_\_\_\_. Facing up to the problem of consciousness. In: SHEAR, J. (Ed.) *Explaining Consciousness — the 'hard problem'*. Cambridge MA: The MIT Press, 1997.
- CHURCHLAND, P. *Matter and consciousness*. Cambridge MA: MIT Press, 1990.
- DAVIDSON, D. Mental events. In: ROSENTHAL D. (Ed.) *The nature of mind*. New York: Oxford UP, 1991, p. 247-56.
- DENNETT, D. *Consciousness explained*. Boston: Little, Brown and Company, 1991.

\_\_\_\_\_. Review of Searle, the rediscovery of the mind. *Journal of Philosophy*, v. 60, n. 4, p. 193-205, 1993.

\_\_\_\_\_. *Kinds of mind*. New York: Basic Books, 1996. Edição brasileira: *Tipos de Mente*. RJ: Rocco, 1997.

\_\_\_\_\_. *Darwin's dangerous idea*. Nova York: Simon & Schuster, 1995. Edição brasileira publicada pela Rocco.

\_\_\_\_\_. Dennett, D. C. In: GUTTENPLAN, S. (Ed.) *A companion to the Philosophy of Mind*. Oxford: Blackwell, 1995, p. 236-244.

DRETSKE, F. *Naturalizing the mind*. Cambridge MA: The MIT Press, 1997.

EDELMAN, G. *Bright air, brilliant fire: on the matter of the mind*. Londres: Penguin, 1992.

FRENCH, P. et al. *Philosophical naturalism*. (Midwest Studies in Philosophy, vol. 19). Notre Dame Indiana: University of Notre Dame Press, 1994.

GODFREY-SMITH, P. *Complexity and the function of mind in nature*. Cambridge MA: Cambridge UP, 1998.

\_\_\_\_\_. On the evolution of representational and interpretive capacities. *The Monist*. v. 85. n. 4, p. 50-69, 2002.

\_\_\_\_\_. On Folk Psychology and mental representation (Comunicação pessoal).

GOLDMAN, A. Epistemologia naturalista e confiabilismo. In: ÉVORA, F.; ABRANTES, P. (Ed.) *Cadernos de História e Filosofia da Ciência*. Campinas, série 3, v. 8, n. 2, p. 109-145, 1998. Publicado originalmente em 1994 com o título 'Naturalistic epistemology and reliabilism'.

GUTTENPLAN, S. (Ed.) *A companion to the Philosophy of Mind*. Oxford: Blackwell, 1995.

HOOKE, C. *Reason, regulation and realism*. Albany: SUNY University Press, 1995.

HUMPHREY, N. *A history of the mind*. Londres: Harper Collins, 1993. Edição brasileira publicada pela Campus.

HUMPHREYS, P. Emergence, not supervenience. *Philosophy of Science*, v. 64, n. 4, 1997a. PSA96. Suplemento.

\_\_\_\_\_. How properties emerge. *Philosophy of Science*, v. 64, n.1, 1997b.

KIM, J. *Philosophy of Mind*. Colorado: Westview Press, 1996.

\_\_\_\_\_. *Mind in a physical world*. Cambridge MA: The MIT Press, 1998.

KIM, J.; SOSA, E. *A companion to metaphysics*. Oxford: Blackwell, 1995.

KITCHER, P. The naturalists return. *The Philosophical Review*, v. 101, n.1, p. 53-114, 1992. Traduzido para o português In: ÉVORA, F.; ABRANTES, P. (Eds.) *Cadernos de História e Filosofia da Ciência*. Campinas, série 3, v. 8, n. 2, p. 27-108, 1998. Publicado originalmente em 1994 com o título 'Naturalism: Metaphysical and Epistemological'.

KORNBLITH, H. (Ed.) *Naturalizing Epistemology*. Cambridge MA: The MIT Press, 1987.

\_\_\_\_\_. Naturalismo: Metafísico e Epistemológico. In: ÉVORA, F.; ABRANTES, P. (Eds.) *Cadernos de História e Filosofia da Ciência*. Campinas, série 3, v. 8, n. 2, p. 147-169, 1998. Publicado originalmente em 1994 com o título 'Naturalism: Metaphysical and Epistemological'.

McGINN, C. *The problem of consciousness*. Oxford: Blackwell, 1993.

\_\_\_\_\_. Can we ever understand consciousness? *New York Review of Books Archives*, 10/06/1999.

Millikan, R. *Language, thought and other biological categories*. Cambridge MA: The MIT Press, 1995.

NAGEL, T. What it is like to be a bat? In: ROSENTHAL D. (Ed.) *The nature of mind*. New York: Oxford UP, 1991, p. 422-8.

OLIVEIRA, M.B. Rumo a uma teoria dialética de conceitos. In: ABRANTES, P. (Org.) *Epistemologia e cognição*. Brasília: Ed. da UnB, 1995, p. 25-69.

PAPINEAU, D. *Philosophical naturalism*. Oxford: Blackwell, 1993.

POPPER, K. Três pontos de vista sobre o conhecimento humano. In: Popper, K. *Conjecturas e refutações*. Brasília: Ed. da UnB, 1972, p. 125-146.

QUINE, W.V.O. Two dogmas of empiricism. In: QUINE, W.V.O. *From a logical point of view*. New York: Harper and Row, 1963. A primeira edição é de 1953, pela Harvard University Press.

\_\_\_\_\_. Epistemology naturalized. In: KORNBLITH, H. *Naturalizing Epistemology*. Cambridge MA: The MIT Press, 1987a. Tradução em Português: Coleção "Os Pensadores".

\_\_\_\_\_. Natural Kinds. In: KORNBLITH, H. *Naturalizing Epistemology*. Cambridge MA: The MIT Press, 1987b. Tradução em Português na Coleção "Os Pensadores".

\_\_\_\_\_. Two dogmas in retrospect. *Canadian Journal of Philosophy*, v. 21, n.3, p. 265-274, 1991.



ROSENBERG, A. A field guide to recent species of naturalism. *British Journal for the Philosophy of Science*, v. 47, p. 1-29, 1996.

ROSENTHAL, D. (Ed.) *The nature of mind*. New York: Oxford UP, 1991.

SEARLE, J. *Mente, cérebro e ciência*. Lisboa: Edições 70, 1984.

\_\_\_\_\_. *The rediscovery of the mind*. Cambridge MA: MIT Press, 1992.  
Tradução em Português: *A redescoberta da mente* pela Martins Fontes.

\_\_\_\_\_. *The mystery of consciousness*. New York: The New York Review of Books, 1997. Edição brasileira: *O mistério da consciência*. Paz e Terra.

\_\_\_\_\_. *Mente, linguagem e sociedade*. Rio de Janeiro: Rocco, 2000.

SELLARS, W. Empiricism and the philosophy of mind. In: SELLARS, W. *Science, Perception and Reality*. Londres: Routledge & Kegan Paul, 1963, p. 127-196.

SOBER, E. Psychologism. *J. Theory Soc. Behavior*, v.8, n.2, p. 165-191, 1978.

STERELNY, K. *Thought in a hostile world*. Malden MA: Brackwell Publishing, 2003.

STICH, S. *The fragmentation of reason*. Cambridge MA: The MIT Press, 1993.

STICH, S.; LAURENCE, S. Intentionality and Naturalism. *Midwest Studies in Philosophy*, v. 19, p. 159-182, 1994. Philosophical Naturalism.

STICH, S.; NISBETT, R.E. Justification and the psychology of human reasoning. *Philosophy of Science*, v. 47, p. 188-202, 1980.



# O reducionismo impiedoso de John Bickle e seus problemas<sup>1</sup>

Paulo Roberto Margutti Pinto<sup>2</sup>

## Observações Iniciais

Recentemente, John Bickle publicou um estimulante livro, intitulado *Philosophy and Neuroscience – A Ruthlessly Reductive Account* (2003). Ali, este autor faz uma defesa enfática da redução dos conceitos da psicologia tradicional aos da neurociência de ponta. Fomos convidados pelos organizadores do V Encontro Brasileiro-Internacional de Ciência Cognitiva para apresentar uma avaliação do livro de Bickle. O presente texto traz os resultados desse trabalho.

Para realizar a tarefa de maneira adequada, fazemos, na seção II, uma exposição dos principais aspectos do reducionismo de Bickle. Com base nisso, efetuamos, na seção III, uma discussão de seis pontos que consideramos importantes para a realização de uma avaliação apropriada da proposta de Bickle. As principais conclusões do debate se encontram listadas na seção IV. Embora Bickle insista muito na apresentação dos detalhes técnicos ligados aos resultados obtidos pela neurociência de ponta, alegando que é o desconhecimento dos mesmos que tem motivado o anti-reducionismo da maioria dos pensadores contemporâneos, mantivemos a discussão num plano mais abstrato. A razão disso está no cerne do próprio argumento que levantamos contra Bickle: mesmo se conseguirmos uma explicação neurocientífica satisfatória em termos de puras seqüências de configurações moleculares, ainda assim estaremos muito longe de realizar uma "redução impiedosa". Por esse motivo, acreditamos que a discussão que se segue constitui uma avaliação justa da proposta de Bickle, sem envolver uma traição ao seu espírito.

---

<sup>1</sup> Este texto constitui uma versão modificada do trabalho apresentado no V EBICC, incluindo as respostas a algumas das intervenções feitas.

<sup>2</sup> Departamento de Filosofia, UFMG.

## O programa da nova onda reducionista

Em seu livro, John Bickle afirma que a posição anti-reducionista da maioria dos filósofos decorre da má compreensão do que vem a ser o reducionismo científico na prática atual (2003, xiv). Ele argumenta que também busca a "visão sinótica" da neurociência que subjaz implicitamente no seu cerne de principal corrente molecular e celular (p. xv). Em um sentido, Bickle deixa as coisas dependerem inteiramente da ciência, mas, em outro, tenta reunir as partes fornecidas por ela para explicitar a "figura mais ampla" que a maior parte dos cientistas deixa implícita. E isso é "filosofia" o bastante para Bickle (p. xv).

No primeiro capítulo, ele afirma que seu livro é sobre neurociência contemporânea. Os filósofos neurocientificamente astutos se concentraram na "neurociência cognitiva", de caráter interdisciplinar. Mas devemos pesquisar em outros ramos da neurociência atual para buscar conseqüências e implicações filosóficas, porque a corrente principal-dessa área está em outro lugar, está na fisiologia celular e na biologia molecular (p. 2). O propósito do livro é revelar os detalhes científicos de algumas das ligações mente-moléculas já realizadas e avaliar o potencial explicativo desse "reducionismo impiedoso" (p. 5).

O conceito de *redução* utilizado por Bickle baseia-se no modelo geral de Hooker. De acordo com este último, a redução interteorética envolve um procedimento dedutivo, no qual a teoria redutora ou de base (Tb) funciona como uma das premissas. Mas a conclusão da derivação não é a teoria a ser reduzida (Tr), e sim uma *imagem* (Ib), especificada no interior do quadro conceitual e do vocabulário da teoria redutora Tb. O esquema da redução proposta por Hooker é o seguinte:

|                         |                      |                      |
|-------------------------|----------------------|----------------------|
| Teoria de base (Tb)     |                      |                      |
| associada a             |                      |                      |
| suposições limitadoras, |                      |                      |
| permite derivar         |                      |                      |
| Imagem (Ib) =>          | Que é "análoga" à => | Teoria reduzida (Tr) |
| Teoremas de Ib:         |                      | Teoremas de Tr:      |
| (x)(Ax => Bx)           | $\Leftrightarrow$    | (x)(Jx => Kx)        |
| (x)((Bx & Cx) => Dx)    | $\Leftrightarrow$    | (x)((Kx & Lx) => Mx) |

O quadro acima ilustra a "relação análoga" entre Ib e Tr. Essa relação é tal que um conjunto de teoremas de Ib, que constitui um conjunto restrito de teoremas de Tb, como, por exemplo, os representados do lado esquerdo do quadro acima, é relevantemente isomórfico a ou "análogo a" um conjunto de teoremas de Tr, como, por exemplo, os representados do lado direito do mesmo quadro. Nessa perspectiva, as *suposições limitadoras* (também chamadas *condições de fronteira*) indicadas no esquema têm o objetivo de restringir a aplicabilidade das generalizações de Tb de modo a isolar as possíveis falsidades em Tr. Suponhamos, por exemplo, a redução da física galileana à mecânica newtoniana. Os princípios desta última poderiam ser associados à suposição contrafactual que limita a aplicabilidade das leis newtonianas a corpos móveis que caem por distâncias muito pequenas. A partir desse complexo redutor, formado pela mecânica newtoniana (Tb) mais o contrafactual (suposição limitadora), seria possível deduzir uma imagem (Ib) capaz de emular o alcance explicativo da física galileana (Tr). Como se pode ver, o modelo não precisa recorrer aos problemáticos *princípios-ponte* ou a *regras de correspondência* ligando Tb a Tr. A correlação é feita entre a imagem Ib e a teoria reduzida Tr. Isso significa que não há vocabulários diferentes a serem atravessados no percurso que vai das premissas (Tb mais suposições limitadoras) à conclusão (Ib). Assim, Tr nunca é deduzida a partir de Tb, mas constitui o alvo de um complexo procedimento de mímica através de Ib, que, ela sim, é deduzida da conjunção de Tb com as suposições limitadoras. A imagem derivada Ib, que mimetiza o alcance explanatório de Tr no domínio de aplicação desta última, é fortemente análoga na estrutura lógica a Tr e sua derivação a partir de Tb exige poucas suposições limitadoras contrafatuais (BICKLE, 2003, p. 16-18)<sup>3</sup>.

Para explicar a "relação análoga" entre Ib e Tr, Bickle substitui a abordagem "sintática" de Hooker por uma abordagem "semântica". Essa abordagem foi desenvolvida no livro *Psychoneural Reduction: The New Wave*, de 1998. Com base nela, a relação entre Ib e Tr é entendida como a construção de uma imagem conjuntista da estrutura de modelos da teoria reduzida Tr no interior do conjunto compreendendo a teoria redutora Tb, o que tem a vantagem

<sup>3</sup> Em nossa apresentação oral das idéias de Bickle durante o V EBICC, afirmamos que Ib é deduzida a partir da conjunção de Tb com *princípios-ponte*. Ora, o que o modelo de Hooker pretende evitar é justamente as dificuldades geradas pela adoção de tais princípios, substituindo-os pelas *suposições limitadoras*. Cometemos naquele momento um erro de interpretação pelo qual nos penitenciamos agora.

de fornecer explicações precisas, semi-formais da "quantidade de correção" e "localização no espectro de redução interteorética". Nessa perspectiva, o reducionismo psiconeural da nova onda é a predição segundo a qual à medida que as teorias maduras se desenvolverem em psicologia (Trs) e neurociência (Tbs), imagens (lbs) da primeira serão construtíveis no interior dos modelos da última. A reformulação por redução interteorética do tradicional problema mente-corpo, feita pela abordagem da nova onda de redução, fornece recursos viáveis para articular e defender o reducionismo psiconeural. Essa abordagem é herdeira de uma tradição que começa com Place e Smart, passando por Nagel e Feyerabend e chegando a Schaffner, Hooker e ao casal Churchland (BICHLE, 2003, p. 27-28).

O livro de Bickle é um exercício no que ele denomina *metaciência da nova onda*. Trata-se de efetuar uma filosofia da ciência "de baixo para cima", que faz emergir o sentido da redução a partir das investigações detalhadas extraídas da prática científica recente, ao invés de "impor" um conceito geral de redução sobre elas, "de cima para baixo". Só assim seremos capazes de ver *como* a neurociência atual está "ligando a mente às moléculas" e só então poderemos especular de maneira responsável sobre o potencial explicativo e o alcance desses recursos para a conduta e a cognição em geral (2003, p. 31).

Um aspecto importante da metaciência da nova onda está em que ela despreza toda preocupação tradicional com ontologia e "metafísica". Embora os cientistas falem uma linguagem que soa "realista", não devemos interpretar suas falas como referentes a questões "externas" às práticas de um empreendimento científico dado. A tarefa da metaciência da nova onda é simplesmente iluminar conceitos como *redução* à medida que eles impregnam a prática científica efetiva. A finalidade não é obter alguma nova maneira de se referir a questões externas reformuladas sobre a existência e natureza de uma "ontologia independente da teoria", mas sim entender as práticas "internas" a importantes empreendimentos científicos correntes e o alcance de sua aplicação e desenvolvimento potencial. Nesse sentido, Bickle está desenvolvendo uma metaciência da psicologia e da neurobiologia contemporâneas, não alguma "ontologia da mente" (2003, p. 31-32)<sup>4</sup>.

---

<sup>4</sup> Este autor toma emprestados os termos *interno* e *externo* de Carnap. Questões *internas* devem ser respondidas através de investigações empíricas. Questões *externas* são mais problemáticas, pois se referem à realidade do próprio mundo-coisa, como, por exemplo, *esta mesa realmente existe?*



A metaciência da nova onda acarreta mudanças na metodologia da filosofia em um outro sentido. Para os filósofos em geral, o *status* da *redução* parece ter algo a ver com a *explicação*. Eles julgam que qualquer defesa do reducionismo exija uma explicação da explicação (*account of explanation*). Na perspectiva da metaciência da nova onda, porém, este não é o caso. Na verdade, os cientistas tendem a se sair bem apenas com uma compreensão grosseira e pronta (*rough-and-ready understanding*) do que conta como "explicação" e o que distingue uma que seja "boa" de uma que seja "pobre" (BICKLE, 2003, p. 36-37).

No segundo capítulo, para ilustrar a redução na prática da principal corrente da neurociência, Bickle se propõe a apresentar um exemplo detalhado das recentes descobertas relativas aos mecanismos moleculares da potenciação de longo prazo (LTP – *long term potentiation*), um tipo importante de plasticidade sináptica dirigida pela experiência, e aos dados comportamentais que esses mecanismos explicam. Com isso, Bickle pretende mostrar que sabemos muito sobre como trabalham pelo menos os componentes fundamentais do cérebro e sobre como podemos manipulá-los para produzir efeitos comportamentais específicos e mensuráveis (BICKLE, 2003, 43-44).

Na opinião de Bickle, o programa de investigação da LTP e seus mecanismos está entre as mais belas e acuradas pesquisas na ciência contemporânea. Acuidade e estética à parte, porém, o propósito de Bickle é apresentar um exemplo paradigmático de uma redução psiconeural acabada que possa constituir um modelo para reduções adicionais (BICKLE, 2003, p. 45).

Mas a proposta discutida é limitada à redução ("ligação") da *consolidação* da memória aos mecanismos moleculares de LTP. Esta é a extensão na qual a redução é oferecida, pelo menos na literatura científica séria. Além disso, foram feitas muitas distinções entre tipos de sistemas de memória, de tal modo que os cientistas virtualmente nunca afirmam a hipótese global de que "a memória é LTP" (BICKLE, 2003, p. 45).

A consolidação da memória possui duas características básicas que constituem os alvos explanatórios da busca reducionista por mecanismos moleculares: a memória de curto prazo e a de longo prazo (2003, p. 47).

---

Estas questões não podem ser completamente respondidas pelos métodos usuais de investigação (análise lógica e investigação empírica).

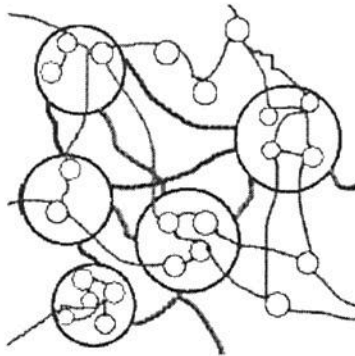
Deixando de lado, por razões de espaço, os detalhes técnicos da exposição de Bickle, podemos dizer que ele chega à seguinte constatação:

Mesmo o simples reflexo da espinha dorsal envolve milhares de neurônios interagindo – sensórios, intra-espinhais e motores. Milhares de outros estão envolvidos mesmo na simples inibição cortical do reflexo da espinha. Estes números aumentam dramaticamente quando consideramos as sensações mais complexas como cognição e comportamento. Mas todo neurônio envolvido está operando ao longo dos princípios básicos esboçados nesta seção [mecanismos moleculares em células individuais]. Estes são, simplesmente, os fatos materiais (2003, p. 59).

Resultados desse tipo levam Bickle a afirmar que a atual neurociência impiedosamente reducionista sabe  *muito* sobre "como o cérebro trabalha", pelo menos quanto aos constituintes básicos e seu modo de interação. Além disso, ela possui instrumentos de descoberta que podemos confiantemente esperar que aumentem nosso conhecimento (2003, p. 95).

As entidades, propriedades e eventos postulados nas explicações psicológicas são tipicamente caracterizadas *de maneira funcional*, puramente em termos de suas *causas e efeitos*, com pouco ou nenhum interesse explícito pelos eventos e processos neurológicos subjacentes que produzem esse perfil funcional (2003, p. 95). Na genética molecular corrente, contudo, esses conceitos puramente funcionais são ligados a seqüências elaboradas de vias moleculares e bioquímicas de transcrição, tradução e recombinação (2003, p. 96)

Bickle oferece, como metáfora visual da ligação interteórica envolvida, a seguinte figura, contendo duas redes superpostas:



Cada uma das redes se compõe de anéis ligados por traços. Os anéis representam as suposições da teoria, enquanto os traços que os ligam representam as conexões causais entre estas suposições. Os anéis da teoria reduzida são os maiores, em negrito; os da teoria redutora são os menores, sem negrito; os traços que ligam os conceitos da teoria reduzida são mais longos, em negrito; os que ligam os conceitos da teoria redutora são mais curtos, sem negrito. Isso reflete a maneira mais grosseira, puramente funcional, pela qual a teoria reduzida "disseca o mundo". É importante notar que os anéis e traços da teoria reduzida não pertencem à teoria redutora. Mesmo assim, a superposição das redes nos mostra quais grupos de anéis e traços da teoria redutora estão colocados debaixo de quais anéis e traços da teoria reduzida. Aplicada ao caso da consolidação da memória, podemos ver *como* (e não apenas *que*) a hipótese celular/molecular emergente dispõe seus constituintes em uma estrutura sequencial e combinatorial abstratamente similar aos conceitos funcionais mais toscos da psicologia. É verdade que as metáforas visuais são cruas e possuem limitações óbvias, mas Bickle pensa que a apresentada acima captura alguns aspectos importantes do reducionismo psiconeural efetivo que ele está anunciando (BICKLE, 2003, p. 100-101). Assim, com respeito à consolidação da memória declarativa, a neurociência "impiedosamente reducionista" já desceu ao nível das "vias moleculares" (p. 102).

Depois destas considerações, Bickle discute, no terceiro capítulo, as implicações para a filosofia da mente e para uma área de proeminência crescente na neurociência atual. Em todos os casos, a redução desempenha um papel central. No caso da filosofia da mente, dois temas importantes são afetados: o problema da causação mental e o argumento da realização múltipla (2003, p. 107).

No que diz respeito ao problema da causação mental, Bickle pensa que a prática científica atual mostra que as explicações psicológicas perdem seu *status* inicial mecânico-causal quando comparadas com uma explicação celular/molecular acabada. Todas as tentativas de "salvar" a causação mental pressupõem que a explicação psicológica permanece como tal, pelo menos em certos contextos, mesmo à luz da explicação neurobiológica dos mesmos dados comportamentais. Em contraste, Bickle afirma que a explicação psicológica *se torna ociosa* quando a explicação celular/molecular é atingida, como no exemplo da LTP. Não há necessidade de evocar explicações psicológicas causais

quando explicações neurobiológicas efetivas são oferecidas. Os filósofos que neguem isto estão sendo em geral guiados por concepções ultrapassadas da prática neurocientífica efetiva. A explicação em um nível menos elevado não exclui a explicação em um nível mais elevado, mas torna-a supérflua, assim como qualquer tentativa de encontrar sucessores empiricamente melhorados no mesmo nível mais elevado (BICKLE, 2003, p. 110).

Isso tudo não significa, porém, que a ciência corrente despreze todas as explicações psicológicas. Bickle oferece uma nova forma de eliminativismo, com base nas seguintes razões. Em primeiro lugar, muitas explicações psicológicas conservam seu prestígio atualmente porque as explicações em nível celular/molecular estão apenas começando a adquirir forma. Dessa maneira, as explicações psicológicas ainda desempenham um *papel heurístico importante* na geração e teste de hipóteses neurobiológicas, sugerindo onde procurar por mecanismos celulares e moleculares e como construir testes comportamentais para essas hipóteses. Em segundo, há vários contextos em que o poder explicativo da explicação neurocientífica não é importante e não exige elaboração. É certo que os contextos determinam as respostas que damos às questões sobre "como" e "por quê". Mas seria contrário à prática científica dizer que os contextos estabelecem os parâmetros do que conta como uma explicação causal suficiente. Na verdade, os contextos determinam se uma explicação mecanicista-causal, dado o estado presente do conhecimento científico, é genuína (BICKLE, 2003, p. 114). Em terceiro, a prática científica pelo menos sugere um destino para os mecanismos celulares e moleculares que é semelhante ao dos mecanismos psicológicos. Com o progresso da investigação, as explicações neurobiológicas poderão dar lugar a explicações bioquímicas, que, por sua vez, poderão dar lugar a explicações termodinâmicas ou eletrodinâmicas ou outras (p. 115).

Para avaliar a posição da neurociência cognitiva, que realiza suas investigações em um nível "mais elevado" do que o celular/molecular e leva em conta propriedades emergentes, Bickle distingue duas abordagens relativas ao problema dos níveis. A primeira delas parte da filosofia da ciência e explica uma relação interteórica abstrata, aplicando-a a questões sobre níveis científicos específicos. Tal abordagem leva a teorias e disputas familiares sobre relações como redução, mecanismo, superveniência, emergência, realização e instanciação. A segunda trabalha *a partir do interior da própria ciência*

*empírica* e procura empregar os diferentes métodos experimentais e técnicas de análise de dados usadas nos níveis cuja relação está em questão. Os resultados são projetos de pesquisa *transdisciplinares* que abordam alguns fenômenos usando recursos de uma variedade de níveis. Bickle considera esta abordagem muito mais frutífera para tratar das questões relativas a níveis de investigação (BICKLE, 2003, p. 116).

Bickle está participando de um projeto de pesquisa transdisciplinar que segue esse programa e busca os mecanismos celulares dos aspectos sequenciais da cognição mais elevada (2003, p. 116-117). O projeto se baseia em resultados obtidos no estudo de certos movimentos dos olhos que continuamente relocalizam a área da retina com a mais elevada acuidade visual em direção a diferentes aspectos daquilo que está sendo observado. Trata-se de processos não cognitivos e não conscientes, mas cujos resultados (*outputs*) se revelaram sequencialmente organizados. Além disso, os componentes e circuitos envolvidos são bem compreendidos no nível fisiológico da célula individual e foram caracterizados por um bem sucedido modelo neurocomputacional, sendo que algumas das suposições puramente computacionais desse modelo foram verificadas biologicamente. Esse modelo frutífero sugere hipóteses testáveis para futura pesquisa no sentido de descobrir os mecanismos celulares dos aspectos sequenciais da cognição mais elevada e da experiência consciente (2003, p. 118, 129). Comparando essa redução do cognitivo ao celular com a "autonomia" das metodologias que postulam níveis mais elevados, Bickle argumenta que projetos transdisciplinares como o dele, empregando métodos e resultados de uma grande variedade de níveis, enfrentam a questão empiricamente. Tais projetos usam modelamento neurocomputacional e construção funcional de imagens neurais (*functional neuroimaging*) para responder a questões de difícil abordagem somente no nível da célula individual. Para Bickle, se isso é tudo a que se reduz a "autonomia do nível mais elevado", então essas técnicas são tratadas como "metodologicamente autônomas" com respeito à fisiologia celular. Mas-isso não é o que os proponentes da "autonomia do nível mais elevado" querem. Eles buscam a autonomia da teoria, afirmando que os mecanismos do nível mais elevado são independentes dos detalhes do nível menos elevado. Essa autonomia não faz parte do projeto transdisciplinar da ciência na prática. As teorias de níveis mais elevados têm apenas um papel heurístico a desempenhar na busca por mecanismos de níveis inferiores. Quando executam essa função, eles devem

ser abandonados, como a escada de Wittgenstein. Os aspectos do todo, do sistema, da população nada mais são do que seqüências complexas, combinações e interações que ocorrem entre os componentes individuais. Quem vai contra isso é um dualista de níveis (BICKLE, 2003, p. 130).

Passando para o segundo problema importante na filosofia da mente contemporânea, o argumento da realização múltipla, Bickle o caracteriza como segue: sua premissa é que um tipo psicológico (propriedade, estado ou evento) é realizado por distintos tipos físicos. Essa premissa se manteve mesmo quando o funcionalismo deu lugar ao fisicalismo não reducionista. Mas Bickle a rejeita, alegando que, na corrente neurociência molecular, a realização múltipla no nível dos sistemas dá lugar a mecanismos evolutivamente conservados que são partilhados por muitas espécies diferentes. Os mecanismos moleculares determinando a atividade e a plasticidade neuronal são os mesmos, dos invertebrados aos mamíferos. As ligações emergentes entre moléculas e mente lançam dúvidas sobre a realização múltipla. Por mais incrível que possa parecer para o filósofo de escritório, os neurocientistas celulares e moleculares estão descobrindo "estados físico-químicos" que servem como mecanismos partilhados para eventos psicológicos partilhados através dos vários filos em biologia (BICKLE, 2003, p. 132).

Bickle trata do problema da consciência no quarto capítulo de seu livro. Ali, ele lembra que, de acordo com a ortodoxia em filosofia da mente, a consciência está além do poder explicativo da neurociência reducionista. Mas Bickle alega já existir muito trabalho nessa área, particularmente no nível celular, e que tal trabalho possui repercussão filosófica. Embora sejam desconhecidos pelos filósofos, os resultados da neurociência fornecem um caso convincente contra a posição ortodoxa, constituindo uma abordagem científica progressiva que objetiva dar conta da consciência através da explicação parte a parte de suas características básicas (BICKLE, 2003, p. 164).

Para ilustrar sua tese, Bickle discute os três aspectos seguintes: a) os experimentos e resultados que revelam os "campos de memória atuante" (*working memory fields*) de neurônios individuais pré-frontais de primatas; b) os efeitos da atenção explícita sobre os perfis potenciais de ação em neurônios visuais individuais; c) os estudos de micro-estimulação sobre grupos diminutos de neurônios visuais e somatossensoriais que induzem experiências fenomenológicas em primatas. Embora não seja possível apresentar aqui todos



os resultados técnicos por questões de espaço, a exposição de Bickle consegue mostrar que, contra as expectativas dos filósofos da mente em geral, em todos os casos a fisiologia de células individuais desempenha um papel fundamental (2003, 164-206). Até mesmo no caso daquilo que Chalmers chama o "problema difícil da consciência", qual seja, o dos *qualia* e da experiência subjetiva, fica caracterizada a base neural (p. 188-189). Bickle deixa claro também que os procedimentos de micro-estimulação são capazes de intervir diretamente em mecanismos de níveis menos elevados para gerar efeitos comportamentais específicos e mensuráveis. Os neurocientistas conseguem ativar diretamente diminutos grupos de neurônios sensoriais e gerar um comportamento observável que se supõe ser guiado normalmente por experiências visuais ou somatossensoriais. Essas manipulações celulares diretas obedecem à exigência explicativa da neurociência contemporânea: não devemos simplesmente encontrar uma atividade neuronal ou intra-neuronal correlacionada com uma tarefa cognitiva, mas sim *invocar* o comportamento específico pela manipulação direta dos mecanismos celulares ou moleculares que constituem a hipótese (p. 206). Nesse sentido, os filósofos não são os únicos proprietários dos "problemas dos *qualia* e da experiência subjetiva", pois os neurocientistas reducionistas também estão no campo (p. 190).

### **Avaliação da proposta reducionista de Bickle**

Conforme mencionado, o livro de Bickle dá continuidade às idéias por ele expressas em *Psychoneural Reduction. The New Wave* (1998). Ali, sua proposta fundamental consiste em reformular o tradicional problema da relação mente-corpo como um problema de redução interteórica, de acordo com o programa de Patrícia Churchland (p. 41). No livro de Bickle que estamos discutindo, a nova onda reducionista parece apresentar-se como uma reafirmação extremada da teoria da identidade mente-cérebro, que vem anunciada através de um discurso inspirado no materialismo eliminativo. Essa versão extremada rejeita não apenas a noção de propriedade emergente em sentido forte, mas inclusive a tese da realizabilidade múltipla do funcionalismo, embora este último seja compatível pelo menos com a teoria da identidade baseada em instâncias. Se, conforme pensa Bickle, a neurociência molecular contemporânea substitui a realização múltipla no nível dos sistemas por

mecanismos moleculares evolutivamente conservados que são partilhados por muitas espécies diferentes, estendendo-se dos invertebrados aos mamíferos, então temos uma volta à teoria da identidade baseada em tipos. Seria de se esperar que Bickle caminhasse em direção à negação dos *qualia*, seguindo a tendência dos partidários da teoria da identidade, mas ele adota a posição mais flexível de aceitá-los, tentando estabelecer um ponto de partida neuronal para explicá-los. Embora os resultados da nova onda ainda se refiram a detalhes neuronais do funcionamento do cérebro que "reduzem" apenas uma pequena área coberta pelos conceitos da psicologia, partiremos do pressuposto otimista de que a promessa feita por Bickle é realizável a longo prazo, para desenvolver a discussão que segue. Por motivos de espaço, analisaremos apenas dois pontos: o conceito de *metaciência* e a viabilidade da "redução impiedosa".

No que diz respeito ao conceito de *metaciência*, gostaríamos de observar que ele possui um *status* ambíguo em Bickle. De um lado, este autor apresenta a redução da nova onda como funcionando a partir da prática científica efetiva, a qual permite uma compreensão grosseira não apenas do que conta como *explicação*, mas também do que pode ser considerado uma explicação *boa* ou *pobre*. Não há dúvida de que aqui estamos diante de uma "metaciência de baixo para cima", em total conformidade com a proposta de Bickle. De outro lado, contudo, ele explica o conceito de *redução* através de um modelo teórico bastante elaborado (o de Hooker). Este modelo é acrescido de uma abordagem semântica para explicar a relação de "analogia" entre a imagem e a teoria reduzida (no caso, entre a neurociência e a psicologia). Isso certamente envolve uma explicação interteórica abstrata, que, nas palavras do próprio Bickle, parte da filosofia da ciência e leva a disputas sobre relações como *redução*, *mecanismo*, *emergência*, *realização*, etc. Nesse caso, não podemos negar que estamos diante de uma autêntica "metaciência de cima para baixo", no sentido contrário ao proposto por Bickle, abrindo perigosamente o flanco para disputas estéreis sobre questões que ele chamaria de "externas" ou "pragmaticamente infrutíferas".

Um outro aspecto da atitude ambígua de Bickle em metaciência se revela quando ele declara que o seu fisicalismo pretende ser uma simples *descrição* das práticas reducionistas. Isso envolve a metaciência de baixo para cima. Mas a abordagem de Bickle não constitui uma mera descrição, pois

envolve também a *prescrição* de um programa de pesquisa, que ele considera mais frutífero que outros. E isso envolve a metaciência de cima para baixo.

Bickle deixa inclusive transparecer em algumas passagens de seu livro que está trabalhando nos dois sentidos indicados, embora não reconheça explicitamente este fato. Em uma delas, ele manifesta sua ênfase na metaciência de baixo para cima quando declara pretender deixar as coisas dependerem da ciência e quando tenta reunir as partes fornecidas pela ciência para explicitar a "imagem maior". Em outra passagem, porém, ele aceita implicitamente a metaciência de cima para baixo, ao informar que o projeto transdisciplinar de que participa usa modelamento neurocomputacional para responder a questões que é difícil abordar somente no nível da célula singular. Isso o leva a reconhecer que há técnicas "metodologicamente autônomas" com respeito à fisiologia celular e, portanto, que existe algo como a "autonomia do nível mais elevado", embora não seja esta a autonomia defendida pelos seus adversários. Assim, apesar de sua ênfase na direção de baixo para cima e de sua obstinada tendência a coletar informações relevantes no nível celular e molecular, Bickle está sendo aqui obrigado a operar também na direção contrária, sempre ocupando um novo degrau da escada científica que, diferentemente da escada wittgensteiniana, nunca pode ser abandonada.

Essa dificuldade se manifesta em Bickle porque, na verdade, não há como privilegiar uma abordagem cujo movimento vai de baixo para cima em detrimento de outra cujo movimento vai de cima para baixo. Não podemos escolher entre essas abordagens alternativas, pois elas são complementares na investigação científica. O privilegiar qualquer uma delas, em detrimento da outra, leva a um tipo inadequado de investigação científica. A ênfase na abordagem baseada na filosofia da ciência, que parte de conceitos abstratos e tenta aplicá-los na análise da atividade científica, leva ao vício que Bacon atribuiu aos racionalistas: como as aranhas, eles tecem teias admiráveis, mas sem solidez. A ênfase na abordagem que trabalha no interior da própria ciência experimental, que procura empregar os métodos e técnicas de análise de dados usados nos níveis cuja relação está em questão, leva ao vício que Bacon atribuiu aos empiristas: como as formigas, eles reúnem os fatos sem ordem e sem escolha. A solução para o dilema, ainda em termos baconianos, seria a das abelhas, que recolhem a seiva das flores, mas com elas segregam o mel. Dessa maneira, os trabalhos das formigas e das aranhas seriam conciliados (BACON,

1979, § 95, p. 63). Utilizando as duas abordagens de maneira complementar, fazemos ciência da maneira mais adequada e completa possível.

Desse modo, a proposta de Bickle, se realmente consistisse apenas no trabalho de reunir de baixo para cima os resultados parciais da neurociência experimental de modo a formar a "imagem maior", possuiria um caráter perigosamente indutivista, dominado pelo espírito das formigas, e correria o risco de ficar na mera acumulação desordenada de dados, sem condições de formular uma síntese. Isso só não acontece porque ele se arrisca a construir uma explicação interteórica abstrata, no espírito das aranhas e na contramão da metaciência de baixo para cima. Assim, queira Bickle ou não, os dois aspectos opostos acabam igualmente contemplados no reducionismo impiedoso da nova onda. A dificuldade da metaciência de Bickle está no enfatizar um deles em detrimento do outro, em virtude da influência do ideal reducionista.

Vejamos agora a viabilidade da "redução" proposta por Bickle. Sabemos que ele se inspira no modelo de Hooker, adaptando-o para o caso da neurociência. As vantagens desse modelo em relação aos demais parecem evidentes. Com efeito, a teoria reduzida *Tr* não é deduzida a partir da teoria redutora *Tb*, o que torna desnecessários os problemáticos princípios-ponte. Além disso, a correspondência é feita entre a imagem *Ib* (obtida por dedução a partir de *Tb* mais as suposições limitadoras) e a teoria reduzida *Tb*. Essa correspondência pode ser estabelecida a partir de uma semântica baseada na teoria dos conjuntos. Em que pese essas vantagens, o procedimento como um todo parece artificial demais para garantir a existência de uma efetiva "redução". Com efeito, o máximo que obtemos é uma *imagem* da teoria reduzida, imagem esta que, por ser deduzida do complexo formado pela associação entre a teoria redutora e as suposições limitadoras, altera o significado e o alcance da teoria redutora. No caso da redução da física galileana à mecânica newtoniana, obtivemos uma imagem que não mais corresponde à teoria redutora original, embora tenha sido deduzida a partir dela. Além disso, esta imagem newtoniana apenas *mimetiza* as explicações da física galileana, que não possuem o mesmo significado que elas. É certo que o procedimento semântico proposto por Bickle permite estabelecer uma correspondência entre os termos da imagem newtoniana e os da física galileana. Mas a natureza dessa correspondência revela-se problemática, uma vez que os termos de cada teoria, ao serem "ligados" através dela, não possuem necessariamente o mesmo significado.



Como a física galileana e a mecânica newtoniana possuem muitas afinidades, o caso que estamos tentando defender fica melhor explicitado se considerarmos as relações entre a mecânica newtoniana e a relatividade estrita. Essas teorias apresentam um parentesco menos acentuado. Suponhamos um ponto  $M$ , definido em um sistema  $O'$  de referência através do grupo de coordenadas  $x', y', z', t'$ . O sistema de referência  $O'$  está se deslocando com velocidade constante  $v$  em relação a um outro sistema  $O$  de referência, cujo grupo de coordenadas é dado por  $x, y, z, t$ . Como definir as coordenadas de  $M$  em  $O$  a partir de suas coordenadas em  $O'$ ? Para facilitar o raciocínio, suponhamos que  $O'$  se desloca ao longo do eixo  $Ox$  do sistema  $O$ . De acordo com a mecânica newtoniana, a medida da coordenada  $x'$ , pertencente a um sistema de referência  $O'$ , que se desloca ao longo do eixo  $Ox$  de um sistema de referência  $O$ , em repouso, é dada pela equação  $x' = x - vt$ . De acordo com a relatividade estrita, a medida da mesma coordenada  $x'$  na mesma situação é dada pela equação  $x' = (x - vt)/(1 - v^2/c^2)^{1/2}$ .<sup>5</sup> Nesta última, se a velocidade  $v$  é muito pequena, o valor da fração  $v^2/c^2$  é praticamente nulo, uma vez que  $c$  é muito grande (300.000 km/s). Ora, isso torna  $1 - v^2/c^2 = 1 - 0 = 1$ . Como  $1^{1/2} = 1$ , segue-se que  $x' = (x - vt)/(1 - v^2/c^2)^{1/2} = (x - vt)/1^{1/2} = (x - vt)/1 = x - vt$ , temos novamente o mesmo valor obtido pela mecânica newtoniana. Isso não ocorre nos casos de valores muito grandes para  $v$ . Nos termos de Bickle, haveria uma redução envolvida aqui, pois, com o acréscimo de uma suposição limitadora ligada a pequenas velocidades, é possível construir uma "imagem relativística" da mecânica newtoniana. Fica claro, entretanto, que esta "imagem", apesar de ter sido deduzida a partir da teoria da relatividade, não mais coincide com ela. E a correspondência entre os conceitos pertencentes à "imagem redutora" e os conceitos da teoria reduzida tem de ser tal que a expressão  $(1 - v^2/c^2)^{1/2}$ , fundamental na teoria redutora, desapareça no cômputo final no interior da "imagem" obtida. Temos, desse modo, uma "imagem" excessivamente artificial da teoria reduzida, em que a expressão  $(1 - v^2/c^2)^{1/2}$  se revela inútil, pois *sempre desaparece* no resultado final. Como se isso não bastasse, os conceitos de *espaço* e *tempo*, pertencentes a tal "imagem", diferem radicalmente dos conceitos respectivos em mecânica newtoniana. Nesta última, o espaço e o tempo não podem ser deformados. Na relatividade estrita, eles podem ser

<sup>5</sup> Para maiores detalhes, inclusive ilustrações, ver Duquesne, 1961, 22-25.

deformados a grandes velocidades. Em outras palavras, apesar de haver coincidência de resultados mensuráveis para pequenas velocidades, os conceitos utilizados para realizar as mensurações já não mais significam a mesma coisa. Além disso, o tipo de correspondência entre a "imagem newtoniana redutora" e a "física galileana reduzida" é diferente do tipo de correspondência entre a "imagem relativística redutora" e a "mecânica newtoniana reduzida". Essas considerações sugerem que o modelo de Hooker, semanticamente aprimorado por Bickle, não é suficiente para dar conta nem do contraste semântico entre os conceitos da "imagem redutora" e os da "teoria reduzida" nem da diversidade de maneiras pelas quais uma determinada "imagem redutora" pode ser relacionada à "teoria reduzida". O máximo conseguido por Bickle parece ter sido a mera *superposição extensional de dois domínios de descrição*, conforme ilustrado por sua "metáfora visual".

Aplicando este raciocínio ao caso da redução da nova onda em neurociência, a conclusão é a mesma: tudo não passa de mera superposição extensional de domínios de descrição. Essa discussão sugere que a imagem neurocientífica obtida por Bickle não parece corresponder a uma *redução efetiva* dos conceitos da psicologia tradicional aos conceitos da neurociência contemporânea.

Se, porém, essa redução fosse possível, o que aconteceria quando a explicação celular/molecular acabada tivesse sido atingida? De acordo com Bickle, embora a explicação celular/molecular, de nível menos elevado, não exclua a explicação psicológica, de nível mais elevado, a primeira torna a segunda supérflua. Ora, pensamos que a relação entre os dois tipos de explicação pode ser caracterizada de maneira diversa. Defendemos a posição de que elas se excluem mutuamente e, em virtude disso, uma jamais tornará a outra supérflua. Vejamos a justificação disso.

Consideremos, por um momento, a relação entre a linguagem de programação e a linguagem de máquina. A primeira utiliza conectivas lógicas como *e*, *ou*, e *se-então*, que, embora sejam tratadas de maneira puramente extensional, possuem um significado eminentemente funcional. Essas conectivas são utilizadas pelo programador na construção de algoritmos que só possuem significado e eficiência porque pertencem a um contexto mais amplo, de caráter totalizante (por exemplo, o programa X está sendo feito para resolver um certo problema prático Y, de modo que o pressionar de uma dada tecla pelo usuário



deverá produzir uma resposta específica por parte da máquina, de acordo com as condições impostas pelo problema e sua possível solução). Já a linguagem de máquina trabalha de maneira diferente. Apropriada para um aparato mecânico que executa operações lógicas, essa linguagem se compõe de um conjunto de instruções mecânicas simples, baseadas no ciclo *busque-execute*, que devem ser realizadas uma de cada vez. Um certo ciclo pode envolver, por exemplo, a tarefa de buscar algum dado em um determinado endereço da memória do computador e depois a tarefa de executar alguma operação com esse dado. Esse ciclo pode ser sucedido por outro e assim por diante, até o término do conjunto de tarefas. O relógio do computador indica quantas dessas tarefas simples (sem distinguir o *buscar* do *executar*) são realizadas serialmente a cada segundo. O resultado final é dado pelo somatório das execuções de cada uma das instruções mecânicas simples e equivale ao resultado final que se obtém quando da execução das operações exigidas para obedecer a uma instrução complexa correspondente. Nessa perspectiva, a linguagem de máquina transforma cada instrução complexa da linguagem de programação em um conjunto discreto de quatro ou cinco instruções mecânicas simples.

Se isso é verdade, então parece que a linguagem de máquina constitui uma "redução impiedosa" da linguagem de programação, bem no espírito da proposta de Bickle. Esse, porém, não é o caso. Suponhamos que tenha havido a efetiva redução de uma linguagem à outra. Ora, isso envolve a seguinte constatação perturbadora: se a linguagem de programação pode ser reduzida à linguagem de máquina, então as duas se equivalem logicamente e esta última também pode ser reduzida à primeira. É perfeitamente possível dizer que a linguagem de programação *também constitui uma redução* da linguagem de máquina. No frigar dos ovos, cada uma dessas linguagens é capaz de "reduzir" a outra.

O que explica isso é o fato de que as duas linguagens estão descrevendo uma mesma entidade ou grupo de eventos, embora realizem abordagens e recortes diferentes da realidade descrita. Nesse sentido, a linguagem de máquina nada mais é do que uma sequência de operações mecânicas simples, sem articulações funcionais, enquanto a linguagem de programação envolve uma lista elaborada de operações lógicas funcionalmente articuladas. Desse modo, uma lista de comandos em linguagem de máquina, como *busque instrução 1, execute instrução 1, busque instrução 2, execute instrução 2, busque instrução*

3, execute instrução 3, busque instrução 4, execute instrução 4, é completamente diversa de uma expressão "equivalente" em linguagem de programação, como, por exemplo, *se A, então B*. É certo que o resultado obtido a partir da aplicação de cada uma das duas expressões é o mesmo, mas a maneira pela qual cada uma delas lida com a mesma realidade é fundamentalmente diferente. Uma descrição mecânica não é uma descrição lógica. Assim, em oposição a Bickle, afirmamos que essas linguagens se excluem mutuamente, em virtude de seus respectivos objetivos e recortes específicos. Nessa perspectiva, elas são complementares. A adoção de uma delas *não torna* a outra ociosa. Isso é verdadeiro no caso considerado, pois a linguagem de máquina não tornou supérflua a linguagem de programação. Por motivos pragmáticos, a maioria dos programadores prefere trabalhar com os aspectos funcionais da linguagem mais complexa de programação a ter de enfrentar os meandros pormenorizadamente sequenciais da linguagem de máquina.

Coisa semelhante acontece no caso do famoso debate em torno das descrições definidas em posição de sujeito nas sentenças declarativas. Suponhamos a sentença

(1) *O rei da França é sábio.*

Frege argumenta que ela pressupõe semanticamente a sentença

(2) *Existe um e somente um rei da França.*

Para Frege, a descrição definida *o rei da França* constitui um nome próprio complexo. A sentença (2) não faz parte daquilo que é afirmado por (1), embora constitua uma condição semântica para a afirmação de (1). Assim, quando (2) é verdadeira, ou seja, quando existe um e somente um rei da França, então (1) e sua negação possuem valor-verdade. Quando, porém, (2) é falsa, ou seja, quando não existe um e somente um rei da França, então nem (1) nem sua negação possuem valor-verdade. Isso constitui uma quebra da lei do terceiro excluído e só ocorre porque a linguagem natural é imperfeita (FREGE, 1966, p. 21-41).

Em oposição a Frege, Russell elaborou a famosa *teoria das descrições*. De acordo com esta última, a sentença (1) deve ser lida da seguinte maneira:

(3) *Existe um x tal que x ocupa o trono francês, x é único e x é sábio.*

Para Russell, a descrição *o rei da França* não corresponde a um nome próprio complexo, como quer Frege, mas constitui um *símbolo incompleto*, que desaparece após a análise lógica apropriada. A sentença (2) não é semanticamente pressuposta por (1), mas faz parte do conteúdo afirmado por (1). Assim, quando (2) é falsa, (1) também é falsa. Não há quebra do princípio do terceiro excluído, embora a linguagem natural seja enganadora, ao sugerir que uma descrição definida funcione como um nome próprio complexo, quando na realidade esse não é o caso (RUSSELL, 1956, p. 39-56).

Essa discussão foi retomada por Strawson e depois passou por uma série de refinamentos com os trabalhos de Donnellan e Kripke.<sup>6</sup> O debate passou quase sempre ao largo de uma importante questão, que foi levantada inicialmente por Alston (1958)<sup>7</sup>. Mais tarde, essa mesma questão foi levantada por Crispin Wright, ao discutir a teoria das descrições: se a paráfrase proposta por Russell, representada aqui por (3), equivale logicamente à sentença contendo a descrição definida em posição de sujeito, representada aqui por (1), então não temos como decidir qual dos lados do bicondicional tem prioridade ontológica (WRIGHT, 1983, p. 31-32)<sup>8</sup>. Em outras palavras, dada a equivalência entre (1) e (3), podemos privilegiar (3), alegando que a descrição definida *o rei da França* é potencialmente enganadora, uma vez que, como quer Russell, ela corresponde a um símbolo incompleto que desaparece depois da análise; mas podemos também inverter o argumento e privilegiar (1), alegando que a paráfrase russelliana é potencialmente enganadora, uma vez que, como quer Frege, a descrição definida *o rei da França* corresponde a um nome próprio autêntico. Em nosso trabalho de doutorado, defendemos a tese de que a controvérsia ligada às descrições definidas em posição de sujeito nas sentenças declarativas é indecidível no nível puramente semântico (PINTO, p. 93 et seq.)<sup>9</sup>. Essa tese é perfeitamente compatível com a posição que estamos defendendo aqui. Ela significa que, em última instância, podemos rejeitar o bicondicional, ficando com duas descrições independentes. Cada uma delas será privilegiada de acordo com nossos interesses. Como se pode ver, no que diz respeito à possibilidade de redução, o

<sup>6</sup> Ver Strawson, 1950; Donnellan, 1966; Kripke, 1981.

<sup>7</sup> Alston, W. Ontological commitment. *Philosophical Studies*, v. 6, p. 8-17, 1958.

<sup>8</sup> Wright, C. *Frege's conception of numbers as objects*. Aberdeen: Aberdeen University Press, 1983.

<sup>9</sup> Pinto, P. R. *Wittgenstein and semantic presuppositions generated by definite descriptions in subject-position*. PhD Thesis. Edinburgh: University of Edinburgh, Dept of Philosophy, p. 93 ss.



debate em torno das descrições definidas envolve a mesma conclusão liberal da discussão sobre as relações entre linguagem de máquina e de programação.

Essas considerações podem ser aplicadas na avaliação da redução da nova onda. Suponhamos que a explicação celular/molecular tivesse alcançado um nível em que fosse possível uma superposição adequada entre uma imagem gerada por ela e a explicação psicológica, nos moldes sugeridos pela metáfora visual de Bickle. Nesse caso, as duas explicações seriam logicamente equivalentes e não teríamos senão motivos arbitrários para privilegiar qualquer um dos lados do bicondicional. As duas descrições seriam igualmente válidas e a ênfase em uma ou em outra dependeria de considerações pragmáticas. Não haveria espaço para uma verdadeira redução aqui.

Isso pode ser ilustrado através de um dos exemplos estudados por Bickle. Nele, o comportamento criminoso é considerado (pelo menos em parte) uma questão de contrações orquestradas de fibras musculares. Desse modo, se a pobreza constitui uma causa do comportamento criminoso, ela também deve ser (pelo menos em parte) uma questão de contrações orquestradas de fibras musculares. Poderíamos apresentar uma versão ainda mais impiedosamente reducionista que a de Bickle, dizendo que um dado comportamento criminoso é totalmente descritível em termos de uma sequência elaborada de configurações moleculares. Nesse caso, se uma dada situação de pobreza constitui uma causa desse comportamento, ela também deverá ser totalmente descritível em termos de uma sequência elaborada de configurações moleculares. Algum dia seremos muito provavelmente capazes de formular adequadamente essas sequências. Isso significa, porém, que as descrições do comportamento criminoso e da pobreza em termos sociológicos, por exemplo, se tornarão ociosas porque puderam ser "reduzidas" a descrições de sequências moleculares? De modo algum. Em primeiro lugar, as descrições de sequências moleculares teriam também sido "reduzidas" às correspondentes descrições sociológicas. Em segundo lugar, o uso de sequências moleculares para descrever situações em perspectiva sociológica será tão complicado e despropositado que ninguém o levará adiante.

A razão mais profunda para isso está na própria natureza dos domínios de descrição envolvidos. Nos bicondicionais considerados, cada lado expressa uma propriedade contraditória com respeito ao outro lado. Quando dizemos, por exemplo, que a *dor* é o mesmo que *excitação das fibras C*, temos, de um lado desta equivalência, um evento mental (*dor*) a que estamos atribuindo uma

propriedade não mental (excitação das fibras C) e, do outro lado, um evento físico (excitação das fibras C) a que estamos atribuindo uma propriedade não física (dor). Para evitar esta contradição, Feyerabend, um dos fundadores do materialismo eliminativo, sugeriu a total redefinição dos conceitos mentais, já que os mesmos se revelaram incompatíveis com as descobertas empíricas, que apontam em direção ao fortalecimento dos conceitos neuronais (1963a, 1963b). Temos aqui um caso de teoria da identidade mente-cérebro, em que a explicação da "mente" é reduzida à explicação do "cérebro", com as conhecidas perplexidades que isso produz. Do ponto de vista lógico, Feyerabend aceita o bicondicional e tenta eliminar a contradição através da ênfase em um dos lados da equivalência.<sup>10</sup> Mas ele poderia ter-se deixado fascinar um pouco menos pelas descobertas empíricas e adotado a atitude de Bohr, por exemplo, que se viu às voltas com as descrições empiricamente contraditórias do elétron como partícula e como onda e admitiu estar tratando da mesma realidade, a qual possuía contudo uma estrutura suficientemente complexa para interagir conosco de forma a serem observadas propriedades mutuamente excludentes. Diferentemente de Feyerabend, Bohr passou a tratar o elétron-partícula e o elétron-onda como dois domínios de descrição mutuamente excludentes mas complementares. Isso lhe permitiu lidar com a complexidade do elétron através de duas descrições independentes, sem cair em contradição e, o que é mais importante, sem privilegiar uma delas em detrimento da outra. Temos aqui um caso de *identidade sem redução*: na equivalência onda-partícula, a explicação como onda e a explicação como partícula são igualmente válidas, sem que uma tenha de ser necessariamente "reduzida" à outra. Do ponto de vista lógico, Bohr rejeita o bicondicional, trocando-o por duas descrições independentes que o livram da contradição. Seguindo o espírito da proposta de Bohr e aplicando-o à

---

<sup>10</sup> Em sua argumentação contra as idéias de Bickle, expostas em *Psychoneural reduction*, de 1998, Achim Stephan enfatiza o outro dado da equivalência, perguntando o que aconteceria se tomássemos as atitudes proposicionais como sendo aquilo que efetivamente experimentamos. Isto lhe permite concluir que o reducionista da nova onda não consegue explicar como é que algum disparo neuronal em nossa cabeça nos dá a vivência de um estado mental correspondente (por exemplo, a impressão de estarmos nos referindo a coisas no mundo exterior, o gosto de uma bebida descendo pela nossa garganta, a sensação de agir livremente, etc.). O reducionista da nova onda não responde a esta questão, mas nos pede para abandonar os elementos essenciais de nossa concepção ligada às atitudes proposicionais. Para Stephan, isto é um sinal de seu fracasso (Stephan, p. 280-281). Para nós, como será mostrado mais adiante, é um sinal de que podemos adotar *posturas descritivas* diferentes e independentes umas das outras.



nossa discussão, obteremos uma perspectiva pluralista que, ao contrário do que ocorre com Feyerabend, não inspira qualquer forma de eliminativismo.

Um exemplo da perplexidade produzida pela abordagem eliminativista está na atitude de Dennett para com o que ele chama de *postura intencional*. Com o objetivo de explicar a pluralidade de descrições possíveis do comportamento dos seres vivos, este autor sugere a existência de três modos diferentes pelos quais podemos observar uma entidade. Primeiro, temos a *postura física*, que consiste em considerar o comportamento da entidade com base nos princípios da física. Nesse caso, a entidade é tratada como um objeto que reage de acordo com suas propriedades físicas (por exemplo, quando predizemos que uma pedra ao ser largada da mão de alguém irá cair ao chão). Segundo, temos a *postura de projeto (design)*, que consiste em considerar o comportamento da entidade com base no funcionamento de sua estrutura como um todo. Nesse caso, a entidade é tratada como um objeto que reage de acordo com o modo pelo qual foi projetado para reagir. Não precisamos ter familiaridade com as leis físicas envolvidas (por exemplo, sabemos que um certo alarme cujos botões foram pressionados de um modo determinado fará um barulho depois de algumas horas, embora não precisemos conhecer as leis físicas envolvidas pelo alarme para realizar essa ação). Terceiro, temos a *postura intencional*, que consiste em considerar o comportamento da entidade com base em suas escolhas racionais. Nesse caso, a entidade é tratada como um *sistema intencional* que escolhe uma dada linha de ação em função de seus fins (por exemplo, poderíamos considerar o alarme como nosso criado que recebeu o comando de nos acordar a uma certa hora; estaríamos então confiando em sua capacidade de compreender nossa ordem e reconhecer a hora exata de nos acordar). Podemos prever o comportamento do alarme como se ele fosse um agente racional. A postura intencional é um atalho lingüístico útil em tal caso e revela suas vantagens quando a entidade envolvida é mais complexa do que um alarme – digamos, um computador ou uma pessoa (DENNETT, 1978, cap. 1).

Até aqui, temos o lado "liberal" da apresentação que Dennett faz das três posturas. Mas suas inclinações eliminativistas *à la* Feyerabend o levam a propor que a "intencionalidade" não deve ser tomada demasiadamente a sério. Ela deve ser tratada como uma ficção útil e não como uma propriedade real da consciência. De acordo com Dennett, um ser onisciente poderia prever tudo o que ocorre no mundo adotando apenas a postura física. Como, porém, não somos

oniscientes, precisamos usar os conceitos decorrentes da postura intencional, que seria apenas um outro nome para a psicologia popular. Nessa perspectiva, a postura intencional é extremamente útil, embora, por vezes, se apresente incompleta ou até mesmo falsa. Dennett se esquece de que a *intencionalidade* não constitui uma ficção maior do que o *elétron* e insiste em considerá-la meramente uma postura, um modo de observar as coisas. Isso faz com que, no balanço geral, sua posição a respeito da postura intencional seja ambígua. Ele gasta muito tempo discutindo, por um lado, a sua utilidade<sup>11</sup> e, por outro, as suas limitações<sup>12</sup>. Embora pareça acreditar que um dia a postura intencional será objeto de redução pela neurociência<sup>13</sup> sua ambigüidade no tratamento da questão aponta em direção à nossa idéia de que, em última instância, as duas linguagens são complementares, como ocorre nos casos da linguagem de programação e de máquina e do debate em torno das descrições definidas.

A abordagem baseada na pluralidade descritiva se tornará mais clara se nos lembrarmos de que esta nossa época "quiniana" está marcada pela crítica de Sellars ao "mito do dado". Esta última nos revela dramaticamente que estamos em relações *causais* e não representacionais com o mundo. Com base em Sellars, sabemos agora que uma crença só pode ser *justificada* por outra crença, embora possa ser *causada* por alguma experiência sensorial. Essas idéias estão em sintonia com o modelo biológico do conhecimento proposto pelo chileno Humberto Maturana, que trabalha com dois aspectos complementares: o fechamento operacional do sistema vivo e suas interações causais com o mundo.<sup>14</sup> No que se segue, estaremos supondo que esse modelo é filosoficamente adequado, embora, por motivos de espaço, não possamos nos dar ao trabalho de justificá-lo aqui, remetendo o leitor para o texto relevante<sup>15</sup>. Uma das características do modelo é permitir a conclusão de que um mesmo

<sup>11</sup> DENNETT, 1978, cap. 1; 1987, cap. 2-4,7.

<sup>12</sup> DENNETT, 1978, Cap. 3 e 11; 1987, Cap. 5.

<sup>13</sup> DENNETT, 1995.

<sup>14</sup> Maturana, H. & Varela, F. *The Tree of Knowledge. The Biological Roots of Human Understanding*. Revised edition. Transl. by R. Paolucci. Boston and London: Shambhala, 1998, cap. 1, 2 e 6.

<sup>15</sup> Ver Maturana, H. & Varela, F. *The Tree of Knowledge. The Biological Roots of Human Understanding*. Revised edition. Transl. by R. Paolucci. Boston and London: Shambhala, 1998. Para uma discussão do significado filosófico da teoria de Maturana e sua convergência com as idéias de Sellars, ver Margutti Pinto, P. R. Ceticismo, Pragmatismo e a Crítica de Sellars ao Mito do Dado. In: Dutra, L. H. de A. & Smith, P. (Orgs.). *Ceticismo. Perspectivas Históricas e Filosóficas*. Florianópolis: NEL/UFSC, 2000, p. 137-168.

evento pode ser objeto de múltiplas descrições, que vão desde um retrato puramente fisicalista até um quadro envolvendo considerações morais. As descrições são múltiplas porque o potencial expressivo da linguagem o permite e porque cada uma delas pode ser causada por um tipo diferente de interação com o mundo.

Embora ligada a um contexto diferente do que aqui adotamos, a importante intuição da possibilidade de múltiplas descrições foi formulada pela primeira vez por Elizabeth Anscombe, em seu livro *Intention*. Ali, o simples fato de um homem estar movendo uma alavanca para baixo e para cima pode ser objeto de um grande número de retratos possíveis, que começam pela descrição das configurações moleculares ligadas aos movimentos do braço humano e da alavanca, passam pela descrição da operação de bombear água por uma pessoa para a caixa d'água de uma casa, transitam pela descrição do processo de envenenamento gradativo efetuado pelo caseiro dos perigosos nazistas que moram na propriedade e chegam finalmente à descrição da tentativa de salvar a humanidade do domínio nazista (ANSCOMBE, 1991, § 23, p. 84 et seq.). O evento é o mesmo, mas cada descrição foi feita a partir de interações e objetivos específicos. A abordagem de Anscombe se encaixa nos termos da proposta que enfatiza as interações causais e não representacionais com o mundo. Ora, a aceitação desta hipótese implica a rejeição do reducionismo da nova onda, porque leva à conclusão de que não há uma descrição privilegiada a ser colocada em posição de destaque. O que temos, na verdade, é uma rede de descrições complementarmente interligadas. O uso da linguagem nos permite uma grande variedade de *posturas descritivas*, que não estão necessariamente limitadas apenas às três mencionadas por Dennett. Todas são igualmente válidas na complexa tarefa de lidar com o mundo. E cada uma delas se tornará privilegiada, quando selecionarmos o ponto de vista correspondente. A seleção do ponto de vista depende de considerações eminentemente pragmáticas.

A posição que estamos defendendo aqui é compatível com a de Rorty e Davidson. Inicialmente impressionado pelo materialismo eliminativista de Feyerabend, Rorty desenvolveu aquilo que então chamou de *teoria do desaparecimento da mente* (RORTY, 1965). Aos poucos, porém, sob a influência de Davidson, adotou a posição mais equilibrada do *fisicalismo não reducionista*. De acordo com essa abordagem, um mesmo evento pode ser descrito igualmente bem em termos psicológicos e em termos fisiológicos. A

abordagem é fisicalista porque admite que todo evento pode ser descrito em termos micro-estruturais ligados a partículas elementares. Ela é não-reducionista porque a possibilidade de descrição em termos fisicalistas decorre do caráter da linguagem e não da ontologia. Assim, descrições psicológicas se referem a estados neurológicos sob uma descrição mental; descrições neurológicas se referem a estados psicológicos sob uma descrição física. Mesmo podendo ser reduzida a uma descrição física, a linguagem psicológica não se torna ociosa. Poderemos continuar a falar indefinidamente sobre o mental (RORTY, 1991).

Numa linha análoga de raciocínio, Horgan e Woodward defendem a tese de que a psicologia popular é um sistema de conceitos que possui algum tipo de autonomia. Para eles, é improvável que a neurociência possa algum dia tornar falsos os conceitos da psicologia popular, principalmente se aceitarmos o monismo anômalo de Davidson. Este último é simultaneamente materialista e naturalista, sem ser reducionista ou eliminativista. Se Davidson está certo, então a psicologia não será redutível à ciência física (HORGAN; WOODWARD, 1985).

A própria "metáfora visual" de Bickle pode ser reinterpretada nesta direção. Na figura apresentada, como vimos, os anéis e traços da teoria reduzida são maiores, ao passo que os da teoria redutora são menores, refletindo a "maneira mais grosseira", "puramente funcional", pela qual a teoria psicológica reduzida "disseca o mundo". As expressões usadas por Bickle denunciam a sua postura fisicalista reducionista, que o leva a avaliar depreciativamente os resultados da psicologia tradicional. Não estivesse ele contaminado por esse viés e poderia reconhecer que a maneira "puramente funcional" pela qual a psicologia "disseca o mundo" não é "grosseira", mas representa um *outro tipo de abordagem*, que vê o ser humano como agente intencional e que por isso se torna capaz de identificar aspectos incomensuráveis com meras seqüências de configurações moleculares e a elas irredutíveis. Na abordagem que estamos propondo, a descrição neuronal de seqüências moleculares e a descrição psicológica de eventos mentais correspondem a descrições mutuamente excludentes de uma mesma realidade, cuja complexidade ultrapassa as limitações de um único tipo de descrição. Nessa perspectiva, até mesmo o célebre argumento de Searle sobre o *quarto chinês*, que mostra a insuficiência de uma explicação meramente sintática (a escolha de sentenças em língua chinesa com base apenas em semelhanças de símbolos abstratos) para dar conta de uma



operação de caráter semântico (o diálogo inteligente em chinês com uma pessoa fora do quarto) pode ser considerado o resultado de um equívoco decorrente da tentativa de mesclar dois domínios opostos de descrição.

Como se pode ver, as posições dos autores mencionados convergem com a nossa, a qual envolve um aprimoramento através da adição do conceito de *pluralismo descritivo*. Isso nos permite introduzir alguns reparos à posição de Davidson e Rorty. Com efeito, a não ser pelas próprias origens históricas do problema e pela evolução do debate por ele gerado, não parece haver uma razão forte para qualificar tal posição como um *fisicalismo*. Em primeiro lugar, este termo parece excluir a qualificação de *não-reducionista*. Não foi à toa que Jaegwon Kim acusou de contraditório o materialismo não-reducionista, alegando que não se pode ser simultaneamente materialista e não-reducionista (1993, cap. 11). Em segundo lugar, como mostramos na discussão acima, o fato de todo evento poder ser descrito em termos micro-estruturais não é motivo para prestar uma homenagem especial ao fisicalismo. Isso ocorre apenas porque ele envolve o nível menos elevado de descrição, sendo, por isso, comum a praticamente todos eventos e entidades do nosso mundo. O exemplo analisado por Anscombe mostra com clareza que outras descrições não-físicas de um mesmo evento, capazes de captar aspectos diferentes deste último, também são válidas. Assim, no espírito da complementaridade das descrições envolvidas, esta abordagem seria melhor denominada *pluralismo descritivo*.

Com o objetivo de esclarecer um pouco mais a posição que estamos defendendo, gostaríamos de lembrar que a neurociência e a psicologia lidam com seres vivos, que, na abordagem de Maturana, constituem sistemas auto-organizados. Ora, de acordo com este autor, todas as vezes que a auto-organização está sendo considerada, temos pelo menos dois domínios de descrição envolvidos. O primeiro deles diz respeito à *fisiologia* do sistema, ou seja, à descrição das operações que ocorrem no seu interior e que garantem a continuidade de sua organização. O segundo diz respeito à *conduta* do sistema, ou seja, à descrição de suas interações causais com o meio circundante. Esses domínios de descrição não podem ser mesclados, sob pena de cairmos em contradição. Agindo dessa maneira, estaremos realizando aquilo que Maturana denomina a *odisséia epistemológica*, que consiste em navegar no estreito caminho entre Cila, o monstro do representacionismo, e Caribde, o redemoinho do solipsismo (MATURANA; VARELA, 1998, p. 134). Como se pode ver, este



modelo é compatível com a idéia de *pluralismo descritivo*, embora tal fato não tenha sido explicitamente reconhecido por Maturana em seu texto.

Para ilustrar, suponhamos uma ameba, um sistema auto-organizado, que se encontre ao lado de uma partícula de alimento. Suponhamos também que ela emita pseudópodos e envolva a partícula, digerindo-a. A descrição que acaba de ser feita se refere à *conduta* da ameba, ou seja, às suas interações causais com o meio circundante. Ao invés dessa descrição, poderíamos dizer que a ameba corresponde a tal e tal conjunto de configurações moleculares que se modificaram de acordo com tal e tal sequência. Essa descrição se refere à *fisiologia* da ameba, ou seja, às operações que ocorrem no seu interior e que garantem a continuidade de sua auto-organização. Estamos sugerindo aqui que essas descrições são igualmente válidas e complementares, querendo com isso significar que não há uma que seja privilegiada com respeito à outra. Mais ainda: os predicados pertencentes à descrição da conduta, comparados com os predicados pertencentes à descrição da fisiologia, envolvem relações funcionais de caráter causal que não são redutíveis aos predicados mecânicos, pertencentes à descrição da fisiologia. Aplicando isto a nossa discussão, vemos que a descrição dos mecanismos neurológicos corresponde à fisiologia, enquanto a descrição dos mecanismos psicológicos corresponde à conduta. Trata-se de dois domínios de descrição diferentes e mutuamente excludentes, embora cada um deles seja indispensável para a explicação completa do ser humano enquanto sistema vivo.

Imaginemos agora o *Deep Blue*, computador dotado com um programa capaz de derrotar um grande mestre de xadrez. A descrição de sua fisiologia nada mais é do que o relato de uma sequência elaborada de configurações moleculares no equipamento do computador, no mesmo estilo da neurociência. Mas a descrição de sua conduta envolve relatos bem mais complexos, do tipo "optou pela abertura siciliana", "decidiu tomar o peão do adversário", "propôs empate", etc. Com base na descrição de sua conduta, vemos que o *Deep Blue* é capaz de realizar uma atividade até agora considerada especificamente humana, ou seja, "jogar xadrez".

Imaginemos também uma outra criação neste domínio. Estamos nos referindo aos pequenos robôs que, ao lado de outros, formam uma equipe e são capazes de participar de uma partida de futebol contra outra equipe de robôs. A descrição da fisiologia de cada um desses robôs continua sendo o relato de

seqüências elaboradas de configurações moleculares no interior das máquinas. Mas a descrição da conduta de cada robô envolve relatos bem mais complexos, do tipo "driblou o adversário", "chutou direto para o gol", "passou a bola brilhantemente para o atacante", "exerceu a função de zagueiro", etc. Com base na descrição de sua conduta, vemos que cada um desses robôs é capaz de realizar, até certo ponto, uma outra atividade até agora considerada especificamente humana, a saber, "jogar futebol".

Nenhuma das máquinas desses exemplos se revela inteiramente auto-organizada, é verdade. Mas, do ponto de vista das atividades realizadas, todas compartilham com os sistemas auto-organizados a propriedade básica do fechamento operacional através de algoritmos apropriados. Nos dois casos, a descrição da fisiologia e da conduta são mutuamente excludentes, mas necessárias para a explicação completa das entidades consideradas: essas descrições são *complementares*, no sentido de Bohr.

A esta altura, alguém poderia replicar que a proposta de Bickle pressupõe a *co-referencialidade* entre os conceitos da imagem redutora e os da teoria reduzida. Este é o sentido básico do *reducionismo*, que procura, no espírito da navalha de Ockham, trabalhar com as descrições provenientes das "imagens redutoras" porque elas são mais "econômicas" do que as descrições provenientes da teoria reduzida. Mas a questão está em saber se duas descrições diferentes são efetivamente co-referenciais. Por exemplo: se uma dada descrição de configurações moleculares e uma certa descrição de processos psicológicos se referem *à mesma coisa*, como podemos determinar se a correspondência entre as descrições é nômica, capaz de fundar uma regra semântica, ou se ela é meramente uma correlação contingente, sem valor explicativo? Este problema também afligiria a tese das descrições complementares. Com efeito, se elas descrevem o mesmo fenômeno, o mesmo referente, então permitem redução no espírito ockhamiano. Se elas não descrevem o mesmo referente, então a complementaridade perde qualquer especificidade: teríamos, no máximo, dois aspectos disjuntos.<sup>16</sup>

Para responder a esta objeção, temos de explicitar o tipo de relação existente entre as descrições complementares. Por um lado, elas possuem o mesmo referente e, nesse sentido, são co-referenciais. Por outro lado, o tipo de

---

<sup>16</sup> Esta objeção importante me foi feita pelo Prof. Marcelo Dascal, durante minha exposição. Espero estar reproduzindo de maneira fiel o seu pensamento.

co-referencialidade envolvido não permite redução, porque os recortes utilizados são mutuamente excludentes: a descrição do elétron como onda exclui a descrição do mesmo como partícula; a descrição neurocientífica da atividade psíquica exclui a descrição da mesma como processo mental. De acordo com o reducionismo da nova onda, se há co-referencialidade, então ela é nômica e fornece uma regra para estabelecer a correspondência semântica. De acordo com a tese da pluralidade descritiva que estamos propondo, a existência de co-referencialidade não é condição suficiente para que ela seja nômica. Isso, porém, não torna meramente acidental a correlação entre as descrições envolvidas. Sabemos que estamos lidando com a *mesma realidade*. O problema é que ela está se revelando complexa demais a ponto de permitir descrições opostas, baseadas em perspectivas mutuamente excludentes e sem possibilidade de redução de uma à outra com base na mera co-referencialidade.

Essa resposta, porém, permite uma outra objeção. Se as descrições são mutuamente excludentes, como sabemos que elas descrevem a *mesma realidade*?<sup>17</sup> Se a abordagem proposta por nós fosse de caráter idealista, essa pergunta ficaria inteiramente sem resposta. Esse, porém, não é o caso, pois estamos procurando superar a tradicional oposição entre *realismo* e *idealismo* através da possibilidade da pluralidade descritiva. Com efeito, estamos nos baseando no fato de que, embora as nossas crenças só possam ser *justificadas* por outras crenças (dimensão "idealista" do modelo), elas podem ser *causadas* pelas nossas interações com o mundo (dimensão "realista" do modelo). Desse modo, sabemos, por exemplo, que as múltiplas descrições de Anscombe se referem a um *mesmo evento* em virtude de nossas interações causais com ele. Da mesma forma, sabemos que a descrição neurocientífica e a psicológica, embora mutuamente excludentes, se referem ao *mesmo evento* em virtude de nossas interações causais com ele.

As considerações feitas até agora sugerem a possibilidade de uma abordagem alternativa para as chamadas *propriedades emergentes*. Elas parecem resultar principalmente do contraste marcante que pode ser observado quando comparamos certas posturas descritivas. Tomemos o caso da descrição da conduta e da fisiologia de um sistema, por exemplo. A primeira se baseia na noção holística de *agente intencional competente*, em oposição à segunda, que se

---

<sup>17</sup> Devo esta objeção também ao Prof. Marcelo Dascal, em uma conversa privada que tivemos a respeito de minha proposta.

funda na noção mecanicista de *configuração molecular singular*. Não há como explicar uma pela outra, ou seja, contra Bickle, não pode haver uma efetiva "redução" aqui. Apesar disso, o pedaço do real que ambas descrevem é suficientemente complexo para que uma explicação completa do mesmo exija o apelo a ambas. Nesse caso, o difícil problema de caracterizar adequadamente as propriedades emergentes não seria ontológico, mas lingüístico. Com efeito, ao observar, por exemplo, a oposição entre a descrição dos mecanismos neurológicos de um ser humano e a descrição de sua vida mental, o pesquisador pode ficar tão maravilhado com a diferença encontrada a ponto de assumir o conceito de *propriedade emergente* para explicar o salto que dá quando vai da neurociência à vida psicológica. Não é à toa que Stephan, por exemplo, propõe seis variedades de emergentismo: o *fraco*, o *sincrônico*, o *diacrônico fraco*, o *diacrônico forte*, o *de estrutura diacrônica* e o *forte de estrutura diacrônica*. O emergentismo sincrônico resulta da adição da tese da *irreducibilidade* ao emergentismo fraco; o diacrônico fraco resulta da adição da tese da *novidade* ao emergentismo fraco; o de estrutura diacrônica resulta da adição da tese da *imprevisibilidade* ao emergentismo diacrônico fraco; etc. Os três primeiros correspondem a formas fracas de emergentismo e são compatíveis com o fisicalismo reducionista, enquanto os três últimos correspondem a formas fortes de emergentismo e se revelam compatíveis com o fisicalismo não-reducionista. Em virtude disso, Stephan conclui que, em filosofia da mente, particularmente no debate sobre *qualia*, há uma necessidade de um conceito forte de emergência. No caso das discussões de redes connexionistas, as noções fracas são suficientes (STEPHAN, 1999)<sup>18</sup>. Ora, pensamos que cada novo tipo de emergentismo da classificação de Stephan surge a partir da adição de alguma tese que, de um modo ou de outro, contradiz a postura própria ao tipo que lhe deu origem: *irreducibilidade*, no caso do emergentismo sincrônico; *novidade*, no caso do emergentismo diacrônico; e assim por diante. Isso significa que cada tipo de emergentismo pode ser tomado como uma *postura descritiva independente* das demais. A compatibilidade ou não de cada tipo com o reducionismo fisicalista ou com o não-fisicalista simplesmente mostra o alcance de cada postura descritiva em função de objetivos pragmáticos específicos.

---

<sup>18</sup> Ver o quadro à p. 55.



Assim, o que estamos sugerindo aqui é que, ao propor o conceito de *propriedade emergente*, o investigador esteja inutilmente tentando encontrar um ponto de contato entre posturas descritivas díspares, ao invés de aceitar a inevitável mudança de perspectiva que decorre da adoção de cada uma delas. Isso envolve a tentativa de *mesclar* dois domínios de descrição mutuamente excludentes, postulando a existência de uma região peculiar da linguagem, de um verdadeiro *limbo lingüístico* em que isso seria possível. Tal região, porém, seria logicamente contraditória. No caso, por exemplo, da tentativa de estabelecer um ponto de contato entre a descrição da conduta e a da fisiologia, teríamos de lidar ao mesmo tempo com os conceitos básicos de *agente intencional competente*, ligado à conduta, e *configuração molecular singular*, ligado à fisiologia. Isso não quer dizer que a realidade considerada seja simples: tudo indica que a realidade psíquica é ontologicamente complexa e possui um duplo aspecto (neurológico e psicológico). O que queremos dizer aqui é que cada aspecto é objeto de uma descrição independente que exclui a descrição do outro, como acontece nos casos anteriormente considerados, envolvendo os seguintes pares: *elétron-onda* versus *elétron-partícula*, *dor* versus *excitação das fibras C*, *descrição como símbolo incompleto* versus *descrição como nome próprio complexo*. Quando falamos de *cérebro* e de *mente*, estamos nos referindo à *mesma coisa*, a qual possui entretanto um duplo aspecto de caráter contraditório e, neste sentido, põe em cheque a nossa capacidade de descrição lingüística. No espírito da pluralidade descritiva, a solução desse dilema não está na busca de uma ponte entre os aspectos opostos, através do problemático conceito de *propriedade emergente*, por exemplo, mas na tentativa de esclarecer os mecanismos envolvidos na descrição independente de cada um deles, sem tentar mesclá-los<sup>19</sup>. Para Bickle, não há propriedades emergentes porque existe

---

<sup>19</sup> Este parágrafo constitui um esclarecimento de minha posição em virtude de uma objeção que me foi feita pelo Prof. Alfredo Pereira Júnior. Ele alega ter ficado com a impressão de que meu argumento às vezes parece insinuar que haveria uma identidade entre cérebro e mente no plano semântico, e que a duplicidade (e complementaridade) de descrições decorreria apenas do modo como funciona a linguagem (i.é., sua sintaxe). Com efeito, durante minha fala em São Vicente, sugeri que a noção de *emergência* (e, implicitamente, também a noção de *redução*) seriam ociosas, pois se obtivermos uma completa descrição de cada aspecto oposto poderíamos traçar correspondências unívocas entre as partes das descrições complementares. Para o Prof. Pereira Júnior, talvez nossa divergência esteja em dois pontos: a) ele pensa que a correspondência unívoca seria apenas entre uma parte da atividade cerebral e a atividade consciente, visto que há atividades cerebrais que não participam diretamente do processo consciente; b) mesmo assim, haveria *a priori* inúmeras formas de se traçar a correspondência unívoca, ou seja, seria ainda preciso



redução impiedosa; para a abordagem que estamos propondo, não há propriedades emergentes porque não há redução impiedosa e porque não podemos cruzar domínios heterogêneos de descrição.

## Observações finais

Esperamos ter conseguido mostrar que a nova onda reducionista de Bickle se filia à teoria da identidade mente-cérebro e se inspira no materialismo eliminativo, rejeitando a noção de *propriedade emergente* em sentido forte e a tese da realizabilidade múltipla.

Na discussão dos dois pontos mais importantes ligados à abordagem de Bickle, procuramos, em primeiro lugar, mostrar não só que o conceito de *metaciência de baixo para cima* é insuficiente para explicar o funcionamento da atividade científica, que também exige considerações na direção inversa, mas também que o próprio Bickle trabalha com os dois aspectos considerados.

Em segundo lugar, fizemos uma crítica à viabilidade da "redução" proposta por Bickle, a qual se baseia no modelo de Hooker e vem acrescida de

---

especificar que ponto ou região de um mapa corresponde a que ponto ou região do outro mapa. Enfim, ele está defendendo uma concepção de complementaridade que se relaciona a uma teoria (semântica, ontológica) de duplo aspecto, enquanto a minha versão da complementaridade seria compatível com uma teoria (semântica, ontológica) da identidade ou da "ipsoreferencialidade". Ambas as posições entenderiam que há co-referencialidade, mas com diferentes conotações. Para a teoria de duplo aspecto, há uma co-referencialidade porque ambas as descrições se referem ao mesmo ser (ou melhor, a aspectos diferentes do mesmo ser); para uma teoria da identidade, ambas as descrições se referem ao mesmo ser, enfocando os mesmos aspectos (daí o neologismo *ipsoreferencialidade*). Espero ter deixado claro em minha exposição que, ao contrário do que pensa o Prof. Pereira Júnior, também estou defendendo uma *teoria da identidade de duplo aspecto*, ou aspectos diferentes do mesmo ser, segundo a qual a realidade se revela ontologicamente complexa a ponto de permitir descrições contraditórias. As nossas interações com as coisas nos levam ao duplo aspecto e a complementaridade é a saída lógica que encontramos para lidar linguisticamente com esta dificuldade. Em virtude disso, a noção de *redução* resultaria da tentativa de explicar tudo a partir do nível físico menos elevado, o das configurações moleculares, desprezando as diferenças evidentes entre as diversas posturas linguísticas, envolvidas no caso de descrições mais complexas da mesma realidade. Quanto à noção de *propriedade emergente*, ela ainda decorreria de um certo fascínio reducionista, já que tenta dar conta de um aspecto novo, localizando-o em um nível mais elevado e explicando-o em parte com base em articulações de aspectos já conhecidos, que são colocados em um nível menos elevado. A propriedade emergente resultaria de uma tentativa equivocada de mesclar domínios de descrição diferentes, baseados em posturas linguísticas diferentes. Neste sentido, ambas as noções seriam ociosas. Não há porém espaço para discutir aqui a oposição entre minha posição e a do Prof. Pereira Júnior. O tempo dirá qual de nós dois tem razão.

uma semântica que liga a "imagem redutora" à "teoria reduzida". No decorrer desta crítica, argumentamos que a "redução" efetuada através de uma "imagem" deduzida a partir da "teoria básica" constitui um mecanismo muito artificial que acaba alterando no interior da "imagem" os conceitos fundamentais da "teoria reduzida". Além disso, as relações entre a "imagem" e a "teoria reduzida" podem variar, tornando assim o modelo de Hooker-Bickle insuficiente para dar conta de todos os casos possíveis. Nessa perspectiva, a "redução da nova onda" não parece capaz de atingir suas metas no caso da psicologia, limitando-se a uma superposição extensional de domínios de descrição. Se a "redução impiedosa" fosse possível, ela *nada reduziria*, pois a linguagem da neurociência e a da psicologia seriam equivalentes e não poderíamos decidir qual delas privilegiar, a não ser com base em critérios pragmáticos. A "redução impiedosa" é de fato impossível, pois os bicondicionais envolvidos possibilitam articulações de propriedades contraditórias. Esses bicondicionais deveriam ser abandonados, pois os domínios de descrição da mesma realidade psíquica são complementares, quer dizer, apresentam-se igualmente válidos, embora sejam mutuamente excludentes. Essa posição, que defende a identidade sem redução e com múltiplas descrições, é convergente com diversos aspectos do *fisicalismo não-reducionista* de Davidson e Rorty. Todavia, corresponde a uma postura que seria mais adequadamente denominada *pluralismo descritivo*.

Além disso, sugerimos que a noção de *propriedade emergente* pode resultar da tentativa equivocada de cruzar domínios complementares de descrição, a partir do fascínio criado pelo acentuado contraste entre as descrições da fisiologia e da conduta do sistema vivo. Assim, as investigações desenvolvidas por Bickle poderão levar a um maior conhecimento dos mecanismos moleculares envolvidos pela atividade do cérebro. Podemos supor até que tais investigações levem a explicações neurocientíficas bastante completas e satisfatórias do que ocorre quando percebemos *qualia* ou quando estamos tendo determinada experiência subjetiva. Isso certamente será muito bem recebido. Mas esses resultados não serão capazes de realizar o programa eliminativista do reducionismo impiedoso da nova onda, já que a mera descrição mecânica de seqüências de configurações moleculares constitui apenas uma face da moeda em questão, que somente será melhor compreendida se a outra face, representada pela descrição holística das ações intencionais, também se revelar completa e satisfatória. A complementaridade entre a descrição neurocientífica e

a psicológica continuará tão necessária quanto o é hoje em dia. Talvez fascinado em excesso pelos resultados da neurociência de ponta, Bickle parece ter-se esquecido dessa possibilidade, apostando numa redução impiedosa que, no final das contas, não "reduzirá" coisa alguma, como já o revela a sua metáfora visual e como procuramos mostrar no decorrer desta discussão.

## **Referências**

- ANSCOMBE, G. E. M. *Intención*. Introdução de J. Mosterín. Trad. De A. I. Stellino. Barcelona: Ed. Paidós Ibérica, 1991.
- BACON, F. *Novum organum*. In: BACON, F. *Novum organum ou verdadeiras indicações acerca da interpretação da natureza*: nova Atlântida. Tradução e notas de J. A. Reis de Andrade. 2.ed. São Paulo: Abril Cultural, 1979.
- BICKLE, J. *Philosophy and neuroscience: a ruthlessly reductive account*. Dordrecht: Kluwer Academic Publishers, 2003.
- \_\_\_\_\_. *Psychoneural reduction: the new wave*. Cambridge, Mass.: MIT Press, 1998.
- DENNETT, D. *Brainstorms*. Cambridge, Mass.: Bradford Books, 1978.
- \_\_\_\_\_. *Darwin's dangerous idea*. New York: Simon and Schuster, 1995.
- \_\_\_\_\_. *The intentional stance*. Cambridge, Mass.: Bradford Books, 1987.
- DONNELLAN, K. S. Reference and definite descriptions. *Philosophical Review*, v.75, p. 201-304, 1966.
- DUQUESNE, M. *Matéria e antimatéria*. São Paulo: Difusão Européia do Livro, 1961.
- FEYERABEND, P. Materialism and the mind/body problem. *The Review of Metaphysics*, v.17, 1963.
- \_\_\_\_\_. Mental events and the brain. *The Journal of Philosophy*, v.55, 1963.
- HORGAN, T.; WOODWARD, J. Folk psychology is here to stay. *The Philosophical Review*, v. 94, p. 197-226, 1985.
- KIM, J. *Supervenience and mind*. Cambridge, UK: Cambridge University Press, 1993.
- \_\_\_\_\_. *Supervenience and mind*. Cambridge, UK: Cambridge University Press, 1993.

KRIPKE, S. Speaker's reference and semantic reference. In: FRENCH, P. A. et al. (Ed.). *Contemporary perspectives in the philosophy of language*. Minneapolis: University of Minnesota Press, 1981. p. 6-27.

MARGUTTI PINTO, P. R. Ceticismo, pragmatismo e a crítica de Sellars ao mito do dado. In: DUTRA, L. H. de A.; SMITH, P. (Org.). *Ceticismo: perspectivas históricas e filosóficas*. Florianópolis: NEL/UFSC, 2000, p. 137-168.

MATURANA, H.; VARELA, F. *The tree of knowledge: the biological roots of human understanding*. Revised edition. Transl. by R. Paolucci. Boston and London: Shambhala, 1998.

RORTY, R. Mind-body identity, privacy and categories. *The Review of Metaphysics*, v.19, 1965.

\_\_\_\_\_. Non-reductive physicalism. In: \_\_\_\_\_. *Objectivity, relativism, and truth..* Cambridge, UK: Cambridge University Press, 1991. p. 113-125. (Philosophical papers v. 1).

SELLARS, W. *Empiricism and the philosophy of mind*. Introduction by Richard Rorty and a Study Guide by R. Brandon. Cambridge, Mass.; Londres: Harvard University Press, 1997.

STEPHAN, A. Varieties of emergentism. *Evolution and Cognition*, v.5, n.1, p.49-59, 1999.

\_\_\_\_\_. How to lose the mind-body problem. *Grazer Philosophische Studien*, v.61, p. 277-81, 2001.

STRAWSON, P. F. On referring. *Mind*, v. 59, p. 320-44, 1950.





## Considerações sobre o perspectivismo e a lógica da Ciência Cognitiva

Elias Humberto Alves<sup>1</sup>

Desejo ressaltar, de início, algumas considerações de Henri Atlan, célebre por suas idéias sobre a noção de auto-organização, considerações essas que aparecem numa entrevista com Guita Pessis-Pasternak, no livro *Do Caos à Inteligência Artificial* (1993). O que se pretende, nesse livro de entrevistas, é constatar uma verdadeira revolução epistemológica, ligada ao surgimento de novas teorias científicas que buscam “alcançar a inteligibilidade do Universo”, com a ajuda de instrumentos conceituais, expressos em palavras como “desordem organizadora”, “complexidade”, “caos” e “auto-organização”.

Ora, Atlan afirma em um de seus últimos livros (*A Tort et à Raison*, 1979), que existiriam várias racionalidades legítimas, embora diferentes, capazes de apreender a realidade. Atribuir o raio à cólera de Júpiter, diz ele, não é menos “racional” do que atribuí-lo à descarga elétrica, mesmo que a eficácia de cada método, na dominação da natureza, seja diferente.

Na entrevista com Pessis-Pasternak, surge uma questão fundamental:

Como apreender a “realidade última”?

A resposta de Atlan é:

O melhor é não acreditar nessa “realidade última”!

Mas, contesta a entrevistadora, a pesquisa científica visa justamente “apreender a realidade”!

Depende, diz Atlan, daquilo que se entende por *apreender*. Se acreditamos que, ao elaborar uma teoria que *funcione*, tenhamos “atingido a realidade”, estamos imersos numa ilusão. Pode sempre haver um maior aprofundamento e, por isso, não há “realidade última”.

Mas, insiste ainda a entrevistadora, então uma teoria como a do Big-Bang que *funciona*, não apreenderia a realidade?

---

<sup>1</sup> Docente do Departamento de Filosofia da UNICAMP e da Faculdade de Filosofia de São Bento. Endereço eletrônico: [ehalves@unicamp.br](mailto:ehalves@unicamp.br)

Trata-se, diz Atlan, de uma teoria mais problemática do que as outras. É cosmogênica, impossível de se reproduzir em laboratório. Na verdade, é uma descrição e, no limite, uma lenda. Constatar que essa teoria *funciona* significa apenas que ela é coerente com outras teorias físicas aceitas na atualidade. Mas não se pode dizer que seja a “apreensão” da “realidade última”.

Atlan parece estar, portanto, manifestando-se a favor de uma posição “perspectivista”. Vou agora, depois dessa motivação, procurar esboçar a posição perspectivista, segundo a formulação de Donald Peterson, no prefácio do seu livro *Forms of Representation* (1996)<sup>2</sup>.

Esse autor afirma que não estamos interessados em saber “se nossas crenças e afirmações *representam* o mundo, num sentido abstrato e quase ótico”. O que queremos, diz ele, é saber “qual é o papel de *nossas representações*, nas *ações mentais* de cálculo e processamento requeridas para gerar crenças, derivar soluções e realizar tarefas”.

É importante, diz Peterson, discutir o problema de como o nosso conhecimento se desenvolve. Ora, uma das maneiras seria através da aquisição e aplicação de uma apropriada *forma de representação*.

Uma *forma de representação* é definida como uma notação, com uma interpretação para essa notação. Exemplos: álgebras, alfabetos, linguagens de programação, diagramas, grafos, modelos mecânicos, notações musicais, formalismos lógicos ou matemáticos, etc., etc. (Enfim, tudo o que se usa, muito comumente na Ciência Cognitiva)<sup>3</sup>.

O que de fato está sendo proposto é uma concepção de conhecimento orientado para a ação. O que devemos investigar são processos *orientados* segundo alguma *forma de representação*, que guia a atuação do sujeito no mundo. Não queremos apenas olhar para o mundo ou para um estoque interno de idéias e “conhecer as coisas”. Ao contrário, nós calculamos, manipulamos e *atuamos sobre as coisas*, em resposta a nossas necessidades e, freqüentemente (embora nem sempre) operamos por meio de notações e suas respectivas interpretações (ou seja, as ontologias conceituais associadas). Como ressalta Peterson, nós realizamos tarefas, resolvemos problemas, conseguimos respostas

<sup>2</sup> Retomo aqui, brevemente, uma caracterização da posição perspectivista que apareceu no artigo “Observações sobre o perspectivismo de Donald Peterson”, *Encontros com as Ciências Cognitivas*, Marília: Unesp, 1998, v. 2.

<sup>3</sup> Estou utilizando a expressão *Ciência Cognitiva* num sentido amplo. Assim, fica inteiramente de lado o problema de saber se devemos falar em *Ciência Cognitiva* ou em *Ciências Cognitivas*.

a questões colocadas por outros ou por nós mesmos. E a efetividade em realizar essas tarefas depende, em parte, das formas de representação que usamos, as quais determinam as diferentes *perspectivas* que adotamos no processo de aquisição do conhecimento.

Trata-se, pois, de uma posição essencialmente metodológica. Contudo, convém assinalar, essa posição assemelha-se à do *pluralismo*: necessitamos de uma pluralidade de formas de representação porque, quando todos os fatores são considerados, cada uma das formas de representação é limitada. De fato, a utilidade de uma forma de representação é relativa a vários fatores e, nesse sentido, é *relativista*. Além disso, esses fatores concernem, em parte, ao usuário – sua tarefa, seus conceitos, poderes cognitivos e circunstâncias. Assim, o tratamento adequado das formas de representação é *subjettivista*. Mas isso, ressalta Peterson, é simplesmente o reconhecimento de que há diversos fatores envolvidos na abordagem perspectivista, no processo de aquisição do conhecimento. Não é, de modo algum, um reconhecimento do *caos epistêmico* do tipo em geral associado à palavra relativismo.

O que significa, então, esse caráter multi-disciplinar das formas de representação? Significa apenas que é muito difícil, para uma única forma de representação, satisfazer todos os nossos objetivos epistêmicos. Além disso, esse próprio fato nos fornece uma maneira de avaliar as diversas formas de representação. Podemos perguntar, dada uma forma de representação, em que dimensões ela é vantajosa e se, nessas condições, ela é melhor do que uma alternativa.

Assim, afirma Peterson, as teses do pluralismo e do caos são diferentes e enquanto o pluralismo é sugerido pela multi-disciplinariedade em questão, o caos é obstruído por essa mesma multi-disciplinariedade.

Tudo isso indica, pois, que a posição perspectivista pode se constituir, talvez, na metodologia adequada à Ciência Cognitiva.

Ora, dentre as formas de representação, estaria aquilo que os lógicos chamam de *sistema formal* ou, simplesmente uma lógica (já que se trata de uma *notação* e uma *interpretação* para essa notação). Cabe, pois, a pergunta: Haveria uma lógica mais adequada do que outras, para as investigações da Ciência Cognitiva?

Vou tentar argumentar na direção de uma resposta afirmativa, sugerindo que um bom candidato seria uma *lógica intensional*, isto é, uma lógica dos

*sentidos e denotações*, segundo a proposta de Frege, no célebre artigo “On Sense and Reference”.

Para Frege, os sentidos e as denotações são componentes fundamentais e irreduzíveis do significado das sentenças (essa distinção é logo ignorada pela maioria dos lógicos, como, por exemplo, Russell, no *Principia Mathematica*).

A partir das idéias de Frege, Carnap (no livro *Meaning and Necessity*, 1956) e Church (no artigo “A Formulation of the Logic of Sense and Denotation”) formularam, com precisão, esse tipo de lógica.

Entretanto, como se sabe, foi somente com Montague que essa lógica foi plenamente desenvolvida, na série de artigos que aparecem no livro *Formal Philosophy* (1974).

A lógica intensional de Montague, além de ser uma lógica intensional, no sentido de Frege, é, ainda, uma lógica extremamente forte, já que contém o sistema *S5* de lógica modal de Lewis e também um sistema de lógica temporal. Além disso, é uma lógica de ordem superior, sendo, portanto, uma extensão da lógica clássica de primeira ordem. Desse modo, é “equivalente”, num certo sentido, à teoria de conjuntos.

O desenvolvimento da lógica intensional se dá, posteriormente, em direção à chamada *lógica ilocucionária*, ou seja, uma tentativa de se obter uma lógica mais ampla que a de Montague, a qual, apesar de tudo, se limitaria às condições veritativas das proposições, excluindo, portanto, como faz a lógica clássica, as ordens, as questões e inúmeras outras cláusulas.

O desenvolvimento se dá a partir das considerações de Austin (*How to do the Things with Words*, 1962) e de Searle (*Speech Acts*, 1969).

Esse último autor, juntamente com D. Vanderveken, constrói uma teoria geral dos atos de fala (*Foundations of Illocutionary Logic*) e o próprio Vanderveken, numa série de trabalhos, desenvolve exaustivamente a lógica ilocucionária<sup>4</sup>.

Vanderveken observa que a lógica ilocucionária tem grande interesse filosófico. Trata-se de uma lógica ampla da *linguagem natural*, que aborda aspectos *transcendentes* da linguagem, como *sentidos, denotações, forças ilocucionárias e atos de fala, verdade e satisfação, analiticidade, validade,*

---

<sup>4</sup> Ver D. Vanderveken, *Meaning and Speech Acts* (1990), onde se mostra que a lógica ilocucionária é uma *extensão conservativa* (no sentido de Shoenfield) da lógica de Montague.

*implicação e consistência*, que fazem parte de qualquer interpretação semântica da linguagem natural.<sup>5</sup>

Tais aspectos transcendentais são *universais* porque são indispensáveis, já que uma linguagem natural desprovida de tais aspectos seria incapaz de cumprir todas as suas *funções essenciais*, quais sejam, fornecer aos falantes humanos meios adequados de expressão e comunicação para os seus pensamentos.

Estamos falando, pois, de uma lógica dos aspectos universais da linguagem, cujo objetivo é descobrir a estrutura profunda, comum a todas as línguas.

Trata-se de procurar encontrar os diferentes princípios de inferências válidas correspondentes a todos os tipos de implicação, não só das implicações veri-condicionais (da lógica clássica), mas também das implicações ilocucionárias.

A partir daí, coloca-se uma questão que, no nosso contexto, é muito importante: *O que podemos aprender, a partir dessas leis ilocucionárias, sobre a natureza da mente humana?*

Vanderveken observa que, sendo a linguagem uma ferramenta necessária na expressão e na comunicação do pensamento, segue-se que todo *pensamento humano conceitual* é, em princípio, expressável, por meio da linguagem, na realização de um ato ilocucionário. De fato, os atos ilocucionários, além de serem as unidades primárias da significação, no uso das línguas naturais, seriam também as unidades do pensamento conceitual<sup>6</sup>.

Ao formalizar as leis da fala, a lógica ilocucionária trata da *razão* e da *racionalidade*. Temos, portanto, desse ponto de vista, que *linguagem, pensamento e razão* são inseparáveis.

É nesse sentido que devemos entender a afirmação de Vanderveken, quando esse autor nos diz que as leis da lógica ilocucionária podem ter um grande impacto sobre a Filosofia da Mente.

De fato, a partir das leis da lógica ilocucionária, na medida em que elas regulamentam as inferências válidas, poderíamos talvez, aprender algo (dentro

---

<sup>5</sup> Ver W.V. Quine, *Philosophy of Logic* (1970), onde é feita uma distinção entre aspectos lingüísticos *transcendentese imanentes*.

<sup>6</sup> Ver J. Searle, *Intentionality* (1963), onde se faz uma distinção entre *pensamento conceitual* e pensamento ligado à percepção e à imaginação.



de certa perspectiva) sobre a natureza da mente humana, contribuindo, dessa maneira, para o desenvolvimento de certas questões em Ciência Cognitiva.

## **Referências**

- ATLAN, H. *A tort et à raison*. Paris: Éditions du Seuil, 1979.
- AUSTIN, J.L. *How to do things with words*. Oxford: Clarendon Press, 1962.
- CARNAP, R. *Meaning and necessity*. University of Chicago Press, 1956.
- CHURCH, A. A Formalization of the logic of sense and denotation. In: HENLE, P.; KALLEN, H.; LANGE, S. (Ed.) *Structure, method and meaning*. New York: Liberal Art Press, 1951.
- FREGE, G. On sense and reference. In: GEACH, P.; BLACK, M. (Ed.) *Translations from the philosophical writings of Gottlob Frege*. Oxford: Blackwell, 1970.
- MONTAGUE, R. *Formal philosophy*. Yale University Press, 1974.
- PETERSON, D. *Forms of representation: an interdisciplinary theme for cognitive science*. Wiltshire: Cromwell Press, 1996.
- PESSIS-PASTERNAK, G. *Do caos à inteligência artificial*. São Paulo: EdUNESP, 1993.
- QUINE, W.V. *Philosophy of logic*. Englewood Cliffs, NJ: Prentice-Hall, 1970.
- SEARLE, J.R. *Intentionality*. Cambridge: Cambridge University Press, 1963.
- \_\_\_\_\_. *Speech acts*. Cambridge: Cambridge University Press, 1969.
- SEARLE, J.R.; VANDERVEKEN, D. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press, 1985.
- SHOENFIELD, J-R. *Mathematical logic*. Reading, MA: Addison-Wesley, 1967.
- VANDERVEKEN, D. *Meaning and speech acts*. Cambridge: Cambridge University Press, 1990. 2v.

# O naturalismo biológico de John Searle e as ambigüidades de uma concepção internalista da mente

Mariana Claudia Broens<sup>1</sup>

Carmen Beatriz Milidoni<sup>2</sup>

No presente trabalho procuramos discutir alguns aspectos da concepção internalista de mente defendida por John Searle. Por internalismo estamos aqui compreendendo, regra geral, a tese segundo a qual os eventos mentais supervêm de eventos internos ao sujeito. Esta tese se contrapõe à tese externalista, a qual postula que também há eventos mentais que resultam (ou supervêm) de eventos ambientais<sup>3</sup>. Procuraremos mostrar que a concepção internalista de mente de Searle é problemática na medida em que pretende, simultaneamente, situar-se num contexto naturalista e preservar uma experiência subjetiva de primeira pessoa.

No texto *A redescoberta da mente* (1997), Searle apresenta inicialmente algumas teses materialistas que considera equivocadas e que, segundo ele, têm em comum o fato de negarem a dimensão subjetiva, interna e de primeira pessoa da vida mental do indivíduo. Segundo Searle, em nome de uma compreensão canhestra da objetividade da investigação científica, recusa-se a possibilidade de que haja algo assim como uma realidade subjetiva quando se trata da vivência mental que cada indivíduo possui.

---

<sup>1</sup> Departamento de Filosofia da Faculdade de Filosofia e Ciências da UNESP – Campus de Marília-SP. Endereço eletrônico: mbroens@uol.com.br

<sup>2</sup> Departamento de Filosofia da Faculdade de Filosofia e Ciências da UNESP – Campus de Marília-SP. Endereço eletrônico: milidoni@marilia.unesp.br

<sup>3</sup> O exemplo clássico de Putnam (1975) diz respeito a uma experiência de pensamento que, segundo ele, evidencia que o significado está no mundo, e não apenas 'na cabeça', ao contrário do que defendem várias teses internalistas. Putnam utiliza o célebre argumento da Terra-Gêmea, segundo o qual é postulada a existência de um planeta, a Terra Gêmea, em tudo igual à Terra, com exceção da composição química da água. Nesse outro planeta, embora a água pareça idêntica à água terrestre externamente, sua composição não é H<sub>2</sub>O, mas XYZ. Quando um habitante da terra gêmea diz 'água', embora o evento mental dele seja idêntico ao de um habitante terrestre que também diz 'água', o significado do termo não será o mesmo na Terra e na Terra Gêmea, uma vez que a composição química de ambos líquidos é diferente.

Searle considera que “os fenômenos mentais são causados por processos neurofisiológicos do cérebro, e são, eles próprios, características do cérebro... Os processos e fatos mentais fazem parte de nossa história natural biológica tanto quanto a digestão, a mitose, a meiose ou a secreção enzimática” (SEARLE, 1997, p. 7). Ao caracterizar assim seu naturalismo biológico, Searle aponta que a investigação dos processos físico-químicos do cérebro que causam a vida mental cabe às neurociências. Mas, ele indica, interrogações sobre a natureza da consciência e de todas suas atividades devem ser respondidas pela Filosofia e pela Psicologia, utilizando uma metodologia investigativa própria desses saberes.

Embora a Filosofia da Mente deva manter uma interlocução constante com as ciências positivas que buscam elucidar os processos fisiológicos responsáveis pela vida mental, para Searle os fenômenos mentais tais como são experienciados pelo indivíduo devem ser abordados por uma perspectiva teórica que considere seu caráter subjetivo.

Para Searle parece evidente, pois, que a explicação científica do funcionamento do cérebro e de como, efetivamente, ele causa a vida mental não alterará o fato de que o indivíduo tem uma experiência subjetiva e interna de sua própria consciência e de seus estados mentais.

Assumir como dado inegável o caráter subjetivo da vida mental do indivíduo não significa, segundo Searle, assumir uma postura cartesiana: afinal pode-se afirmar que há mente porque há um cérebro. A subjetividade a que se refere não é a subjetividade introspectiva do *cogito* que, para ser determinada, precisou abrir mão da interação perceptiva com o mundo físico e precisou, igualmente, abrir mão (mesmo que apenas metodologicamente) do próprio corpo.

Searle argumenta que, se as investigações da Filosofia da Mente encontram-se ainda impregnadas de um ranço cartesiano, este está presente nas escolas materialistas. Ele afirma, com efeito, que boa parte dos erros cometidos pelos diferentes tipos de materialismo deve-se a que, embora o dualismo substancial seja considerado por todos como inaceitável, concepções como monismo anômalo ou dualismo de propriedades, antimentalismo ou mesmo o próprio materialismo mantém, à revelia, o cartesianismo como ponto de partida para suas análises.

Segundo Searle, um exemplo deste erro (talvez o mais grave) surge quando os filósofos da mente entendem que criticar o materialismo necessariamente implica assumir teses do tipo cartesiano ou situar-se na perspectiva do dualismo substancial. Atribuir ao pensamento cartesiano o papel de interlocutor privilegiado, isto é, colocar o cartesianismo num pólo oposto que se deve combater já mostra por si um vício de origem.

Dando seqüência a sua argumentação, Searle afirma que uma das mais recentes escolas materialistas que se recusa a admitir o caráter subjetivo e interno de nossa vida mental é o chamado Externalismo.

Para ele, o procedimento adotado pelo Externalismo consiste em naturalizar a intencionalidade, isto é, considerar que o conteúdo semântico não é apenas interno ao indivíduo, pois precisa de: “[...] um conjunto de relações causais físicas reais com os objetos do mundo” (SEARLE, 1997, p. 75). Uma versão mais forte de “naturalizar a intencionalidade, e que teria servido de base para o Externalismo, foi a do funcionalismo do primeiro Putnam (1975), para o qual “[...] seria preferível uma explicação do conteúdo intencional formulada exclusivamente em termos de relações causais entre pessoas, por um lado, e objetos e estados de coisas no mundo, por outro” (Searle, 1997, p. 75). Em consonância com (parte) deste pensamento, na perspectiva externalista, as crenças, os desejos, etc. não poderiam ser explicados como estados mentais internos, pois, de algum modo, estariam também fora do indivíduo, estariam no mundo.

O erro externalista seria, para Searle, semelhante aos erros cometidos pelas demais abordagens contemporâneas, qual seja, preterir a realidade subjetiva dos estados mentais ao exteriorizar tais estados. Ele observa que o senso comum poderia argumentar contra a tese externalista mostrando que naturalizar a intencionalidade implicaria que a intencionalidade deixasse de ser ela própria e, podemos acrescentar, exteriorizar a mente faria que a mente deixasse de ser mente e passasse a ser meio ambiente. Searle insiste: consciência e intencionalidade têm uma realidade subjetiva.

Adepto de uma terceira via, Searle pretende mostrar que é possível simultaneamente afirmar que os estados mentais têm uma realidade subjetiva e que a mente resulta de processos físico-químicos do cérebro sem repetir os equívocos do materialismo em suas várias versões e do dualismo substancial cartesiano, os quais, a rigor, são, segundo ele, verso e reverso da mesma moeda.

No entanto, parece-nos que a subjetividade de nossa vida mental em geral, defendida por Searle e considerada por ele como uma evidência incontestável, é de alguma forma afetada pelo próprio conceito searleano de *background*.

Searle caracteriza o *background* como sendo constituído: "... pelas capacidades, aptidões e *know-how* geral que possibilitam que nossos estados mentais atuem" (SEARLE, 1997, p. 249). Condição necessária dos estados intencionais, o *background*, segundo Searle, não é, ele próprio, intencional. Os estados intencionais não são autônomos, para que se efetivem é necessário todo um conjunto de atividades de *background*.

Ele sugere que essas capacidades de *background* sejam divididas em dois tipos: um *background* profundo, próprio da espécie e que resulta de nossa história biológica e um *background* local, relativo a práticas culturais ou a formas pontuais de interação com um meio ambiente específico.

O ponto mais significativo da apresentação deste importante conceito consiste no que Searle denomina "leis de funcionamento do *background*". Ressaltaremos duas. A primeira aponta que: "em geral, não há nenhuma ação sem percepção, nenhuma percepção sem ação" (SEARLE, 1997, p. 278). E a segunda considera que: "a intencionalidade ocorre em um fluxo coordenado de ação e percepção, e o *background* é a condição de possibilidade das formas assumidas pelo fluxo" (SEARLE, 1997, p. 278).

Em outras palavras, particularmente no que diz respeito ao *background* local, sendo esta condição para que possam ocorrer atos intencionais, podemos perceber que há uma espécie de "exteriorização" de certos aspectos da vida mental do indivíduo. Em se tratando de certas aptidões e capacidades de *background*, o que se tem é um processo dinâmico não mediado pela consciência: não há nesses casos uma ação prévia consciente que determine a interação do indivíduo com o meio.

Searle parte da noção comum de subjetividade da vida mental e, como apontamos, considera-a uma espécie de dado inquestionável. No entanto, ele acaba por alterar de algum modo essa noção do senso comum ou da experiência individual ao postular o conceito de *background* enquanto condição de possibilidade da interação intencional do indivíduo com o mundo.

Em artigo dedicado ao problema do auto-conhecimento na perspectiva externalista, Akeel Bilgrami (1992) propõe caracterizar o externalismo em geral



como sendo a negação da tese segundo a qual: "... estados intencionais ... não pressupõem a existência de nada externo ao agente possuidor de tais estados" (p. 363).

Considerando esta tese básica, podemos perceber que o conceito searleano de *background* é compatível com a tese externalista. Porém, ao atribuir à consciência uma realidade subjetiva, embora causada pelo cérebro, Searle desperta a desconfiança mesmo no leitor mais generoso.

Mas façamos uma recapitulação dos pontos nodais da concepção de mente searlina. Com o intuito de apontar as propriedades que seriam essenciais aos fenômenos mentais e que serviriam para circunscrever o domínio do mental enquanto correspondente àquilo que satisfizer tais propriedades, em sua obra *Mente, cérebro e ciência* (1987), Searle escolhe, como traços "privilegiados" do mental, os seguintes: *consciência* (o aparecer, ser fenomênico para); *intencionalidade* (a característica de um estado de um sistema de ser acerca de objetos e estados de coisas no mundo diferentes dele próprio); *subjetividade* (a propriedade de um indivíduo de ser consciente de si mesmo e de seus estados mentais "internos" como algo inteiramente distinto da individualidade e dos estados mentais de outras pessoas); *causação mental* (o fato de nossos pensamentos e sentimentos terem algum efeito causal sobre o mundo físico) (SEARLE, 1987, p. 21-2).

Por outro lado, e como já consignamos, os fenômenos mentais seriam explicados como causados por e realizados em processos cerebrais. Esta explicação repousaria sobre um tipo de redução physicalista que Searle denomina *redução causal*. A mesma é caracterizada como sendo "(...) uma relação entre quaisquer dois tipos de coisas que possam ter capacidades causais, em que se demonstra que a existência e, *a fortiori*, as capacidades causais da entidade reduzida são inteiramente explicáveis em termos das faculdades causais dos fenômenos redutores" (SEARLE, 1997, p. 166).

Mas Searle faz questão de esclarecer que esta visão das relações entre mente/cérebro como uma forma de *redução causal* não implica uma *redução ontológica*, no sentido de que os fenômenos mentais seriam *nada exceto* processos cerebrais (SEARLE, 1997, p. 163 e seguintes). Todavia, e para que não restem dúvidas quanto à sua ontologia do mental, Searle sublinhará sua adesão à tese mentalista, chamada por ele de "ingênua", que sustenta a

existência de fenômenos especificamente mentais, tais como crenças, intenções, desejos (SEARLE, 1987, p. 33).

Esta tese, perfeitamente clara e sustentável como opção filosófica, se vê, no entanto, obscurecida quando Searle se ocupa da ontologia da consciência. Para este autor, a consciência não pode ser definida em termos de “realidade”, porque a consciência consiste nas próprias aparências. Diz Searle: “*Onde a aparência está envolvida, não podemos fazer a distinção aparência-realidade porque a aparência é a própria realidade*” (SEARLE, 1997, 176, grifo do autor). E perguntamos: que tipo de redução causal dos fenômenos mentais seria adequado para explicar causalmente a consciência? Não haveria aqui uma espécie de dissonância entre o naturalismo biológico searleano e uma abordagem “fenomenológica” da consciência?

Se articularmos agora os diferentes planos de aproximação ao universo do mental acima considerados (planos lógico, causal e ontológico) teremos que admitir a maestria de Searle ao conseguir conservar a tese mentalista clássica de que haveria estados propriamente mentais dotados de propriedades que os fariam ontologicamente irreduzíveis a qualquer outra coisa e ao mesmo tempo superar a opção entre dualismo e materialismo reducionista (tanto no sentido ontológico quanto lingüístico da redução), opção essa que, para nosso filósofo, seria uma falsa opção.

Sem embargo, devemos ressaltar que a validade dessas conseqüências depende da noção de causação, do nível micro (processos cerebrais) para o nível macro (fenômenos mentais), e que a dita noção não é suficientemente esclarecida por Searle.

Voltando à temática da situação do pensamento searleano em termos da oposição “externalismo/internalismo”, julgamos que seria necessário, para afastar desse pensamento qualquer suspeita de incompatibilidade com a opção externalista, determinar se existe algum sentido em que os processos cerebrais possam ser considerados “externos” com relação a seus efeitos, quais sejam, os fenômenos mentais de um certo sujeito. Esta não parece ser tarefa fácil, já que Searle rejeita – se bem que perifrásticamente – a concepção humeana de causalidade, que pressupõe a exterioridade da causa relativamente a seu efeito. (SEARLE, 1987, p. 25-6).

Para concluir, entendemos que Searle corretamente denuncia a pesada carga semântica de nosso vocabulário, ainda repleto de entulhos cartesianos.

Resta indagar se, a despeito de todo seu esforço e a despeito da busca de objetividade presente no conceito de *background*, Searle não acaba também por sucumbir diante do mesmo problema.

Há, efetivamente, um conflito entre as concepções internalistas e externalistas da mente. Entre o internalismo radical cartesiano, para o qual nossa vida mental independe de qualquer interação com o mundo físico, e o externalismo radical, para o qual não há sequer representações mentais que cumpram o papel mediador do indivíduo com o meio ambiente, talvez não seja possível assumir, como pretende Searle, uma via intermediária conciliatória.

Se, de um lado, algumas teses externalistas e suas implicações causam um certo desconforto justamente porque contrariam algumas aparentes evidências intuitivas de nossa experiência mental, por outro lado parece inegável que, à luz das investigações contemporâneas a respeito da natureza de nossos processos mentais, é cada vez mais difícil permanecer internalista.

Entendemos, além disso, que postular uma subjetividade de primeira pessoa implica estabelecer artificialmente um recorte entre mente (interna/subjetiva) e ambiente (externo/objetivo). Postular tal recorte significa criar artificialmente uma cisão entre o agente e o meio em que ele age. Dissociar agente e meio é altamente contestável da perspectiva de nossa história evolutiva: se os seres humanos são dotados de certas aptidões cognitivas, como a denominada 'consciência', isso parece ter ocorrido como resultado de uma dinâmica de interação com o meio ambiente ao longo do tempo. Podemos indagar, por fim, até que ponto a postulação de uma experiência subjetiva de primeira pessoa como o principal componente da vida mental humana não implica o esquecimento dessa dinâmica. É precisamente a relevância da interação agente/meio que vem sendo defendida por uma vertente da Ciência Cognitiva denominada *Cognição situada e incorporada*.

Na perspectiva dos teóricos de tal vertente<sup>4</sup>, as leituras tradicionais da Filosofia da Mente e da Ciência Cognitiva tendem a preterir, para dizer o mínimo, a relevância da corporeidade e suas complexas interações ambientais na constituição das habilidades cognitivas dos seres vivos. Isto ocorre, em especial, quando os tradicionalistas criam modelos funcionais das aptidões ditas

---

<sup>4</sup> Podemos citar, por exemplo, Chiel, H.J. & Beer, R.D. (1997), Clark, A. (1997), Thelen, E. (2000), Knoblich, G. (2002), Haselager, W.F.G., de Groot, A.D. e van Rappard, J.F.H. (2003).

“racionais” do ser humano, isto é, da habilidade de resolver problemas aplicando regras lógicas.

Adotando um ponto de vista diverso, que consideramos da mais alta relevância, a cognição incorporada e situada entende que as aptidões cognitivas dos seres vivos não podem ser reduzidas à mera obediência a regras lógicas. Isto porque a ação inteligente, em especial a que se destina a garantir a sobrevivência do organismo, é frequentemente *reativa* diante de desafios imprevistos postos pelas mais variadas condições ambientais e não *lógico-dedutiva*, que resultaria de um prévio planejamento mental interno.

## Referências

- BILGRAMI, A. Can externalism be reconciled with self-knowledge? In: *Philosophical Topics*, v.15, p. 191-226, 1992.
- CHIEL, H.J. e BEER, R.D. The brain has a body: adaptive behavior emerges from interactions of nervous system, body and environment. In: *Trends in Neurosciences*, 20(12), 553- 557, 1997.
- CLARK, A. *Being there: putting brain, body and world together again*. Cambridge, Massachusetts: MIT- Press.
- HASELAGER, W.F.G., de GROOT, A.D. e van RAPPARD, J.F.H. Representationalism versus anti-representationalism: a debate for the sake of appearance. *Philosophical Psychology*, 16(1), 5-23, 2003.
- LUDLOW, P.; MARTIN, N. *Externalism and self-knowledge*. Califórnia: Center for the Study of Language and Information – CSLI Publications, 1998.
- PUTNAM, H. The meaning of 'meaning'. In: GUNDERSON, K. (Ed.) *Language, mind and knowledge*. Minnesota: University of Minnesota, 1975.
- SEARLE, J. *A redescoberta da mente*. Tradução de Eduardo Pereira E. Ferreira. São Paulo: Martins Fontes, 1997.
- \_\_\_\_\_. *Mente, cérebro e ciência*. Tradução de Artur Mourão. Lisboa: Edições 70, 1987.
- THELEN, E. Grounded in the world: developmental origins of the embodied mind. In: *Infancy*, 1(1), 3-28, 2000.



# Bergson: consciência, corpo e ação

*Jonas Gonçalves Coelho<sup>1</sup>*

## I

Que interesse haveria, além do histórico, em retomar as idéias apresentadas pelo filósofo Henri Bergson no final do século XIX e início do século XX sobre a relação entre corpo e consciência, considerando-se o grande desenvolvimento da ciência – Psicologia Cognitiva, Neurociência, Inteligência Artificial, etc. – e as contribuições da Filosofia da Mente posteriores a esse período?

Entendemos que o interesse está na maneira original pela qual Bergson relaciona o surgimento e os vários modos de consciência, da mais simples percepção às mais sofisticadas operações da inteligência, à estrutura e modo de funcionamento do sistema nervoso e seus respectivos corpos, ao longo da evolução dos seres vivos. Consideramos que a originalidade bergsoniana consiste na importância concedida à ação na produção dos processos conscientes. Como veremos, a consciência, tal como descrita por Bergson, é uma consciência incorporada, intimamente associada à inserção pragmática dos corpos no mundo que os cerca; em outras palavras, há um íntimo entrelaçamento entre percepção consciente e ação cuja forma elementar é o movimento no ambiente.

Procurando desenvolver essa idéia, torná-la mais clara, apresentaremos a concepção bergsoniana sobre a relação entre consciência, corpo e ação na evolução da vida sobre a Terra. Para Bergson, a consciência constitui-se nos organismos nos quais a mobilidade prevalece sobre a fixidez e a ação livre prevalece sobre o automatismo. Consideraremos, em separado, cada uma dessas oposições, estabelecendo a sua relação com os reinos vegetal e animal e, no caso dos animais, a diferença entre os modos de consciência dos predominantemente instintivos e dos predominantemente inteligentes.

---

<sup>1</sup> Professor Assistente Doutor de Filosofia do Departamento de Ciências Humanas da Unesp de Bauru e do Programa de Pós-Graduação Interdisciplinar em Filosofia da Unesp de Marília.



## II

Para Bergson, o mundo vegetal e o mundo animal constituem-se como dois desenvolvimentos divergentes da vida. Embora os dois reinos não sejam absolutamente distintos, há uma diferença de ênfase, de predominância de determinadas características que tendem a se acentuar no decorrer da evolução, definindo, de maneira mais precisa, cada um dos grupos. “Não há manifestação da vida que não contenha, em estado rudimentar, latente ou virtual, as características essenciais da maioria das demais manifestações. A diferença está nas proporções” (BERGSON 1991, p. 107). Deve-se, assim, procurar identificar as tendências predominantes nos vegetais e nos animais. É, então, sob o ponto de vista segundo o qual as mesmas características essenciais estão presentes em todas as manifestações de vida, embora em proporções diferentes, que Bergson considera, primeiramente, o desenvolvimento divergente dos vegetais e dos animais.

Um dos principais aspectos que os distingue é o modo de alimentação. Os vegetais tomam os elementos indispensáveis à sua sobrevivência — carbono e nitrogênio — diretamente do ar, da água e da terra, diferentemente dos animais que se apropriam desses elementos através de outros organismos nos quais estão fixados — vegetais e outros animais. Essa diferença não é absoluta, existindo várias exceções de um lado e de outro — plantas insetívoras de um lado e cogumelos de outro —, mas, como indicado anteriormente, o relevante para Bergson é menos a presença excepcional de determinadas características do que a tendência predominante em cada um dos reinos.

Essa diferença no modo de alimentação dos animais e vegetais está geralmente associada a uma outra diretamente relevante para o surgimento da consciência: a mobilidade no espaço. Os animais, diferentemente das plantas, devem se movimentar para obter o seu alimento, ou seja, se no caso das plantas os elementos fornecidos continuamente pelo meio podem ser extraídos sem movimento, maquinalmente, no caso dos animais, são obtidos dos “organismos que já os fixaram”, por meio de uma “ação descontínua concentrada em alguns instantes ... ciente” (BERGSON 1991, p. 115).

Segundo Bergson, as tendências à fixidez e à mobilidade diferenciadoras dos reinos animal e vegetal, são “indícios superficiais de tendências mais profundas”. Essas são, do lado dos animais, a consciência — desperta — e a

sensibilidade e, do lado dos vegetais, a inconsciência (consciência adormecida), a insensibilidade<sup>2</sup>. Isso porque a atividade motora dos organismos tornou possível o surgimento da consciência, enquanto sua falta a atrofia ou adormece.

Bergson postula, portanto, que há uma tendência à inconsciência nos vegetais associando-a a sua imobilidade constitutiva. A presença de uma membrana de celulose envolvendo o protoplasma imobiliza o organismo vegetal simples e o protege das excitações externas. Ao mesmo tempo, esses vegetais fabricam diretamente as substâncias orgânicas a partir dos minerais sem que necessitem locomover-se para nutrir-se: “esta aptidão o dispensa, em geral, de mover-se e, por isso mesmo, de sentir” (BERGSON 1991, p. 113).

Também nesse caso não se pretende que fixidez e mobilidade sejam atributos que distinguem, de maneira absoluta, esses dois gêneros de vida. Afinal, não se pode ignorar a existência de plantas trepadeiras e insetívoras e dos parasitas no mundo animal. Assim, essa distinção em termos de mobilidade não é válida universalmente visto que também se podem observar, do lado dos animais, os crustáceos, cuja fixidez e parasitismo acompanham a “degeneração e o quase desaparecimento do sistema nervoso” e nos vegetais, a locomoção dos zoósporos e das algas.

Para Bergson, o fato de as tendências divergentes, mobilidade/consciência e fixidez/inconsciência, características dos dois reinos, coexistirem nos mesmos, diferindo apenas a proporção, indica que as células animais e vegetais tiveram uma origem comum e que os primeiros organismos “oscilaram entre a forma vegetal e a forma animal, participando de uma e de outra ao mesmo tempo” (BERGSON 1991, p. 113).

Via de regra, uma das duas tendências encobre ou esmaga a outra, mas, em circunstâncias excepcionais, esta se separa e reconquista o lugar perdido. A mobilidade e a consciência da célula vegetal não estão adormecidas a tal ponto que não possam despertar quando as circunstâncias o permitem ou o exigem. E, por outro lado, a evolução do reino animal foi, sem cessar, retardada, detida ou arrastada para trás pela tendência que conservou à vida vegetativa. Por mais plena, por mais transbordante que possa de fato parecer a atividade de uma espécie animal, o torpor e a inconsciência a espreitam. Ela só mantém

---

<sup>2</sup> Em seu ensaio *Le possible et le réel* Bergson diz: “o vivente é consciente de direito; ele torna-se inconsciente, de fato, aí onde a consciência dorme, mas, até nas regiões onde a consciência dorme, no vegetal, por exemplo, há uma evolução regulada, progresso definido, envelhecimento, enfim, os sinais exteriores da duração que caracteriza a consciência [...]” (BERGSON 1993b, p. 101).

seu papel por um esforço, ao preço de um cansaço. Ao longo do caminho no qual o animal evoluiu, ocorreram fraquezas sem número, fracassos que se devem, na maior parte, a hábitos parasitários; são como os desvios ferroviários na vida vegetativa. Assim, tudo nos faz supor que o vegetal e o animal descendem de um antepassado comum que reunia, no estado nascente, as tendências de um e de outro. (BERGSON 1991, p. 114).

Assim, para o filósofo, a fixidez no animal parece, em geral, um torpor em que a espécie caiu, uma “recusa em evoluir”, enquanto que os movimentos dos vegetais não são tão freqüentes, variados e amplos como nos animais, parecendo mais o despertar de uma atividade adormecida. É nesse sentido que se postula que a mobilidade e a fixidez sejam apenas uma tendência dominante ou a característica essencial, embora não exclusiva, de cada um dos reinos:

[...] se a mobilidade e a fixidez coexistem no mundo vegetal como no mundo animal, o equilíbrio é manifestamente rompido em favor da fixidez num caso e da mobilidade do outro. Essas duas tendências opostas são tão evidentemente diretrizes das duas evoluções que se poderia já definir mediante elas os dois reinos. (BERGSON 1991, p. 110)

Essa mesma oposição entre mobilidade e fixidez que permite a Bergson distinguir os vegetais (adormecidos em sua imobilidade) dos animais, pode, como anteriormente indicado, ser também encontrada no próprio mundo animal. A fauna dos tempos primitivos mostra, o que para Bergson é impressionante, animais aprisionados “num invólucro mais ou menos rígido”, que prejudicava e até impedia os seus movimentos: os moluscos eram mais freqüentemente dotados de conchas, os artrópodes providos de carapaças, e até os peixes possuíam invólucro ósseo extremamente rígido. Essa couraça tinha, certamente, a função de proteger esses animais, tornando-os, organismos flácidos que eram, “tanto quanto possível, indevoráveis” (BERGSON 1991, p. 131). Pelo fato de esses invólucros prejudicarem os movimentos e até os imobilizarem, esses animais enclausurados condenaram-se a uma sonolência, à inconsciência. Para Bergson, os equinodermos e mesmo os moluscos vivem até hoje nesse torpor. Já os artrópodes e os vertebrados escaparam da inconsciência pelo fato de privilegiarem a locomoção, na busca de seu alimento, vegetais ou outros animais, ou na fuga de seus predadores. Enquanto os peixes substituíram sua couraça por escamas, os insetos também desembaraçam-se da sua, defendendo-se pela agilidade que lhes permite atacar e fugir no momento oportuno. Para

Bergson, é esse fato — “feliz circunstância” — que possibilitou o “atual desabrochar das mais elevadas formas de vida” (BERGSON, 1991, p. 132). Como diz o filósofo, na evolução conjunta da vida, “os mais retumbantes êxitos couberam àqueles que aceitaram os maiores riscos” (BERGSON, 1991, p. 133). No processo evolutivo, acabou prevalecendo no reino animal o fator que o separou do mundo vegetal, o movimento e, conseqüentemente, a consciência, apesar de entorpecimentos localizados.

Como veremos, a seguir, o movimento é uma condição necessária mas não suficiente para o surgimento da consciência. Um outro fator fundamental privilegiado por Bergson é a liberdade, entendida como oposta ao automatismo. Isso significa, como também veremos, que existe uma íntima relação entre a riqueza da percepção consciente e a variedade de ações possíveis.

### III

Segundo Bergson, a relação entre percepção consciente e ação está associada à história do desenvolvimento do sistema nervoso, desde sua forma mais simples na série animal até o homem. Essa relação mantém-se com o aumento de complexidade do sistema nervoso, isto é, a complexidade crescente do sistema nervoso não produz mudanças essenciais no que diz respeito à relação entre percepção e ação. Tal relação já pode ser encontrada na massa protoplasmática que recebe a influência de estímulos externos e responde a eles através de reações mecânicas, físicas e químicas. Está também presente nas séries de células nervosas agrupadas em sistemas, nos quais se reage à excitação exterior por movimentos variados. É também constitutiva dos organismos superiores, nos quais se radicaliza a distinção entre o automatismo relacionado à medula e a atividade voluntária articulada com o cérebro. Desse modo, a diferença entre as formas mais simples e as mais complexas é que o estímulo recebido, entra em relação com uma variedade cada vez mais considerável de aparelhos motores, o que faz com que seja esboçado simultaneamente um número cada vez maior de ações possíveis.

Ainda segundo o filósofo, o desenvolvimento do sistema nervoso desde a monera, passando pelos insetos mais bem dotados, até os vertebrados mais inteligentes, consiste no desenvolvimento e na conciliação de duas tendências antagônicas: “adaptação mais rigorosa dos movimentos” e “maior margem



deixada ao ser vivo para escolher entre eles”. No organismo humano, e no organismo animal em geral — em proporções variáveis —, tem-se um “número considerável de mecanismos motores” montados na medula e no bulbo os quais podem ser combinados de diversas maneiras. Um cérebro desenvolvido tem um maior número de mecanismos motores a serem escolhidos. Uma encruzilhada mais complicada onde se cruzam as vias motoras permite os movimentos mais variados. “O organismo se conduz cada vez mais como uma máquina para agir que se reconstruísse inteiramente a cada ação nova, como se ela fosse de borracha e pudesse, a qualquer instante, mudar a forma de todas as suas peças” (BERGSON 1991, p. 253). É nesse sentido que se considera o ser vivo como um “centro de ação”, e, conseqüentemente, “determinada soma de contingência introduzindo-se no mundo”. A quantidade e qualidade da ação possível, bem como a extensão da escolha das ações variam com o desenvolvimento do sistema nervoso o qual esboça os caminhos flexíveis de ação.

Consideremos mais detalhadamente a estrutura e modo de funcionamento dos sistemas nervosos mais desenvolvidos, tal como descritos por Bergson. Primeiro, observamos que as funções perceptivas estão distribuídas entre a medula e o cérebro. Não há uma diferença de natureza, mas apenas de “complicação” de grau entre as funções reflexas da medula espinhal — “automatismo puro” — e os processos perceptivos cerebrais — “atividade voluntária”. O papel da medula é transformar os estímulos recebidos em “movimentos executados”. Em outras palavras, o “movimento centrípeto comunicado pela excitação reflete-se imediatamente, por intermédio das células nervosas da medula, num movimento centrífugo determinando uma contração muscular” (BERGSON, 1990, p. 25).

Em sua conferência *La conscience et la vie*<sup>3</sup>, Bergson diz ser a medula o lugar em que estão montados mecanismos que contêm, já prontas para serem movimentadas, ações que o corpo pode realizar, ou seja, um “grande número de respostas já prontas”. Se esses mecanismos são acionados diretamente a partir do mundo externo, o corpo executa “imediatamente, como resposta à excitação recebida, um conjunto de movimentos coordenados entre si” (BERGSON, 1993a, p. 8).

---

<sup>3</sup> Conferência feita em inglês na Universidade de Birmingham em 29 de maio de 1911.



Mas, muitas vezes, o estímulo externo dirige-se, primeiramente, ao cérebro, este colocando-se como um “intermediário”, antes de se dirigir às células motoras da medula. Quanto mais desenvolvido o cérebro, maior é o número de ações motoras possíveis de serem realizadas:

[...] estas células das diversas regiões ditas sensoriais do córtex, células interpostas entre as arborizações terminais das fibras centrípetas e as células motoras do sulco de Rolando, permitem ao estímulo recebido atingir *à vontade* este ou aquele mecanismo motor da medula espinhal e escolher assim seu efeito. Quanto mais se multiplicarem essas células interpostas, mais elas emitirão prolongamentos amebóides capazes sem dúvida de se aproximarem diversamente, mais numerosas e variadas serão também as vias capazes de se abrirem ante um mesmo estímulo vindo da periferia e, conseqüentemente, haverá mais sistemas de movimentos que uma mesma excitação deixará à escolha. (BERGSON, 1990, p. 25)

Para Bergson, o cérebro funciona como uma “encruzilhada” onde estímulos vindos de “qualquer via sensorial” podem se ligar a “qualquer via motora”. A partir do estímulo recebido, há uma escolha e possível acionamento do mecanismo motor, da mais apropriada dentre as respostas que a medula coloca à disposição. É sob esse aspecto que o filósofo em *La conscience et la vie* define o cérebro como um “órgão de escolha” (BERGSON, 1993a, p. 9). Comparando-o a uma central telefônica, Bergson afirma que o papel do cérebro se limita a “efetuar a ligação ou fazê-la aguardar”, a “transmitir e a repartir movimentos” (BERGSON, 1990, p. 26). O cérebro não acrescenta nada ao que recebe, constituindo-se como um centro onde os estímulos provenientes dos órgãos dos sentidos são colocados em contato com mecanismos motores da medula e do bulbo raquidiano. A partir de excitações recebidas, o cérebro prepara movimentos apropriados. Ele conduz o movimento a um “órgão de reação escolhido” ou abre a esse “movimento a totalidade das vias motoras para que aí desenhe-se todas as reações que ele pode gerar e para que analise a si mesmo ao se dispersar” (BERGSON, 1990, p. 26). O cérebro “parece um instrumento de análise com relação ao movimento recolhido e um instrumento de seleção com relação ao movimento a executar” (BERGSON, 1990, p. 26).

De acordo com essa concepção, o papel do cérebro, assim como o da medula, é apenas esboçar “uma pluralidade de ações possíveis” ou organizar “uma delas” (BERGSON, 1990, p. 26). A representação da matéria “é a medida

de nossa ação possível sobre os corpos” (p. 35) não sendo produzida pelo cérebro, embora seja por ele acionada<sup>4</sup>:

[...] o cérebro recebe dos órgãos dos sentidos, por intermédio dos nervos centrípetos, a indicação de certos atos possíveis, transmite aos órgãos locomotores, por intermédio dos nervos centrífugos, a ordem de executar ou esboçar tal ou tal ato efetivamente, mas se limita, em suma, a escolher entre várias ações possíveis, e a preparar uma certa ação real: como tudo isso diz respeito apenas à *ação*, e que entretanto a *representação* se produz, é preciso admitir que a representação não é criada pelo fenômeno cerebral, que ela é simplesmente ocasionada ou movimentada por ele, e que, desde então, o cérebro se limita a recortar sem cessar, para apresentar à consciência, tal ou tal porção de uma percepção inconsciente do Todo”. (BERGSON, 1972, p. 806)

É em função da relação entre a estrutura do órgão perceptivo e a riqueza de ações por ele possibilitada que se deve compreender a riqueza da percepção consciente. Quanto mais desenvolvido o sistema nervoso, mais rica é a percepção, ou seja, os mecanismos motores se tornam “cada vez mais complexos” e o campo de relação com os objetos exteriores se amplia, abrangendo um maior número de objetos e em uma maior distância. Em decorrência, as ações podem se tornar cada vez menos imediatas, menos necessárias. O desenvolvimento do sistema nervoso possibilita, então, um aumento da indeterminação da ação, refletindo, portanto, o grau de indeterminação da mesma e a riqueza da percepção, também voltada para a ação; a percepção, enquanto um dos componentes associados àquele desenvolvimento, simboliza “a parte crescente de indeterminação, deixada à escolha do ser vivo em sua conduta em face das coisas” (BERGSON, 1990, p. 27). A extensão da percepção consciente está vinculada à “intensidade da ação de que o ser vivo dispõe” (BERGSON, 1990, p. 28). Ela está ausente sempre que um estímulo se prolongue em reação necessária — automatismo; ela se submete à relação entre necessidade e distância: “à medida que a reação se torna mais incerta, que ela deixa mais lugar à hesitação, aumenta também a distância na qual se faz sentir ... a ação do objeto” (BERGSON, 1990, p. 28).

As considerações anteriores parecem sugerir uma continuidade, uma diferença apenas de grau entre o homem e o animal, decorrente da semelhança

---

<sup>4</sup> Tendo em vista o objetivo principal desse artigo, não apresentaremos a teoria bergsoniana sobre a origem e natureza das representações, tema amplamente tratado em *Matéria e memória*..

de constituição de seus cérebros, e associada à diferença de volume e complexidade entre eles. Contrariando essa expectativa, Bergson pretende que haja efetivamente uma diferença muito mais profunda entre esses cérebros; uma diferença de natureza, a diferença entre o “limitado” e o “ilimitado”. Para Bergson, é apenas no ser humano que o número de mecanismos motores que se podem montar, e, conseqüentemente, “o número dos detonadores que têm a função de acionar o mecanismo motor entre os quais oferece a escolha é infinito” (BERGSON, 1991, p. 264). Essa diferença é relevante o suficiente para estabelecer a possibilidade da liberdade e, conseqüentemente, de uma consciência livre e abstrata apenas na esfera humana e a sua impossibilidade ou, quando muito, uma alternância entre consciência meramente perceptiva e inconsciência, no caso dos animais e, em decorrência do automatismo.

Quando se considera o funcionamento do cérebro do animal e do homem, está-se, no primeiro caso, diante de um “mecanismo que absorve a atenção” e, no segundo, de um “mecanismo de que se pode desviar”. No animal, os mecanismos motores montados pelo cérebro — “os hábitos que sua vontade adquire” — apenas realizam movimentos “armazenados nesses organismos” — “esboçados nesses hábitos”. Só no homem “o hábito motor pode ter um segundo resultado incomensurável com o primeiro. Ele pode impedir outros hábitos motores e, com isso, disciplinando o automatismo, pôr em liberdade a consciência:

A primitiva máquina a vapor, tal como a concebeu Newcomen, exigia a presença de uma pessoa exclusivamente encarregada de manobrar as torneiras, seja para introduzir o vapor no cilindro, seja para nele lançar a chuva fria destinada à condensação. Conta-se que um menino empregado nesse trabalho, muito entediado com a obrigação de o fazer, teve a idéia de ligar as manivelas das torneiras, por cordões, ao pêndulo da máquina. Desde então, a máquina abria e fechava por si mesma as suas torneiras; ela funcionava sozinha. Ora, um observador que comparasse a estrutura dessa segunda máquina com a da primeira, sem se ocupar dos dois meninos encarregados da vigilância, só teria verificado entre elas uma ligeira diferença de complicação. É tudo o que se pode perceber, com efeito, quando só se olham as máquinas. Mas se dermos uma olhadela nos meninos, veremos que um está absorvido por sua vigilância, e que o outro está livre para divertir-se à vontade, e que, por esse aspecto, a diferença entre as duas máquinas é radical, a primeira mantendo a atenção prisioneira, a segunda lhe dando livre trânsito. (BERGSON, 1991, p. 185)

Essa diferença entre o homem e o animal nos leva de volta à questão da relação entre a consciência e a matéria no processo evolutivo. Bergson associa a consciência à capacidade de escolha, ela é “sinônimo de invenção e de liberdade”. Ocorre que no animal a invenção é uma mera variação do rotineiro, a iniciativa individual podendo ampliar os hábitos, mas caindo num novo automatismo: “as portas de sua prisão se fecham tão logo abertas; esforçando-se por escapar de sua corrente só consegue alongá-la” (Bergson 1, p. 264). A vida antes do homem é um “esforço da consciência para sacudir a matéria” e um “esmagamento da consciência pela matéria” (BERGSON, 1991, p. 264).

A consciência, definida algumas vezes por Bergson como uma “*exigência de criação*”, só se manifesta a si mesma onde houver a possibilidade de criação. Ela “adormece” onde houver automatismo e “desperta” onde há “possibilidade de uma escolha” (BERGSON, 1991, p. 262). Nos “organismos desprovidos de sistema nervoso”, é o “poder de locomoção e de deformação” que determina a variação de consciência; já nos “animais dotados de sistema nervoso, ela é proporcional à complicação da encruzilhada onde se encontram as vias chamadas sensoriais e as vias motoras, isto é, do cérebro” (BERGSON, 1991, p. 262). E é apenas no homem que a consciência consegue quebrar a corrente, libertar-se.

#### IV

A relação entre automatismo e liberdade dos movimentos e a constituição e modo de ser da consciência pode ser melhor compreendida quando se consideram as duas formas de consciência inerentes aos artrópodes e aos vertebrados, a consciência instintiva e a consciência inteligente. Bergson postula uma evolução divergente culminando de um lado nos himenópteros, nos quais prevalece a consciência instintiva, e de outro no homem no qual prevalece a consciência intelectual. O fundamental para a constituição desses dois tipos de consciência é o privilégio do automatismo ou da liberdade dos movimentos, ou seja, a consciência se manifesta onde houver liberdade, espaço para a escolha, e se anula onde prevalece o automatismo. O que acontece é que nas atividades maquinais a representação do ato é impedida pela sua execução, a representação é obstruída pela ação. O ato é tão perfeitamente semelhante à representação e nela se insere tão exatamente — adequação perfeita da representação e da ação



— que nenhuma consciência pode mais transbordar. Para Bergson, a consciência se manifesta se houver um obstáculo à realização do ato, ou seja, ela estava presente, “mas neutralizada pela ação que ocupava o lugar da representação” (BERGSON, 1991, p. 145)<sup>5</sup>. É nesse sentido que o filósofo define a consciência como a “inadequação do ato à representação” (BERGSON, 1991, p. 145), ou como uma “*diferença aritmética entre a atividade virtual e a atividade real*”, medindo “*a distância entre a representação e a ação*” (EC, 145). Para Bergson, há uma incompatibilidade entre consciência e automatismo, este definido como “atividade real”. A consciência envolve “escolha”, “hesitação”, “ações possíveis” ou “atividade virtual”.

Deduz-se, a partir daí, que a inteligência tende à consciência, enquanto que o instinto tende à inconsciência. Onde prevalece o instinto há pouco lugar para a “hesitação” e a “escolha”, já que a natureza organiza o “instrumento a manejar”, fornecendo o “ponto de aplicação” e o próprio “resultado a obter”. Nesses casos, a consciência é rara, só aparecendo onde houver “contrariedades” ao instinto, não estando, portanto, relacionada à própria natureza do instinto. Já no caso da inteligência, o que define a sua essência é o “sofrer contrariedades”, sendo, portanto, o déficit o seu estado normal: “Tendo por função primitiva o fabricar instrumentos inorganizados, ela deve, atravessando mil dificuldades, escolher para esse trabalho o local e o momento, a forma e a matéria. E ela não pode satisfazer-se plenamente, porque toda satisfação nova cria novas necessidades” (BERGSON, 1991, p. 146).

Para Bergson, a diferença entre os conhecimentos do instinto e da inteligência é de natureza e não de grau, embora, cada um deles, sob um determinado aspecto, possui conhecimento inato — virtual, natural — o qual é “desempenhado e inconsciente” e recai sobre as coisas, ou seja, “atinge os objetos determinados em sua própria materialidade”. No caso do instinto, o conhecimento está implícito, exteriorizando-se em “desempenhos precisos”, ou seja, está “implicado na ação realizada”. Trata-se de um conhecimento limitado que se dirige a um determinado objeto ou parte dele, possuindo a respeito desse um conhecimento “interior e pleno”. Bergson apresenta um exemplo de

---

<sup>5</sup> Como vimos anteriormente, o surgimento da consciência depende também da constituição corpórea. Ou seja, os modos de ser das consciências instintiva e inteligente dependem das correlatas constituições corporais.



comportamento instintivo, no qual a conduta esboça um conhecimento não aprendido: o sítaris:

Esse coleóptero deposita seus ovos na entrada das galerias subterrâneas que uma espécie de abelha, a antófora, cava. A larva do sítaris, após longa espera, espreita a antófora macho ao sair da galeria, agarra-se a ele, fica aferrada a ele até o “vôo nupcial”; nessa ocasião, ela aproveita a ocasião para passar do macho à fêmea, e espera tranquilamente que esta ponha seus ovos. Salta então sobre o ovo, que lhe vai servir de sustento no mel, devora o ovo em alguns dias, e, instalada na concha, sofre a primeira metamorfose. Organizada agora para flutuar sobre o mel, ela consome esse suprimento de alimentação e torna-se ninfa, depois inseto perfeito. Tudo acontece *como se* a larva do sítaris, desde sua eclosão, soubesse que o antóforo macho sairia da galeria primeiro, que o vôo nupcial lhe forneceria o meio de se transportar para a fêmea, que esta a conduziria a um depósito de mel capaz de alimentá-la quando se transformasse, e que, até essa transformação, ela tivesse devorado aos poucos o ovo do antóforo, de modo a se nutrir, a se sustentar na superfície do mel, e também eliminar o rival que tenha saído do ovo. E tudo acontece também *como se* o próprio sítaris soubesse que sua larva saberá todas essas coisas. (BERGSON, 1991, p. 147)

## V

Consideremos mais detalhadamente a natureza e o modo de funcionamento da inteligência, a relação entre liberdade e consciência nas suas formas mais abstratas. Para Bergson, a inteligência é, essencialmente, uma faculdade de fabricação. O aparecimento do homem — “parte culminante da evolução dos vertebrados” — sobre a Terra é associado ao período em que se fabricaram as primeiras armas, os primeiros instrumentos. Um instrumento fabricado é a concretização da invenção a qual define melhor a inteligência que a inferência, visto que essa atividade pode ser considerada como uma modalidade, um componente da própria invenção: “Sem dúvida, há inteligência sempre onde houver inferência; mas a inferência, que consiste no aproveitamento da experiência passada no sentido da experiência presente, é já um começo de invenção (BERGSON, 1991, p.139).

A inteligência, mesmo em sua forma mais simples, “aspira a fazer com que a matéria atue sobre a matéria”. Tendo em vista essa ligação entre a

inteligência e a ação, compreende-se que as invenções mecânicas permitem utilizar cada vez mais a matéria. A invenção mecânica, a fabricação de objetos artificiais como as ferramentas, foi o primeiro empenho essencial da inteligência humana, o elemento em torno do qual gravita a vida social. Daí Bergson considerar legítimo definir o homem tendo em vista essa atividade: “Se pudéssemos nos despir de todo orgulho, se, para definir nossa espécie, nos ativéssemos estritamente ao que a história e a pré-história nos apresentam como a característica constante do homem e da inteligência, talvez não disséssemos *Homo sapiens*, mas *Homo faber*” (BERGSON, 1991, p. 140).

O instrumento produzido pela inteligência é imperfeito mas flexível; exige esforço para ser obtido; é de “manejo penoso”; mas por ser feito de matéria inorganizada, “pode assumir uma forma qualquer, servir a qualquer fim, livrar o ser vivo de qualquer dificuldade nova que surja e lhe conferir uma quantidade ilimitada de poderes” (BERGSON, 1991, p. 141). Assim, embora os instrumentos da inteligência sejam aleatórios, eles podem produzir conquistas ao infinito. Influencia quem o fabricou, levando-o a exercer uma nova função; ao satisfazer uma necessidade, ele cria uma nova, o que faz com que, diferentemente do instinto que fecha o círculo de ação em que o animal se move automaticamente, ele “abre a essa atividade um campo infinito, impelindo-a cada vez mais além e tornando-a cada vez mais livre” (BERGSON, 1991, p.142). Assim, a inteligência é uma faculdade voltada para a fabricação de instrumento inorganizado — artificial. Como ela não possui esse instrumento, ela diversifica a sua fabricação conforme as “circunstâncias” e as “dificuldades”.

Um exemplo paradigmático de instrumento flexível criado pela inteligência e diretamente responsável por uma consciência abstrata é a linguagem. A linguagem, se por um lado depende de uma organização física complexa, é também inerente à vida social humana. Como a função — fabricação/ação — não está previamente determinada pela estrutura dos indivíduos, sendo aprendido o papel social de cada um deles. Impõe-se, assim, “uma linguagem que permita, a todo instante, passar do que se sabe ao que se ignora ... uma linguagem cujos signos — que não podem ser em número infinito — sejam extensíveis a uma infinidade de coisas” (BERGSON, 1991, p. 142). Por isso, a característica fundamental da linguagem humana não é a generalidade, mas a mobilidade do signo, sua “tendência a se transladar de um objeto a outro” (BERGSON, 1991, p. 142).

Bergson atribui à “mobilidade das palavras” a “libertação” da inteligência, permitindo a essa ir além dos próprios objetos materiais que lhe interessam. Inicialmente, a inteligência se adapta à “forma da matéria bruta”, ou seja, a linguagem designa apenas coisas; é graças à mobilidade das palavras que ela pode voltar-se para um “objeto que não é coisa”, e perceber a si mesma como “faculdade de representação em geral”, isto é, é a linguagem que torna possível a consciência de si.

A palavra, feita para ir de uma coisa a outra, é, de fato, essencialmente, deslocável e livre. Ela poderá, pois, estender-se, não apenas de uma coisa percebida a outra coisa percebida, mas ainda da coisa percebida à lembrança dessa coisa, da lembrança precisa a uma imagem mais fugidia, de uma imagem fugidia, contudo representada ainda, à representação do ato pelo qual se a representa, isto é, à idéia. Desse modo, vai abrir-se aos olhos da inteligência, que olhava de fora, um mundo interior, o espetáculo de suas próprias operações (BERGSON, 1991, p.160).

## **Referências**

- BERGSON, H. *L'évolution créatrice*. 5.ed. Paris: PUF, 1991.  
\_\_\_\_\_. *L'énergie spituelle*. 4.ed. Paris: PUF, 1993a.  
\_\_\_\_\_. *Mélanges*. Paris: PUF, 1972.  
\_\_\_\_\_. *Matière et mémoire*. 3.ed. Paris: PUF, 1990.  
\_\_\_\_\_. *La pensée et le mouvante*. 3d. Paris: PUF, 1993b.

**PARTE II — AUTO-ORGANIZAÇÃO, INFORMAÇÃO E  
O PROBLEMA DAS REPRESENTAÇÕES NAS  
CIÊNCIAS COGNITIVAS**





# O mal estar do representacionismo: sete dores de cabeça da Ciência Cognitiva<sup>1</sup>

*Willem Ferdinand Gerardus Haselager<sup>2</sup>*

## Introdução

Há um crescente número de cientistas cognitivos que enfrentam um mal-estar com a abordagem representacionista na explanação do comportamento inteligente. No entanto, dizer exatamente o que há de errado com o representacionismo é uma tarefa muito difícil, até porque a utilização de representações não é completamente inadequada. Mais precisamente, entendemos que a Ciência Cognitiva tornou-se dependente de representações. Nosso principal objetivo neste texto é discutir alguns dos sintomas dessa dependência.

## Ciência Cognitiva *versus* Behaviorismo

O principal objetivo da Ciência Cognitiva é explicar o comportamento inteligente. Embora essa ciência possa ser retratada como um estudo da mente, cabe notar que estados mentais ou internos são considerados necessários por causa de seus efeitos, direta ou indiretamente, no comportamento externo. Mesmo nos primórdios da Inteligência Artificial, McCarthy e Hayes (1969, p. 463) declaravam explicitamente que era seu objetivo criar “um programa computacional capaz de agir inteligentemente no mundo”.

Como sabemos, antes do surgimento da Ciência Cognitiva, o behaviorismo era a abordagem predominante na Psicologia. Uma característica fundamental do behaviorismo era sua rejeição de estados mentais internos na explicação do comportamento. Skinner, por exemplo, defendia a hipótese de que o uso de estados mentais nas explicações do comportamento apenas revelava a

---

<sup>1</sup> Texto traduzido do inglês por Mariana Claudia Broens.

<sup>2</sup> Artificial Intelligence, Nijmegen Institute of Cognition and Information (NICI), Radboud University, Nijmegen, Holanda e Pós-Graduação em Filosofia, UNESP, Marília, SP, Brasil. Endereço eletrônico: [w.haselager@nici.ru.nl](mailto:w.haselager@nici.ru.nl). Home page: [www.nici.ru.nl/~haselag](http://www.nici.ru.nl/~haselag).

lacuna de nosso conhecimento: usamos conceitos mentais apenas quando as reais causas ambientais do comportamento são obscuras para nós. Obviamente, o problema do behaviorismo era que suas próprias explicações limitavam-se ao comportamento e a processos de aprendizagem relativamente simples. A cognição mais abstrata, como a encontrada na linguagem, no raciocínio e no planejamento, por exemplo, permanecia fora de seu alcance.

A Ciência Cognitiva assumiu, no lugar do behaviorismo, uma posição predominante na área. Ela pretende explicar a cognição mais abstrata tendo como base a suposição de que estados mentais internos são causalmente relevantes para o comportamento. Esse efeito causal no comportamento pode ser entendido em termos de processos computacionais sobre representações internas. A despeito de que a noção de representação é uma das pedras fundamentais da Ciência Cognitiva (o outro é a computação), seu significado preciso tem se tornado constantemente mais obscuro. Tomaremos a definição a seguir, freqüentemente utilizada nos debates em torno ao representacionismo, como nosso ponto de partida:

Um sistema sofisticado (um organismo) projetado (tendo evoluído) para maximizar algum fim (a sobrevivência, por exemplo) deve, em geral, ajustar seu comportamento às características, estruturas ou configurações ambientais que lhe são específicas de formas não inteiramente pré-configuradas em seu *design*. [...] Mas se as características relevantes não estão sempre presentes (ou não são detectáveis), então elas podem, pelo menos em alguns casos, ser representadas; isto é, algo diferente pode ocupar seu lugar se tiver o poder de guiar o comportamento. Aquilo que ocupa o lugar de algo dessa maneira é uma representação; aquilo cujo lugar é ocupado é seu conteúdo; e ocupar o lugar de um conteúdo é representá-lo (HAUGELAND, 1991, p. 62).

As duas características mais importantes das representações são que elas se colocam no lugar de algo e que o sistema usa as representações com o objetivo de guiar seu comportamento. De acordo com a Ciência Cognitiva tradicional, então, as representações desempenham um duplo papel: carregam um conteúdo e causam o comportamento. Mesmo se a Ciência Cognitiva clássica e o Conexionismo discordam a respeito do efetivo formato das representações, eles têm esse pressuposto em comum.

A perspectiva geral é a de que o meio ambiente apresenta problemas para o sistema, ao quais devem ser resolvidos através do processamento interno

de informação. As representações são postuladas para atuar como intermediárias entre a percepção e a ação, especificando informações sobre o mundo, freqüentemente de forma independente do contexto – e da ação.

Certamente, a credibilidade das representações internas foi consideravelmente fortalecida pela existência de modelos computacionais que efetivamente desempenham tarefas que em seres humanos requerem inteligência. Como afirmou Fodor: “Sem representações não haveria computação; sem computação não haveria modelagem” (FODOR, 1975, p. 31). E poderíamos acrescentar que sem modelagem não haveria Ciência Cognitiva. Além disso, um forte argumento a favor das representações sempre foi a indagação retórica: “E de que outro modo poderíamos explicar a cognição e o comportamento inteligente?”.

### **As dores de cabeça representacionistas**

A seguir, pretendemos criticar o uso de representações na Ciência Cognitiva, mas não queremos fazê-lo exageradamente. Para evitar, desde o início, mal entendidos, ressaltamos que não estamos sugerindo que a Ciência Cognitiva deva abandonar inteiramente sua posição representacionista. Não consideramos exagerado dizer que a Ciência Cognitiva derrotou o behaviorismo porque o uso de representações internas tornou possível estudar fenômenos cognitivos que estavam fora do alcance deste último. Não estamos afirmando que o behaviorismo estava certo nem que, por outro lado, o uso explanatório de representações internas está equivocado. No entanto, consideramos que a Ciência Cognitiva exagerou em sua reação contra o behaviorismo e que isso a impediu de avaliar criticamente seu principal instrumento explanatório. Como colocado, entendemos que a Ciência Cognitiva tornou-se dependente (ou adicta) do representacionismo. E, como ocorre freqüentemente em casos de dependência, isso significa que a Ciência Cognitiva utiliza representações rapidamente e com excessiva freqüência. Além disso, essa enorme pressa nem sempre permite avaliar com clareza o uso de representações. Em suma, entendemos que o uso acrítico de representações originou várias dores de cabeça à Ciência Cognitiva.

## **1 “Atue normalmente!” O problema do *frame***

Um dos mais importantes problemas enfrentados pela Ciência Cognitiva é como fazer com que computadores apresentem senso comum e executem tarefas do cotidiano. Essa questão não se refere ao desempenho eficiente de tarefas especializadas (derrotar o campeão mundial de xadrez ou detectar onde se encontram reservatórios ocultos de petróleo, por exemplo). Trata-se de ações de nosso dia-a-dia: preparar o café da manhã, rir de uma piada, conversar durante o café da manhã, escolher um caminho para ir a um lugar específico na cidade onde moramos, etc.

O problema é que, quando o examinamos de perto, o “agir normalmente” conta com uma enorme quantidade de conhecimento comum. Precisamos saber, por exemplo, que a porta deve ser aberta antes de entrar num lugar ou que a água é necessária para fazer café. Do mesmo modo, precisamos saber que as pessoas gostam, em geral, de perder peso, mas não querem fazê-lo cortando seus braços ou pés, ou que pessoas não entram em ruas sem saída em circunstâncias normais, etc.

A questão central aqui é que as pessoas parecem ter uma enorme quantidade de conhecimento comum e são capazes de utilizar esse conhecimento rapidamente e com eficiência. Já foi demonstrado, inclusive, que é muito difícil formular exatamente o que conhecemos (ver, por exemplo, o Cyc-project: GUHA; LENAT, 1990, 1993; DREYFUS, 1992, p. p.xvi-xxx). Mas é ainda mais difícil modelar nosso uso eficiente e flexível de partes relevantes daquilo que conhecemos. Até o momento, as tentativas de modelagem do raciocínio do senso comum, tendo por base uma grande quantidade de conhecimento interno representado, fracassaram. Representações em demasia resultam em apatia computacional, pois o sistema se perde em seu próprio armazenamento de informações. Um sistema representacional capaz de agir inteligentemente no mundo é rapidamente sufocado por seus próprios recursos representacionais.

Um dos modos em que esta questão se manifesta é no assim chamado ‘problema do frame’, postulado por McCarthy e Hayes (1969); ver também PYLYSHYN, 1987; HASELAGER, 1997). O problema do frame contribuiu para a pesquisa de métodos alternativos de representar o conhecimento, tais como as representações distribuídas do Conexionismo. Em outras ocasiões, procuramos mostrar que o problema não foi resolvido dessa maneira, mas



apenas foi transformado em outro problema diferente, mas igualmente sério (HASELAGER, 1999; HASELAGER; van RAPPARD, 1998).

Uma vez que aparentemente estamos hoje tão longe de resolver o problema quanto estavam McCarthy e Hayes no final dos anos 1960, uma questão se coloca com toda força: o conhecimento comum precisa realmente ser representado? Assim, o problema do frame constitui uma das principais razões para assumir uma postura crítica em relação ao representacionismo.

## **2 “Dois candidatos, ambos precários” O problema da estrutura representacional.**

Nos bons velhos tempos da Ciência Cognitiva, quando o behaviorismo havia sido recém superado, todos sabiam o que eram representações: seqüências de símbolos bem estruturadas de acordo com regras sintáticas e semânticas adequadas. Sistemas formais automatizados simulavam processos cognitivos humanos que se supunha ocorrerem numa “Linguagem do Pensamento” (LP; FODOR, 1974). Então, no início dos anos 1980, os problemas com que se deparou a abordagem tradicional tornaram-se mais e mais evidentes. O formato simbólico-representacional passou a ser considerado excessivamente lingüístico e muito frágil para dar conta da cognição humana. A alternativa conexionista, baseada na representação distribuída, foi recebida com entusiasmo. Conjuntos de pesos e padrões flexíveis de ativação de unidades neurônio-símile (UNS) substituíram as estruturas de blocos de representações do LP. No entanto, os problemas da representação distribuída rapidamente também se tornaram evidentes. Uma estrutura representacional (por exemplo, quem faz o quê com quem e com o quê, como em “Maria bate em João com a colher”) é possível, por meio de um árduo trabalho, apenas em domínios limitados. O desenvolvimento de representações distribuídas utilizáveis é excessivamente dependente de arranjos pré-fixados, e em si mesmos frágeis, de conjuntos de unidades, cujo treinamento requer uma espantosa quantidade de computações.

Certamente os conexionistas tiveram sucesso ao apontar as fraquezas do esquema representacionista clássico, enquanto que os proponentes da tendência clássica indicam as dificuldades experimentadas pelos conexionistas (por exemplo, Aizawa, 1997; Chalmers, 1993; Churchland, 1989; Fodor; Pylyshyn, 1988; Hadley, 1994; Haselager; van Rappard, 1998; Horgan; Tienison, 1997;



Smolensky, 1988). O ponto importante, e frustrante, é que os dois lados têm razão. É preocupante que os fundamentos da Ciência Cognitiva, que se alicerçam nas representações, sejam objeto de acalorados debates sobre sua correta estrutura. Esta questão se torna mais séria ainda quando as críticas a ambas as estruturas representacionais estão basicamente corretas. A questão que surge em decorrência disso é: será a pedra fundamental da Ciência Cognitiva tão sólida quanto se pensava?

### **3 “O que isso significa?” O problema das raízes do significado**

A principal razão pela qual as representações são consideradas úteis é, obviamente, que elas ocupam o lugar de alguma outra coisa, elas se referem a algo ausente no ambiente. No entanto, é altamente problemático explicar como as representações adquirem significado. Como podem as representações de sistemas computacionais tornar-se significativas para os próprios sistemas? É importante notar que esta interrogação não diz respeito ao significado das representações de um modelo para seu programador, mas sim do que elas significam para o próprio modelo. Como Harnad (1990) ressaltou, o problema é fazer com que o significado das representações se torne intrínseco ao sistema e não parasitário do significado que está na cabeça de seu idealizador. Para acentuar a diferença entre sistemas naturais e sistemas computacionais foram cunhadas expressões como “significado genuíno” ou “Intencionalidade intrínseca” (Searle, 1980a; 1980b). A principal dificuldade que este problema levanta é que parece difícil que um sistema representacional conecte efetivamente suas representações internas com o mundo, fundamentando-as durante o processo. Assim, a principal característica das representações (‘ocupar o lugar de algo’) permanece um mistério para a perspectiva computacional.

### **4 “Representações? Onde?” O problema da operacionalização**

Ultimamente tem sido desenvolvido um número crescente de sistemas que aprendem a efetuar tarefas de um modo mais ou menos autônomo, como, por exemplo, os que utilizam algoritmos genéticos. Frequentemente, esses sistemas são analisados apenas depois de terem aprendido como realizar a tarefa. Assim, esta linha de trabalho é diferente da abordagem tradicional na qual os

recursos representacionais são previamente especificados e programados no sistema. É interessante notar que pesquisadores trabalhando nessa área levantam a questão de como determinar se os sistemas autônomos de aprendizagem são representacionais e, se forem, que representações eles desenvolveram:

Investigamos máquinas simples de sistemas visuais desenvolvidas por evolução artificial no lugar das técnicas de *design* da engenharia tradicional, e note-se que a tarefa de identificar representações internas torna-se difícil pela ausência de uma definição operacional de representação no plano da causação mecânica. Conseqüentemente, questionamos a natureza, e até mesmo a existência, de representações que se postula serem utilizadas nos sistemas visuais naturais (isto é, nos animais) (CLIFF; NOBLE 1997, p.1156; ver também BEER 1995; BROOKS, 1991; 1995).

Um modo usual de estabelecer se um sistema físico constitui um sistema computacional e representacional é indicando um isomorfismo entre eles (o sistema físico e uma descrição representacional e computacional). No entanto, Putnam (1988) criticou severamente o uso abusivo de isomorfismos. De acordo com sua tese da instanciação universal, todo sistema físico aberto pode efetuar qualquer computação. Embora essa tese tenha sido questionada por Chalmers (1995, 1996) e Chrisley (1995), permanece o fato de que o isomorfismo está longe de ser considerado uma noção não problemática. Entendemos que é altamente problemático que numa ciência haja incertezas a respeito da operacionalização adequada de uma de suas noções centrais.

## **5 “Este comportamento parece inteligente, então deve ser representacional”. O problema do observador**

Todos estamos familiarizados com a tendência humana de projetar nossas vontades e desejos nos outros. Fazemos isso não apenas em relação ao conteúdo dessas vontades e desejos (como, por exemplo, quando dizemos “você realmente quer que eu saia, não quer?”), mas também em relação ao próprio esquema conceitual das crenças e desejos em geral. Não apenas atribuímos crenças e desejos a animais como cachorros e gatos, mas fazemos o mesmo, muitas vezes sem nos dar conta, com computadores (“por que você não quer fazer isto?”) e até com carros. Em especial, robôs parecem estimular ainda mais nossa disposição de atribuir a outros seres uma vida mental. Encontramos

exemplos interessantes dessa disposição em relação a robôs no trabalho de Braitenberg (1984); e a empresa Sony está tentando explorar economicamente essa inclinação humana com o AIBO e outros projetos de robôs domésticos.

Consideramos que os cientistas cognitivos têm essa mesma tendência em relação às representações: sempre que observam o comportamento de um sistema, manifestam essa compulsão de atribuir-lhe representações internas. No entanto, um dos fundadores da Ciência Cognitiva alertou-nos contra essa tendência nos primórdios dessa área do saber. Simon (1969/1996, p. 51) deu o exemplo de uma formiga na praia:

Ela se move em frente, vira à direita para facilitar a escalada de um montículo íngreme, desvia-se de uma pedra, pára por uns segundos para trocar informações com outro membro da colônia. Assim ela vai traçando seu caminho tortuoso de volta a seu lar.

Embora seu comportamento pareça complicado e difícil de descrever, Simon argumenta que a complexidade não reside na formiga, mas surge da complexidade da superfície da praia. Simon (1969/1996, p. 53) coloca que o mesmo pode ser dito de seres humanos:

Seres humanos, considerados como sistemas comportamentais, são bastante simples. A aparente complexidade de nosso comportamento ao longo do tempo é em grande parte um reflexo da complexidade do ambiente em que nos encontramos.

Consideramos que esse alerta de Simon foi negligenciado devido aos muitos sucessos do representacionismo. Ao primeiro sinal de inteligência (e muitas vezes até sem isso, como mostraremos a seguir), começamos quase automaticamente a procurar a complexa estrutura interna que torna tal comportamento possível. Entendemos que esta disposição é prejudicial.

## **6 “Vou primeiro buscar uma ~~bebida~~ representação...” O problema da dependência (primeira parte)**

Tim van Gelder (1995) oferece uma boa ilustração do problema da dependência (embora ele não utilize esta expressão). Seu exemplo data de 1788, no período da revolução industrial inglesa, quando máquinas a vapor eram utilizadas para tecer o algodão usado em roupas e tapeçarias de qualidade superior. Para obter a mais alta qualidade, a pressão do vapor precisava ser

transferida para o movimento gradual da roda que controlava os teares. Uma válvula podia ser usada para regular a pressão do vapor, mas como controlar a válvula?

Um cientista cognitivo com uma boa formação, observa van Gelder (1995, p. 348), imediatamente pensaria mais ou menos o seguinte:

1. Medir a velocidade da roda.
2. Comparar a velocidade atual com a velocidade desejada.
3. Se não há discrepância, voltar ao primeiro passo. Se houver:
  - a. medir a pressão usual do vapor;
  - b. calcular a mudança desejada na pressão do vapor;
  - c. calcular o ajuste necessário na válvula.
4. Fazer o ajuste necessário na válvula reguladora da pressão.
5. Voltar para o primeiro passo.

Trata-se de uma perfeita solução representacional e computacional e não há nada de errado com ela na perspectiva da Ciência Cognitiva. Vários componentes do sistema são identificados (um para medir a velocidade da roda, outro para medir a pressão do vapor, uma unidade central de processamento) e estes componentes se comunicam por meio de representações até que eventualmente alguma ação tenha que ser realizada.

A solução para o problema da regulação da pressão do vapor, tal como proposta por James Watt, é completamente diferente. Watt constrói uma máquina, o “Regulador Watt” (RW), que possui um eixo vertical conectado à roda. Dois braços com esferas metálicas na ponta são anexados ao eixo com dobradiças. Quando a roda se movimenta, a força centrífuga empurra as esferas para o lado e para o alto, o que provoca o fechamento da válvula, reduzindo a quantidade de vapor e movendo lentamente a roda para baixo. Assim é produzido o movimento gradual dos teares.

Van Gelder aponta que não há troca de representações entre os componentes do sistema, inexistindo o processamento computacional de representações para determinar o que deve ser feito. Ao invés disso, há um acoplamento físico direto entre os componentes do sistema. O objetivo de van Gelder, que consideramos acertado, é ressaltar que caímos na tentação representacionista muito rapidamente e sem necessidade. Entendemos que a

razão pela qual os cientistas cognitivos postulam soluções representacionistas decorre de sua superexposição a esse tipo de solução à qual se apegam tão automaticamente quanto alcoólatras se apegam a sua garrafa.

## **7 “Será tudo um sistema representacional?” O problema da dependência (segunda parte)**

Entendemos que o quadro que traçamos até agora já parece suficientemente problemático, mas como reação ao exemplo do ‘Regulador Watt’ de van Gelder, Bechtel (1998) argumenta que o RW é, apesar de tudo, um sistema representacional<sup>3</sup>. Bechtel explicita que um sistema pode ser considerado representacional se: 1) um objeto externo (x) é representado internamente por um sistema (z); 2) um estado interno (y) ocupa o lugar do objeto (x) e 3) o sistema (z) coordena seu comportamento em relação ao objeto representado (x) por meio da representação interna (y). Entendemos que esta colocação está de acordo com a definição de Haugeland (já discutida). Esta análise, observa Bechtel, pode ser aplicada ao RW sem nenhum problema: o ângulo dos braços de metal representa (‘ocupa o lugar de’) a velocidade da roda e é usado pelo RW para regular a válvula de pressão (isto é, para coordenar o comportamento do sistema com a velocidade da roda). Em outros termos, para Bechtel, o RW é um sistema representacional, mesmo se se trata de “um caso muito pouco interessante”, como ele é o primeiro a admitir.

Por mais surpreendente que possa parecer a conclusão de Bechtel, ela dificilmente poderá ser evitada. Embora isso possa ser interpretado como uma amarga derrota imposta ao anti-representacionismo de van Gelder, entendemos que a conclusão de Bechtel é, antes de tudo, uma ilustração de uma dependência que foge ao controle dos cognitivistas e que faz perder a capacidade de discernimento (alcoólatras graves ingerem outros líquidos que contêm álcool como se fossem bebidas alcoólicas). Contudo, entendemos que, se todo sistema pode ser considerado representacional, o próprio conceito de representação corre

---

<sup>3</sup> Não faço justiça a Bechtel ao referir-me a ele em relação à dependência representacionista. Longe de ser dependente, Bechtel tem um enfoque muito sutil, e respeitável, sobre os temas relativos ao representacionismo. No artigo que citamos, ele apenas pretende explicar que, de acordo com a interpretação representacionista padrão, o RW é um sistema representacional. Ao fazer isso, ele forneceu um exemplo muito útil para o objetivo geral deste trabalho. O mesmo ocorre com o trabalho de Churchland e Sejnowski que serão discutidos a seguir.



o risco de esvaziamento. Afinal de contas, a idéia original era que a Ciência Cognitiva podia superar o behaviorismo porque era capaz de modelar a cognição abstrata. Desse modo, se até mesmo mecanismos simples, como o Regulador Watt, podem ser considerados representacionais, então a noção de representação perde as vantagens explicativas que possuía quando combatia o behaviorismo. Note-se, a propósito, que a situação atual é muito diferente da situação dos primórdios da Ciência Cognitiva, uma vez que ainda estava claro naquela época o que era (ou não) uma representação. Afinal, tendo por base a Tese da Linguagem do Pensamento, o Regulador Watt não pode ser considerado um sistema representacional.

Um exemplo semelhante em relação à noção de computabilidade pode ser encontrado no trabalho de Churchland e Sejnowski (1992, p. 61-69). O exemplo que apresentam é o da máquina de peneirar grãos de Garret construída em 1851 que não explicaremos neste texto. Os autores alegam que há um claro isomorfismo entre as descrições funcionais e físicas da peneira mecânica, suficiente para garantir uma interpretação computacional. Os mesmos autores observam ainda que:

Esta caracterização geral de computação pode ser criticada justamente por ser excessivamente geral. Isto porque, num sentido amplo, mesmo um coador ou uma peneira mecânica pode ser considerado um computador desde que se ordene seus inputs em tipos; e se desejarmos investir nosso tempo nisso, podemos até descrever o comportamento dos inputs e outputs. Embora esta observação seja correta, não é propriamente uma crítica: trata-se principalmente de uma avaliação legítima sobre a abrangência da noção de computação (CHURCHLAND; SEJNOWSKI, 1992).

Dada a relevância atribuída neste trabalho à questão da dependência, não podemos deixar de notar a semelhança da última frase da citação acima com a jovial colocação: "Eu não tenho um problema com a bebida porque nunca tenho problema para beber". A conotação positiva da abrangência da computação sugerida por Churchland e Sejnowski pode talvez ser entendida se consideramos que eles utilizaram tal noção no sentido amplo e universalmente aceito na Ciência Cognitiva: no interior do debate entre a IA e o Conexionismo não estão em questão a existência e o papel das representações e da computação, mas sua natureza e formato. No entanto, com o surgimento da abordagem dinâmica da cognição, a situação mudou; é justamente no debate com os anti-

representacionistas que não se pode mais considerar como auto-evidente que a cognição se resume a operações computáveis ou estados representacionais. Considerando as mudanças ocorridas no debate, o apreço manifestado por Churchland e Sejnowski pela amplitude das noções de representação e computação não pode mais ser sustentado sem bons argumentos.

## **Conclusão**

Os problemas discutidos acima não constituem, certamente, argumentos decisivos contra o uso de representações. No entanto, consideramos que os problemas levantados fornecem bons motivos para indagar seriamente se as representações mentais são sempre necessárias ou úteis e se é possível produzir modelos cognitivos e explicações sem elas. Desse modo, a máxima metodológica que propomos não é “abaixo as representações!”, mas “utilizar representações em explicações e modelagens apenas quando for necessário”.

Em si mesma esta máxima está baseada em um princípio muito sensato, a ‘navalha de Ockam’, que postula a não multiplicação desnecessária de entidades. Representações são pressupostos adicionais pelos quais paga-se um preço. Entendemos que uma atitude mais ‘sóbria’ seria invocar representações quando existir um bom motivo ao invés de assumi-las como um ponto de partida.

Durante muito tempo, é claro, o problema era que a alternativa anti-representacionista (ou representacionista moderada) não parecia viável (a menos que se quisesse voltar ao behaviorismo, o que os cientistas cognitivos não desejavam fazer). Não existia o suporte empírico de modelos que possuíssem o mesmo grau de rigor e detalhamento que são rotineiros na Ciência Cognitiva tradicional. Não é sem razão que a pergunta: “de que modo pode-se explicar a cognição sem representações?” foi considerada quase retórica durante muitos anos. No entanto, esses dias acabaram. Uma abordagem que enfatize a importância da interação corporal com o ambiente (‘Teoria de Cognição Incorporada e Situada’), empiricamente sustentada pelas ferramentas da Teoria dos Sistemas Dinâmicos, da Teoria da Auto-Organização e de várias tendências da robótica faz da opção por uma modelagem não representacional um tópico a ser considerado seriamente (ver AGRE, 1997; CLANCEY, 1997; CLARK, 1997; GONZALEZ, 1996; HASELAGER; GONZALEZ, 2004; JUARRERO,

1999; KELSO, 1995; PORT; van GELDER, 1995; THELEN; SMITH, 1994). Considerando as dores de cabeça apresentadas acima, esta abordagem pode oferecer um certo alívio à Ciência Cognitiva.

### **Agradecimentos**

O autor agradece a Jelle van Dijk, Iris van Rooij, Piet Smit, Eunice Gonzalez e Mariana Broens pelos comentários construtivos durante a preparação deste artigo, a Mariana Broens pela excelente tradução, e à FAPESP e ao NICI, por financiarem a pesquisa que é a base deste artigo.

### **Referências**

AGRE, P.E. *Computation and human experience*. Cambridge: Cambridge University Press. 1997.

AIZAWA, K. Explaining systematicity. *Mind & Language*, v. 12, p. 115-136. 1997.

BECHTEL, W. Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science. *Cognitive Science*, v. 22, p. 295-318. 1998.

BEER, R. A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, v.72, p. 173-215, 1995.

BRAITENBERG, V. *Vehicles: experiments in synthetic psychology*. Cambridge, MA: MIT Press, 1984.

BROOKS, R. Intelligence without representation. *Artificial Intelligence*, v.47, p.139-159. 1991.

BROOKS, R. Intelligence without reason. In: STEELS, L. & BROOKS, R. (Ed.), *The artificial life route to artificial intelligence: Building embodied, situated agents*. Hillsdale, NJ: Lawrence Erlbaum, 1995. p. 25-81.

CHALMERS, D. J. Connectionism and compositionality: why Fodor and Pylyshyn were wrong. *Philosophical Psychology*, v. 6, p. 305-319. 1993

CHALMERS, D.J. On implementing a computation. *Minds and Machines*. v. 4, n. 4, p. 391-402, 1995.

CHALMERS, D.J. Does a rock implement every finite-state automaton? *Synthese*, v.108, p. 309-333, 1996.

CHRISLEY, R.L. Why everything doesn't realize very computation. *Minds and Machines*, v.4, n.4, p. 403-420, 1995.

CHURCHLAND, P. M. *A neurocomputational perspective: the nature of mind and the structure of science*. Cambridge: MIT Press, 1989.

CHURCHLAND, P. S.; SEJNOWSKI, T. J. *The computational brain*. Cambridge, MA: MIT Press, 1992.

CLANCEY, W.J. *Situated cognition: on human knowledge and computer representation*. Cambridge: Cambridge University Press. 1997.

CLARK, A. *Being there: putting brain, body and world together again*. Cambridge, Massachusetts: MIT-Press, 1997.

CLIFF, D.; NOBLE, J. Knowledge-based vision and simple visual machines. *Philosophical Transactions of the Royal Society: Biological Sciences*, v.352, n.1358, p. 1165-1175, 1997.

DREYFUS, H. *What computer's still can't do*. Cambridge, MA: MIT Press, 1992.

FODOR, J.A. *The language of thought*. Cambridge: Harvard University Press, 1975.

FODOR, J.A.; PYLYSHYN, Z. W. Connectionism and cognitive architecture: a critical analysis. *Cognition*, v. 28, p.3-72, 1988.

GONZALEZ, M.E.Q. Ação, caudalidade e ruído nas redes neurais auto-organizadas. In: DEBRUN, M.; GONZALEZ, M.E.Q.; PESSOA JÚNIOR, O. (Ed). *Auto-organização: estudos interdisciplinares..* Campinas: Unicamp, 1996. p. 273-297. (Coleção CLE, v. 18).

GRUSH, R. The architecture of representation. *Philosophical Psychology*, v.10, n.1, p. 5-23, 1997.

GUHA, R.V.; LENAT, D.B. Cyc: a midterm report. *AI Magazine*, p.32-59. 1990.

\_\_\_\_\_. Re: Cycling paper reviews. *Artificial Intelligence*, n. 61, p. 149-174, 1993.

HADLEY, R.F. Systematicity in connectionist language learning. *Mind & Language*, v.9, p.247-272, 1994.

HARNAD, S. The symbol grounding problem. *Physica D*, v.42, p.335-346, 1990.

HASELAGER, W.F.G. *Cognitive Science and folk psychology: the right frame of mind*. London: Sage, 1997.

\_\_\_\_\_. On the potential of non-classical constituency. *Acta Analytica*, v. 22, p. 23-42, 1999.

\_\_\_\_\_.; van RAPPARD, J.F.H. Connectionism, systematicity and the frame problem. *Minds and Machines*, 8, p. 161-179, 1998.

\_\_\_\_\_.; De GROOT, A.D.; van RAPPARD, J.H. Representationalism versus anti-representationalism: A debate for the sake of appearance. *Philosophical Psychology*, v. 16, n.1, p. 5-23, 2003.

\_\_\_\_\_. (2004). Auto-organização e comportamento comum: Opções e problemas. In: SOUZA, G.M.; D'OTTAVIONO, I.M.L.; GONZALEZ, M.E.Q. (Ed.). *Auto-organização: Estudos interdisciplinares*. Campinas: Unicamp, 2004. p. 213-235.(Coleção CLE, v. 38).

HAUGELAND, J. Representational genera. In: RAMSEY, W.M., STICH, S.P.; RUMELHART, D.E. (Ed.) *Philosophy and connectionist theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1991, p. 61-90.

HORGAN, T.; TIENSON, J. *Connectionism and the philosophy of psychology*. Cambridge, Mass.: MIT-Press, 1997.

HUBEL, D.H.; WIESEL, T.N. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, London, v.195, p.215-243, 1968.

JUARRERO, A. *Dynamics in action: intentional behavior as a complex system*. Cambridge, MA: MIT Press, 1999.

KELSO, J.A.S. *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, Mass: MIT-Press, 1995.

MCCARTHY, J.; Hayes, P.J. Some philosophical problems from the standpoint of artificial intelligence. In: MELTZER, B.; MICHIE, D. (Ed.). *Machine intelligence*. Edinburgh: Edinburgh University Press, 1969.

NEWELL, A. Physical symbol systems. *Cognitive Science*, v. 4, p. 135-183, 1980.

PORT, R.F.; VAN GELDER, T. *Mind as motion: explorations in the dynamics of cognition*. Cambridge, MA: MIT Press, 1995.

PUTNAM, H. *Representation and reality*. Cambridge, MA: MIT Press, 1988.

PYLYSHYN, Z.W. (Ed.). *The robot's dilemma*. Norwood: Ablex, 1987.

SEARLE, J. Minds, brains, and programs. *Behavioral and Brain Sciences*, v.3, p. 417-424, 1980a.



\_\_\_\_\_. Intrinsic Intentionality. *Behavioral and Brain Sciences*, v. 3, p. 450-456, 1980b.

SIMON, H. *The sciences of the artificial*. Cambridge, MA: MIT-Press, 1969/1996.

SMOLENSKY, P. On the proper treatment of connectionism. *Behavioral and Brain Sciences*, v. 11, p. 1-74, 1988.

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

VAN GELDER, T. What might cognition be, if not computation? *Journal of Philosophy*, v. 92, p. 345-381, 1995.

# Informação e auto-organização: tramas conceituais revisitadas

*Maria Luísa Bissoto*

## **Apresentação**

A partir das décadas de 1970-80, o conceito de auto-organização vem ganhando crescente importância<sup>1</sup>; fato que pode ser atribuído ao caráter predominantemente conexcionista e dinâmico que tem, mais recentemente, pautado as pesquisas ligadas à vida e à inteligência artificial e à intensificação da investigação dos fenômenos do vivo pela óptica dos sistemas complexos. Tais pesquisas também renovaram o vigor das discussões teóricas quanto ao conceito de informação, por muito tempo predominantemente ligadas a concepções de caráter sintático. Com a intenção de colaborar com o estudo do conceito de auto-organização, este artigo trará correlações traçadas entre os conceitos de auto-organização e de informação, por diferentes vieses teóricos, considerando-se que o modo como o conceito de informação é concebido, conjugado a como se compreende que um sistema “contacta” informação, determina diferentes compreensões do conceito de auto-organização, e assim o alcance e as possibilidades desse conceito como recurso teórico para a explicação dos diferentes fenômenos aos quais ele é/está cada vez mais associado.

## **Considerações referentes ao conceito de auto-organização**

A definição elaborada por Michel Debrun guiará a abordagem conceitual a ser aqui utilizada em relação à auto-organização:

Há auto-organização cada vez que, a partir de um encontro entre elementos realmente (e não analiticamente) distintos, desenvolve-se uma interação sem supervisor (ou sem supervisor onipotente) — interação essa que leva eventualmente à constituição de uma “forma”

---

<sup>1</sup> Shalizi (2001b), contabilizando o número de entradas para o conceito de auto-organização em resumos de dissertações nos EUA, encontrou 4 entradas antes de 1970 e 593 entradas entre os anos 1999-2000.

ou à reestruturação, por 'complexificação', de uma forma já existente<sup>2</sup>. (1996a, p. 13)

Antes que um evento possa ser considerado como resultante de auto-organização há, portanto, princípios delimitadores a serem observados. Tais princípios ajudam a diferenciar um estado auto-organizado, aquele em que houve diferenciação na complexificação do sistema, daqueles, mais comuns, nos quais ocorreram somente simples rearranjos, ou *reorganizações* (COLLIER; HOOKER, 1999). Esses princípios envolvem:

- \* a não existência de coações externas, conduzindo, direcionando, peremptoriamente, os rumos organizativos do sistema. Isso diz respeito também às condições iniciais, que "disparam" o processo auto-organizativo: longe de determinarem a seqüência e previsibilidade desse processo, o que contrariaria a idéia de organização *por si* imbuída no termo "auto", as condições que desencadeiam a auto-organização são antes importantes por "impulsionarem" o sistema a uma ruptura com o estado no qual esse se encontra.

- \* da mesma maneira, como assertado por M. Debrun (1996, p. 35) "a auto-organização não se reduz à simples maturação ou explicitação (passagem aristotélica da "potência" ao "ato") de uma estrutura inata". É um *momento* diferenciado de criação, propício à emersão de uma *nova* forma, que atende à situação *presente* do sistema existente, ou daqueles elementos que formarão um sistema, e não o "desenrolar, a explicitação de um sistema já dado." (M. DEBRUN, 1996, p. 53).

- \* não se coaduna com o conceito de auto-organização a existência de um controle centralizador permanente, mesmo que esse seja inerente ao sistema. (Se um tipo onipotente de auto-consciência nos possibilitasse o controle de todos os nossos atos, não haveria possibilidade de auto-organização no organismo humano, por exemplo);

- \* dinamismo, entendido como busca, como procura, por estados configuracionais mais satisfatórios para o sistema como um todo, a partir da "percepção" de diferentes opções de estados; o que se traduz em considerar um

---

<sup>2</sup> Ao primeiro caso, "à constituição de uma "forma", Debrun aplica a denominação de auto-organização primária, e ao segundo caso, de "reestruturação" de uma forma já existente, de auto-organização secundária. Maiores detalhes quanto às especificidades de cada uma podem ser encontrados em Debrun, 1996a, 1996b e 1997.

sistema auto-organizativo operando sempre em instabilidade e capaz de quebra de simetria;

- \* possibilidade de alcançar situações limite, ou de criticalidade, que favoreçam a mudança de fase/estado;

- \* possibilidade de emergência de novos estados a partir de interações localizadas num dado espaço do sistema;

- \* manutenção da estabilidade do sistema perante variações exteriores;

- \* a abertura do sistema a fluxos de matéria/energia, ou seja, o sistema deve poder processar energia/matéria; não pode estar num fluxo zero de trocas ou equilibrado termodinamicamente, pois isso indicaria a extinção do sistema.

- \* a memória do sistema deve ser/estar *distribuída* por todo o sistema. (CILLIERS, 1998)

O conceito de organização é de difícil explicitação; entretanto, o mais importante a ser entendido em relação a esse termo — e é daí que advém a dificuldade em defini-lo — é que a concepção de que há algo organizado está *intrinsecamente ligada àquele que observa*; é dependente da, ou relativa à, figura do observador. Assim, a delimitação de variáveis, que permitem discernir se um todo se organiza/está organizado não está — somente — nas propriedades pertinentes à coisa organizada, mas sim, na relação que se estabelece entre ambos: a mudança de observador, ou no observador, acenará à possibilidade de diferentes concepções do que é/está organizado, do que é todo, e do que/quais são as partes<sup>3</sup>.

Empregado em abordagens de objetos muito distintos — podendo variar da matemática aplicada à psiquiatria, dos pregões da bolsa ao modo de vida de cupins e de formigas, e aos estudos da robótica, entre outros —, o conceito de auto-organização acompanha aqueles que se referem às concepções de matéria<sup>4</sup> e ao modo como essa se organiza: como essa se arranja/movimenta, de como as partes, se as há, se ordenam e interagem entre si, para formar o "substrato real", em que vivemos.

---

<sup>3</sup> Ver ROSS ASHBY, 1962 e HUMBERTO MATURANA, 1999.

<sup>4</sup> E assim também as concepções de informação, já que informação pode ser concebida como "trechos" de conteúdos materiais que representam, caracterizam ou desvelam a "realidade" circundante, possíveis de serem transmitidos, "captados" ou interpretados numa relação transcorrida entre emissor/receptor ou entre sistema/entorno, dependendo das correntes teóricas que definam esse termo.

Nesse sentido, o conceito de auto-organização se desenvolveu, firmando-se enquanto nomenclatura científica, somente no século XX, no bojo dos estudos dos sistemas cibernéticos: daqueles envolvidos no surgimento das primeiras redes neurais (MCCULLOCH; PITTS, 1943) e dos primeiros computadores, na década de 1940. Em virtude desse histórico, as principais áreas da ciência que suportaram o desenvolvimento do conceito de auto-organização foram a física e a ciência da computação; sendo ainda importantes "incubadores" teóricos a "descoberta" da matemática não linear, da matemática dos fractais, da teoria do caos e da teoria dos sistemas.

Na história mais recente do conceito de auto-organização os nomes de William Ross Ashby (1903-1972)<sup>5</sup>, e de Heinz Von Foerster (1911-2002), físico vienense, aparecem como precursores na utilização e nos esforços de fundamentação desse conceito, já num segundo momento da cibernética. Momento no qual a concepção prevalecente não era mais tão fortemente caracterizada pela óptica de rigidez na organização do sistema, própria do primeiro momento da cibernética (cibernética de primeira ordem), mas, sim, na importância da autonomia (auto-gerenciamento) dos sistemas e no papel exercido pelo observador (pelo conhecedor) na constituição desses: a construção de modelos computados a partir de outros modelos cibernéticos (cibernética da cibernética, metacibernética ou cibernética de segunda ordem). Cabe porém a ressalva de que esse segundo momento da cibernética não representou uma ruptura total com a concepção, mais reducionista, presente na cibernética de primeira ordem; sendo marcadamente influenciada pelos pressupostos dessa (HEYLIGHEN; JOSLYN, 2001).

É nesse cenário cibernético que a questão da informação — primeiramente nos moldes de uma teoria da comunicação, como concebida por C. Shannon, de caráter eminentemente sintático —, imbrica-se fortemente ao estudo da autonomia organizativa dos sistemas: grosso modo, a informação seria o fluxo de *dados* (ou de sinais) provenientes da exterioridade ou de partes inerentes ao próprio sistema, que "alimentariam" as regulações a serem (auto) efetivadas por esse, de modo a obter e manter uma performance otimizada,

---

<sup>5</sup> Ashby, psiquiatra inglês, foi um dos principais cibernéticos das décadas de 50 e 60, destacando-se pelo aprofundamento lógico matemático que realizou sobre as redes neuronais, explicando o funcionamento cerebral exclusivamente pelos mecanismos físico-químicos de integração neuronal, trabalho que levantou questões que puseram em xeque os modelos até então estabelecidos, redirecionando os rumos das pesquisas cognitivas. (BISSOTO, 1999).



frente às condições encontradas. Concepções de informação e de auto-organização, que embora ainda presentes, passaram e têm passado por importantes transformações, levando a outras concepções de auto-organização.

### **Auto-organização e informação: quatro vieses teóricos**

Em *Introduction to Cybernetics*, 1956, embora ainda não utilizando o termo auto-organização<sup>6</sup>, Ashby desenvolve uma teoria matemática *da regulação e do controle*, anunciando pressupostos que permitissem estudar sistemas complexos: aqueles "que são tão dinâmicos e interconectados que a alteração de um fator imediatamente age como causa a evocar alterações em outros, talvez mesmo em um grande número desses" (1956, p. 6). O importante, segundo Ashby, seria entender a extensão em que o sistema está sujeito a determinar e controlar esses fatores. Exemplos de sistemas complexos seriam colônias de formigas, sociedades humanas, o funcionamento cerebral, a economia, entre outros.

Os mecanismos de regulação e controle estudados por Ashby se referem às restrições (coerções), dinamicamente baseadas, oriundas do próprio sistema, que limitam o acesso — ou a chegada — desse a certos espaços (ou fases) de estado. Isso ocorre pelo bloqueio do fluxo de informação, repleto de "ruído", direcionando o sistema para aquelas variáveis essenciais à sua manutenção (1956, p. 196). Por "dinamicamente baseada" devem ser entendidas as relações de feedback (de retroalimentação, de "comunicação" de estados) transcorridas no interior do sistema. Os "mecanismos" de controle e a regulação também comparam condições relevantes de estado do sistema com condições de referência, podendo corrigir os sinais de feedback mais adequadamente. Toda e qualquer aprendizagem particularizada (entendida aprendizagem como modificação de comportamento) efetivada por um sistema é tributária desses mecanismos de regulação e controle da *informação* (essa devendo ser entendida, em Ashby, como o que remove, diminui, a incerteza do sistema. Ver Ashby, 1956, secções 7/22 e 13/7).

Ashby, apoiado na teoria da comunicação de Shannon, imbrica à presença de organização uma restrição, uma delimitação, quanto ao rol de

---

<sup>6</sup> Usa o termo auto-catálise, que é por vezes empregado como termo correlato à auto-organização por teóricos da Biologia. (MIT Encyclopedia of Cognitive Science, 1999).

estados possíveis de serem assumidos por uma/numa relação de acoplamentos. Entretanto, posto que um tal número de estados, enquanto abstrações, excede aqueles existentes no mundo físico real, uma delimitação desses estados inelutavelmente ocorre quando esse acoplamento é confrontado por um observador, no mundo físico real. A teoria da organização se mostra assim, em grande parte, "concernente a *propriedades que não são intrínsecas à coisa, mas são relacionais entre observador e coisa*" (ASHBY, 1962, p. 258, grifo do autor).

Interligadamente, se a organização é (mas não só, no sentido de pura idealização) relativa ao observador, a definição do que é parte e do que é todo *idem*, e da mesma maneira as concepções de complexidade de um sistema, das qualidades negativas ou positivas de uma organização e da própria auto-organização. São fenômenos dependentes da "revelação", do (d)escrutínio, da *seleção*, transcorrida no contexto observador/observado. Isso equivale, entretanto, ao questionamento da possibilidade (auto)-organizativa de um sistema, pois a condição de *organizado* — e de sistema, por suposto — estaria, *a priori*, na mente do observador, e não no sistema em si.

Em relação ao conceito de auto-organização propriamente dito, a análise desenvolvida por Ashby retira desse conceito qualquer coerência. O autor atribui ao conceito de auto-organização dois possíveis significados: o primeiro deles se refere à formação advinda de partes que, anteriormente separadas — isto é, comportamentalmente independentes —, se conjugam, mudando do "não organizado para o organizado" (ASHBY, 1962, p. 267). Embora Ashby se refira a esse tipo de configuração como uma forma "honesta" de auto-organização, considera também que denominá-la como auto-conectante seria mais simples (menos ambíguo), pois a ação "real" praticada pelas partes, ou seja, a mudança de um comportamento independente para outro marcado pela condicionalidade devida ao enlace entre partes, foi efetuar uma conjunção. Qualquer outra consideração a respeito dessa organização é inferência do observador

O segundo sentido se refere àquele sistema que, já organizado, passa por uma alteração (novidade, criação) de *comportamento*. Na argumentação de Ashby, não é possível pensar que um sistema altere sua programação no âmbito de seu próprio domínio; isso geraria uma sucessão causal infinita de meta-programas logicamente superiores. Se há, entretanto, alteração na organização de um sistema, e na impossibilidade dessa ser gerada exclusivamente no interior

desse próprio, isso deve ser atribuído ou a) à interferência (input) de um agente exterior ao sistema, ou b) à inabilidade do observador que, adotando uma óptica mais restrita, julga como novidade o que já era pertinente ao sistema. Novamente, a concepção de *auto-organização* está em xeque, pois, no primeiro caso, um sistema só poderia ser entendido como auto-organizado à medida que é parcialmente organizado *pelo outro*: a entrada de informação (input) tenderia a influenciar a ação organizativa de um sistema num certo rumo e não em outros; e no segundo caso, a alteração no sistema estaria, antes, no observador.

As questões — e comprovações — levantadas por Ashby em relação ao conceito de auto-organização estremeceram muitas das fundamentações do pensamento cibernético até então desenvolvido, marcando um ponto de transição nesse pensamento. O segundo momento da cibernética seria caracterizado pela elaboração de modelos teóricos, que buscavam "solucionar" as objeções por eles postas em relação à possibilidade de autonomia dos sistemas.

Heinz von Foerster, em *Observing Systems* (1984), já na esteira das discussões levantadas por Ashby, traz outro enfoque referente à questão auto-organização-informação, também enfatizando a relevância que o papel do observador assume na constituição do "real". Repetindo ao longo dessa obra a frase "o meio (o entorno) não contém informação; o entorno é como ele é", o autor define um sistema auto-organizativo como uma classe restrita de sistemas que, estando em contato estreito e constantemente interativo com um entorno que disponha de *energia e ordem* (estrutura), tem seu potencial (termodinâmico) de desordem (ou de ampliação de probabilidades de estados) maximizado; logrando manter e incrementar sua organização interna, no transcorrer do tempo, exatamente por encontrar como viver às expensas dessa disponibilidade.

Nesse âmbito de pensamento, o entorno é definido como uma representação do sistema, fundada no processamento — interno — de relações (inferenciais ou dedutivas) por ele elaboradas; representações que são pertinentes apenas ao acoplamento desse sistema ao meio, não existindo anteriormente a esse acoplamento. Esse processamento é determinante para a auto-organização do sistema por se relacionar com o aporte, transferência e realocação de *energia* entre as diferentes partes do mesmo, de acordo com as solicitações recursivas desse como um todo; e se confundem com a própria concepção de cognição defendida por von Foerster: é o que o sistema faz para *transformar* seu entorno em seu próprio benefício, viabilizando sua



sobrevivência. Os estados de coerência interna (ou de maior resistência a perturbações), que o sistema logra alcançar por meio desse jogo cognitivo, são denominados como *Eigenbehaviors*; e enquanto pontos de estabilidade dinâmica tendem a circularmente influenciar, sempre juntamente com a fisiologia própria do sistema, a natureza dos futuros acoplamentos sistema/entorno. O *auto* assim se explica: é a marca da autonomia de um sistema que regula suas próprias regulações.

Também nas décadas de 1950-1960, há que se destacar o trabalho de Herman Haken (1927- ) um dos teóricos pioneiros no estudo da auto-organização, tendo por base pesquisas realizadas com o raio laser. Subjacente à perspectiva de auto-organização presente nos estudos desse autor se encontra a idéia de que um sistema pode incrementar sua organização por ser capaz de gerar informação a partir de condições iniciais randômicas (não determinísticas).

Em linhas muito gerais, a hipótese desenvolvida por Haken foi a de que o incremento na energia de um fluxo de raios laser adviria de uma *sincronização* (de uma ação cooperativa entre elementos ocupando um mesmo estado de fase) dos componentes atômicos desses, o que ocorreria quando um certo patamar máximo de inputs de energia fosse atingido. Esse incremento de energia levaria esses componentes a amplificarem, pela repetição e retransmissão, um certo tipo de informação — uma mensagem ou sinal, percebida como parâmetro de ordenação pelo sistema — criando um campo num feixe de luz, "escravizando" e forçando a movimentação desses átomos ritmicamente no tempo; gerando assim uma nova estruturação organizacional. Houve a geração de uma informação de mais alta complexidade (ordem), produzida pelas relações de cooperação "comunicacionais" estabelecidas entre os componentes do sistema. Todo esse fenômeno está implicado na definição de auto-organização proposta pelo autor (2000, p. 11): "a aquisição de uma estrutura espacial, temporal ou funcional, por parte de um sistema, sem interferência específica do exterior"; o sistema seria "provocado" pelo exterior, mas não de uma maneira determinista.

Haken estende vários dos pressupostos da teoria sinérgica por ele desenvolvida para sistemas biológicos de maior complexidade, como os seres humanos e as sociedades; embora deixe clara a impossibilidade de responder se fenômenos tão complexos quanto aqueles ligados à biologia poderiam ser entendidos exclusivamente pelo modelo teórico por ele proposto. Entretanto,

esses pressupostos, enquanto relacionados ao movimento (auto)organizativo de estruturas altamente “instáveis” (afastadas do equilíbrio termodinâmico) e não apontando uma relação de causa-efeito direta entre a entrada de uma *dada* informação e a re-organização do sistema, podem ser vistos como precursores teóricos de um “alargamento” do conceito de informação e de auto-organização, mormente em relação àqueles anteriormente aqui citados. E é em relação a possíveis extensões do conceito de auto-organização encontrado em Haken, para a compreensão dos fenômenos ligados ao vivo, que algumas considerações são traçadas abaixo.

A primeira delas se refere às afirmações de Haken quanto a *pontos de instabilidade*. Haken parte do princípio de que sistemas são sempre, continuamente, constrangidos por parâmetros externos, que delimitam seu raio e possibilidades de atuação. Esses parâmetros podem ser a quantidade de energia/alimento disponível para um sistema em seu entorno, condições de habitat, normatizações sociais... Pontos de instabilidade são alcançados sempre que os parâmetros que constroem mais significativamente os sistemas mudam. Para além desses pontos de instabilidade, o sistema pode adquirir diferentes possíveis estados e necessita de informação adicional quanto a qual estado escolher (HAKEN, 2000, p. 27). Essa informação adicional pode ser desencadeada pela interação do meio com o código genético, por limites advindos da mobilização “interna”, mais particular, do sistema ou do contexto de inserção do sistema, “que julga o valor do tipo de estado a ser estabelecido” (HAKEN, 2000, p. 27).

Com base no exposto, argumentamos então que, no caso dos seres humanos, os limites constrangedores quanto às possibilidades de ser se encontram, em muito, estabelecidos na rede de engajamentos, na *circunstância*, constituída pelos humanos ao se porem no mundo, ao viverem. Em assim sendo, ao atingir pontos de instabilidade, novas possibilidades organizacionais serão buscadas — escolhidas — nessa rede de engajamentos, e selecionadas congruentemente aos juízos de valor aí construídos. Tendo em vista que a efetivação dessas escolhas engendrará novos sinais (informações) ordenadores, que serão retransmitidos, amplificados e difundidos, criando um campo de inter-relações sinérgicas, que “canalizará” o fluxo de ação nesse campo, o *tipo* de possibilidades de escolhas que conduzirão o sistema a novas oportunidades organizacionais se reveste de fundamental importância. Que tipos de



possibilidades organizacionais podem emergir de uma rede de engajamentos centrada na exacerbação da individualidade, na cristalização de condutas quanto à diversidade de viveres, tão marcante do modo de vida atual, chamada aliás, de Sociedade da Informação?

Nas palavras de Haken (2000, p. 29):

Em minha opinião, o estudo da informação em sistemas biológicos é também de interesse para a sociedade moderna, cujo próprio funcionamento repousa na adequada produção, transferência e processamento de informação. Talvez o aspecto mais importante que tem emergido é aquele da causalidade circular, que resulta em um estado coletivo, que na sociologia pode representar um clima social, uma opinião geral pública, uma democracia ou uma ditadura.

Uma segunda "ponte": Haken afirma que sistemas complexos, enquanto sistemas afastados do equilíbrio termodinâmico, portanto, com trajetórias organizativas indeterminadas e irreversíveis no tempo, só são cognoscíveis em relação aos meandros de seu "operar" mediante o estabelecimento de um *processo de aprendizagem*; processo esse que revele caminhos para que se consiga lidar com tal classe de sistemas. A chave para essa aprendizagem é a compreensão de que não é o entendimento do comportamento de partes individuais que deve ser buscado, mas, antes, da *orquestração* das partes. Deixando claro que a visão de orquestração numa concepção sinérgica não remete a relações causais diretas, nem a um resultado esperado para tal orquestração, nem a um controlador onipotente a reger tal orquestração — embora haja níveis hierárquicos de organização em um sistema.

Argumentamos que a concepção de que é preciso *aprender* a lidar com sistemas complexos abre espaço para a reflexão de que a relação entre seres humanos não é, de saída, espontânea, natural. Enquanto sistemas altamente complexos, há também um *aprender* a ser humano, dependendo desse aprendizado os caminhos pelos quais trilhará a espécie. Da mesma forma, há que se buscar a compreensão dessa humanidade em estruturas outras que não o ser individual: não é a frágil barreira da pele que limita o ser, é o campo relacional, no qual esse está/se coloca. É para esse campo que o olhar investigativo deve se dirigir.

Uma terceira e última "ponte": a construção de sentido — o "tratamento" da informação que reduzirá a redundância do sistema, — requer um alto grau de cooperatividade por parte do sistema e uma repetida interação sistema/meio. A

informação não é, portanto, no entender de Haken (2000), uma propriedade estática, mas um *processo* no transcurso do qual o contexto e a relevância de novas organizações são checadas, reforçadas ou restringidas, várias vezes. O que pode remeter diretamente à idéia de que o aprender – entendido nesse artigo como a construção/emersão de novas configurações organizativas num sistema —, é movimento de *ajuste*, de idas e vindas imanentes ao engajamento vivencial do sistema: processo de longo prazo e continuidade em seu sempre checar, ajustar, orquestrar as partes.

### **Considerações finais**

O incremento de explicações — científicas — de mundo fundadas sobre o conceito de auto-organização, mesmo observando-se as diferenças de enfoque, vem permitindo o uso de outras "lentes" conceituais, principalmente no que se refere: a) à concepção das irregularidades da natureza, que passam a não mais serem vistas como anomalias, mas sim como probabilidades possíveis de ocorrer no âmbito do vivo, b) à valorização da diversidade, c) à busca por processos dinâmicos, substituindo a busca por concepções estáticas de mundo e d) à crítica à visão de que é o meio que controla, de fora para dentro, a manutenção da vida, a partir de uma adaptação do sistema ao meio: antes, cabe ao sistema *selecionar*, no âmbito do seu engajamento com o meio, as provocações que lhe permitirão manter sua ontogenia. Processo que leva a conceber informação como *agir continuamente interpretativo*, posição já defendida, por exemplo, nos estudos da biossemiótica ou por teóricos da percepção, como J. J. Gibson.

Contudo, por mais que a auto-organização seja um conceito sedutor, principalmente em seu viés de aplicabilidade aos fenômenos do vivo, pelas oportunidades aventadas de construir outras bases para a compreensão do humano, há que se reconhecer que, ao mesmo tempo, esse é um conceito de difícil teorização. As razões para isso são várias. As já aqui expostas, como as armadilhas semânticas do termo *auto*, que põem em xeque a própria possibilidade de emprego científico do conceito, a relativa precariedade quanto à definição de termos-chave como organização, padrão, ordem, se juntam outras, talvez ainda mais nebulosas: o *como* a interação mantém um campo, ou numa visão mais clássica, um sistema, coeso; o *quanto* de diversidade a auto-organização de um campo (ou sistema) pode suportar, ou, dito de outro modo,

que tipo de diversidade "dispara" a auto-organização? Além disso, como definir um conceito que é processo, que é melhor compreendido na ausência de substancialidade, sem acabar — sem tornar nulo e incoerente — com o próprio conceito?

Mas, parafraseando H. Letiche e J. Maier (2001), talvez mais do que um conceito, a auto-organização seja um *compromisso*: a procura por elaborar compreensões não reducionistas do mundo-e-do-humano; impulso para a construção de "realidades" melhores.

### Referências

ASHBY, R. Principles of the self-organizing system. In: VON FOERSTER, H.; ZOPF, G. (Ed.) *Principles of self-organization*. Pergamon Press, 1962. p. 255-277.

\_\_\_\_\_. *An introduction to cybernetics*. London: Chapman & Hall, 1956. Versão on line: 1999Disponível em: <<http://pcp.vub.ac.be/books/IntroCyb.pdf>> Acesso em: nov. 2002.

BISSOTO, M. L. *Deficientemente humano?: uma incursão na relação cognição/normalidade*. 1999. 185 f. Dissertação (Mestrado em Educação) – Universidade Metodista de Piracicaba, 1999.

CILLIERS, P. *Complexity & Postmodernism: understanding complex systems*. London: Routledge, 1998.

COLLIER, J.D.; HOOKER, C.A. Complexly organised dynamical systems. *Open Systems and Information Dynamics*, n. 6, p. 241-302, 1999.

DEBRUN, M. A dinâmica da auto-organização primária. In: DEBRUN, M; GONZALES, M.E.Q.; PESSOA JÚNIOR, O. (Org.) *Auto-organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*, Campinas/SP: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, v.18, p. 25-59, 1996b. (Coleção CLE) 450p.

\_\_\_\_\_. A idéia de auto-organização. In: DEBRUN, M; GONZALES, M.E.Q.; PESSOA JÚNIOR, O. (Orgs) *Auto-organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*, Campinas/SP: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, v.18, p. 3-23, 1996a. (Coleção CLE) 450p.

\_\_\_\_\_. Auto-organização e ciências cognitivas. In: GONZALES, M.E.Q. et al (Orgs) *Encontro com as ciências cognitivas*. 2ed., Marília: EdUnesp, 1997. 275p.

GIBSON, J.J. A preliminary description and classification of affordances. In: REED, E.S.; JONES, R. (Ed.). *Reasons for Realism*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1982. Cap. 4.9, pt.2, p. 403-406.

HAKEN, H. *Information and self-organization: a macroscopic approach to complex systems*. 2ed. Berlin: Springer-Verlag, 2000. 220p.

HEYLIGHEN, F.; JOSLYN, C. Cybernetics and second order cybernetics. In: MEYERS, R.A. (Ed.) *Encyclopedia of physical science & technology*. 3ed. New York: Academic Press, 2001.

LETICHE, H.; MAIER, J. *The 'self' in self-organization*. Disponível em: [www.imagilab.org/working\\_papers\\_pdf/WP2001-5b.pdf](http://www.imagilab.org/working_papers_pdf/WP2001-5b.pdf) Acesso em: out. 2002.

MATURANA, H. et al (Org). *Ontologia da realidade*. 2.ed. Belo Horizonte: Ed. da UFMG, 1999. 350p.

SHALIZI, C. *Causal architecture, complexity and self-organization in time series and cellular automata*. Disponível em: <http://www.santafe.edu/projects/CompMech/> Acesso em: dez. 2002. (livro, data original 2001b, 170 p.)

\_\_\_\_\_. *Notebook on self-organization*. 2001b. Disponível em: <http://cscs.umich.edu/~crshalizi/notebooks/self-organization.html> Acesso em: ago. 2002.

SHANNON, C. E.. A Mathematical theory of communication. *The Bell System Technical Journal*, v. 27, p.379-423, 623-656, July, October, 1948.

von FOERSTER, H. *Observing systems*. 2.ed., California: Intersystems Publications, 1984. 329p.

WILSON, R.; KEIL, F. (Ed) *MIT: encyclopedia of cognitive science*. Disponível em: <http://cognet.mit.edu/MITECS/>





# Atividade cerebral em múltiplas escalas e a experiência consciente

*Alfredo Pereira Júnior<sup>1</sup>*

## Introdução

O progresso recente das neurociências tem mostrado que os processos cognitivos no cérebro ocorrem simultaneamente em múltiplas escalas espaciais e temporais, envolvendo diferentes tipos de sinais físico-químicos. Os neurobiólogos que receberam o Prêmio Nobel de Medicina em 2001 e 2004 (respectivamente, P. Greengard, A. Carlsson e E. Kandel; L. B. Buck e R. Axel) estudaram níveis moleculares e bioquímicos de articulação das funções cognitivas no cérebro. Essas pesquisas levaram a um melhor entendimento dos mecanismos sinápticos e dos processos intra-celulares de transdução de sinais que suportam a cognição.

As duas últimas décadas de pesquisa nas neurociências têm conduzido a uma mudança de paradigma, na qual os processos cognitivos são correlacionados com processos que ocorrem em nível microscópico, na escala temporal de milissegundos, rompendo assim com pressupostos ingênuos de que tais mecanismos ocorreriam apenas no mesmo nível de organização da experiência consciente humana (isto é, nível macroscópico e escala temporal acima de 100 ms).

Entretanto, as descobertas que conduziram os neurobiólogos ao Nobel não implicam a conclusão de que os mecanismos subjacentes aos processos cognitivos seriam apenas químicos e moleculares locais; pelo contrário, nos defrontamos com um panorama de complexidade, sugerindo que um entendimento adequado das funções cognitivas no cérebro deve cobrir simultaneamente diversos níveis de organização e suas interações.

Tomando como exemplo o sistema visual dos mamíferos, notamos que diversos tipos de processos, distribuídos em diferentes escalas espaciais e temporais, contribuem para que tenhamos experiências visuais. A retina possui células especializadas para detectar padrões de informação em populações de

---

<sup>1</sup> Professor Adjunto do Instituto de Biociências — UNESP/Botucatu. Endereço eletrônico: [api@ibb.unesp.br](mailto:api@ibb.unesp.br)

fótons, oriundos do ambiente, que a atingem. Tal informação é transduzida para padrões de atividade elétrica, que são conduzidos através do feixe de axônios daquelas células (isto é, pelo nervo óptico) até o tálamo. O tálamo está conectado com outras estruturas subcorticais, como a formação reticular, que liberam transmissores e moduladores capazes de controlar a atividade do sistema límbico e do neocórtex, estruturas cruciais para a geração dos conteúdos da consciência. Portanto, o tálamo e as outras estruturas subcorticais alimentam o sistema límbico e o neocórtex de uma variedade de padrões elétricos oriundos da interface com o ambiente, e de substâncias químicas reguladoras; aqueles sistemas, por sua vez, realizam diversos processos a partir desses sinais, vindo a construir o mundo da experiência consciente.

### **Correlatos macroscópicos**

O estudo científico da consciência humana se faz através da busca de correlações entre medições da atividade cerebral, medições comportamentais e relatos (verbais e não-verbais) das experiências subjetivas (vide FRITH et al., 1999; PEREIRA JR., 2003a). Os padrões de atividade cerebral que acompanham uma determinada atividade consciente são intitulados "correlatos neurais da consciência".

Com certeza, parte da operação do sistema límbico e neocórtex se manifesta na escala macroscópica, em termos de padrões de atividade eletromagnética, medidos pelo eletroencefalograma (EEG) e magnetoencefalograma (MEG), correspondendo a padrões de fluxo sanguíneo para suas diferentes partes, conforme medido pela ressonância magnética funcional (fMRI) e pela tomografia por emissão de pósitrons (PET).

Para muitos neurocientistas, tais padrões macroscópicos de atividade constituiriam os correlatos últimos da experiência consciente (vide, por exemplo, MCFADDEN, 2002). Contudo, essas correspondências entre atividade neuronal e atividade consciente são até o presente pouco precisas; elas abrangem apenas os diferentes *estados de consciência* (por exemplo, a distinção entre estado de vigília, sono profundo e sono REM), e os *tipos de cognição* (modalidades perceptuais, mecanismos de atenção, funções executivas e motoras), mas não dizem respeito aos *conteúdos da consciência* (isto é, não permitem inferir o que

se passa na consciência de uma pessoa através de dados sobre sua atividade cerebral na escala macroscópica).

Contrariamente à crença de que o cérebro constitua uma máquina elétrica, os correlatos elétricos da consciência são escassos. Quando se acompanha a trajetória dos sinais aferentes até o córtex sensorial, constata-se que, após ativar as áreas primárias, os sinais "desaparecem", ou seja, não é possível medir com o EEG ou o MEG atividades elétricas diretamente correlacionadas com o conteúdo consciente. Por outro lado, aumentos de amplitude dos potenciais de ação podem ocorrer durante a anestesia ou durante o sono "slow wave", quando a consciência se encontra ausente ou extremamente reduzida (vide discussão em SEWARDS; SEWARDS, 2001).

Com a utilização de aparelhos de produção de imagens da atividade cerebral (PET-scanners e fMRI) é possível obter boas correlações entre atividade cerebral e processos conscientes, inclusive nas áreas corticais associativas, porém existe uma controvérsia na literatura a respeito do tipo de processo fisiológico realmente medido por esses aparelhos. Em ambos os casos (considerando-se a modalidade "BOLD" do fMRI), esses aparelhos medem o fluxo de sangue arterial para as diversas regiões cerebrais, mas não se sabe se o aumento de consumo de glicose e oxigênio nessas áreas refletiria principalmente a atividade elétrica (geração de potenciais locais), ou a atividade metabólica no interior dos neurônios e glia em escalas espaciais microscópicas (por exemplo, atividades desencadeadas pelo íon de cálcio).

Uma característica central dos padrões de atividade macroscópicos relacionados com os processos cognitivos conscientes é a existência de *coerência* (conceito que foi proposto como princípio de auto-organização cortical por SINGER, 1990). Isso significa que a distribuição espacial da excitação neuronal, ao longo do sistema nervoso central (SNC), apesar de ser extremamente dinâmica e variada ao longo do tempo, não tem caráter completamente redundante nem aleatório (para uma abordagem matemática, vide TONONI et al., 1998). A cooperação entre regiões relativamente distantes do SNC parece constituir requisito necessário, não só para a ocorrência de processos conscientes, como também para a capacidade mais geral de resolução de problemas com base no raciocínio inteligente.

Retornando ao exemplo do sistema visual, sabe-se que seu funcionamento no SNC dos primatas requer a cooperação de, ao menos, quatro subsistemas:

a) as áreas visuais primárias, que recebem os sinais elétricos oriundos do tálamo, levando à ativação de suas redes neuronais na forma de um "mapa", que reproduz as relações espaciais entre os itens informacionais tal como estes atingem a retina (ou seja, existe um isomorfismo espacial entre tal mapa — chamado de *mapa retinotópico* — e o campo visual da retina);

b) as áreas visuais secundárias, caracterizadas por neurônios especializados em detectar aspectos específicos, e relativamente simples, da informação oriunda dos sensores periféricos, como cor, direção de movimento, aspectos da forma, etc.;

c) as áreas associativas parietal e temporal: a primeira é especializada em identificar *onde* estão os objetos percebidos (e por isso é chamada de "where pathway"), e a segunda especializada em identificar formas complexas, como formas de rostos de pessoas (e por isso é chamada de "what pathway");

d) áreas específicas do córtex pré-frontal, como a área dorsolateral, que codifica conteúdos da memória de trabalho ("working memory"), e a área de controle do movimento ocular ("frontal eye field"), que está envolvida com o processo de atenção visual.

Diversos estudos recentes têm mostrado que a cognição visual envolve a cooperação desses quatro subsistemas (e ainda outras estruturas subcorticais, como o colículo superior); em termos de um processamento distribuído da informação, vindo a gerar uma experiência visual unitária, na qual os diversos aspectos processados em diferentes partes do SNC surgem integrados uns aos outros. Dois importantes indicadores do papel da coerência na produção da consciência estão bem estabelecidos no estudo de potenciais evocados humanos:

a) as oscilações sincrônicas nas faixas *gama* e *teta*, concomitantemente a processos conscientes. As oscilações dizem respeito a uma sincronização dos potenciais de ação de milhões ou bilhões de neurônios, podendo ou não gerar disparos axonais (quando não geram, são chamados de *potenciais subliminares*, uma terminologia traduzida do inglês "subthreshold"; vide Engel et al., 2001). Diversas confirmações da existência deste fenômeno no ser humano foram obtidas (vide RODRIGUEZ et al., 1999), mas a interpretação de seu significado está em aberto. Podemos notar que tais mecanismos que geram coerência seriam



importantes para a integração do conteúdo consciente, porém não constituem, eles mesmos, os veículos informacionais destes conteúdos (vide discussão em PEREIRA JR. e ROCHA, 2000);

b) os ERPs ("event related potentials"), que são variações (positivas ou negativas) bem definidas no EEG, em intervalos de tempo de algumas centenas de milissegundos (p.ex., 300 ou 700 ms) após a apresentação de um estímulo, que se supõe seja processado conscientemente. Acredita-se que os ERPs se correlacionem com o processamento cognitivo consciente, quando algum aspecto do estímulo gera surpresa ou estado de alerta (p.ex., uma frase contendo erro gramatical).

Para se entender melhor como o conteúdo consciente é processado pela atividade coletiva e coordenada dos neurônios, é necessário enfocar o que ocorre nas unidades participantes dessas redes, ou seja, é preciso considerar os processos informacionais nos neurônios individuais, revelando como e por que a cooperação referida acima seria necessária para que determinados eventos venham a ocorrer nos mesmos. Neste trabalho, argumentamos pela necessidade de enfocar tanto a informação molecular, presente nas proteínas que compõem as cadeias de sinais celulares, quanto, segundo nossa hipótese, a informação quântica carregada por íons biologicamente relevantes.

### **A informação molecular é relevante para o estudo da consciência?**

O trabalho de E. Kandel e colegas, entre outros estudos contemporâneos, tem mostrado que os mecanismos elétricos ainda estão longe de constituir o "fim da linha" do processamento de informação cerebral. Concomitantemente à produção dos potenciais de ação existem níveis de organização nos quais ocorrem outros tipos de processamento, que terminam por se entrelaçar de modo circular, envolvendo até a expressão de genes.

Thagard (2003), entre outros, argumenta que os sinais moleculares também devem ser considerados como correlatos dos processos cognitivos. Tal abordagem molecular não seria necessariamente reducionista, podendo ser entendida como parte de um modelo multi-dimensional da atividade cerebral. O papel de íons, moléculas e macromoléculas deve ser pensado em termos da dinâmica do cérebro como um todo, em suas interações com o corpo e com o ambiente (PEREIRA JR., 2001a).



Quando um sinal bioelétrico é enviado pelo axônio de um neurônio, e atinge seu terminal sináptico, faz com que íons de cálcio ali presentes rompam vesículas contendo substâncias transmissoras, as quais se projetam através da fenda sináptica e se ligam aos receptores de membrana do neurônio pós-sináptico. Neste momento recomeça um novo tipo de processamento de informação.

É preciso distinguir, no funcionamento sináptico e no metabolismo intra-neuronal, entre os mecanismos bioquímicos costumeiramente enfocados nos manuais de fisiologia e os propriamente informacionais, que dizem respeito à transmissão da informação biológica molecular e que são enfocados na literatura de genética molecular. Os primeiros geralmente dizem respeito às transações moleculares nas quais ocorrem reações químicas, ou seja, há troca ou compartilhamento de elétrons. Já nos processos moleculares propriamente informacionais, há um outro tipo de interação, baseada na configuração tridimensional das macromoléculas biológicas, e em campos eletrostáticos microscópicos, sendo expressa na metáfora da "chave e fechadura".

As macromoléculas possuem sítios ativos (i. e., contendo um micro-campo eletromagnético), que adotam uma determinada conformação tridimensional, possibilitando ou não "encaixes" com outros sítios de outras macromoléculas. Ocorrendo um encaixe, há uma interação dos campos eletromagnéticos, provocando alterações em uma das macromoléculas (chamamos de *efetores* as macromoléculas que causam a alteração, e de *receptores* aquelas que são afetadas). Tais alterações constituem a *transição alostérica* (vide MONOD; CHANGEUX; JACOB, 1963), a qual segue a seguinte regra: todos os sítios de uma proteína devem ficar em um mesmo estado, ou no estado *relaxado* ou no estado *tenso*. Portanto, basta que um efector altere um sítio para que os demais adotem o mesmo estado.

Os receptores localizados na membrana do neurônio pós-sináptico (em uma estrutura chamada de *densidade pós-sináptica*) são proteínas alostéricas (macromoléculas "grandes", com sítios ativos) que se ligam a efetores, os quais podem ser neurotransmissores (macromoléculas de tamanho médio) ou neuromoduladores (macromoléculas de tamanho pequeno, também chamadas de hormônios). Em uma explicação pouco rigorosa, os transmissores seriam os portadores da informação oriunda do neurônio pré-sináptico, enquanto os moduladores controlariam o ritmo da atividade sináptica.

A primeira classe de receptores de membrana é a dos receptores *ionotrópicos*, que controlam o fluxo de íons através da membrana, constituindo assim um mecanismo que irá determinar a atividade eletromagnética, ou seja, a produção dos potenciais de ação. Os íons mais utilizados para tal função são os cátions de sódio e potássio. Os transmissores mais comuns que agem sobre os receptores ionotrópicos são o glutamato e a acetilcolina.

Dentre os ionotrópicos, o receptor glutamatérgico NMDA (n-methyl-d-aspartato) se diferencia por contribuir menos para tal atividade eletromagnética e mais como porta de entrada dos cátions de cálcio, os quais, devido à complexidade de sua estrutura eletrônica (não encontrada nos demais íons acima citados), carregam sinais biológicos (vide GHOSH; GREENBERG, 1995; ALKON et al., 1998; LOEWENSTEIN, 1999; JAISWAL, 2001), controlando circuitos intra-celulares de processamento de informação chamados de "caminhos de transdução de sinais" (do inglês *signal transduction pathways*, ou STPs).

Outros receptores, chamados *metabotrópicos*, também têm sua ação dirigida para o controle de STPs intra-neuronais, se ligando a G-proteínas e ativando a produção de cAMP (adenosina monofosfórica cíclica), que é fonte de energia para os STPs. Eventualmente, receptores metabotrópicos também controlam o funcionamento de receptores ionotrópicos, por meio de STPs internos ao neurônio (vide ilustrações em PEREIRA JR.; LUNGARZO, 2005). Os neuromoduladores mais comuns que se ligam aos receptores metabotrópicos são a dopamina, serotonina e noradrenalina, mas alguns destes receptores também podem ser ativados pelos transmissores glutamato e acetilcolina.

Como se pode facilmente constatar, devido à existência de dezenas de diferentes tipos de transmissores, moduladores e receptores nas sinapses, o "código" informacional molecular que opera nesse nível do processamento é bem mais rico que o código utilizado na transmissão dos sinais elétricos entre os neurônios. Portanto, seria correto afirmar que as sinapses não "refletem" ou "representam", mas antes *interpretam* os sinais oriundos do neurônio pré-sináptico, incluindo os sinais provenientes do mundo externo ao sistema nervoso central. Desse modo, o estudo dos STPs se aproxima da semiótica.

O conhecimento sobre STPs é bastante recente, e ainda limitado a partes desses processos. Algumas simulações computacionais têm sido feitas, a partir do conhecimento parcial existente, para se tentar entender o que ocorre nesse

nível de organização (vide BHALLA e IYENGAR, 1999). Sabe-se que os sinais moleculares e iônicos que ativam STPs incluem, além dos íons de cálcio que entram pelos canais NMDA e pelos canais de cálcio regulados por voltagem, e dos receptores metabotrópicos associados a G-proteínas, também elementos compartilhados com os sistemas endócrino e imune, e fatores de crescimento celular (chamados de "fatores tróficos").

Há um envolvimento conjunto de neurotransmissores, receptores e moduladores no processo gerador da consciência e suas alterações. Por exemplo, o anestésico ketamina ministrado em dose sub-anestésica causa uma hipoatividade dos receptores NMDA no córtex sensorial, coexistindo com uma hiperatividade dos receptores AMPA. Os efeitos da ketamina também estão relacionados à diminuição da inibição GABAérgica no sistema límbico, levando ao aumento da disponibilidade da dopamina, que interferiria na atividade das áreas frontais e também no controle estriatal da liberação tálamo-cortical da acetilcolina, favorecendo a ocorrência de distorções perceptivas e alucinações (vide PEREIRA JR.; JOHNSON, 2003).

Diversos correlatos da atividade consciente têm sido encontrados em nível molecular, mais especificamente na atividade de determinadas enzimas que ocupam posições-chave em processos de transdução de sinais intra-neuronais. Diversas descobertas apontam para um papel da família de proteínas quinase no processamento cognitivo e emocional (vide discussão em PEREIRA JR.; LUNGARZO, 2005). Estima-se que existam cerca de 512 tipos de enzimas no "quinoma" (proteoma das quinases), que correspondem a 2% do proteoma humano. Considerando-se uma média de 15.000 proteínas por célula, teríamos 300 diferentes tipos de quinases operantes em cada neurônio. Já foram detectadas as seguintes correlações:

a) processos de aprendizagem e formação de memória declarativa (consciente) envolvem a proteína quinase A (PKA), a proteína quinase C (PKC), a proteína quinase calmodulina-dependente II (CaMKII), e a proteína quinase mitógeno-ativada (MAPKb);

b) processos psicóticos, envolvendo de suscetibilidade a alucinações, incluem a proteína quinase creatina (CK);

c) processos relacionados à dor envolvem a proteína quinase C (PKC);

d) processos relacionados à aprendizagem do medo (relativamente a determinados estímulos ou contextos) envolvem a proteína quinase A (PKA), proteína quinase calmodulina-dependente IV (CaMKIV) e LIMK-1;

e) processos relacionados ao estresse, com efeitos sobre a memória associativa, envolvem a quinase ciclina-dependente 5 (Cdk5).

Outros processos moleculares envolvidos na co-produção das funções cognitivas do cérebro e na interação cérebro-(restante do) corpo seriam:

a) componentes dos STPs atingem o retículo endoplasmático e núcleo dos neurônios, levando à ativação de genes que possivelmente reporiam as proteínas necessárias para a manutenção dos próprios processos (em termos de sinapse, membrana e STPs) que conduziram à ativação desses genes (vide o modelo proposto por BAILEY; KANDEL, 1995, a partir de um trabalho pioneiro com o neurônio gigante de moluscos). Criam-se assim ciclos de interação proteínas-genes, que podem durar por muito tempo na vida de um organismo individual, e que seriam a base da memória de longo termo (p.ex., as lembranças que temos de fatos que ocorreram durante nossa infância);

b) em alguns STPs, ou mesmo em nível do retículo ou do núcleo, seriam ativados processos de produção dos neuropeptídeos, que atuam não só como moduladores da atividade das sinapses, como também atravessam a barreira cérebro-corpo, difundindo-se via circulação para outras partes do corpo (p.ex., já foram encontrados no sistema digestivo de roedores), possivelmente controlando um novo nível de processamento, responsável por reações psicossomáticas.

### **Íon de cálcio: um portador de informação física com papel central na cognição**

Um dos modelos existentes (vide ROCHA; PEREIRA JR.; COUTINHO, 2001; PEREIRA JR., 2001, 2003; ROCHA; MASSAD; PEREIRA JR, 2005) para se explicar a relação entre os padrões de atividade macroscópica (eletroquímica e sanguínea) e a atividade celular microscópica dos neurônios e células gliais se baseia na atividade fisiológica do íon de cálcio, o qual é considerado como um portador de informação (vide GHOSS; GREENBERG, 1995; ALKON et al., 1998; LOEWENSTEIN, 1999; JAISWAL, 2001).

Como se sabe, o íon de cálcio tem diversos papéis no funcionamento cerebral (vide CARAFOLI, 2002):



a) como carregador de sinais através da astroglia, compondo as chamadas “ondas de cálcio”, que dão suporte a várias funções;

b) como agente que controla a já mencionada abertura de vesículos contendo neurotransmissores, nos terminais axônicos, iniciando a transmissão sináptica;

c) como *primeiro mensageiro*, trazendo informação de fora para dentro do neurônio, quando entra através dos canais de cálcio — sejam os canais regulados por voltagem, sejam os canais controlados pelo receptor NMDA. Este último também é controlado por ligantes (os transmissores glutamato e glicina), tendo sido bastante estudado por seu possível envolvimento nos mecanismos de aprendizagem e memória;

d) como *segundo mensageiro* dentro do neurônio, controlando diversas funções, inclusive podendo interferir na regulação da expressão gênica (vide Alkon et al., 1998).

Na hipótese teórica desenvolvida por Rocha, Pereira e Coutinho (2001), e ampliada por Rocha, Massad e Pereira (2005), a função de primeiro mensageiro do íon de cálcio é relacionada com o processo de consciência, o que explicaria, entre outras coisas, a estreita relação entre a consciência, a aprendizagem e a formação de memórias. Para que o canal NMDA se abra, é preciso que o neurônio receba sinais excitatórios sucessivos, em um certo intervalo de tempo (na escala de milissegundos). Por essa razão, o canal NMDA é considerado um “detector de coincidência”, que seria de especial valor para os processos de aprendizagem associativa. Tais sinais seriam providos por meio da atividade cooperativa entre neurônios piramidais no córtex sensorial (onde tal receptor recebe conjuntamente sinais aferentes e reentrantes), o que demonstra a possível relação desse mecanismo com os padrões de coerência mencionados acima. Com tal hipótese, o papel das oscilações sincrônicas para a geração da consciência seria a coordenação de uma população de detectores de coincidência, ou seja, de receptores NMDA, para a geração de estados coerentes em populações de íons de cálcio espacialmente distribuídas.

Esta atividade coordenada resultaria na produção de estados coerentes em populações de  $\text{Ca}^{++}$ , fluindo de modo temporalmente sincronizado através dos canais NMDA nos neurônios sensoriais, interagindo com a calmodulina e sendo confinadas nas espículas (estruturas presentes nos dendritos de sinapses excitatórias no neocórtex; vide Sabatini et al., 2001; 2002). Deste modo,



entendemos que tornariam possível um processo cerebral de computação quântica. O mesmo raciocínio pode se aplicar ao canal NMDA no sistema hipocampal, operando na faixa *teta*. A computação quântica constituiria um novo nível de processamento de informação, que permitiria uma comunicação não-local entre neurônios de diversas partes do cérebro, dando suporte aos estados unitários de consciência. Ressalte-se que tal hipótese (por mim defendida, juntamente com os autores acima citados), de que os processos conscientes seriam suportados pela informação e computação quânticas, é controvertida, não sendo aceita pela maioria dos neurocientistas, que trabalham em paradigmas nos quais os mecanismos explicativos adotados se restringem à física clássica.

Os receptores metabotrópicos atuam em uma escala temporal mais dilatada que o NMDA (isto é, na escala de segundos), que seria inadequada para suportar os processos perceptivos. Contudo, sabemos que estes receptores estão envolvidos na produção do humor, assim como em estados afetivos e emocionais (sensações de dor e prazer, medo e ansiedade, tristeza ou felicidade, etc.). Nos sistemas cerebrais relacionados com a geração destes estados, como o giro do cíngulo, a amígdala, o hipotálamo e o colículo inferior, podemos conjecturar que os receptores metabotrópicos, conjuntamente com neuropeptídeos liberados a partir da interação com o corpo e o ambiente, ativariam STPs que modulam os processos geradores da consciência perceptiva, para gerar estados conscientes unitários nos quais os conteúdos afetivos e emocionais se ligam aos conteúdos perceptivos.

### **Uma hipótese de ocorrência de informação e computação quânticas no cérebro, dando suporte aos processos conscientes**

Para se discutir a hipótese levantada, é preciso entender as características do íon de cálcio, tanto no contexto da atividade biológica cerebral, quanto no contexto dos computadores quânticos artificiais. Esses dois contextos têm características diferentes, mas também têm características comuns, que possibilitam a utilização de determinadas técnicas de realização experimental na computação quântica (em particular, a espectroscopia Raman, que já é largamente utilizada no estudo da estrutura de proteínas) para estudos 'in vivo' da dinâmica informacional do íon de cálcio no cérebro.

Constitui uma característica importante deste íon (mas não dos demais íons com atividade biológica, os quais não têm uma estrutura eletrônica suficientemente complexa), a capacidade de participar de processos de transmissão de informação molecular, atuando como um intermediário entre proteínas, “como se fosse um hormônio” (na expressão de LOEWENSTEIN, 1999). Uma de suas vantagens é a capacidade de atravessar poros nanométricos e percorrer distâncias relativamente grandes com rapidez, o que não é possível para as proteínas. Além disso, ele é suficientemente pequeno para participar de processos de computação quântica, o que explica, ao lado de sua relativa complexidade, a escolha feita pelos engenheiros que planejaram o computador quântico (mais especificamente, o modelo experimental de *armadilha de íons*; vide CIRAC; ZOLLER, 2000; NIELSEN; CHUANG, 2000).

A informação portada pelo íon de cálcio diz respeito a seus estados vibratórios e respectivos arranjos eletrônicos, tendo sido elucidada em estudos recentes de computação quântica. Os computadores de armadilha de íons têm sido ampliados para trabalhar com populações de íons (KIELPINSKI et al., 2002), o que indica a possibilidade de uma aproximação teórica com a dinâmica de populações de íons no cérebro. Mais importante ainda, em um sistema dinâmico com característica de auto-similaridade, como se acredita ser o cérebro, durante os estados coerentes típicos do processamento consciente, podemos levantar a possibilidade de que a informação codificada na escala microscópica reflita os estados globais do sistema, ou seja, os estados globais do corpo e suas relações com o ambiente.

Cada íon de cálcio, presente, entre outros lugares, nas terminações neuronais, na astroglia e nas espículas, pode assumir diversos estados informacionais, não havendo qualquer razão ‘a priori’ para se supor que esteja restrito a estados binários. Nos computadores artificiais, planejados para realizar operações binárias, alguns destes estados são “proibidos”. Experimentalmente, a indução de transições entre os estados permitidos, assim como a leitura dos mesmos, se faz através de *lasers* estruturados (utilizando-se por exemplo as técnicas de *Oscilações Rabi* ou *Espectroscopia Raman*; vide HUGHES et al., 1997; SMITH; DENT, 2005). O laser na faixa de frequência ultravioleta entra em ressonância com o padrão vibratório externo do íon de cálcio, causando uma transição entre estados internos.

Tal padrão vibratório é chamado de *fonônico*, tendo em vista que o fônon é o *quantum* de energia envolvido na agitação térmica, vibração molecular e formação de ondas sonoras. O estudo pioneiro do papel dos fônons em sistemas biológicos foi desenvolvido por Herbert Fröhlich, enfocando sistemas vibracionais presentes em macromoléculas e estruturas biológicas (vide Pokorny e Wu, 1998). A aplicação desse conceito para o íon de cálcio ocorreu mais recentemente, no contexto da realização experimental de computadores quânticos (vide HUGHES et al., 1997; NIELSEN; CHUANG, 2000).

Um sistema (como o íon de cálcio) no qual a atividade vibracional externa está acoplada com estados informacionais internos (por exemplo, com duas possíveis distribuições de energia na estrutura eletrônica) é chamado de *sistema de dois níveis* (que abreviarei por TLS, do inglês “two level systems” — vide Nagerl et al., 2000; o hamiltoniano que relaciona o estado vibracional externo com o estado informacional interno também é descrito em Nielsen e Chuang, 2000, p. 317-19). Para que a comunicação e computação quânticas ocorram, é preciso que haja um sistema maior composto por dois ou mais TLS, no qual se forme um *emaranhado* (do inglês “*entanglement*”) de estados dos TLS (para uma discussão filosófica, vide PEREIRA JR., 2001, 2003).

No cérebro, mecanismos macroscópicos, como os circuitos recorrentes de excitação elétrica, as oscilações sincrônicas, os canais de membrana e os compartimentos intra-neuronais (como as espículas), controlam a atividade de populações do íon de cálcio. Seu estado informacional seria espontaneamente determinado pela interação com proteínas como a S-100 (presente na astroglia) e a calmodulina (abundantemente presente na densidade pós-sináptica). Quando os mecanismos macroscópicos acima citados induzem a formação de um emaranhado quântico em uma população, a informação molecular previamente transmitida pela ligação com proteínas seria integrada, produzindo-se um momento de consciência (p. ex., com duração de 300 ms). A geração de uma seqüência de momentos de consciência se encadearia de modo quase contínuo, formando o fluxo de eventos que caracteriza a experiência consciente.

A interação entre processos macroscópicos e quânticos está prevista na constituição dos computadores quânticos artificiais, nos quais uma interação dos estados microscópicos com o nível macroscópico é necessária não só para introduzir e resgatar informações dos processos quânticos, mas também como condição prévia para produzir um estado global coerente em uma população de

TLS. No cérebro, tal interação é necessária não só para gerar os estados coerentes nos quais os íons de cálcio formam um emaranhado quântico, como também para se fazer uma leitura dos resultados do processamento quântico, que corresponderia ao momento em que os íons ligados à CaM interagem com a CaMKII. Desse modo, o processo consciente não seria epifenomenal, mas poderia influenciar o comportamento, a partir dos efeitos produzidos pela CaMKII nos processos intra-neuronais e também nos potenciais de membrana (p.ex., através do controle da CaMKII sobre o receptor AMPA).

Para se testar a hipótese de trabalho acima levantada, poder-se-ia utilizar a espectroscopia Raman 'in vivo' (vide HANLON et al., 2000), um procedimento não invasivo, para interferir com a informação quântica em uma população de íons de cálcio em uma determinada região cerebral, procurando-se não afetar outros tipos de atividade (como a amplitude ou a frequência dos potenciais de ação, ou a taxa de atividade metabólica). A previsão é que alterando-se o padrão de informação quântica em uma população de íons de cálcio, e apenas isso, haveria um efeito nos conteúdos da consciência. Não ocorrendo qualquer efeito, a hipótese seria rejeitada.

### **Comentários finais**

Os modelos reducionistas, que pretendem explicar a atividade do cérebro a partir de um nível privilegiado de descrição, têm encontrado dificuldades frente às evidências de que o cérebro simultaneamente atua em diferentes níveis organizacionais, e em interação com o corpo e com o ambiente. Um novo tipo de explicação da atividade cerebral se faz necessário, enfatizando a existência de complexas interações entre os diferentes níveis de processamento, e com fatores somáticos e ambientais. Possivelmente tais interações apresentam características observadas em outros sistemas dinâmicos, como a existência de valores críticos para a interferência da atividade de um nível de organização sobre outros.

Desse modo, os processos mentais correspondem à auto-organização do sistema cerebral, em suas interações com o corpo e o ambiente, sendo que, no tocante ao cérebro, os diversos níveis de organização — e respectivos modos de processamento da informação — seriam relevantes para se entender os diversos aspectos da cognição. A experiência consciente se correlaciona com atividades

presentes em todas as escalas de atividade cerebral. Considerando que, a partir da formação de estados coerentes globais, tais atividades venham a apresentar a característica de auto-similaridade, faz sentido a hipótese apresentada, de que os conteúdos da experiência consciente sejam suportados pela informação e computação quânticas em populações de íons.

Como o íon de cálcio tem um papel central na fisiologia do sistema nervoso central, e como (diferentemente dos demais íons com atividade biológica) também tem capacidade de codificar informação por meio da flexibilidade de seu arranjo eletrônico, as populações deste íon envolvidas em atividades coerentes poderiam ser o suporte físico dos conteúdos da consciência. Tal hipótese, por mim levantada, deverá ser testada, de modo não-invasivo, por meio de emissões estruturadas de laser 'in vivo', gerando interferências com os padrões fonônicos do íon de cálcio (e respectivas transições entre estados informacionais) para provocar efeitos conscientes, sem alterar as demais atividades cerebrais. Só a partir desse teste encontraremos (ou não) o procurado apoio experimental para aquilo que a hipótese presume, ou seja, a existência de computação quântica no cérebro.

Um dos principais campos para a interpretação e aplicação desses resultados é a psiquiatria. As relações entre neurociência e psiquiatria têm se estreitado ultimamente, e constituem objeto de interesse para ambas as partes. Um melhor entendimento da multi-dimensionalidade da consciência humana, incluindo não só a complexidade neurobiológica como também as atividades humanas no contexto social, é de grande importância para o entendimento do processo de saúde/doença mental. Neste ensaio, enfoquei principalmente a complexidade das escalas espaciais e temporais nas quais se desenrola a atividade cerebral, ficando em segundo plano a questão da complexidade temporal da experiência humana, que foi enfocada em um outro trabalho (PEREIRA JR.; GUERRINI, 2004).

**Agradecimentos:** Ao CNPq, por concessão de Bolsa de Produtividade, em suporte a esta pesquisa; ao parecerista anônimo e aos editores do livro, pelas correções e importantes sugestões apresentadas.



## Referências

- ALKON, D.L., et al. Time domains of neuronal  $\text{Ca}^{2+}$  signaling and associative memory: steps through a calyculin, ryanodine receptor,  $\text{K}^{+}$  channel cascade. *Trends in Neuroscience*. V.21, n.12, p.529-537, 1998.
- BAILEY, C.H.; KANDEL, E.R. Molecular and structural mechanisms underlying long-term memory. In: GAZZANIGA, M. (Ed.) *The cognitive neurosciences*. Cambridge: MIT Press, 1995.
- BHALLA, U.S.; IYENGAR, R. Emergent properties of networks of biological signaling pathways. *Science*, v. 283, p.381-387, 1999.
- CARAFOLI, E. Calcium signaling: a tale for all seasons. *Proc. Natl. Acad. Sci.*, v.99, n.3, p.1115-1122, 2002.
- CIRAC, J.I.; ZOLLER, P. A scalable quantum computer with ions in an array of microtraps. *Nature*, v.404, p.579-581, 2000.
- ENGEL, A.K.; FRIES, P.; SINGER, W. Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews Neuroscience*, v. 2, p.704-716, 2001.
- FRITH, C.; PERRY, R.; LUMER, E. The neural correlates of conscious experience: an experimental framework. *Trends in Cognitive Sciences*, v.3, n.3, p.105-114, 1999.
- GHOSS, A.; GREENBERG, M. E. Calcium signaling in neurons: molecular mechanism and cellular consequences. *Science*, v.268, p.239-247, 1995.
- HANLON, E.B. et al. Prospects for in vivo raman spectroscopy. *Phys Med Biol.*, v.45, n.2, p.1-59, 2000.
- HUGHES, R. J. et al. The los alamos trapped ion quantum experiment. 1997. Publicado eletronicamente em *quant-ph/9708050* 27 Aug 1997; republicado em *Fortschritte der Physik* v.46, p.329-361, 1998. Disponível em: <[http://kh.bu.edu/qcl/pdf/hughes\\_r19977f150e1f.pdf](http://kh.bu.edu/qcl/pdf/hughes_r19977f150e1f.pdf)>
- JAISWAL, J.K. Calcium: how and why? *J. Biosci.*, v. 26, n.3, p.357-363, 2001.
- KIELPINSKI, D., MONROE, C.; WINELAND, D.J. Architecture for a large-scale ion-trap quantum computer. *Nature*, v.417, p.709-711, 2002.
- LOEWENSTEIN, W.R. *The touchstone of life: molecular information, cell communication and the foundations of life*. New York: Oxford University Press, 1999.

- MCFADDEN, J. Synchronous firing and its influence on the brain's electromagnetic field: evidence for an electromagnetic field theory of consciousness. *Journal of Consciousness Studies*, v.9, n.4, p. 23-50, 2002.
- MONOD, J.; CHANGEUX, J.P.; JACOB, F. Allosteric proteins and cellular control systems. *J. Mol. Biol.*, v.6, p.306-329, 1963.
- NAGERL H.C., et al. Cavity QED-Experiments: atoms in cavities and trapped ions. In: BOUWMEESTER D.; EKERT, A.; ZEILINGER, A. (Ed.) *The physics of quantum information*. Berlim:Springer-Verlag, 2000.
- NIELSEN, M. A.; CHUANG, I.L. *Quantum computation and quantum information*. Cambridge: Cambridge University Press, 2000.
- PEREIRA JR., A.; ROCHA, A.F. Temporal aspects of neuronal binding. In BUCCHERI, R. (Org.): *Studies on the structure of time: from physics to psychopathology*. New York: Kluwer Academic: Plenum Publishers, 2000. p.97-106.
- \_\_\_\_\_. Coexisting spatial and temporal scales in neuroscience. *Minds and Machines*, v.11, n.4, p. 457-465, 2001a. Disponível em: <<http://www.springerlink.com/media/MH9XDNYTQP3WUERGXLDV/Contributions/Q/4/2/1/Q42185815996N557.pdf>>
- \_\_\_\_\_. The quantum mind/classical brain problem (versão original). *Informação e Cognição*, v.3, 2001b. Disponível em: <<http://www.marilia.unesp.br/atividades/extensao/revista/v3/artigo3.html>>
- \_\_\_\_\_. Uma abordagem naturalista da consciência humana. *Trans/Form/Ação* v.26, n.2, p.109-141, 2003a.
- \_\_\_\_\_. The quantum mind/classical brain problem (versão corrigida). *Neuroquantology*, v.1, p.94-118, 2003b. Disponível em: <<http://www.neuroquantology.com/2003/01/pp94.118.htm>>
- \_\_\_\_\_.; GUERRINI, I. A. Unidade e multiplicidade do tempo: uma abordagem transdisciplinar. *Interface*, v.8, n.15, p.247-256, 2004.
- \_\_\_\_\_.; JOHNSON, G. Toward an explanation of the genesis of ketamine-induced perceptual distortions and hallucinations. *Brain and Mind*, v.4, n.4, p. 307-326, 2003.
- \_\_\_\_\_.; LUNGARZO, C. A framework for the computational approach to cellular metabolism supporting neuronal activity. *International Journal of Computational Cognition*, v.3, n.3, 2005. Disponível em: <<http://www.yangsky.com/ijcc33b.pdf>>

POKORNY, J.; WU, T-M. *Biophysical aspects of coherence and biological order*. Berlin: Springer, 1998.

ROCHA, A.F.; PEREIRA JR, A.; COUTINHO, F.A. NMDA channel and consciousness: from signal coincidence detection to quantum computing. *Progress In Neurobiology*, v.6, n.64, p.555-573, 2001.

ROCHA, A.F.; MASSAD, E.; PEREIRA JR, A. *The brain: from Fuzzy Grammar to Quantum Computing*. Berlin: Springer, 2005.

RODRIGUEZ, E. et al. Perception's shadow: long-distance synchronization of human brain activity. *Nature*, v.397, p.430-433, 1999.

SABATINI, B. L.; MARAVALLI, M.; SVOBODA, K.  $Ca^{2+}$  signaling in dendritic spines. *Current Opinion in Neurobiology*, v.11, n.3, p.349-356, 2001.

SABATINI, B. L.; OERTNER T.G.; SVOBODA K. The life cycle of  $Ca^{2+}$  ions in dendritic spines. *Neuron*, v.33, n.3, p. 439-452, 2002.

SEWARDS, T.V.; SEWARDS, M.A. On the correlation between synchronized oscillatory activities and consciousness. *Consciousness and Cognition*, v.10, p. 485-495, 2001.

SINGER, W. Search for coherence: a basic principle of cortical self-organization. *Concepts in Neuroscience*, v.1, n.1, p.1-26, 1990.

SMITH, E.; DENT, G. *Modern raman spectroscopy: a practical approach*. J. Wiley and Sons, 2005.

THAGARD, P. How molecules matter to mental computation. *Philosophy of Science*, v.69, p.429-446, 2002.

TONONI, G.; EDELMAN, G.M.; SPORNS, O. Complexity and coherency: integrating information in the brain. *Trends in cognitive science*, v.12, n.2, p.474-484, 1998.

# A postura da mão influencia a compatibilidade estímulo-resposta numa tarefa de reconhecimento da lateralidade

*Lameira, A. P.<sup>a</sup>; Gawryszewski, L. G.<sup>a</sup>; Carvalho, A. A.<sup>a</sup>;  
Soeiro-Santos, R.<sup>a</sup>; Guimarães-Silva, S.<sup>a</sup>; Ferreira, F.M.<sup>b</sup>; Pereira Jr,  
A.<sup>c</sup>; Vargas, C. D.<sup>d</sup> e Umiltà, C.A.<sup>e 1</sup>*

## Introdução

A mão direita é uma imagem especular da mão esquerda, ou seja, as mãos são objetos enantiomorfos. A discriminação de imagens especulares é particularmente difícil para os animais e para as crianças pré-alfabetizadas. Segundo Gross e Borstein (1978), esta é uma dificuldade adaptativa, pois a discriminação de figuras especulares só é necessária para humanos alfabetizados que têm de distinguir letras especulares, tais como, "d" de "b" ou "p" de "q".

Não seria de se esperar que a discriminação da lateralidade da figura de uma mão apresentada numa tela de computador fosse uma tarefa difícil para adultos alfabetizados e, na verdade, não o é. Mas é surpreendente como a determinação da lateralidade da figura de uma mão envolve não só o sistema visual, mas também os sistemas somestésico e motor. E que a decisão seja tomada através da formação da imagem mental de uma das mãos, sua projeção para a tela e a verificação da congruência ou não entre a figura na tela e a imagem da mão (PARSONS, 1987b; PARSONS; FOX, 1998).

Neste artigo, iremos analisar alguns fatores que podem afetar o tempo de reação necessário para determinar a lateralidade da figura de uma mão.

---

<sup>1</sup> <sup>a</sup> Programa de Neuroimunologia, Dep. Neurobiologia- UFF, CAIXA POSTAL 100.180, Niterói, RJ-24.001-970, Brazil; <sup>b</sup> Departamento de Fisiologia, UFPA e Department of Neurobiology, Duke University, USA; <sup>c</sup> Instituto de Biociências da Unesp/Botucatu; <sup>d</sup> Instituto de Biofísica, UFRJ, Brazil; <sup>e</sup> Dipartimento di Psicologia Generale, Università di Padova, Italy.



## **Compatibilidade estímulo – resposta**

O termo Compatibilidade Estímulo – Resposta (CER ou SRC-Stimulus Response Compatibility) é usado para descrever as interações entre elementos do estímulo e elementos da resposta que influenciam (podem facilitar ou inibir) a velocidade e a acurácia com que uma tarefa é executada. Ou seja, a presença de um certo atributo do estímulo faz com que os sujeitos tenham uma forte tendência para selecionar uma resposta particular em vez de outra resposta. Assim, a interação entre o estímulo e a resposta que produz latências menores e baixas taxas de erro é dita mais compatível do que a interação que produz latências maiores e altas taxas de erro (UMILTÁ; NICOLETTI, 1990).

Kornblum, Hasbroucq e Osman (1990) postularam a existência de vários tipos de CER baseados em dois conceitos importantes da relação estímulo – resposta. O primeiro é a noção de Sobreposição Dimensional, que se refere ao fato de o estímulo compartilhar ou não atributos categóricos com a resposta. Esta sobreposição não está restrita somente aos atributos físicos do estímulo e da resposta, mas também se estende para as suas representações mentais. Por exemplo, se o sujeito for instruído a pressionar uma tecla verde quando o estímulo verde aparecer e pressionar uma tecla vermelha quando o estímulo vermelho aparecer, a Sobreposição Dimensional entre o estímulo e a resposta ocorrerá através da cor, pois tanto o estímulo quanto a resposta compartilham esse atributo categórico. Se a instrução for apertar tecla verde quando aparecer um estímulo vermelho e a tecla vermelha quando aparecer um estímulo verde, o tempo de reação nesta condição (incompatível) será maior do que na condição anterior (compatível).

O segundo aspecto é a Relevância Dimensional, ou seja, a distinção entre as dimensões relevante e irrelevante do estímulo. A dimensão será relevante quando a resposta depender do valor do estímulo naquela dimensão, e irrelevante se o valor do estímulo nessa dimensão não estiver relacionado com a resposta requerida. Por exemplo, o sujeito é instruído a pressionar a tecla à direita quando o estímulo verde aparecer e a pressionar a tecla à esquerda quando o estímulo vermelho aparecer, sendo que os estímulos podem aparecer no hemisfério visual direito ou esquerdo. A dimensão relevante é a cor do estímulo, pois é este atributo que define qual resposta será executada. A dimensão irrelevante é o lado em que o estímulo aparece, pois isso influencia na



resposta do sujeito, mas não tem nenhuma relevância para a tarefa, ou seja, a resposta ao estímulo vermelho (verde) será mais rápida quando o estímulo aparecer à esquerda (direita) do que quando aparecer à direita (esquerda).

A combinação das noções de Sobreposição Dimensional e Relevância Dimensional formam quatro tipos de Compatibilidade Estímulo – Resposta (CER):

TIPO 1: Caracteriza-se por não existir nenhuma Sobreposição Dimensional em nenhuma das dimensões relevante ou irrelevante. Por exemplo, o sujeito é instruído a pressionar a tecla direita (ou esquerda) em resposta a um estímulo mostrado acima (ou abaixo) do ponto de fixação central. Para sujeitos destros, o tempo de reação é menor quando respondem com a tecla esquerda para o estímulo abaixo do ponto de fixação e com a tecla direita para o estímulo acima do que quando a situação é reversa (LÁDAVAS, 1987).

TIPO 2: Ocorre pela presença de Sobreposição Dimensional somente na dimensão relevante. Por exemplo, se um estímulo é apresentado no hemisfério visual direito (esquerdo) e o sujeito responde pressionando a tecla direita (esquerda), as latências serão menores do que quando o sujeito deve responder com a tecla esquerda (direita) quando o estímulo vem à direita (esquerda), ou seja, quando o estímulo e a tecla de resposta estiverem em lados opostos (ANZOLA et al, 1977).

TIPO 3: A dimensão relevante do estímulo não tem nenhum tipo de sobreposição com nenhuma dimensão da resposta, diferente da dimensão irrelevante. Por exemplo, são apresentadas aos sujeitos figuras geométricas e eles são instruídos a pressionar uma tecla esquerda em resposta a uma figura, e a tecla direita em resposta a outra figura. Os estímulos são apresentados no hemisfério visual direito e esquerdo. Embora o lado onde o estímulo aparece seja irrelevante para a tarefa, pois o sujeito tem que reconhecer a forma para determinar a resposta, as latências serão menores quando o lado da tecla de resposta corresponder com o lado onde o estímulo aparece do que quando estímulo e resposta estiverem em lados opostos. O mesmo é observado quando se usam estímulos com a mesma forma mas com cores diferentes (Umiltá e NICOLETTI, 1990).

TIPO 4: Existe sobreposição entre o estímulo e a resposta em ambas as dimensões relevante e irrelevante. Por exemplo, o sujeito é instruído a mover uma alavanca para cima em resposta a um som de alta frequência ou para baixo

em resposta a um estímulo de baixa frequência. Os estímulos eram emitidos aleatoriamente por dois alto-falantes posicionados um acima do outro. Os tempos de reação serão menores quando a frequência do som (frequência alta, por exemplo) coincidir com a posição do alto-falante (alto-falante de cima) do que quando não coincidir (frequência alta emitida pelo alto-falante de baixo). A frequência do som é a dimensão relevante e a posição dos alto-falantes é a dimensão irrelevante (SIMON et al., 1976)

A correspondência espacial é uma poderosa ferramenta para se obter a Sobreposição Dimensional entre os elementos do estímulo e os elementos da resposta. A Sobreposição Dimensional entre o estímulo e a resposta na dimensão relevante (TIPO 2), ocorrendo através da correspondência espacial, é chamada de Compatibilidade Espacial Estímulo-Resposta. Uma outra forma de Sobreposição Dimensional entre o estímulo e a resposta através da correspondência espacial, ocorrendo na dimensão irrelevante (TIPO 3), é o Efeito Simon. Então, a posição espacial do estímulo pode influenciar na seleção da resposta em ambas as dimensões, relevante e irrelevante, determinando assim latências menores quando o estímulo e a resposta estiverem no mesmo lado (condição compatível) do que quando eles estiverem em lados opostos (condição incompatível).

### **Reconhecimento da lateralidade das mãos**

Segundo Parsons (1987a, 1987b), numa tarefa de reconhecimento da lateralidade das mãos, o primeiro passo realizado pelos sujeitos é uma rotação mental da sua própria mão correspondente, seguida de uma comparação da sua mão mental com a mão-estímulo para a realização do julgamento “direito” ou “esquerdo”. Desta forma os sujeitos transformam mentalmente uma representação interna da própria parte do corpo para compará-la com o estímulo visual externo e assim comparar a lateralidade deles e fazer o julgamento (PARSONS, 1987a, 1987b).

O processo de rotação mental é um componente essencial para a tarefa de julgamento da lateralidade das mãos e, também, para o planejamento de um movimento real das mãos. Isso foi deduzido porque o tempo necessário para se fazer o julgamento da lateralidade das mãos, a rotação mental da representação interna da mão e o correspondente movimento real da mão são similares

(PARSONS, 1987b). Devido a essa relação entre mão imaginada e o movimento real da mão, o julgamento da lateralidade das mãos é influenciado pelas limitações anatômicas que dificultam o movimento real, ou seja, movimentos que são difíceis de executar, na realidade, são difíceis de reproduzir mentalmente e com isso dificultam o julgamento da lateralidade das mãos (PARSONS, 1994).

Parsons (1994) desenvolveu um modelo de reconhecimento da lateralidade de figuras das mãos (mão-estímulo) que compreende cinco passos: a) reconhecimento pré-atencional da lateralidade da mão-estímulo apresentada; b) processamento da orientação da própria mão representada mentalmente; c) planejamento da rotação mental da própria mão representada internamente para a orientação da mão-estímulo; d) simulação mental da rotação planejada; e) o “encaixe confirmatório” entre a orientação da mão representada internamente e a orientação da mão-estímulo. O estágio pré-atencional seria um conhecimento implícito precoce da lateralidade da mão-estímulo, e após o movimento imaginado da mão interna e do “encaixe confirmatório” o sujeito faria a tomada de decisão consciente da lateralidade desse estímulo.

Dessa forma, o reconhecimento da lateralidade da mão depende de vários fatores. Alguns parecem ser específicos para o uso da figura da mão como estímulo a ser discriminado, e outros dependem dos vários fenômenos de compatibilidade estímulo-resposta descritos acima. No experimento realizado, uma figura da mão (direita ou esquerda) é apresentada na tela de um computador. A figura pode aparecer no hemisfério esquerdo, no centro ou no hemisfério direito. A vista pode ser da palma ou do dorso da mão. Além disso, as mãos que respondem podem estar na postura prona (palma da mão para baixo) ou na postura supina (palma da mão para cima), pressionando interruptores localizados à direita e à esquerda da linha média do corpo.

## **Objetivos**

Levando em consideração que o desenho da mão aciona códigos visuais e somestésicos-motores, nossos objetivos são verificar se:

1. A compatibilidade estímulo-resposta que tem sido observada quando se usam estímulos geométricos (efeito Simon) também está presente quando se usam figuras de partes do corpo (mão esquerda e mão direita).

2. A postura da mão que responde influencia a latência para o reconhecimento da lateralidade da figura da mão, bem como se influencia os fenômenos de compatibilidade estímulo-resposta.

## **Materiais e métodos**

### **Sujeitos**

Oito estudantes universitários participaram do experimento como voluntários (cinco homens e três mulheres). Todos os sujeitos eram destros, apresentando quociente de lateralidade de 64% (um sujeito) e 95% (sete sujeitos) de acordo com o questionário de Oldfield (OLDFIED, 1981) e apresentavam acuidade visual normal. Os estudantes estavam numa faixa etária de 20 a 33 anos. Antes do início do experimento, o sujeito era informado detalhadamente sobre as tarefas a serem executadas e sobre como as respostas seriam medidas e avaliadas. Todavia, a hipótese de trabalho não era comunicada ao sujeito, já que isso poderia influenciar o seu desempenho. Finalmente, o sujeito era informado que poderia interromper o teste e desistir de participar do experimento, no momento em que assim decidisse. Este projeto foi aprovado pelo Comitê de Pesquisa da Pró-Reitoria de Pesquisa e Pós-Graduação da Universidade Federal Fluminense.

### **Aparato experimental**

O experimento foi conduzido em uma sala escura e com atenuação sonora. Um computador foi utilizado para apresentar os estímulos e registrar as respostas dos sujeitos. Os participantes apoiavam a cabeça em um suporte de frente e mento e eram orientados a mantê-la o mais estável possível.

O computador utilizado nas sessões era um PC 486, e a sequência de eventos durante o teste era determinada pelo programa MEL (Micro Experimental Laboratory, versão 2,0). Os estímulos eram exibidos num monitor VGA de 21 polegadas localizado a uma distância de 57 cm. Os estímulos eram desenhos do dorso e da palma da mão direita e da mão esquerda com um tamanho de 9,3° de largura por 15,5° de altura e eram apresentados a 7,5° à direita ou à esquerda do ponto de fixação central ou no meridiano vertical

(Figura 1). As respostas eram feitas pressionando dois interruptores localizados à direita e à esquerda da linha média do corpo.

Os sujeitos tinham o movimento ocular monitorado através de um sistema (Eye Track, Modelo 210- Applied Science Laboratories) que utiliza um emissor de raios infravermelhos, os quais eram refletidos pela esclera e captados por dois detectores localizados à direita e à esquerda do emissor. Esse sistema era fixado no apoiador de frente e mento, de forma que os detectores e o emissor fossem facilmente ajustados para cada pessoa. Somente a posição do olho direito foi monitorada através do sistema acima, que permitia que a posição horizontal do olho fosse acoplada à imagem da tela do monitor vista por uma câmera e acompanhada numa tela de televisão em uma sala adjacente. A posição do olho e a imagem da tela do monitor eram também gravadas em fita de vídeo para uma eventual análise posterior.

### **Procedimento experimental**

O experimento foi realizado em uma única sessão experimental que foi precedida por um curto treinamento. Nesse treinamento, consistindo de 10 a 20 testes, os sujeitos tomavam contato com o tipo de tarefa a ser realizada no bloco de testes a seguir.

A sessão era composta por quatro blocos de 78 testes, resultando num total de 312 testes por sessão. Em dois blocos, os sujeitos respondiam com a palma da mão voltada para baixo (postura prona) e, nos outros dois, a resposta era feita com a palma da mão voltada para cima (postura supina). O mesmo acontecia com a apresentação dos estímulos. Em dois blocos eram apresentados o desenho do dorso da mão e, nos outros dois blocos, o desenho era o da palma da mão. Desta forma, a postura da mão e o desenho da mão eram constantes no bloco de 78 testes. Os estímulos eram apresentados aleatoriamente à esquerda, no centro ou à direita e durante a realização dos testes, o indivíduo era instruído a:

- 1- manter o olhar fixo no ponto de fixação (PF) central até o aparecimento do estímulo, mas piscar brevemente depois de responder manualmente;
- 2- não olhar para o estímulo (desenho da mão) e sim continuar fixando o ponto de fixação;



- 3- responder o mais rápido possível à ocorrência do estímulo visual, pressionando o interruptor direito se o desenho fosse o da mão direita ou pressionando o interruptor esquerdo se o desenho fosse o da mão esquerda.

As condições experimentais empregadas foram as seguintes :

- 1) Dorso da Mão – Postura Prona (DP).
- 2) Palma da Mão – Postura Prona (PP).
- 3) Dorso da Mão – Postura Supina (DS).
- 4) Palma da Mão – Postura Supina (PS).

Variamos a distribuição das quatro condições de modo a facilitar a execução das respostas e a eliminar os efeitos de treinamento ao longo da sessão e de cansaço, ou seja, os sujeitos realizavam dois blocos de teste com a mesma postura (prona ou supina). Em um dos blocos, a figura era do dorso da mão e no outro, a figura era da palma da mão. Por exemplo, o sujeito 1 seguiu a sequência (DP) / (PP) / (PS) / (DS), o sujeito 2 seguiu a sequência (PP) / (DP) / (DS) / (PS), o sujeito 3 fez / (PS) / (DS) (DP) / (PP) e o sujeito 4 fez (DS) / (PS) / (PP) / (DP). Os outros quatro sujeitos duplicaram as seqüências acima.

As medianas dos tempos de reação manual (TRM) para as figuras das mãos aparecendo no hemisfério direito e esquerdo da tela foram usadas numa Análise de Variância (ANOVA) com os seguintes fatores:

- a) Vista da Mão: palma da mão e dorso da mão
- b) Hemisfério Visual de Estimulação: hemisfério visual direito e esquerdo
- c) Postura da Mão: postura Prona e postura Supina
- d) Compatibilidade: compatível (desenho da mão direita aparecendo no hemisfério visual direito ou vice versa) e incompatível (desenho da mão direita aparecendo no hemisfério visual esquerdo ou vice versa).

Realizamos também uma ANOVA para os TRM quando as figuras das mãos foram apresentadas no centro da tela. Os seguintes fatores foram utilizados:

- a) Vista da Mão: palma da mão e dorso da mão
- b) Postura da Mão: postura Prona e postura Supina
- c) Mão: desenho da mão direita e da mão esquerda

Além disso, quando necessário, empregamos o método de Newman-Keuls para testar e realizar a análise post-hoc. O nível de significância adotado foi  $p < 0,05$ .

## **Resultados**

Quando analisamos as respostas para os estímulos lateralizados (hemicampo esquerdo e direito), observamos que somente o fator Hemicampo Visual foi significativo. O TRM para um estímulo aparecendo no hemicampo visual direito (464ms) é menor ( $p=.050$ ) do que o TRM para um estímulo aparecendo no hemicampo visual esquerdo (477ms). Além disso, há também uma interação (FIGURA 2) entre Compatibilidade e Postura da Mão ( $F_{1,7}=8,047$ ,  $p=0,025$ ). A análise post-hoc, utilizando o método de Neuman-Keuls, mostrou que na postura prona, o TRM na condição compatível (448ms) é menor ( $p<0,05$ ) do que na incompatível (472ms). Todavia, na postura supina, não há diferença significativa entre as condições compatível (485ms) e incompatível (478ms). Ou seja, a única condição significativamente menor que as demais é o TRM na condição compatível (448 ms) na postura prona. As diferenças entre os TRMs na condição supina não são estatisticamente diferentes ( $p>0,05$ ) do TRM na condição prona incompatível (472 ms).

Para a ANOVA com os TRM quando as figuras das mãos eram apresentadas no centro da tela, encontramos que somente o fator MÃO foi significativo. O tempo de reação usando a mão direita em resposta à figura da mão direita (412 ms) é menor ( $F_{1,7}=5,847$ ,  $p=0,046$ ) do que o tempo de reação usando a mão esquerda em resposta à figura da mão esquerda (427,5 ms). É importante salientar que a resposta para a figura da mão esquerda (direita) é realizada com a mão esquerda (direita). Dessa forma, não é possível separar se esta diferença sem deve à discriminação da figura ou à mão que executa a resposta.

## **Discussão**

Nossos resultados mostram que o tempo de reação para a discriminação da lateralidade de figuras da mão aparecendo no hemicampo visual direito é menor do que para figuras aparecendo no hemicampo visual esquerdo. Esta diferença é independente da lateralidade da figura da mão. Ou seja, não existe diferença entre o tempo de reação para uma figura da mão direita (resposta com a mão direita) e para uma figura da mão esquerda (resposta com a mão esquerda). É possível que este efeito do hemicampo se deva ao fato de só

utilizarmos sujeitos destros. Todavia, não sabemos se esta diferença entre os hemicampos é resultado dos mesmos mecanismos responsáveis pela diferença entre os TRM obtidos com a mão esquerda e direita quando o estímulo vem no centro da tela, pois a resposta para a figura da mão direita é feita com a mão direita e para a figura da mão esquerda, com a mão esquerda.

A interação que encontramos foi entre a compatibilidade e a postura da mão. Assim, quando o sujeito responde com as mãos na postura prona, o tempo de reação para uma figura da mão esquerda (direita) aparecendo à esquerda (direita) é menor do que o tempo de reação para uma figura da mão direita (esquerda) aparecendo à esquerda (direita). Todavia, este efeito desaparece quando o sujeito responde com as mãos na posição supina.

Alguns resultados negativos merecem ser comentados. Não existiu nenhuma interação entre a vista da mão (palma ou dorso) e postura da mão (prona ou supina), nem para estímulos lateralizados, nem para estímulos no centro da tela. Na verdade, esperávamos que aparecesse um efeito de Compatibilidade Estímulo-Resposta entre estes fatores e que a resposta para a vista da palma da mão com a postura supina (palma para cima) fosse mais rápida do que com a postura prona (palma para baixo) e vice-versa para a vista do dorso da mão. Mas essa interação não foi significativa ( $P > 0,05$ ).

O efeito de Compatibilidade Estímulo-Resposta que existiu estava relacionado com a correspondência espacial e ocorreu na dimensão irrelevante. Porém, esse efeito ocorreu apenas quando os sujeitos respondiam na postura prona, ou seja, os tempos de reação eram menores quando a mão-estímulo correspondente aparecia no mesmo lado da tecla de resposta (mão direita aparecendo no hemisfério visual direito, por exemplo) do que quando a mão-estímulo e a tecla de resposta estavam em lados opostos. Esse efeito da correspondência espacial sobre a Compatibilidade Estímulo-Resposta ocorrendo na dimensão irrelevante desapareceu quando os sujeitos realizaram a tarefa respondendo com a mão na postura supina. A dimensão relevante do estímulo era a natureza direita ou esquerda da mão, a qual determinava a seleção da resposta.

Dessa forma, a dimensão relevante do estímulo (natureza direita e esquerda) se sobrepõe com a da tecla de resposta (tecla direita e esquerda) e a dimensão irrelevante (localização espacial) também. Portanto, segundo a taxonomia de Kornblum et al. (1990), o presente experimento seria uma tarefa

de Compatibilidade Estímulo-Resposta do TIPO 4, pois existe sobreposição entre o estímulo e a resposta na dimensão relevante e irrelevante.

A influência da postura sobre o processamento da informação irrelevante no nosso experimento parece estar relacionada ao fato de que o estímulo, por representar uma parte do corpo, não é apenas “visto” mas também pode ser “sentido”. Em tarefas de julgamento da lateralidade das mãos parece existir uma forte sobreposição entre o estímulo visual e a mão que responde (efetor), no sentido de que existe um movimento imaginário do efetor para a orientação da mão-estímulo e subsequente decisão consciente da lateralidade do estímulo (PARSONS, 1987a 1987b). Essa sobreposição mão-estímulo e mão que responde acontece na dimensão relevante, ou seja, no momento em que o sujeito deve avaliar o valor do estímulo para determinar a resposta.

Levando em consideração que existe uma sobreposição entre o estímulo visual e a mão que responde (efetor), além da sobreposição dimensional relevante e irrelevante entre a mão-estímulo e a tecla de resposta, podemos dizer que os códigos espaciais somestésicos-motores fornecidos na dimensão relevante interferem no processamento da informação espacial irrelevante do estímulo. Talvez, o fato de a postura prona ser uma postura bastante comum facilite a reprodução mental do movimento da mão de resposta, sendo fácil imaginar a mão que responde se movimentando para a orientação da mão-estímulo na tela, facilitando o julgamento da lateralidade do estímulo. Todavia, a postura supina, mais incomum, dificulta a reprodução mental do movimento da mão de resposta e, conseqüentemente, se torna difícil imaginar a mão se movimentando para a orientação da mão-estímulo que aparece na tela, dificultando o julgamento da lateralidade do estímulo e interferindo no processamento da informação espacial irrelevante do estímulo. É importante observar que, na posição supina, os tempos de reação (nas condições compatível e incompatível) não diferem dos tempos de reação observados na condição incompatível com a mão na postura prona. Ou seja, a postura prona parece facilitar a resposta na condição compatível, o que não ocorre na postura supina. Ainda são necessários mais experimentos para podermos determinar a real razão pela qual a postura da mão do sujeito interfere no processamento da informação espacial irrelevante do estímulo.

Em resumo, os nossos resultados mostram uma interação entre os fenômenos de compatibilidade espacial e a postura da mão. Estes resultados são



relevantes não só para os estudos dos mecanismos de reconhecimento de formas e padrões nos seres humanos e nas máquinas, mas também para os estudos sobre a interação entre os mecanismos sensoriais e motores em seres humanos e em robôs.

## **Referências**

ANZOLA, G.P. et al. Spatial compatibility and anatomical factors in simple and choice reaction time. *Neuropsychologia*, v.15, p.295-382.

GROSS, C. G.; BORNSTEIN, M. H. Left and right in science and art. *Leonardo*, v.11, p.29-38, 1978.

KORNBLUM, S.; HASBROUCQ, T.; OSMAN, A. Dimensional overlap: Cognitive basis for stimulus-response compatibility: a model and taxonomy. *Psychological Review*, v.97, p. 253-270, 1990.

LÁDAVAS, E. Influence of handedness on spatial compatibility effects with perpendicular arrangement of stimuli and response. *Acta Psychologica*, v.64, p.13-23, 1987.

OLDFIELD, R. C. The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia*, v.9, p.97-113, 1981.

OTTOBONI, G. et al. Is handedness recognition automatic? (no prelo).

PARSONS, L. M. Imagined spatial transformation of one's body. *Journal of Experimental Psychology: Human Perception and Performance*, v.116, p.172 - 191, 1987a.

\_\_\_\_\_. Imagined transformation of one's hands and feet. *Cognitive Psychology*, v.19, p.178-241, 1987b.

\_\_\_\_\_. Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance*, v.20, p.709-730, 1994.

\_\_\_\_\_; FOX, P. T. The neural basis of implicit movements used in recognising hand shape. *Cognitive Neuropsychology*, v.15, p.583-615, 1998.

SIMON, J. R. et al. Processing auditory information: interaction of two populations stereotypes. *Journal of Applied Psychology*, v.60, p.354-358, 1976.

UMILTÁ, C.; NICOLETTI, R. Spatial stimulus: response compatibility. In: PROCTOR, R. W.; REEVE, T.G. (Ed.) *Stimulus: response compatibility: An integrated perspective*. Amsterdam: North - Holland, 1990, p.89-116.



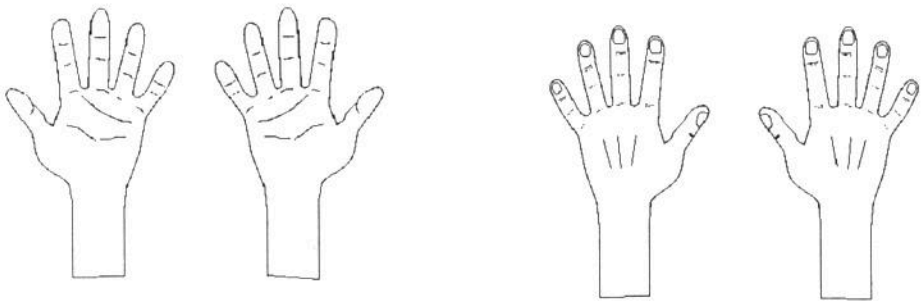


FIGURA 1- Desenhos das mãos que eram apresentados na tela (vista palmar à esquerda e vista dorsal à direita).

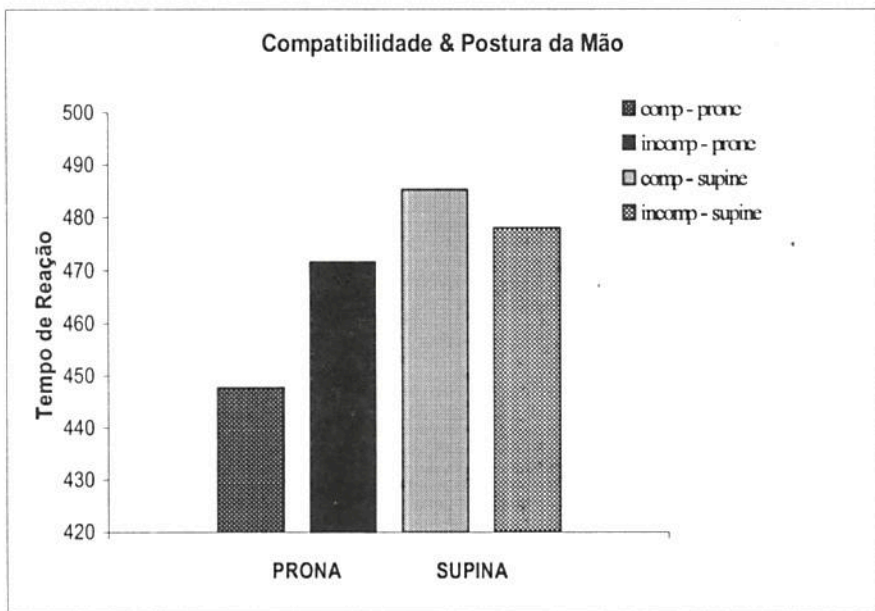


FIGURA 2 – Efeito das posturas prona e supina sobre os tempos de reação a estímulos ocorrendo no mesmo lado da resposta (condição compatível) ou no lado oposto ao da resposta (condição incompatível).



## O amplo conceito peirceano da mente: sua relevância para a biologia, inteligência artificial e cognição

*Lucia Santaella<sup>1</sup>*

Peirce fez algumas afirmações a respeito da mente e pensamento que têm continuamente perturbado seus comentadores. Entre essas afirmações, aquela que tem sido mais citada é a seguinte:

O pensamento não está necessariamente conectado a um cérebro. Ele aparece no trabalho das abelhas, nos cristais e espalhado pelo mundo puramente físico; e ninguém pode negar que ele esteja realmente lá, tanto quanto estão as cores, as formas etc. dos objetos. Se aderir consistentemente a essa negativa sem garantia, você será levado a alguma forma de nominalismo idealista próximo ao de Fichte. Não apenas o pensamento está no mundo orgânico como lá se desenvolve (CP 4.451).

A primeira conclusão que vem às nossas mentes, depois de ler essa passagem aparentemente estranha, é a de que há pensamento nos cristais (ver JOHANSEN 1993: 190). Mas Peirce foi ainda mais longe quando disse: “Definirei a essência da mente e a lei da causação final junto com sua aplicação a fenômenos não biológicos” (CP 7.374). Uma vez que essa visão do pensamento e da mente, como não restritos a uma moldura antropológica, é uma constante nos escritos de Peirce, especialmente depois de 1900, o primeiro passo a ser dado, para se entender o que ele pretende dizer com essas duas palavras – pensamento e mente – é reconhecer o fato de que temos de nos livrar dos seus significados convencionais e buscar pelo novo campo semântico no qual Peirce as inseriu. A primeira pista para realizar essa tarefa pode ser encontrada na sua Filosofia da mente (CP 7.362-7.688) que tem início com a crítica peirceana aos psicólogos seus contemporâneos, crítica, aliás, que ainda se mantém atual para a maioria das correntes da psicologia atual.

Para começar, os psicólogos ainda não esclareceram sobre o que é a mente. Não quero me referir ao seu substratum; mas eles nem tornaram claro o

---

<sup>1</sup> Universidade Católica de São Paulo. E-mail: lbraga@pucsp.br

que é um fenômeno psíquico. Ainda menos do que isso, nenhuma noção de mente foi estabelecida ou geralmente reconhecida que se possa comparar, por um instante, em precisão, com o conceito dinâmico de matéria. Quase todos os psicólogos ainda nos dizem que mente é consciência. Mas, tanto quanto sei, Hartman provou conclusivamente que existe mente inconsciente. O que se quer significar por consciência não é nada em si mesmo senão sentimento. Gay e Hartley [...] acharam que deve haver, e provavelmente há, algo da natureza geral do sentimento quase em todo lugar, contudo, o sentimento, em algum grau verificável, é uma mera propriedade do protoplasma, talvez apenas do material nervoso. Ora, acontece que os organismos biológicos, e especialmente o sistema nervoso são condicionados favoravelmente para exibir o fenômeno da mente também; e assim não é surpreendente que a mente e o sentimento sejam confundidos. Mas não acredito que a psicologia possa se endireitar enquanto a importância do argumento de Hartman não for levada em conta e seja visto que o sentimento não é senão o aspecto interior das coisas, enquanto a mente, ao contrário, é essencialmente um fenômeno externo (CP 7.364).

Está novamente enfatizado na passagem acima o conceito alargado da mente, de modo a ser aplicável a uma grande variedade de fenômenos, nesse caso, o sistema nervoso. Junto com a mente e conectada a ela, o conceito de sentimento também é grandemente dilatado. Do modo como está, a passagem ainda está longe de ter deixado a questão clara. Mas nos dois parágrafos que dão prosseguimento a essa passagem, explanações iluminadoras começam a emergir, pois consciência, sentimento e mente ficam mais claramente diferenciados.

O que os psicólogos estudam é a mente, não a consciência exclusivamente. O equívoco sobre esse ponto produziu resultados singularmente desastrosos, visto que a consciência é algo muito simples. Apenas tome cuidado para não fazer a confusão de supor que é a auto-consciência que está sendo referida. Então se verá que a consciência não é outra coisa senão sentimento, em geral, — não sentimento no sentido alemão, mas mais geralmente, o elemento imediato da experiência generalizado ao extremo. A mente, ao contrário, é uma coisa muito difícil de analisar, quando você capta a verdade de que ela não é consciência, nem proporcional de qualquer modo a esta. Não estou falando de alma, o substrato metafísico de mente (se é que haja algum), mas de mente **entendida** fenomenicamente. Para se chegar a uma concepção de mente ou fenômeno mental, tal como a ciência da dinâmica nos fornece de matéria ou de

eventos materiais, é um assunto que só pode ser resolvido através de investigação científica resoluto. Mas os psicólogos foram impedidos de fazer essa investigação devido à ilusão de que mente é apenas consciência, uma questão simples, na medida em que se trata do mero fenômeno, sobre o qual não há lugar para dúvida ou erro (CP 7.365).

Os psicólogos dizem que a consciência é essencialmente um atributo da mente; e que o propósito é apenas uma modificação especial. Sustento que propósito, ou melhor, causação final, da qual o propósito é uma modificação consciente, é o assunto essencial dos estudos dos psicólogos; e que a consciência é um acompanhamento especial e não universal da mente (CP 7.366).

E aqui chegamos ao coração do conceito peirciano da mente: o conceito coextensivo de causação final.

### **Causação final e causação eficiente**

A relevância do conceito peirceano de causação final foi investigada por Potter (1967), Ransdell (1977, 1981, 1983) e Johansen (1993a e b). Pape (1993, p. 593-594) mostrou que a causalidade final é a chave geral para a compreensão da unidade sistemática da filosofia peirciana e também a chave especial para o entendimento da semiose. "A definição geral peirceana de signo ou representamen é a descrição mais geral da estrutura interna da causação final[...]. Não pode haver um processo signico sem haver uma causa final ativamente envolvida nele". Na mesma linha de pensamento, desenvolvi vários estudos com ênfase na relevância da causação final não apenas para o entendimento do sistema semiótico de Peirce, mas também como um conceito unificador da sua fenomenologia, semiótica e metafísica (SANTAELLA 1992, 1994, 1996). No momento, estou convencida de que a causação final, junto com a sua contraparte, a causação eficiente, e seus opostos complementares, acaso e sentimento, podem levar à compreensão de uma das noções peirceanas mais enigmáticas, a saber, a indissolubilidade entre mente e matéria.

Para começar, devo chamar atenção para o fato de que o conceito peirceano de causação final é extremamente original. De acordo com Peirce, há dois tipos de ação no universo, a ação diádica, que "é bruta, não inteligente e desligada do resultado que pode advir dela" (CP 6.332), e a ação triádica, que é ação governada por lei.



Peirce estava certamente familiarizado com os quatro tipos de causas aristotélicas que entram em ação para determinar um fato: a causa material, a formal, a eficiente e a final (CP 6.66). Desconsiderando a causalidade material e a formal, ou talvez equacionando a material com a eficiente e a formal com a final, Peirce reteve apenas o que ele chamou de dois grandes ramos da causação: a eficiente ou causação de força como ação diádica; e a ideal ou final como ação triádica. A causação final, de acordo com a sua descrição, é:

[...] aquele modo de fazer os fatos acontecerem de acordo com o qual uma descrição geral ou resultado acontece independentemente de qualquer compulsão para que ele aconteça deste ou daquele modo particular; embora os meios possam ser adaptados aos fins. O resultado pode ocorrer num momento de um determinado modo e em outro momento de outro modo. A causação final não determina de que modo particular ele irá acontecer, mas apenas que o resultado tenha um certo caráter geral (CP 1.211).

A causação eficiente, por outro lado, "é uma compulsão determinada pela natureza particular das coisas", de modo a fazer com que uma situação comece a mudar de um modo perfeitamente determinado; e qual possa ser o caráter geral do resultado não diz respeito à causação eficiente (CP 1.212).

Embora Peirce estivesse aparentemente muito próximo da "verdade da afirmação de Aristóteles" (CP 1.211), um estudo mais acurado revela que não apenas os conceitos aristotélicos de causação final e eficiente, mas também a inter-relação de ambos, recebeu de Peirce uma nova interpretação.

### **A nova interpretação peirceana da causalidade**

Como Aristóteles, Peirce não limitou a causação final a processos conscientes, nem a tomou como sendo um tipo atual, mas geral. Como Aristóteles, Peirce considerou os processos de causação final como sendo observáveis e não os considerou capazes de agir sem a cooperação da causação eficiente. Diferente de Aristóteles, entretanto, Peirce não atribuiu a influência das causas finais à perfeição, nem ao bem, nem à fonte primeva e pura da atividade (SHORT, 1981, p. 369-371). Além disso, de acordo com Pape (1993, p. 603, n. 3)

A diferença entre Peirce e Aristóteles depende em última instância do *insight* peirceano, que ele foi o primeiro na história a formular, de que as

possibilidades de desenvolvimentos irreversíveis das distribuições do acaso é uma condição necessária para todos os tipos de causas finais, incluindo os propósitos, mesmo que eles sejam escolhidos devido ao seu bem.

Peirce foi levado a reinterpretar a doutrina aristotélica da causação devido a inconsistências históricas e à visão estreita de causalidade que foi adotada por seus contemporâneos, a saber, "que a causa é um evento de um tal tipo de modo a ser necessariamente seguido por um outro evento que é o efeito" (CP 6.66). Ele se referia aqui ao conceito de causação – se A ocorre, então B ocorre – que se tornou largamente aceito a partir de Hume.

Peirce discordou dos dois conceitos mais influentes de causação, o conceito determinista, de um lado, e o conceito teleológico de Kant, como seqüência invariante e incondicional, de outro lado (ver 6.590, 6.600). O fundamento geral da doutrina peirceana de causação estava nas suas categorias fenomenológicas e metafísicas. De fato, é sobre esse pano de fundo que sua teoria da causalidade deve ser considerada.

### **Causalidade e as categorias**

Pode parecer surpreendente encontrar uma dicotomia, tal como a de causação final e eficiente, em um sistema inteiramente triádico como o de Peirce. As ações diádica e triádica estão certamente conectadas às categorias de secundidade e terceiridade. Não há ação na primeiridade, apenas talidade e caráter. Assim sendo, as duas espécies de ações não contradizem a triadicidade (ver CP 6.331).

Começemos com o entendimento que Peirce teve de causação eficiente, visto que esta noção é mais crua e simples do que a causação final. Como uma instância de mera secundidade, uma diáda é um fato descontínuo que ocorre "aqui e agora". Não há nenhuma generalidade nela. Como uma causação eficiente, ela é força bruta ou compulsão, uma ação efetiva, cega, não racional, singular na sua ocasião, apenas uma compulsão factual no aqui e agora de um evento. Como tal, ela não pode ser confundida com causalidade determinista. A causação eficiente é mais primitiva do que a necessidade positivista do tipo "se A, então B". Ela também é diferente da causação mecânica. No entanto, essas

diferenças sutis só podem ser compreendidas depois da investigação da noção mais complexa de causação final.

Grosseiramente falando, causação final é ação exercida por leis. Ora, uma lei é algo geral e, por essa razão, ela não é uma força. "Pois força é compulsão; e compulsão é *hic et nunc*. Ou é isso, ou a compulsão não é nada. A lei sem força para exercê-la, seria uma corte sem um xerife; e todos os seus dicta não passariam de vapores". Assim, a relação da lei como uma causa, para a ação da força como seu efeito, é causação ideal ou final, e não causação eficiente (CP 1.213). Em muitas passagens, Peirce enfatizou a indissolubilidade das causações final e eficiente. Causação final sem causação eficiente é desesperançada, mas causação eficiente sem causação final é pior do que desesperançada; "pois ela é mero caos, e caos, sem causação final, não é nem mesmo algo como caos, mas é um nada completo (CP 1.220). Não é difícil compreender a interdependência das causações final e eficiente, mas na moldura do sistema triádico de Peirce, a inclusão da secundidade, dentro da terceiridade é apenas um lado da questão. O outro lado, quer dizer, a inclusão da primeiridade dentro da secundidade e dentro da terceiridade, é mais complexo.

Algo dessa complexidade é evidente, por exemplo, quando Peirce afirmou que, sem causa final, na forma da lei, a causação eficiente não teria sequer eficiência. "Ela pode se exercer e algo deve seguir *post hoc*, mas não *propter hoc*, pois *propter* implica em regularidade potencial. Ora, sem lei, não há regularidade; e sem a influência de idéias não há potencialidade" (CP 1.213). A referência a idéias e potencialidade em conexão com a lei introduz um tópico inesperado que exige explicações.

Peirce não negou que há causas reais, mas compreendeu-as como causas finais. Assim, o tipo de necessidade envolvida nas leis da natureza (que diferem das generalizações empíricas que fazemos dessas leis) devem ser analisadas à luz do conceito de tendencialidade, no reino da terceira categoria, isto é, da generalidade, continuidade, tempo, mudança e evolução, ou melhor, semiose. Por outro lado, casos de necessidade nas leis da natureza, que podem ser expressas em proposições do tipo "Se A então B", devem ser vistos como casos no limite de uma tendência, casos nos quais a tendência já está completamente rígida, muito pouco aberta à interferência do acaso e muito pouco submetida à mudança. A explicação de Pape (1993: 589-590) a essa questão é especialmente clara: ... as leis da mecânica clássica não são leis da natureza de modo algum. De



fato, Peirce repetidamente enfatizou que elas são meramente fórmulas. Correspondentemente, forças mecânicas não são causas no sentido estrito, pois 'pode-se razoavelmente objetar ao dito de que a força mecânica é a causa do movimento, em vez de chamá-la de inclinação temporal do movimento' (MS 1343: 29). No caso da ação, tal como foi descrita pela mecânica clássica, o estado final de um sistema deve-se completamente a uma transformação isomórfica daquilo que foi dado em termos das posições iniciais das partículas na configuração final das posições. Para cada configuração inicial das partículas e distribuição de forças, há exatamente um modo de produção do estado final pelas acelerações correlatas de partículas. No caso da ação mecânica, não faz sentido se falar de um estado final definido que teria de ser alcançado em um número de diferentes modos.

Diferentemente da rigidez acima descrita, o conceito peirceano de lei como "poder vivo" (POTTER, 1967) pode ser traduzido como a tendencialidade do universo a adquirir novos hábitos, o que só se faz entender à luz do conceito de causação final, a ação do signo ou ação inteligente. Assim sendo, a lei determinista que se expressa no conceito positivista de lei é apenas um caso no limite de uma tendencialidade que perdeu seu frescor. A causação bruta ou eficiente, por outro lado, não deve ser confundida com essa rigidez, pois ela diz respeito apenas ao instante singular no qual a força bruta age cegamente. É dessa ação *hic et nunc* que a ação inteligente depende para poder agir.

Por isso mesmo, em uma outra passagem, Peirce acrescentou: "Por gênese deve ser entendida não a ação eficiente que produz o todo ao produzir as partes, mas a ação final que produz as partes porque elas são necessárias para realizar o todo (o todo chama por suas partes) (CP 1.220). Gênese é produção a partir de idéias. Pode ser difícil compreender como isto é verdadeiro no mundo biológico, embora haja prova suficiente de que é assim" (1.227). Se Peirce tivesse então o conhecimento que se tem hoje da célula como uma maquinaria codificadora, essa dificuldade teria deixado de existir.

Tanto quanto posso ver, as questões levantadas por essas teses podem ser respondidas no contexto mais amplo da filosofia peirceana, no qual o conceito de causação final abraça uma lógica única e complexa — aquela da triadicidade — as idéias coextensivas de tempo, pensamento, inteligência, vida, crescimento e evolução. Desse modo, a chave para a causalidade final está nos conceitos relacionados com a terceiridade, isto é, continuidade, generalidade, lei,

mente, lei da mente e hábito. Entretanto, como as categorias peirceanas são onipresentes e inter-relacionadas (CP 5.436, 7.532), considerar a causalidade final como terceiridade isolada da secundidade ou causação eficiente seria um equívoco sério tanto quanto seria isolá-la da primeiridade, a categoria do acaso e do sentimento. “O acaso não é senão o aspecto exterior daquilo que dentro de si mesmo é sentimento” (CP 6.265).

### **Causas finais aquém do limiar do vivo**

A forma prototípica de causação final é aquela da mente. “A mente tem seu modo universal de ação, a saber, por causação final... Ser governada por causas finais é a própria essência do fenômeno psíquico em geral (CP 1.269, ver também 2.66 e 7.559). Até aqui, não parece haver nada de novo nessa idéia. Que a mente almeja atingir certos fins, parece óbvio. Entretanto, as afirmações de que os princípios da ação inteligente podem ser deduzidos dos caracteres primitivos do protoplasma (CP 6.283) e de que mesmo uma rã decapitada raciocina (CP 6.286, 6.144, 2.711) são menos óbvias e mostram claramente que a concepção peirceana de inteligência e mente estão longe de serem convencionais.

De fato, para Peirce, as funções da mente não se restringem aos seres humanos (CP 1.269), nem se limitam ao reino biológico, pois não apenas qualquer organismo biológico tem mente, mas também as máquinas, tais como os computadores podem ter mente. Como seu conceito de mente está subordinado aos princípios gerais da causação final, haverá mente onde houver triadicidade; onde houver tendência para aprender, para o crescimento, ou evolução, aí haverá mente, não importando quão rudimentar essa ação possa ser.

No manuscrito ainda não publicado, MS 1343, Peirce discutiu todas as gradações que vão do nível mais baixo da causalidade bruta até as formas mais complexas da causalidade final. Pape (1993) apresentou uma análise detalhada de cada um dos seguintes graus:

- (a) ação mecânica;
- (b) comparação entre ação despropositada e ação quase-propositada;
- (c) ação governada pela mente, na qual o exercício da mente é tão estúpido quanto possível;



- (d) comparação entre sistemas quase inteligentes até a inteligência animal;
- (e) da inteligência animal aos sistemas inteiramente inteligentes.

Assim, a causação final abrange desde as suas formas mais desenvolvidas, ou seja, o ato humano intencional de raciocinar sob a orientação do auto-controle e da auto-crítica, até suas formas mais rudimentares, quer dizer, onde quer que haja ação visando a fins definidos no mundo puramente físico.

A causação final envolve intencionalidade, que é a versão psicológica autoconsciente da causação final, mas não se limita a ela. A causalidade final também envolve propósito, mas ainda não se limita a ele. “É um erro largamente suposto pensar que causa final é necessariamente propósito. Ora, um propósito é meramente aquela forma de causação final que é mais comum à nossa experiência” (CP 1.211), quer dizer, “um propósito é a modificação consciente da causação final” (CP 7.366).

“Um propósito é um desejo operativo. Ora, o desejo é sempre geral” (CP 1.205, ver também MS 1343), o que significa que, enquanto o desejo é predominantemente final e um terceiro, o propósito, por ser operativo, e por servir como um meio para um fim, deve ser o protótipo de uma conexão perfeita entre a causalidade eficiente e final. Além disso, os propósitos podem também ser inconscientes, e, no mundo biológico, Peirce sustentou que o propósito pode ser meramente “uma ação virtualmente direcionada para a remoção do estímulo” (CP 5.563, ver também 1.392 e 6.281). Não obstante, qualquer que possa ser o tipo de propósito, “uma causa final deve ser concebida como operando sem ter sido o propósito de qualquer mente” (CP 1.204). Vem daí que Peirce tenha proposto que, além do estudo de “como a causação final opera no desenvolvimento de espécies biológicas” e de como ela pode ser aplicada a indivíduos biológicos, deveria haver um estudo prévio para definir a essência da mente e a lei da causação final, junto com sua aplicação a fenômenos não-biológicos.

Em síntese, a causação final é inerente a qualquer atividade direcionada para um fim. Trata-se da forma geral de um processo, a tendência para um estado final, “o traço geral de tal tendência em qualquer meio que possa ocorrer” (RANSDELL, 1977, p.163). Que os processos vivos exemplificam algumas dessas formas foi plenamente reconhecido sob rótulos tais como “cibernética”,

“homeostase”, e particularmente “teleonomia”. De fato, quando Peirce descreveu o nascimento como “as relações de algo com suas causas finais originárias” (CP 1.244), essa definição está muito próxima da noção moderna de teleonomia, a saber, aquela que está implícita na noção cibernética de propósito.

### **Causas finais e as ciências contemporâneas**

Para sintetizar, deve ser retido que, quando se fala de ação inteligente, no contexto do pensamento de Peirce, não se deve entender esse adjetivo dentro de limites antropocêntricos. Semiose ou ação do signo é um termo técnico geral que recobre o campo semântico de termos tais como inteligência, mente, pensamento — que não são privilégios da espécie humana. Onde houver tendência para aprender, para processos de autocorreção, mudanças de hábito, onde houver ações direcionadas por um propósito, haverá inteligência, onde quer que ela ocorra: no grão do pólen que fertiliza o óvulo de uma planta (W1, p. 333), no vôo de um pássaro, no sistema imunológico, na perversidade do inconsciente, ou na razão humana. É por isso que a causalidade final deve ser compreendida lado a lado com conceitos cibernéticos, tal como feedback, e biológicos, tais como morfogênese, teleonomia, autopoiesis, ou mesmo conceitos naturais, tais como estruturas dissipativas, sistemas auto-organizativos (SHORT, 1983; RANSELL, 1983).

Peirce limitou-se, portanto, a descrever um processo lógico, o processo da ação do signo, que foi técnica e formalmente descrito nas suas inúmeras definições dos signos. Não há nada exclusivamente antropológico nisso, visto que essa lógica é capaz de descrever processos biológicos e mesmo físicos de qualquer espécie contanto que sejam irreversíveis e apresentem uma tendência assintótica para a finalização de um estado de coisas (EMMECHE, 1991; HOFFMEYER, EMMECHE, 1991). Se não é o bem potencial que move esses processos, como queria Aristóteles, então o que os guia? A resposta de Peirce começou com a análise de ações não-conservativas.

Essas ações não conservativas, que parecem violar a lei da energia e que os físicos explicam como sendo devidas à ação do acaso entre trilhões de moléculas, são todas elas marcadas por dois traços essenciais. O primeiro é que elas agem em uma direção determinada e tendem assintoticamente para a realização de um estado de coisas final. Se teleológica é uma palavra muito forte para ser aplicada a elas,

podemos inventar a palavra *finious*, para expressar sua tendência a um estado final. O outro traço das ações não conservativas está na sua irreversibilidade (CP 7.471).

Peirce prossegue com o reconhecimento de que a causa eficiente (força) é incapaz de explicar a irreversibilidade. “Tendências incertas, estados instáveis de equilíbrio são condições *sine qua non* para a manifestação da mente”, ele disse (CP 7.381), chegando à conclusão de que só um tipo geral de natureza auto-reprodutiva seria capaz de governar a atualização de particulares. Isto é o que ele chamava de lei, mente ou ação *signica*, ação inteligente. O modo pelo qual os particulares são atualizados sob o poder dessa ação pode variar enormemente, mas o processo guiado para um fim, que os governa, tem uma única base lógica. Essa base é aquela da ação triádica que, mesmo nas suas formas mais rudimentares, exibe pelo menos uma gota de inteligência. Trata-se invariavelmente da tendência de um processo para chegar a um resultado de um certo tipo geral.

Muito provavelmente, toda ação *signica* tem algo de antropomórfico, na medida em que envolve sempre a causação através de abstrações ou formações gerais, cuja forma típica é encontrada no autocontrole que a mente humana pode exercer sobre a conduta. Mas o que surge aqui é simplesmente uma das formas típicas da causação final, talvez a mais complexa, mas não sua forma exclusiva. Assim, a liberalização por que Peirce fez passar termos como “mente”, “inteligência” e “pensamento” teve por finalidade tornar evidente a continuidade que existe entre a mente humana e outros processos movidos por um propósito; estes exibem alguma forma de mentalidade, que inclui, entre outras, o comportamento de micro-organismos, a evolução biológica, e mesmo o crescimento dos cristais. Esta postulação está perfeitamente de acordo com teoria das estruturas dissipativas (PRIGOGINE, STENGERS, 1984), nas quais a causação final reside na tendência para a ordem que Prigogine encontrou até mesmo em tipos rudimentares de reações químicas.

A diferença de grau entre os processos de causação depende do autocontrole, maior ou menor, mas sempre relativo, que pode ser exercido sobre o estado final. É por isso que esses processos podem ser guiados pela autocorreção, a escolha e o acaso ao mesmo tempo. O reconhecimento do acaso objetivo como tendo um papel importante a desempenhar no processo, e a conseqüente ocorrência de erros, reduz o determinismo da causação final, mas

apenas na natureza do resultado, pois não interfere nos modos pelos quais o resultado pode ser atingido. Acaso, pura possibilidade, e ação dinâmica ou bruta, mediadas por um princípio geral, um princípio guia, unindo o primeiro infinito, acaso, com o segundo infinito, ação cega, constitui a triade dentro da qual a causação final se move.

Como se pode ver, o conceito peirceano de mente é muito amplo e liberal. Mas é justamente essa liberalização que coloca esse conceito em sintonia com algumas das mais recentes preocupações na física, biologia e inteligência artificial. O modelo lógico básico da semiose, que se expressa na definição de signo, não é apenas um modelo para a descrição da mente, pensamento, inteligência, continuidade e crescimento, ele é também, conseqüentemente, um modelo para o entendimento da evolução, desde que Peirce acreditou que processos evolutivos em geral são manifestações da mente, entendida no sentido alargado que ele deu a essa palavra. O que, na época de Peirce, soava como um aparente absurdo é exatamente aquilo que está soando como mais atual no debate contemporâneo de idéias renovadas.

## **Referências**

EMMECHE, C. A semiotical reflection on biology, living systems and artificial life. *Biology and Philosophy*, v.6, p.325-340, 1991.

\_\_\_\_\_. Defining life as a semiotic phenomenon — on replicators and autopoiesis as emergent signs of life and the notion of onto definitions. Preliminary manuscript presented at the VIth International Congress of the IASS, Semiotics bridging nature and culture. Guadalajara, Mexico, July 1997.

\_\_\_\_\_.;HOFFMEYER, J. From language to nature: the semiotic metaphor in biology. *Semiotica*, v.84, n.1/2, p.1-42, 1991.

HOFFMEYER, J.; EMMECHE, C. Code duality and the semiotics of nature. In: MYRDENE, A.; FLOYD, M. (Ed.). *On semiotic modeling*, Berlin: Mouton de Gruyter, 1991, p.117-166.

JOHANSEN, D. J. *Dialogic semiosis: an essay on signs and meaning*. Bloomington: Indiana University Press, 1993a.

MATURANA H. R.; VARELA F. J. *The tree of knowledge: the biological roots of human understanding*. Boston: New Science Library: Shambhala, 1987.

PAPE, H. Final causality in Peirce's semiotics and the classification of the sciences. *Transactions of the Charles S. Peirce's Society*, v.29, n.4, p.581-607, 1993.

PEIRCE, C. S. *Collected papers*. Edited by Charles Hartshorne and Paul Weiss v.1-6; Arthur Burks (v.7-8). Cambridge: MA, Harvard University Press, 1931-1958. MS refers to the unpublished manuscripts as paginated by the ISP, Lubbock, Texas.

POTTER, V. *Charles Sanders Peirce: On norms and ideals*. Boston: University of Massachusetts Press, 1967.

PRIGOGINE, I.; STENGERS, I. *Order out of chaos*. London: Heinemann, 1984.

RANDELL, J. Some leading ideas of Peirce's semiotics. *Semiotica*, v.19, n.3/4, p.157-178, 1977.

\_\_\_\_\_. Semiotic causation: a partial explication. In: KETNER et al. (Eds.), 1981. p.102-106.

\_\_\_\_\_. Peircean semiotics. Draft of a work in progress, 1983.

SANTAELLA, B.L. *A assinatura das coisas: Peirce e a literatura*. Rio de Janeiro: Imago, 1992.

\_\_\_\_\_. Peirce's broad concept of mind. *S. European Journal for Semiotic Studies*, v.6, n.3/4, p.399-411, 1994.

\_\_\_\_\_. Semiosphere: the growth of signs. *Semiotica*, v.109, n.1/2, p.173-186, 1996.

SHORT, T.. Peirce's concept of final causation. *Transactions of the Charles S. Peirce's Society*, p.368-382. , Fall 1981

\_\_\_\_\_. Life among the legisigns. *Transactions of the Charles S. Peirce Society*, v. 18, n.4, p.235-310, 1982.

\_\_\_\_\_. Teleology in nature. *American Philosophical Quarterly*, v.20, n.4, p.311-319, 1983.

\_\_\_\_\_. The growth of symbols. *Cruzeiro Semiótico*, v.8, p.81-87, 1983.





## Notas sobre a lei da mente peirceana

Max Rogério Vicentini<sup>1</sup>

*"That phrase 'a gentle force which commonly prevails'  
describes the phenomenon to perfection"*

(CP 7, 389)

Procuramos neste artigo fazer uma exposição e análise da idéia peirceana de 'lei da mente', com a intenção de revelar elementos que permitam, ao mesmo tempo, uma melhor compreensão do pensamento de Peirce e uma possível contribuição para o debate contemporâneo a respeito da noção de auto-organização. A discussão sobre a noção de auto-organização é hoje uma constante no meio intelectual que busca resolver os impasses criados por uma ciência da mente.

O objetivo deste artigo é, a partir da exposição de elementos da filosofia de Charles S. Peirce, em particular da lei da mente, pensar elementos que possam enriquecer a noção de auto-organização. Se alcançarmos sucesso na tentativa de mostrar que há muitas similaridades e convergências entre o princípio de lei da mente, tal qual pensado por Peirce, e as principais características hoje atribuídas à classe de fenômenos descrita pela noção de auto-organização, poderemos indicar como a abordagem peirceana aponta e sugere aspectos que ainda permanecem estranhos ao contexto de discussão atual, mas que, na elaboração realizada pelo autor, são de profunda relevância para o delineamento de uma visão adequada do crescimento da ordem no universo, isto é, da origem e aumento de complexidade dos sistemas.

### Notas introdutórias

Ainda que de maneira breve, fazem-se necessárias algumas considerações acerca do contexto teórico que permitiu a elaboração desta hipótese de investigação, isto é, de que a lei da mente apresenta similaridades

---

<sup>1</sup> Professor de Filosofia do Departamento de Ciências Sociais da UEM. Endereço eletrônico: [mrvicentini@uem.br](mailto:mrvicentini@uem.br).

com o processo de auto-organização ou, ao menos, que possua elementos que permitam fazer aproximações esclarecedoras para ambas as teorias.

A idéia de auto-organização surge no contexto de busca por explicações que considerem a natureza complexa dos fenômenos. Segundo este novo paradigma, haveria eventos resistentes aos modelos de explicação tradicionais. Estes não seriam definidos tão somente por suas partes, mas ainda pela totalidade constituída a partir das “interações” entre elas.

Michel Debrun (1996) fornece uma circunscrição filosófica do conceito de auto-organização. Esse autor parte de uma definição intuitiva desse conceito, segundo a qual formas poderiam surgir ou se reestruturar principalmente a partir delas próprias. Debrun reconhece a dificuldade de se definir rigorosamente um termo que, em princípio, parece autocontraditório. Insiste, contudo, que haveria um sem número de “formas” que seriam mais ou menos impermeáveis aos tratamentos tradicionais, “e cujos processos de formação parecem ser melhor explicados pela abordagem da auto-organização” (DEBRUN, 1996, p. xxxiv). Diante da vagueza da idéia inicial, a estratégia seguida é a de buscar o enriquecimento desse conceito a partir de sua aplicação nos diversos campos do conhecimento humano. Com esse objetivo, Debrun propõe alguns critérios que, supõe, podem orientar a busca por definições mais precisas do conceito de auto-organização.

Para Debrun, o fenômeno da auto-organização está presente em vários eventos que presenciamos cotidianamente. Sua marca distintiva consistiria no fato de serem processos que surgem ou se reestruturam a partir do próprio processo, dependendo apenas minimamente das condições de partida ou de contorno. Exemplos desse processo são encontrados na Física, Biologia, Política, mas também em situações comuns como um jogo de futebol. O autor traça uma distinção entre auto-organização primária e secundária. Na primeira, teríamos a constituição de uma forma a partir da interação casual de elementos distintos. Já na auto-organização secundária, haveria um aumento na complexidade de formas já existentes devido ao trabalho do sistema sobre si próprio. Nesta modalidade, é possível o domínio de uma parte do sistema sobre a outra, contudo, não absoluto, pois, nesse caso, teríamos a instanciação de um processo de hetero-organização. No caso do processo de auto-organização humana, há surgimento de uma face sujeito, responsável por um certo direcionamento do processo, ainda que não plenamente determinado.

Debrun parece insistir que o processo de auto-organização não pode prescindir do fator temporal, uma vez que é somente com o desenvolvimento no tempo de um determinado sistema que poderemos, *a posteriori*, identificar os elementos que compõem a auto-organização. Uma definição precisa desse conceito, entretanto, parece contrapor-se à sua própria idéia, que tem nos elementos mais ou menos soltos do início do processo uma de suas condições fundamentais, bem como em uma certa orientação da dinâmica de desenvolvimento do sistema que se atem às formas que lhe são próprias.

A partir dessa rápida caracterização dos processos auto-organizados, buscaremos na filosofia de Charles S. Peirce elementos que possam ajudar-nos a pensar esse processo de desenvolvimento da ordem. A filosofia evolutiva de Peirce parece mobilizar um rol de conceitos e idéias que se aproximam dessa visão de auto-organização. A ênfase no acaso, bem como a classificação de diferentes tipos de evolução, parecem propor uma visão do mundo e do seu desenvolvimento compatível, em muitos aspectos, com a proposta pelos estudos de auto-organização.

A procura de soluções para problemas clássicos da filosofia conduz Peirce a defender uma posição bastante original quanto ao estatuto da realidade e seu modo de ser. A compreensão de suas idéias sobre o evolucionismo representa uma tarefa central para todo aquele que pretende avaliar a novidade que seu vigoroso pensamento apresenta.

Em primeiro lugar, deve ser lembrado que Peirce pretende que a sua filosofia seja de cunho evolucionário; que procura estabelecer princípios que possam explicar a totalidade do que existe no universo como resultante de um processo de crescimento contínuo. E ainda, que esse processo de crescimento, ou evolução, deva ser, ele próprio, explicado como estando submetido aos mesmos princípios a que todos os demais fenômenos estão. Assim, faz-se necessário encontrar uma lei que tenha na sua aplicação o motor de seu fortalecimento, ou seja, uma lei que possa crescer devido à sua própria natureza: esta é a lei da mente. Como explicita Peirce: “a única tendência que pode crescer por sua própria virtude, a tendência de todas as coisas a adquirirem hábitos” (CP 6.100). Em um universo de pura possibilidade, todas as atualizações necessariamente extinguir-se-iam na medida em que nada garantiria a sua subsistência. Apenas a realização de algo que tivesse, em sua própria realização, o motor de seu fortalecimento poderia perdurar, essa é a lei a mente, afirma Peirce.

Diante das evidências do que percebemos do cosmos, contudo, uma nova exigência faz-se necessária. A possibilidade de se pensar um universo evolutivo requer a existência de uma matriz ontológica indeterminista que tenha no acaso seu elemento fundante. Como afirma Peirce, “o tiquismo<sup>2</sup> deve dar origem a uma cosmologia evolucionária” (CP 6.102). Com essa tese, Peirce afirma que há no mundo um caráter de real espontaneidade, não devido à nossa subjetividade ou incapacidade cognitiva, mas à sua própria constituição ontológica. Esse elemento nega a possibilidade do desenvolvimento necessário dos sistemas determinados pelas leis mecânicas e condições iniciais que só podem ser entendidas como estatísticas. Explicando, dessa maneira, que pela atuação do acaso surgem a novidade e a variedade do universo.

Peirce argumenta longamente em defesa da realidade da existência desse princípio de absoluta indeterminação no seio da natureza. Contra os necessitaristas, o autor afirma que a hipótese da existência do acaso absoluto, além de ser muito mais natural face à observação da diversidade das coisas que aparecem – como esclarece Peirce, basta abrir os olhos para notar a incrível variabilidade existente em todos os fenômenos da natureza – também oferece uma vantagem explicativa na medida em que não deixa sem resposta a pergunta pela origem das próprias leis, que, de outra maneira, adquiriram o *status* de incognoscíveis, transgredindo, dessa maneira, a primeira lei da lógica, qual seja, a de não bloquear a via da investigação. Desse modo, o acaso, na evolução, é inseparável e contribui para o crescimento da lei.

As razões que conduzem o autor a defender a idéia de que a diversidade do mundo tem crescido continuamente são dadas na sequência:

1. A natureza explicada segundo leis mecânicas prescinde do curso do tempo. Aqui, Peirce sustenta sua idéia de crescimento contínuo do universo, apelando para as ciências que têm no tempo uma variável imprescindível para as suas elaborações e descobertas. Basta inquirir à Paleontologia ou à Geologia ou, até mesmo, à Astronomia para se constatar que há um crescimento e aumento da complexidade.

2. A diversidade e novidade existentes no mundo não podem ser adequadamente explicadas pela atuação de leis mecânicas absolutas. Pois é a essência de uma lei mecânica que causas idênticas sempre produzam efeitos

---

<sup>2</sup> O termo Tiquismo vem do radical grego *tyche*, que significa acaso.



idênticos. Desse modo, uma hipótese razoável é que sejam os resultados das pequenas variações ocorridas no momento da atuação da lei os responsáveis pelo crescimento da complexidade existente.

3. A hipótese da existência do acaso absoluto, manifesto nas pequenas variações a partir da aplicação das leis, permite explicar como a diversidade surge. Permite, ainda, a justificação da atuação de um outro tipo de causalidade, a da lei da mente, que explica a formação da uniformidade constatada no universo. Afirmar, por outro lado, que a diversidade e irregularidade bem como uniformidade e regularidade existentes no universo não podem ser explicadas e, nem mesmo, deveriam ser questionadas, como fazem os deterministas, significa ferir a primeira lei da lógica, que consiste em não bloquear a investigação.

4. A partir da introdução da causação mental ou causação teleológica no esquema explicativo, Peirce afirma ser possível explicar diversos fenômenos que, de outra maneira, adquiririam o caráter de fantasmagórico, como por exemplo, a consciência e correlatos.

5. A mais importante razão, que permanecerá não desenvolvida no texto em análise, é a de que “a hipótese do acaso-espontaneidade é daquelas cujas conseqüências inevitáveis são capazes de ser traçadas com precisão matemática em detalhe considerável” (CP 6.62).

Temos, desse modo, uma defesa da presença do acaso no mundo que não apenas recebe o suporte das observações e evidências que estão disponíveis para quem as desejar constatar, como parece se impor como uma evidência para todo aquele que esteja familiarizado com a prática científica. E que, ainda, fornece uma hipótese explicativa para os fenômenos naturais, tanto da diversidade como da uniformidade do mundo, que impulsiona o desenvolvimento de futuras pesquisas sem postular o incompreensível no seio da natureza.

Sumariando, podemos dizer que contra a argumentação peirceana os deterministas só poderiam responder de uma das três maneiras elencadas abaixo:

- (1) O acaso é algo absolutamente inteligível.
- (2) Não há qualquer fenômeno no qual a atuação do acaso possa ser observada de forma inequívoca.
- (3) O acaso não pode ser uma *vera causa*.

Quanto à objeção (1), Peirce indica que sua hipótese do acaso absoluto, diferentemente da tese determinista, fornece uma explicação para o surgimento tanto das irregularidades quanto das uniformidades existentes no mundo. O que parece ser um ganho considerável uma vez que desobstrui os caminhos da investigação científica. Um outro aspecto que fornece um suporte extra a esta hipótese é o fato de que ela dá lugar à elaboração da lei da mente. O acaso uma vez manifestado desenvolve-se segundo determinadas maneiras e não de uma maneira qualquer. Essa maneira de desenvolvimento segue a lei da mente, também chamada lei da aquisição de hábito ou, simplesmente, lei do hábito.

A objeção do tipo (2) é, na visão de Peirce, bastante ingênua, pois parece ignorar todos os avanços científicos que se verificaram nas áreas de Astronomia, Biologia ou Geologia, permanecendo cega para os fenômenos de crescimento da diversidade ou evolução. Esses fenômenos são totalmente refratários à abordagem mecanicista no entender desse autor. Composto com esses fenômenos aparece a consciência como outro empecilho às pretensões dos mecanicistas, que se fiam somente na idéia de regularidade para compor o seu quadro explicativo. Com essa atitude a noção de regularidade serve apenas "*to block the road of inquiry*" (CP 6.64).

À acusação (3), de que o acaso não pode ser considerado uma *vera causa*, Peirce responde que a doutrina em questão não pode ser aplicada a concepções elementares. Se levarmos a sério essa acusação, seria preciso reavaliar a própria crença na existência do universo material, uma vez que a variedade é um fato que não pode ser ignorado.

Tendo assim defendido a existência do acaso absoluto, Peirce acredita que não resta dúvida que as pesquisas devem ser refeitas e direcionadas para novas bases que levem em consideração o caráter estatístico das regularidades, bem como o indeterminismo reinante no universo que, em última instância, é a fonte de diversidade e novidade existente no mundo.

### **O continuum**

Se de um lado temos no acaso um dos constituintes fundamentais da realidade, do outro, encontramos no *continuum*, ou nas leis, outro elemento de igual importância.

A elaboração que Peirce oferece deste tópico está em estreita conexão com a matemática dos infinitesimais, por ele estudada. Nosso interesse aqui não é o de aprofundar tais concepções, procuramos apenas extrair algumas características de sua descrição que suportem a nossa análise da lei da mente. A questão levantada por Peirce e que servirá de guia para a nossa discussão será a de "como pode uma idéia passada estar presente?" (CP 6.109). A resposta, por ele proposta, é de que devemos estar diretamente conscientes das idéias do passado imediato. Caso houvesse uma lacuna entre nosso conhecimento presente e as idéias do passado, todo o nosso conhecimento do passado não seria mais do que uma ilusão. Assim, como consequência, devemos estar conscientes do passado "através de um intervalo de tempo infinitesimal" (CP 6.110). Nessa mesma linha aponta Hausman que "nós somos levados a concentrar nosso exame da experiência de ser consciente através de um fluxo de tempo que deve ser compreendido em termos de componentes infinitesimais" (HAUSMAN, 1993, p. 179).

Não temos, por conseguinte, uma apreensão completa e definida de uma idéia em um determinado instante do tempo. A cognição se dá no *continuum*, apenas através de um processo de imbricação das idéias que ocorre por meio de uma sucessão de intervalos infinitesimais no tempo. Durante tais intervalos, dirá Peirce, "nós percebemos diretamente a seqüência temporal de seu início, meio e fim, não naturalmente, como uma reconhecimento, pois a reconhecimento é unicamente do passado, na maneira de um sentimento (*feeling*) imediato" (CP 6. 110, 111).

Se, em um primeiro instante, temos acesso ao sentimento do imediato, característico da categoria de primeiridade, no segundo momento, temos a percepção da passagem do tempo. Ao ligar esses dois, temos um terceiro que indica uma direção e que requer uma interpretação, uma inferência, na medida em que se obtém uma generalidade.

Peirce continuará dizendo que quando há "um fluxo contínuo de inferências através de um tempo finito" o resultado "será uma consciência objetiva mediata de tempo completo no momento passado" (CP 6.111). Por meio dessa análise, o autor está propondo uma abordagem da relação do fluxo temporal sentido e a interpretação cognitiva do mesmo.

A cognição que é uma experiência mediata requer a consideração dos sentimentos passados, impondo a eles uma certa definição de tal maneira que possam ser comparados, o que se dá segundo um ato interpretativo. Cada

intervalo consiste em um momento da percepção consciente, de tal maneira que se pode afirmar que a própria consciência é contínua.

### **O idealismo objetivo**

Um grande problema para o materialista é o de explicar como as leis da matéria, que são completamente cegas, ou não teleológicas, podem dar origem a uma mente que não age cegamente, mas que possui propósitos; dirá Peirce, o problema é de como a ação teleológica ou propositada pode ser um efeito secundário de uma ação não teleológica.

A solução para a questão segue-se a partir da inspiração darwiniana, por meio da combinação de vários fatores, entre eles: as insensíveis variações fortuitas e a destruição das formas menos aptas, cujas transformações as levam além de certos limites.

O último desses fatores poderia ser entendido como uma lei mecânica, mas o que dizer do primeiro? Pergunta-se Peirce.

A necessidade de não bloquear a via de investigação, como tantas vezes apontadas pelo autor, não é transgredida por essa pergunta irrespondível, uma vez que seria um absurdo a tentativa de explicação do próprio acaso, pois as explicações se dão por meio de generalidades e as variações fortuitas são únicas e auto-referentes. Já a lei da evolução ou lei da mente necessita de uma explicação e pode ser aqui esboçada.

Em uma formulação geral, Peirce distinguiu três tipos de evolução: a evolução por variação fortuita, denominada de evolução ticástica, ou **ticasmo**, a evolução por necessidade mecânica, denominada de evolução anancástica, ou **anancasmo** e a evolução por amor criativo, denominada de evolução agapástica, ou **agapasmo**.

Os três modos de evolução são compostos pelos mesmos elementos gerais, diz Peirce. Os pais passam uma quantidade de "talentos" espontaneamente para a geração seguinte e esta possui a disposição de apanhá-los e desenvolvê-los e, dessa maneira, servir a um propósito geral. Esses elementos mostram-se mais claramente no agapasmo, mas também estão presentes de maneira diversa no ticasmo e no anancasmo. Essas duas formas podem ser vistas como formas degeneradas de agapasmo, afirma o autor.



A evolução agapástica, que por ser mais representativa da atuação da causalidade final, interessa-nos, aqui, mais de perto. Este tipo de evolução considera que a própria idéia possui uma força de atração que conduz o pensamento para si. Como afirma Peirce:

O desenvolvimento agapástico do pensamento é a adoção de certas tendências mentais, não totalmente descuidadas como no ticismismo, nem completamente cegadas pela mera força das circunstâncias ou da lógica, como no anancismo, mas por uma imediata atração da própria idéia, cuja natureza é adivinhada depois que a mente a possui, pelo poder da simpatia, isto é, por virtude da continuidade da mente... (CP 6, 307).

Quando Peirce refere-se ao pensamento, este deve ser entendido como “o princípio primordial para a compreensão não só dos fenômenos do espírito, mas da totalidade do cosmo” (SILVEIRA, 1989, p. 75). A matéria, afirma Peirce,

[...] em obediência ao princípio, ou máxima, da continuidade, de que devemos assumir que as coisas são contínuas tanto quanto possamos, urge que devamos supor uma continuidade entre os caracteres da mente e da matéria, de tal forma que a matéria não seria mais do que mente, que tendo hábitos de tal modo empedernidos é levada a agir com um grau particularmente elevado de regularidade mecânica ou rotina (CP 6.277).

Em decorrência da aceitação da idéia de continuidade entre a mente e a matéria, Peirce se autodenomina em algumas passagens de idealista objetivo, afirmando a sua diferença com relação ao monismo materialista, indica que “o [monista materialista] faz da lei da mente um resultado especial das leis da matéria, enquanto que o [idealista objetivo] faz com que as leis da matéria sejam um resultado especial da lei da mente” (N 1, 200).

A partir de 1890, tendo conseguido vários avanços no estabelecimento das categorias da experiência, Peirce oferece uma interpretação da noção de hábito na qual lhe é conferido um estatuto ontológico, distanciando-se, dessa maneira, tanto da noção humeana, como daquela utilizada pelos psicólogos do início do século XX, como afirma Bortolotti “o hábito passa a ser a categoria universal da lei, dotado de objetividade” (BORTOLOTTI, 2003, p. 1). Segundo Peirce, “um hábito não é uma afecção da consciência, é uma lei geral da ação, de tal forma que em um certo tipo geral de ocasião um homem será mais ou menos apto a agir de um certo modo geral” (CP 2.148). Como uma lei geral, o hábito se entranha na própria constituição do cosmo.



Como aponta Ivo Assad Ibri, referindo-se à filosofia de Peirce, “parece configurar-se um método de investigação que busca remover dificuldades tópicas que obstaculizam o exercício do pensamento” (IBRI, 1992, p. 61). A hipótese de uma lei da mente como necessária à expansão da inteligibilidade do universo, só se faz adequada se o próprio universo for inteligível. O idealismo peirceano tem caráter ontológico. As distinções entre interno e externo, entre mente e matéria são solapadas e, em seu lugar, institui-se um contínuo mente/matéria, como visto na sessão anterior. A concepção de um cosmos inteligível é uma decorrência da própria necessidade de se seguir ao primeiro princípio da lógica e não deixar nenhum elemento exilado para além de toda compreensão possível.

### **A Lei da Mente**

Peirce descreve a lei da mente como o fato de que as “idéias tendem a se dispersar continuamente e afetar certas outras que permanecem em uma relação peculiar de afetabilidade com relação a ela. Nessa dispersão, elas perdem intensidade, e especialmente o poder de afetar as outras, mas ganham em generalidade e se tornam amalgamadas com outras idéias” (CP 6.104).

É natural atribuir a uma lei o caráter de absoluta, como o atribuído às leis da física; a lei da mente, entretanto, não requer uma exata conformidade. Como afirma Peirce, “a conformidade exata estaria em conflito inequívoco com a lei, uma vez que isto cristalizaria instantaneamente o pensamento e impediria as formações posteriores de hábitos” (CP 6.23). A lei da mente apenas torna um certo sentimento mais provável (*likely*). Ou seja, a lei da mente inclina o curso futuro dos eventos a uma certa direção sem, contudo, determiná-lo.

A partir dessas considerações, Peirce conclui que o único tipo de causação primária deve ser a causação final e que esta deve ser considerada uma lei de caráter absoluto. Assim, sob a atuação da lei de aquisição de hábitos ou da mente, há um contínuo crescimento da uniformidade a partir das formas diferenciadas. Mas “as mudanças divergentes da lei estão agindo perpetuamente a fim de aumentar a variedade do mundo, e são controladas por um tipo de seleção natural ou de qualquer outro tipo (pois Peirce julga que a seleção natural não seja, por si só, suficiente) de tal forma que o resultado geral deve ser

descrito como “heterogeneidade organizada”, ou melhor, “variedade racionalizada” (CP 6, 23).

O princípio de aquisição de hábitos, ou lei da mente, é tomado por Peirce como o modo pelo qual as leis do cosmo se constituíram. Em outro momento, o filósofo dirá que “está claro que nada a não ser o princípio do hábito, ele mesmo devido ao crescimento de uma tendência infinitesimal do acaso em direção à aquisição de hábitos, é a única ponte que pode ligar o abismo entre o acaso do caos e o cosmo da ordem” (CP 6.263). Segundo Santaella, “para Peirce, a tendência do universo a adquirir novos hábitos, tendência esta que tem seu expoente na mente humana, é aquilo que permite o contínuo crescimento da potencialidade da idéia” (SANTAELLA, 2004, p. 248-9).

No plano da evolução cosmológica, o hábito, ou a lei, inclina a um desenvolvimento do universo que tem como meta torná-lo mais razoável. Peirce afirma que “lei é *par excellence* a coisa que quer uma razão” (CP 6.12).

O incremento da razoabilidade do universo não é, contudo, um processo que se desenvolva sem tropeços. Para bem compreendermos esse desenvolvimento, faz-se necessário uma incursão, ainda que breve, devido aos limites dessa exposição, na teoria das inferências que Peirce desenvolveu, em particular, no tipo de inferência denominado de abdução ou retrodução, termos que traduzem o termo aristotélico *apagoge*.

Nos trabalhos produzidos entre 1865 e 1901, vemos Peirce aprofundar e desenvolver suas análises e concepções do processo de produção de inferências, culminando com a distinção de três tipos de raciocínios que se integram e concorrem para a realização do conhecer: a indução, a dedução e a abdução. O problema de se determinar a origem das premissas de qualquer raciocínio foi uma constante a conduzir suas investigações nesse período. Conforme ressalta Santaella (2004), a interpretação da abdução tem sido um tópico polêmico entre os comentadores da obra de Peirce. O que distingue esse processo inferencial é o fato de ser, ao mesmo tempo, um processo instintivo e ter a natureza de uma inferência lógica. É por meio da abdução, deixa claro Peirce, que as novidades são introduzidas no processo evolutivo da natureza e do homem. A dificuldade do tratamento e compreensão desse conceito fundamental é bem sintetizada por Santaella ao se perguntar: “se as hipóteses são frutos da maravilhosa faculdade imaginativa humana, como podem elas se acomodar dentro da forma de uma inferência lógica?” (SANTAELLA, 2004, p. 109).

A intuição, entendida como *flash* de criatividade, é a base do processo de abdução que terá como finalidade a produção de uma nova idéia ou forma, que poderá dar lugar a uma inclinação geral ou hábito. Em uma leitura rápida, é tentador compreender esse flash como sendo de natureza intuitiva. Nada mais equivocado, se levarmos em conta que a principal crítica que Peirce endereça a Descartes refere-se justamente ao conceito de intuição. Para Peirce, toda cognição possui uma base inferencial, como é bem conhecido a partir dos textos de 1868-69. Se não há a possibilidade de uma cognição sem a existência de uma cognição que a preceda, como entender o surgimento da novidade que a abdução põe em cena?

A solução parece derivar do caráter habitual do instinto. Peirce afirma que todo instinto tem o caráter de um hábito e, dessa forma, distingue a ação por ele produzida da reação bruta e cega. São derivados da própria evolução a que está submetido o cosmo e o homem como uma parte integrante. Santaella aponta que “do instinto peirceano germina a abdução, fonte de todas as iluminações e criações humanas, mas também o mais frágil de todos os raciocínios, o mais falível, sem nenhum poder de comprovação, necessitando da dedução e da indução para que possa ter qualquer valor de verdade” (SANTAELLA, 2004, p. 113-14).

É desse modo, portanto, que podemos concluir que a grande lei atuante na constituição do cosmo ordenado é a lei da mente ou princípio de aquisição de hábitos, que atuando por meio de uma idéia que permanece como um ser *in futuro*, atrai os acontecimentos de forma a incrementar o grau de razoabilidade no universo. Essa lei, que rege um processo eminentemente falível recebe na descrição Peirceana sua melhor definição: “uma força gentil que comumente prevalece” (CP 2, 389).

### **Considerações finais**

Duas sugestões para o desenvolvimento da área de auto-organização podem ser apontadas a partir da exposição das idéias de Peirce realizada acima. Em primeiro lugar, uma avaliação e análise mais cuidadosa da noção de *continuum* e, em segundo lugar, a consideração da causalidade final como constituinte primário dos fenômenos de crescimento ocorrentes no universo, seja

a própria mente, ou qualquer outro evento, na medida em que a matéria, segundo nosso autor, nada mais é que um tipo especial de mente ou pensamento.

## **Referências**

- BORTOLOTTI, R. G. Características e fundamento do hábito no pensamento de C. S. Peirce. In: V SIMPÓSIO EM FILOSOFIA E CIÊNCIA DA UNESP, 5, 2003, Marília. *Anais...* Marília: FUNEPE, 2003. CD-rom.
- DEBRUN, M., GONZALES, M. E. Q.; PESSOA JR., O. (Org.) *Auto-organização: estudos interdisciplinares*. Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, 1996.(Coleção CLE; v. 18).
- HAUSMAN, C. R. *Charles S. Peirce's evolutionary philosophy*. Cambridge: Cambridge University Press, 1993.
- IBRI, I. A. *Kósmos noétos*. Perspectiva: São Paulo, 1992.
- PEIRCE, C. S. *Collected papers of Charles S. Peirce*. Edited by Charles Hartshorne and Paul Weiss. Cambridge, MA: The Belknap Press of Harvard University, 1934. v.1-6.
- \_\_\_\_\_. *Collected papers of Charles S. Peirce*. Edited by Arthur Burks. Cambridge, MA.: The Harvard University Press, 1958. v.7-8.
- \_\_\_\_\_. *Contribution to the nation*. In: Past Masters, C-rom databases. IntelLex Corporation, 1992.
- SANTAELA, L. *O método anticartesiano de C. S. Peirce*. São Paulo Editora Unesp:, 2004.
- SILVEIRA, L. F. B. Cosmo evolutivo e plano da criação na filosofia peirceana. In: *Trans/Form/Ação*, São Paulo, v.8, p.1-24, 1985.
- \_\_\_\_\_. Charles Sanders Peirce: ciência enquanto semiótica. In: *Trans/Form/Ação*, São Paulo: v. 12, p.71-84, 1989.





# Informação e conhecimento: notas para uma taxonomia da informação

Maria Eunice Quilici Gonzalez<sup>1</sup>

Thiago Carreira Alves Nascimento<sup>2</sup>

Willem F. G. Haselager<sup>3</sup>

*A vida tem um logos que se aumenta a si mesmo*

*Heráclito 115 D*

## Introdução

Vivemos na era da informação, mas sabemos pouco sobre a sua natureza ontológica e epistemológica. Em contraste com o conceito de conhecimento, cujo estudo permeia a história da Filosofia e da Ciência, o conceito de informação passa a ser objeto de estudos sistemáticos apenas nos meados do século XX. No ocidente, sua origem pode ser encontrada nas idéias de Hartley (1928) e Szilard (1929/1972), em especial nas suas reflexões sobre a segunda lei da Termodinâmica. Tais reflexões possibilitaram a formulação de hipóteses sobre a natureza da informação em termos da relação entre ordem e desordem de um sistema no âmbito da comunicação. Idéias semelhantes foram posteriormente desenvolvidas por Shannon e Weaver (1949/1998), as quais se propagaram rapidamente em diversas áreas do conhecimento, com resultados geralmente diversos.

O objetivo central deste capítulo é elaborar um mapa conceitual das principais abordagens filosófico-científicas do conceito de informação, que possibilite uma classificação de suas principais vertentes.<sup>4</sup> Tal classificação é

---

<sup>1</sup> Departamento de Filosofia e Programa de Pós-Graduação em Filosofia da Faculdade de Filosofia e Ciências da Unesp, *campus* de Marília. E-mail: gonzalez@marilia.unesp.br.

<sup>2</sup> Curso de Graduação em Filosofia da Faculdade de Filosofia e Ciências da Unesp, *campus* de Marília.

<sup>3</sup> Artificial Intelligence, Nijmegen Institute of Cognition and Information (NICI), Radboud University, Nijmegen, Holanda e Pós-Graduação em Filosofia, UNESP, Marília, SP, Brasil. E-mail: w.haselager@nici.ru.nl. Home page: [www.nici.ru.nl/~haselag](http://www.nici.ru.nl/~haselag).

<sup>4</sup> O presente capítulo é parte integrante de um projeto de maior amplitude (desenvolvido junto ao Departamento de Filosofia da Unesp/Marília), cujo título é *Informação e auto-organização no processo de aquisição do conhecimento*.

provisória, dado o nosso atual estado de conhecimento sobre a natureza da informação; porém, a partir dela, acreditamos esclarecer alguns problemas contemporâneos concernentes ao estatuto ontológico da informação e sua relação com o conhecimento.

Variando de acordo com os instrumentos tecnológicos da época, as investigações sobre a natureza do conhecimento, que no passado pertenciam essencialmente ao domínio clássico da filosofia, têm sido objeto de estudo interdisciplinar dos cientistas cognitivos. Na Ciência Cognitiva, questões relativas à natureza dos processos de aquisição e desenvolvimento do conhecimento são analisadas a partir de uma perspectiva interdisciplinar que integra, entre outras, a Filosofia, a Psicologia, a Antropologia, a Linguística, a Ciência da Informação e a Inteligência Artificial, seja no seu viés simbólico computacional, seja no viés neurológico computacional do Conexionismo, da Robótica e da Neurociência Cognitiva. O sujeito cognitivo deixa de ser visto, nessa perspectiva, como um sujeito absoluto ou transcendental, situando-se no plano contingente dos sistemas informacionais que se auto-organizam no embate entre elementos físicos, biológicos e sociais.

O estudo interdisciplinar do sujeito cognitivo realizado a partir da perspectiva informacional requer uma taxonomia que possibilite o entendimento da natureza da informação, bem como o estabelecimento de sua relação com os processos de aquisição do conhecimento. Atualmente, a pluralidade de concepções sobre a informação não somente está relacionada aos fins para os quais elas se desenvolvem (em termos de explanação teórica e prática), como também às bases a partir das quais elas se erigiram. Entendemos que a ausência de uma taxonomia comum a respeito da natureza da informação dificulta o diálogo interdisciplinar e gera, muitas vezes, equívocos no tratamento de questões ligadas à relação sujeito/conhecimento e informação.

Com o objetivo de facilitar o diálogo interdisciplinar no estudo de questões que envolvem o tópico da informação, apresentamos, na *Parte I* deste capítulo, um breve histórico do conceito de informação, seguido de uma análise desse conceito no viés da clássica teoria matemática da comunicação.

Na *Parte II*, apresentamos as principais hipóteses constitutivas da vertente que denominamos *Epistemológica e Ontológica da Informação*, a qual fornece subsídio para a nossa proposta de desenvolver uma taxonomia da informação. A partir dessa taxonomia, propomos uma definição do conceito de

informação que julgamos reunir as principais características dos processos auto-organizados, tal como caracterizados por Debrun (1996), Haken (1983), Haken e Wunderlin (1990) e Gonzalez (1996). Segundo essa definição, a *informação* não é uma entidade, coisa ou substância mas, sim, um processo auto-organizado que permite o estabelecimento de padrões de ação para organismos situados em ambientes governados por relações compartilhadas de ordem. Argumentamos que essa definição pode nos auxiliar na compreensão da natureza de aspectos do conhecimento comum.

### Considerações históricas e etimológicas sobre o conceito de informação

De acordo com Capurro (1978), as raízes etimológicas do termo latino *informatio* estariam ligadas a conceitos gregos como *typos*, *idea* e *morphé*, sendo que a sua acepção atualmente conhecida de “dar forma a algo” só foi consolidada entre os séculos XV-XVII. Considerando ainda as raízes etimológicas do termo informação, verificamos que ele é composto pelo prefixo *in* e pelo nome (substantivo) *formae*, o qual designa a noção de forma. Segundo Agostinho (1993), forma, por sua vez, pode ser entendida segundo a concepção platônica de idéia (*ιδέα*). Como aponta Capurro, os termos latinos *informatio/informare* aparecem nas traduções e comentários desses conceitos gregos, apresentando-se sob um paradigma ontológico que, como veremos na Parte II, passa a adquirir atualmente também uma caracterização epistemológica.

O prefixo *in* do termo informação pode ser entendido, enquanto modo acusativo, em três concepções: 1) como expressão da noção de movimento em direção ao lugar para onde se dirige algo; 2) como ordem temporal; de algo que se desenvolve através de processos, e, 3) em sentido figurado; como referência à noção de devir. Considerando essas três acepções do prefixo *in*, podemos compreender o termo informação em relação a um movimento que se desenvolve, através do tempo, rumo ao estabelecimento de algo ainda em formação.

A partir dessa breve indicação da origem e constituição do termo informação, propomo-nos a analisá-lo, historicamente, através de duas vertentes

principais; quais sejam: (I) a da teoria matemática da comunicação, MTC, de caráter tecnológico; e (II) das teorias de cunho epistemológico e ontológico<sup>5</sup>.

Na vertente (I) são preponderantes os problemas formais e técnicos da engenharia sobre a aplicação de teorias no âmbito tecnológico envolvidos na transmissão e recepção de mensagens. As preocupações centrais dos teóricos dessa vertente giram em torno da comunicação, da mensurabilidade e da quantidade de informação gerada em uma fonte, da capacidade do canal que a transmite, bem como da sua confiabilidade e efetividade na transmissão de dados. O elemento comum das teorias constitutivas dessa vertente é o manuseio estatístico de dados, que ressalta o seu caráter probabilístico no interior de estruturas simbólicas devidamente codificadas segundo um padrão de organização. O foco dessa vertente, essencialmente técnica, não se situa, em geral, sobre questões epistemológicas ou ontológicas, mas, antes, no controle eficiente da comunicação entre fonte e receptor, independentemente do significado e da natureza do conteúdo informacional transmitido.

As reflexões desenvolvidas nessa primeira vertente se iniciaram, como mencionamos, com os trabalhos de Hartley (1928) e Szilard (1929/1972) que, investigando a segunda lei da termodinâmica, sugerem uma analogia entre informação, formas de organização e entropia. De acordo com a segunda lei da Termodinâmica, há uma tendência nos sistemas físicos de se tornarem progressivamente desorganizados, sendo a entropia caracterizada pela quantidade de desordem de seus constituintes; analogamente, em sistemas físicos, a informação sobre a localização precisa de seus elementos se perderia progressivamente, sendo entretanto possível obter uma medida da informação, em termos da redução de incerteza na escolha desses elementos, uma vez que se conheça a organização geral do sistema em questão.

Algumas das idéias de Hartley e Szilard foram desenvolvidas e divulgadas no período pós-guerra, principalmente com os trabalhos de Shannon

---

<sup>5</sup> Rafael Capurro (2003), ao desenvolver uma “Epistemologia da informação” apresenta a questão da informação na contemporaneidade relacionando-a com a noção de paradigma, originalmente proposta por Kuhn, classificando cada campo em que a informação é considerada como um “paradigma epistemológico”. São apresentados por ele três paradigmas, a saber: físico, cognitivo e social. Uma similaridade entre nosso esquema e aquele apresentado por Capurro reside somente na tentativa de esquematização da questão da informação na contemporaneidade. Contudo, enquanto Capurro discorre sobre a fundamentação de uma ciência da informação, buscando maneiras de conceituar e justificar o campo de trabalho dessa ciência, nós, por outro lado, buscamos entender, e posteriormente conceituar, o que é isso que chamamos de informação.



e Weaver (1949), que forneceram uma caracterização matemática da medida da informação na transmissão de mensagens em sistemas físicos. Nessa mesma época se iniciara também a Cibernética, que teve Wiener como seu principal mentor, o qual a caracterizou como a ciência do controle por máquinas de processamento de informação. Para ele:

O objetivo da Cibernética é desenvolver uma linguagem e técnicas que nos capacitem, de fato, a lidar com o problema do *controle e da comunicação em geral*, e também a descobrir o repertório de idéias técnicas adequadas para classificar suas manifestações específicas sob a rubrica de certos conceitos" (WIENER, 1968, p.18, *itálico nosso*)

Na obra *The Human use of human beings*, acima citada, Wiener propõe uma análise do conceito de informação a partir da noção de mensagem – entendida como seqüências ordenadas de símbolos – sob a qual se funda a sua teoria da comunicação aplicada às máquinas e à sociedade. Nessa obra ele argumenta que:

A tese deste livro é que a sociedade só pode ser compreendida através de um estudo das mensagens e das facilidades de comunicação de que disponha; e que, no futuro desenvolvimento dessas mensagens e facilidades de comunicação, as mensagens entre o homem e as máquinas, entre as máquinas e o homem, e entre a máquina e a máquina, estão destinadas a desempenhar um papel cada vez mais importante (WIENER, 1968, p. 18).

O estudo da natureza da informação desempenha um papel de extrema importância para essa nova ciência, uma vez que ela constitui o elemento chave para se compreender a comunicação e a inter-relação entre as máquinas, bem como entre os organismos e o meio que os circunda. Das suas várias caracterizações, destaca-se aquela fornecida por Wiener (1968, p.19), segundo a qual *informação*: "[...] é o nome dado ao conteúdo daquilo que permutamos com o mundo exterior ao ajustar-nos a ele, e ao fazer com que nosso ajustamento seja nele sentido".

A questão que logo se destaca nesse período, pela voz de Wiener, diz respeito ao controle e à comunicação, ganhando notoriedade com os trabalhos de Shannon e Weaver, *The Mathematical Theory of Communication* (MTC). Embora o conceito de informação não seja definido em detalhes nessa obra, que focaliza essencialmente a *medida* da informação, fica claro que ele está diretamente relacionado à noção de ordem e ao número de decisões envolvidas



no processo de redução de incerteza na escolha de mensagens. Pelo seu caráter primordialmente quantitativo, a MTC focaliza, como já ressaltamos, aspectos objetivos concernentes à medida da quantidade de informação gerada em uma fonte e à transmissão de mensagens, deixando de lado aspectos de ordem *semântica* ou que envolvam o *significado* da mensagem transmitida. Weaver, discorrendo sobre a noção de informação na MTC resalta que:

O termo informação na teoria da comunicação diz respeito nem tanto àquilo que você diz, mas àquilo que você poderia dizer. Isto é, informação é uma medida da liberdade de escolha quando se seleciona uma mensagem. Em uma situação muito elementar, na qual uma de duas mensagens alternativas deve ser escolhida, então arbitrariamente se diz que a informação, associada a esta situação, é a unidade (SHANNON; WEAVER, 1949/1998, p. 8-9).

Nessa mesma obra, Weaver resalta que, no contexto técnico da MTC, o conceito de informação, diferentemente da noção de significado, não se aplica a mensagens individuais, mas, sim, a situações compreendidas em sua totalidade. Nesse sentido, a unidade de informação indica que em tal e tal situação se dispõe de uma certa liberdade de escolha na seleção de uma mensagem.

A MTC tem como pressuposto o cenário de um universo regido por leis físicas expressas com o auxílio do formalismo probabilístico da Termodinâmica. Não entraremos em detalhes sobre esse formalismo que permite calcular a quantidade de informação em uma fonte, quantidade essa que expressa a redução de incerteza presente no processo de seleção de mensagens (um comentário mais detalhado sobre esse tópico foi desenvolvido em PEREIRA Jr.; GONZALEZ, 1995). Para o presente propósito basta lembramos que, de acordo com a segunda lei da termodinâmica, existe uma tendência dos sistemas fechados a se encaminharem, com o passar do tempo, para a morte térmica, isto é, para o aumento progressivo da desordem em detrimento da presença de algum tipo de ordem. Nessa situação, a homogeneidade será crescente, indicando a independência dos seus elementos.

Analogamente, em uma fonte geradora de informações, se as mensagens forem relativamente independentes ou igualmente prováveis, a escolha de qualquer uma delas será livre e, conseqüentemente, a incerteza envolvida nessa escolha será máxima. Em contraste, se a ocorrência de uma mensagem for mais provável do que outra, em conseqüência de certas relações

de dependência, a redução de incerteza na sua escolha será menor e, de acordo com a MTC, a quantidade de informação gerada será menor do que aquela da situação anterior.

Esse modelo de processamento de informação, esquematizado com fins práticos voltados principalmente à telefonia, vem sendo desde então aplicado, com grande sucesso, nos sistemas de comunicação em geral, desempenhando um papel fundamental não apenas na Cibernética, mas também no desenvolvimento tecnológico que permeia o nosso cotidiano. Wiener, influenciado pelas idéias de Boltzmann e Gibbs sobre entropia, considera que estamos imersos em um mundo que, em sua totalidade, obedece à segunda lei da Termodinâmica: a confusão aumenta e a ordem diminui. A vida, entretanto, expressa a força que se opõe a essa tendência, sendo que *a fonte dessa força é a informação*, representada pela constatação, temporária, de ordem no pano de fundo da desordem. De acordo com essa visão, a principal característica que permite aos organismos se manterem vivos é a sua inter-ação com o meio circundante através da comunicação, sendo a informação o “alimento” que rege essa dinâmica de comunicação.

Ao investigar a dinâmica dos sistemas físicos e biológicos de processamento de informação, Wiener inicia a segunda vertente de estudos da informação que, como veremos, tem uma preocupação de cunho ontológico e epistemológico. Além disso, dados os propósitos técnicos e práticos das reflexões sobre comunicação realizadas pela primeira vertente de estudos da informação, lacunas teóricas foram deixadas em relação ao seu estatuto ontológico e epistemológico: afinal, o que é informação? O que é isso que se mede e se transmite de uma fonte para um receptor segundo um princípio de redução de incerteza? Trata-se de uma forma de organização essencialmente abstrata ou, em contraste, de algo físico que constituiria os tijolos fundamentais dos processos estruturadores do mundo, em geral, e da mente, em particular? Possuiria a informação elementos subjetivos, difíceis de serem conhecidos a partir de uma perspectiva objetiva? Ou, como postularam os adeptos da vertente I, seria ela uma propriedade objetiva do mundo, passível de medição e descrição probabilística? Existiria alguma semelhança entre a noção técnica de informação e aquela utilizada pelo senso comum, que envolve, entre outros, significado?

As questões acima escapam dos interesses centrais da vertente I, mas elas constituem, como veremos a seguir, o ponto de partida e o núcleo da vertente II de estudos da informação.

### **Abordagens ontológica e epistemológica da Informação**

Três tendências podem ser destacadas entre as reflexões atuais sobre a natureza ontológica e epistemológica da informação, as quais serão aqui reunidas, para efeito de classificação, na vertente II de estudos e denominadas, respectivamente (1) realismo informacional, (2) ecologia informacional e (3) semântica informacional.

O *realismo informacional* é caracterizado por teorias que afirmam a realidade física e objetiva da informação. Um dos seus principais representantes é Tom Stonier, que se tornou conhecido pelas obras *Information and the Internal Structure of the Universe* (1990) e, também, *Information and Meaning* (1997). As idéias de Schaeffer (2001, 2004) e Zeiman (1970), entre outros, estão próximas às de Stonier no que diz respeito ao pressuposto da realidade objetiva da informação, ainda que se diferenciem no emprego de formalismo matemático em sua teoria.

Ao defenderem a existência objetiva da informação, os realistas informacionais compartilham da concepção apresentada pela MTC (primeira vertente), mas diferem em um ponto crucial: no que diz respeito ao papel da organização e da ordem na constituição da informação. A ordem é entendida pelos teóricos *realistas* numa acepção muito próxima àquela de “forma”; isto é, como um arranjo não randômico das partes de uma estrutura ou sistema; enquanto que a organização é entendida como um “reflexo da ordem” (STONIER, 1990, p. 26). Um pressuposto fundamental do realismo informacional é que ambas, ordem e organização, constituem a matéria prima da informação que se propaga pelo universo em crescente complexidade.

O pressuposto acima diferencia radicalmente os realistas informacionais dos teóricos da MTC, pois, como indicamos, para estes últimos, quanto maior a quantidade de desordem em um sistema, expresso em termos da relativa independência das mensagens, maior será a quantidade de informação gerada. Já os realistas informacionais argumentam que, quanto maior a

desordem, menor a quantidade de informação; e isso porque a informação – constatada pela presença de ordem nos sistemas – tem que ser algo que esteja de acordo com um todo. Assim, para comentar um exemplo fornecido por Stonier, podemos imaginar uma biblioteca na qual os livros foram parcialmente destruídos, sendo suas páginas rasgadas e espalhadas pelo local. Para os teóricos realistas, muito pouco de informação poderia ser extraída dessa situação randômica; no entanto, já para os teóricos da MTC, essa situação carregaria mais informação, caso se conseguisse decodificar as mensagens nesses livros, do que se a biblioteca estivesse em suas condições normais (isto é, organizada corretamente).

Uma caracterização alternativa que poderia exemplificar a concepção realista da informação nos é dada por Zeiman (1970, p. 157), de acordo com o qual: “A informação é, pois, a qualidade da realidade material de ser organizada (o que representa, igualmente, a qualidade de conservar este estado organizado) e sua capacidade de organizar, de classificar em sistema, de criar”. Tal capacidade organizadora tende a se propagar pelo universo aumentando a sua complexidade.

Um outro pressuposto fundamental do realismo informacional é que a informação é um elemento constituinte *per se* do mundo físico. Stonier (1990), por exemplo, argumenta que a informação é um elemento constituinte do universo, ao lado de matéria e energia; sendo a existência de partículas genuínas de informação por ele denominada de *infons*. Estas, por sua vez, não se reduzem à matéria ou energia, mas possuem uma existência própria – tais partículas são tão reais quanto a matéria e a energia (lembrando-se que, para o realismo informacional, a informação é algo *real* e pode ser constatada pela presença de ordem, independente de qualquer observador atual). A combinação dos *infons* em padrões significativos de informação depende do contexto em que eles se encontrem.

Em resumo, a hipótese do realismo informacional, segundo a qual a informação é um constituinte *per se* do mundo físico, leva-nos à concepção de um universo organizado e hierarquizado em graus de informação, de acordo com uma crescente complexidade, sendo o seu significado decorrente de contextos específicos<sup>6</sup>.

---

<sup>6</sup> Entendemos que, de certo modo, esta concepção está muito próxima da imagem de mundo que o filósofo do século XVII, Leibniz, apresentou em sua *Monadologia*, à medida que há nela um



O segundo sub-grupo da vertente (II), que denominamos *ecologia informacional*, entende a informação como inerente à *relação* do ser com o mundo. Para os teóricos dessa vertente, dentre os quais se destacam Gibson (1966, 1979) e Bateson (1999), a informação está intrinsecamente ligada à ação e à situação dos organismos no meio ambiente. Nesse sentido, ela é um constituinte fundamental, quando não central, da vida; sendo, ademais, aquilo que possibilita enxergar (para utilizarmos uma expressão de Bateson) “o padrão que conecta” as coisas. A informação guia o ser em sua existência no mundo, não sendo considerada como algo (puramente) físico, mas essencialmente relacional e significativo.

Para os teóricos dessa vertente ecológica, o conteúdo significativo que a informação possa carregar é relativo ao tipo de organismo que a experiência em sua atuação no mundo. Nesse sentido, o estudo da informação não se resume à constatação das características físicas organizadoras dos eventos em questão, mas sim, à sua relação para com os organismos que, de certa forma, a constituem no seu agir: o caráter qualitativo da informação, como também significativo, dá-se na inter-ação do organismo (agente) com o mundo. Nesse sentido, segundo a concepção ecológica, a informação ganha não só uma propriedade organicista, como também um aspecto relacional e qualitativo inerente à própria vida.

Para Gibson (1979), por exemplo, a informação ecológica não pode ser caracterizada apenas em termos de comunicação de mensagens, tal como a entendem os teóricos da MTC, pois o ambiente não se comunica com um observador passivo que apenas seleciona mensagens mas, sim, com um organismo ativo que participa do seu universo de escolhas, alterando-o e sendo por ele alterado. Nesse sentido, Gibson cria o termo *affordance* para designar o processo de interação do ser-no-mundo; um neologismo provindo do verbo *to afford* (fornecer, possibilitar). Esse termo não possui uma tradução direta para o português ou mesmo um significado em si, mas designa aquilo que o meio fornece (possibilita ou oferece) para facilitar a ação de um organismo situado em seu ambiente natural. Assim, por exemplo, se para um ser humano uma árvore possibilita descanso à sua sombra (a *affordance* sombra está disponível no ambiente para o ser humano), para um macaco esta mesma situação pode

---

pressuposto sobre a realidade última do universo fundamentado em uma ontologia de partículas genuínas constituintes de um universo hierarquizado segundo uma harmonia pré-estabelecida, que poderíamos atualmente entender em termos *informacionais*.



oferecer a chance de grande agitação propiciada pela oportunidade de subir e colher frutos.

Em sendo relacional, a existência das *affordances* depende do tipo de situação que se configura no meio ambiente e do tipo de organismo que se encontra preparado para captá-las. Apesar de seu caráter relacional, Gibson insiste em que as *affordances* possuem uma existência própria, podendo ser entendidas como informação objetiva, relativa à espécie, independente dos estados subjetivos particulares dos indivíduos que as captam.

A existência objetiva das *affordances* não se deve unicamente aos estímulos físicos do meio ambiente, mas envolve a história evolutiva e a dinâmica intrínseca dos seres que as experienciam. Tal dinâmica está diretamente ligada à estrutura corpórea dos organismos e do meio em que eles se situam. Assim, por exemplo, a estrutura corpórea de um pássaro lhe permite realizar certos movimentos no ar, mas não na água ou na terra, e captar as *affordances* disponíveis no seu nicho. Uma vez estabelecidas, essas *affordances* deixarão traços no ambiente, facilitando as ações futuras dos organismos que compartilham a mesma dinâmica intrínseca; certas *affordances* permanecem inalteradas ainda que mudanças radicais ocorram nos estímulos recebidos por um organismo, como é o caso, por exemplo, de uma cadeira de madeira, de um bloco de metal ou de uma pedra que oferecem a mesma possibilidade de sentar para um ser humano, apesar de suas diferenças materiais.

Para Gibson, as *affordances* constituem informação diretamente relacionada às *invariâncias*, que podem ser entendidas seja na perspectiva de padrões recorrentes de movimento, constituindo as *invariantes transformacionais*; seja como padrões luminosos refletidos pelos objetos devido às suas propriedades físico-químicas constitutivas, denominadas *invariantes estruturais*.

Exemplos de invariantes transformacionais podem ser encontrados nos padrões que distinguem o caminhar de uma pessoa quando ela está com pressa, bêbada, doente, tranqüila, etc. Esses padrões, apesar de diferirem entre si, constituem invariantes que, por estarem disponíveis no meio ambiente sócio-cultural, são imediatamente reconhecidos nas mais diversas situações do cotidiano. Conforme argumentaremos na terceira parte deste capítulo, a apreensão deste tipo de *affordances* dispensa inferências ou representações proposicionais; elas ilustram, no exemplo em questão, informação ecológica

sobre o conhecimento comum relacionado ao andar; informação esta que se encontra disponível em ambientes nos quais pessoas caminham, correm, se embebedam, adoecem, etc.

Exemplos de invariância estrutural, por outro lado, podem ser encontrados em objetos de vidro, madeira, ferro, etc., que são identificados independentemente de suas formas, tamanho ou lugar, graças às suas propriedades físico-químicas (invariantes) refletidas em padrões luminosos. Assim, qualquer pedaço de vidro será reconhecido como vidro, esteja ele na forma de um vaso, copo, vidraça, cacos, etc. Isso porque, em condições normais, os padrões luminosos deles refletidos facilitam a interação dos organismos com o ambiente e, em conjunto com os invariantes transformacionais, delimitam sua ação como ser-no-mundo.

Ao caracterizar a informação ecológica em termos de invariantes estruturais ou transformacionais (que constituem as *affordances*), Gibson a distingue daquela do realismo informacional em seu pressuposto fisicalista-reducionista, mas estabelece o que ele denomina de “realismo ingênuo” (*Naive Realism*). De acordo com o realismo ingênuo gibsoniano, a informação é uma propriedade objetiva do mundo ecológico – do qual a ação dos organismos constitui parte essencial. Nesse sentido, o realismo informacional (segundo Stonier) se diferencia do realismo ingênuo ao colocar em segundo plano o patamar fundamental da ação.

Em resumo, a informação para os teóricos da ecologia informacional, entre os quais Gibson foi tomado como exemplar, é o conjunto de invariantes constitutivos das *affordances* que se estabelecem na relação entre organismo e meio ambiente. Apesar de seu caráter objetivo, ela varia, em geral, de espécie para espécie, dependendo das suas formas de atuação e dos seus respectivos ambientes.

Finalmente, para encerrar esta classificação provisória das principais concepções da informação, vamos investigar as suas características distintivas tais como indicadas pelo sub-grupo (3), que denominamos *semântica informacional*. Entre as abordagens mais significativas deste sub-grupo, destacam-se as tentativas de explicitar a natureza da informação que lhe confere significado, tal como realizadas por Bar-Hillel e Carnap (1953) e Dretske (1981, 1992, 1995).

Carnap, por exemplo, direcionou parte de suas investigações sobre probabilidade indutiva ao desenvolvimento de uma concepção semântica de informação caracterizando-a como um elemento esclarecedor (desvelador) dos fundamentos de qualquer teoria que viesse a utilizá-la. Nesse contexto, a informação semântica versaria sobre as relações de significado do conteúdo carregado por um sinal numa dada linguagem. Segundo Carnap, essa sua teoria estaria próxima do que se poderia chamar de *teoria pragmática da informação*, e seria de grande valor não apenas para a filosofia em geral, mas também para ciências como a Psicologia, uma vez que a mensagem expressa através de proposições não apenas teria um valor de verdade, como também, um significado que estaria atrelado à linguagem em questão. A proposta de Carnap, entretanto, teve pouca repercussão na filosofia e na ciência, ficando engavetada no seu mundo de sonhos positivista.

Uma abordagem contemporânea da *semântica informacional*, amplamente conhecida pelos filósofos da mente e cientistas cognitivos, foi originalmente elaborada por Dretske (1981), que aborda problemas relativos à Teoria do Conhecimento, a partir da perspectiva informacional<sup>7</sup>.

De acordo com Dretske (1981, 1992, 1995), o conhecimento, tradicionalmente entendido como “uma forma justificada de crença verdadeira”, pode ser descrito de modo mais eficiente em termos de “crença fundada em informação”. Aquilo que daria veracidade e justificaria uma crença, culminando em conhecimento empírico, perceptual, seria a informação.

Inspirado na MTC, Dretske desenvolve uma teoria naturalista do conhecimento tendo como pressuposto a hipótese de que a informação é um elemento objetivo, “cuja existência (enquanto informação) é (amplamente) independente das atividades interpretativas de agentes conscientes” (DRETSKE, 1981, p. 55). Contudo, diferentemente dos teóricos da MTC, a preocupação de Dretske reside na tentativa de explicar o aspecto significativo da informação subjacente às crenças que fundamentam o conhecimento vinculado à percepção. Nesse contexto, o seu ponto de partida consiste em explicitar a noção de conteúdo informacional de um sinal, evento ou situação por meio do que ele denomina *definição nuclear* de informação.

---

<sup>7</sup> Parte da análise realizada aqui sobre a semântica informacional Dretskeana foi apresentada na tese de livre docência realizada por Gonzalez (2004).

Segundo a definição nuclear de informação proposta por Dretske: um sinal carrega informação (nuclear) sobre o que ocorre em uma fonte (expressando o seu conteúdo) se ele for capaz de reproduzir factualmente as relações que se estabelecem na fonte, tornando-as acessíveis para qualquer observador que se encontre em condições de recebê-las.

Conforme ressaltado por Gonzalez, em Milidoni, Gonzalez e Broens (2004), a definição dretskeana de informação nuclear pressupõe que os eventos no ambiente impõem *restrições* aos sinais que carregam informação sobre eles, tornando-os necessariamente dependentes. Em outras palavras, Dretske (1981, p. 45) argumenta que um sinal, enquanto candidato à informação no sentido nuclear, deve refletir fielmente aquilo que ocorre na fonte, de tal forma que um receptor, devidamente equipado, poderá conhecer, com probabilidade condicional = 1, ou vir a aprender o que lá ocorre, apenas observando esse sinal.

A correspondência estrita entre o que ocorre na fonte e o sinal que a transmite deve obedecer, segundo Dretske (1981, p. 65), a seguinte condição:

Se existe uma correlação necessária entre os eventos X e Y, de tal forma que uma mudança específica em X produz uma mudança, igualmente específica em Y, com probabilidade condicional = 1, e se tal mudança é registrada em Y, então Y carrega informação sobre X.

Tal exigência, da probabilidade condicional = 1, que expressa uma correlação nômica entre X e Y, constitui para Dretske uma característica intrínseca das relações informacionais. Contudo, o caráter estritamente normativo dessa exigência (da probabilidade condicional = 1) tem sido criticado devido à natureza contingente do mundo em que vivemos: Uma vez que em nosso mundo contingente não podemos estabelecer a ocorrência de um evento com probabilidade = 1, a hipótese informacional dretskeana não pode, em princípio, ser testada. Essa dificuldade será retomada ao final deste capítulo; no momento, interessa-nos analisar apenas a noção de conteúdo informacional.

Segundo Dretske, o conteúdo informacional de um sinal pode ser explicitado através de sua digitalização. A seguinte distinção entre informação digital e analógica nos ajuda a compreender sua hipótese:

[...] um sinal (estrutura, evento, estado) carrega informação que *s* é *F* na forma digital se e somente se o sinal não carrega informação adicional sobre *s*, que não esteja já contido no fato de *s* ser *F*. Se o sinal carrega informação adicional sobre *s*, a qual não está imbricada



no fato de *s* ser *F*, então direi que o sinal carrega essa informação na forma analógica (DRETSKE, 1981, p. 137).

De acordo com a caracterização acima, a informação digital resulta de um recorte na informação originalmente apresentada na forma analógica; uma filtragem deverá ocorrer nesta última para que a condição de digitalização seja satisfeita. Em tal condição, Dretske argumenta que um *conteúdo natural* será transmitido através de sinais ou eventos, conteúdo este que desempenhará a função de um indicador natural para a ação.

A hipótese central de Dretske é que, em princípio, a digitalização de um sinal garante a especificidade da informação percebida através da filtragem da informação analógica disponível no ambiente. Essa operação de filtragem estaria presente também nas atividades cerebrais responsáveis pela nossa percepção do mundo. Contudo, a expressão (e articulação) apropriada dessa digitalização envolveria uma postura intencional:

[...] afirmações que descrevem a informação carregada por um sinal são descrições *intencionais*, e os fenômenos que elas descrevem podem ser qualificados como fenômenos intencionais (DRETSKE, 1981, p. 76).

A hipótese dretskeana, segundo a qual a especificidade da informação carrega consigo um conteúdo natural, cuja articulação apropriada requer uma postura intencional, possui vários desdobramentos, os quais estão detalhados em Milidoni, Gonzalez e Broens (2004). Para o presente propósito é suficiente ressaltarmos que, segundo Dretske, e outros representantes da semântica informacional, a informação significativa não estaria no mundo físico, como querem os defensores do realismo informacional, nem nas *affordances* experienciadas pelos organismos, como querem os realistas ingênuos. O aspecto significativo da informação dependeria da postura intencional, intrinsecamente representacional, de sistemas cognitivos complexos que, a exemplo do ser humano, possuem a capacidade de aprender através da representação e correção de sinais indicadores de regularidades no mundo.

O caráter naturalista da abordagem dretskeana aproxima a sua semântica informacional tanto da MTC quanto da abordagem ecológica da informação. Contudo, eles se diferenciam pela ênfase dada ao aspecto intencional e representacional da informação digital, significativa, que, para



Dretske, possui uma contrapartida no mundo físico, mas não se reduz a elementos físicos.

Conforme procuramos ressaltar, embora os três subgrupos de estudos da informação, constitutivos da vertente II, compartilhem, como na MTC (vertente I), da tese da existência objetiva da informação, cada um deles apresenta um foco distinto sobre a sua natureza, incluindo (ou excluindo) certas propriedades consideradas fundamentais. De forma resumida, entendemos que o pressuposto da objetividade unifica as vertentes I e II no que concerne à natureza da informação. Além disso, ambas as vertentes relacionam informação a ordem/desordem. Contudo, enquanto os proponentes do realismo informacional admitem uma relação direta entre informação e ordem, tanto os teóricos da MTC quanto da Semântica Informacional atribuem à desordem, ao ruído e ao erro um papel importante na caracterização da informação. Os teóricos da Ecologia Informacional, por sua vez, caracterizam a informação em termos de ordem/desordem, centrando-se na dinâmica relacional organismo/meio ambiente na constituição direta (não mediada por representações) de informação intrinsecamente significativa.

Uma diferença central entre as vertentes I e II é que a primeira emprega o termo informação dissociado do seu uso cotidiano enquanto a vertente II emprega esse termo aproximando-o da noção intuitiva, corrente, utilizada pelo senso comum. O esquecimento desta diferença tem produzido inúmeros mal entendidos nos estudos interdisciplinares da cognição que fazem referência ao conceito de informação.

No que se segue, procuraremos indicar como essa taxonomia da informação pode nos auxiliar na investigação de problemas tradicionais sobre a natureza do conhecimento em pauta na Filosofia da Mente contemporânea. Em especial, discutiremos o problema relativo à natureza proposicional/não proposicional do conhecimento.

### **Informação, Conhecimento e Auto-Organização**

Se o caráter objetivo da informação é o elemento comum às várias teorias da informação, como procuramos indicar nas seções anteriores, por que o tratamento do conceito de informação ocupa um lugar tão inexpressivo e fragmentado na ciência e na filosofia? Por que o conceito de informação começa

a fazer parte do vocabulário da Física apenas no século XXI com as pesquisas de ponta? Quais as dificuldades envolvidas no emprego desse conceito na descrição do mundo físico, que há tempos emprega conceitos tão complicados como o de matéria, espaço, massa, energia, entre outros? Ou, alternativamente, quais são as razões que dificultam o emprego do conceito de informação na Filosofia, e nas Ciências Humanas em geral, para explicitar questões referentes à natureza do conhecimento, das crenças e da ação?

Uma resposta para algumas das questões acima formuladas é sugerida por von Baeyer para quem a dificuldade central em aplicar o conceito de informação na Física se resume no fato de que ele é vago e mal definido. Von Baeyer (2003, p. 9-10) reconhece a importância da informação: “Como seres humanos, não apenas adquirimos informação através dos sentidos, mas também nos sentimos compelidos a compartilhá-la entre nós”. Extrapolando os limites do estritamente humano, ele argumenta, como Stonier, que a informação constitui não apenas um ingrediente essencial do mundo que nos cerca, mas o tecido fundamental do universo.

Uma vez que von Baeyer acredita que, no futuro, os cientistas aprenderão a explicar os fenômenos físicos em termos da linguagem da informação, sua principal tarefa na obra *Information* consiste na tentativa de explicitar o conceito de informação, tornando-o menos vago, no domínio da Física. Um objetivo semelhante tem direcionado parte significativa de nossa pesquisa, com a diferença que nossa atenção está voltada nem tanto para a Física mas, principalmente, para a compreensão da relação entre informação, ação e conhecimento.

A natureza do conhecimento tem sido objeto de discussão na filosofia por séculos, remontando pelo menos aos gregos antigos. Tal discussão revela uma série de problemas, a maioria deles diz respeito à possibilidade de fundamentação do conhecimento. Esse problema enfrenta, por sua vez, a dificuldade central de se obter os princípios a partir dos quais o conhecimento pode ser racionalmente justificado, sem recair em um processo de regressão infinita ou recorrer, circularmente, ao conceito de conhecimento cuja natureza se procura explicitar. Conforme mencionamos na Parte II, uma tentativa recente de compreensão dessa dificuldade foi proposta por Dretske, que através de um estudo cuidadoso do conceito de informação nos indica uma possível saída para

a aparente circularidade presente nos processos de justificação racional do conhecimento.

Grosso modo, a proposta de Dretske (1981) envolve duas etapas. Na primeira, apoiando-se na MTC, ele fornece uma caracterização naturalista da informação, tal como esboçada na Parte II: O mundo físico é considerado como estando repleto de relações legiformes que se estabelecem entre elementos, padrões ou estruturas de naturezas diversas. A informação, na sua forma mais básica, é justamente aquele indicador de relações que pode ser objetivamente registrado por um receptor (humano ou não) que esteja devidamente equipado para tal.

Contudo, como ressaltamos na Parte II, Dretske argumenta que existe uma dimensão da informação, concernente ao seu aspecto significativo, que envolve uma postura intencional por parte do sujeito que a manipula. Uma das propriedades fundamentais da postura intencional, característica das crenças, é, segundo Dretske (1995, p. 28), sua propensão ao erro: crenças têm o poder de “dizer” ou “significar” que K é F quando K não é F. Na verdade, elas têm esse poder mesmo quando K não existe.

Longe de ser visto como uma deficiência do sistema que processa informação, o erro é entendido como um dos elementos chave para se compreender o significado presente na informação; a sua existência é fundamental para se distinguir o comportamento inteligente, que envolve aprendizagem, do comportamento instintivo, que supostamente não envolve reflexão. Por meio do erro, e de sua potencialidade para a correção, é possível se discernir as condições favoráveis das desfavoráveis no comportamento. É justamente a partir desse discernimento que se instaura, segundo Dretske, o processo de geração de informação significativa que, por sua vez, dará origem ao conhecimento.

Podemos resumir a sugestão proposta por Dretske (1981, 1995) quanto ao problema da caracterização do conhecimento, fundado em informação significativa, da seguinte maneira: se um organismo tiver acesso a uma fonte de informação e, além da experiência sensorial (que lhe permite captar informação analógica), ele tiver o poder de discernir as condições favoráveis das desfavoráveis ao seu comportamento, então, indicadores informacionais significativos se estabelecerão, os quais servirão como instrumento de aprendizagem e controle do comportamento em questão. Uma vez aprendidos,

esses indicadores instauram o significado no universo informacional e fornecem subsídios para as crenças verdadeiras a respeito do mundo. De posse de tais indicadores, um organismo teria acesso ao conhecimento sempre que:

- (a) possuir crenças sustentadas por informação (**CRI**) e,
- (b) As **CRI** se mostrarem relevantes para a otimização de seu processo de aprendizagem.

A partir de (a) e (b), Dretske caracteriza o conhecimento (perceptual) como um estado do sistema que possui um conteúdo informacional correspondente à informação que o produziu.

Como o erro é naturalmente possível, o organismo possuidor de conhecimento teria que, além de aprender a eliminar as crenças que não se fundam em informação, dispor de critérios de relevância que lhe auxiliem na seleção das crenças verdadeiras relevantes (fundadas em informação). Mas, nesse caso, quais seriam os critérios de relevância adotados para realizar essa tarefa? Dretske nos remete à noção de contexto e à mencionada exigência da probabilidade condicional = 1, que expressa uma correlação nômica entre eventos no meio ambiente, e às condições estruturais do organismo, imerso em um contexto, para a solução dessa dificuldade. Contudo, entendemos que sua teoria, explicitamente representacional, nos deixa com a semente do problema clássico da justificação da crença verdadeira que, inevitavelmente, se perpetuará enquanto buscarmos no mundo contingente em que vivemos regularidades cuja probabilidade condicional de ocorrência seja = 1.

A nossa sugestão é que abandonemos uma tal busca de verdades na fundamentação do conhecimento, adotando uma postura informacional perspectivista que, sem cair no relativismo, possa estar comprometida com a ação, a preservação do meio ambiente e da vida na sua dinâmica criadora. Não entraremos em detalhes aqui sobre o difícil problema de diferenciar o perspectivismo do relativismo (esse tópico foi parcialmente analisado em GONZALEZ, 1998, 2004). Para concluir, propomos uma caracterização do conceito de informação que julgamos reunir as principais propriedades enfocadas nas teorias contemporâneas da informação, argumentando que essa caracterização pode nos auxiliar na compreensão da natureza do conhecimento comum.



Ao tomar como ponto de partida para a análise do conhecimento comum as relações que se estabelecem entre os agentes e o mundo em que se situam e com o qual interagem, queremos ir além daquilo que foi proposto por Dretske. Para isso, concebemos um programa de pesquisa apoiados em duas bases teóricas. A primeira é fornecida pelas hipóteses da Teoria da Auto-Organização (TAO) formuladas por Debrun (1996), Haken (1983), Haken & Wunderlin (1990) e Gonzalez (1996), que oferece subsídios para a compreensão dos processos de formação de padrões espontâneos sem a intervenção de um centro controlador. A segunda base é constituída pelas hipóteses da Teoria Ecológica da Informação apresentada por Gibson (1979), a qual está essencialmente relacionada à história das interações espontâneas que se estabelecem entre o organismo e o meio ambiente. De acordo com este programa de pesquisa, o significado da informação presente no conhecimento comum seria, em última instância, fundado em relações auto-organizadas que se estabelecem entre o organismo e o meio ambiente. Tais relações, quando eficazes no plano da ação ao longo da história dos organismos, atestariam sua própria relevância e significado.

Uma hipótese a ser investigada, conforme observa Gibson, é que o conhecimento de um sistema está em grande parte (senão inteiramente) conectado a suas possibilidades de ação (i.e., aquilo que um organismo normalmente realizaria em seu meio ambiente).

Ainda que, em alguns casos, o conhecimento possa estar desvinculado de sua efetivação comportamental, mesmo assim podemos concebê-lo em termos de contrafactuais ou de estados disposicionais.

A fim de estender a interpretação gibsoniana da percepção de modo a englobar a concepção de conhecimento comum, sugerimos a investigação das seguintes duas hipóteses:

1) O conhecimento comum pode ser visto como a criação, preservação e quebra de hábitos baseados na troca contínua de informação entre o organismo e seu meio. A informação, nesse caso, pode ser entendida como *um processo auto-organizado que permite o estabelecimento de padrões de ação para organismos situados em ambientes governados por relações compartilhadas de ordem.*



2) O comportamento comum pode ser concebido como uma contínua dinâmica que alterna modos efetivos de agir (que se tornam habituais) com as possibilidades de ação disponíveis no mundo (*affordances*).

De acordo com as sugestões esboçadas acima, a aquisição de conhecimento é vista em termos de formação, manutenção e quebra de hábitos de organismos situados em um ambiente repleto de informação potencialmente significativa. Esse programa de pesquisa, ainda incipiente, tem como fio condutor a análise da trama informacional constitutiva do conhecimento comum. Julgamos que as notas aqui apresentadas para uma taxonomia da informação podem vir a nos auxiliar na difícil compreensão do intrincado processo informacional que caracteriza o conhecimento comum, algumas vezes subestimado pelas abordagens mecanicistas da mente.

## Referências

- AGOSTINHO. De ideis (Quaestio XLVI). In: *De Diversis quaestionibus octoginta tribus liber unus*, Edidit Mutzenbecher. Sancti Augustinini Opera Pars XIII, 2. *Corpus Christianorum*, Series Latina 44 A. Turnhout, Brépols, 1975; p. 70-73. Versão em português retirada de *Cadernos de Trabalho Cepame*, v.2, n.1, p. 5-11, 1993. Tradução de Moacyr Novaes.
- BAEYER, H. C. von. *Information: the new language of science*. London: Weidenfeld & Nicolson, 2003.
- BAR-HILLEL, Y.; CARNAP, R. Semantic information. *British Journal of Science*, v.4, p.147-157, 1953.
- BATESON, G. *Steps to an ecology of mind*. Chicago: Chicago University Press, 1999.
- CAPURRO, R.. Information. Ein Beitrag zur etymologischen und ideengeschichtlichen Begründung des Informationsbegriffs. Munich, Germany: Saur. 1978. <http://www.capurro.de/info.html>
- \_\_\_\_\_. On the Genealogy of Information. In Kornwachs, K., Jacoby & K. (Eds.), *Information. New questions to a multidisciplinary concept*. (pp. 259-270). Berlin, Germany: Akademie Verlag, 1996, p.259-270. <http://www.capurro.de/cottinf.hbm>
- \_\_\_\_\_. Epistemologia e Ciência da Informação. Conferência apresentada no V Encontro Nacional de Pesquisa em Ciência da Informação, Belo Horizonte, 10

de novembro de 2003. Disponível em: [http://www.capurro.de/enancib\\_p.htm](http://www.capurro.de/enancib_p.htm), Acesso em: 08 out. 2004.

DEBRUN, M. M. Auto-organização e ciências cognitivas. In: GONZALEZ, M. E. Q. et al. (Org.). Encontro com as Ciências Cognitivas, 1, 1996, Marília. *Anais...* Marília: Faculdade de Filosofia e Ciência. v.1, 1996 . p.29-38.

\_\_\_\_\_. A idéia de auto-organização. In: DEBRUN, M.; GONZALEZ, M. E. Q.; PESSOA JR., O. (Org.). *Auto-organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: UNICAMP, Centro de Lógica, Epistemologia História da Ciência, 1996, p. 3-24.

\_\_\_\_\_. A dinâmica de auto-organização primária. In: DEBRUN, M. et al. *Auto-organização: estudos interdisciplinares em filosofia, ciências naturais e humanas, e artes*. Campinas: UNICAMP, Centro de Lógica, Epistemologia História da Ciência, 1996. p.25-59.

DRETSKE, F.I. *Knowledge and the flow of the information*. Oxford: Blackwell Publisher, 1981.

\_\_\_\_\_. *Explaining behavior: reasons in a world of causes*. Cambridge: MIT Press, 1992.

\_\_\_\_\_. *Naturalizing the mind*. Cambridge: MIT Press, 1995.

FARIA, E. *Gramática superior da língua latina*. Rio de Janeiro: Livraria Acadêmica, 1958.

GIBSON, J. J. *The senses considered as perceptual systems*. Boston: Houghton Mifflin, 1966.

\_\_\_\_\_. *The ecological approach to visual perception*. Boston: Houghton Mifflin, 1979.

GONZALEZ, M.E.Q. Ação, causalidade e ruído nas redes neurais auto-organizadas. In: DEBRUN, M.; GONZALEZ, M.E.Q.; PESSOA JR., O. (Ed.) *Auto-organização: estudos interdisciplinares*. Campinas: UNICAMP, 1996. (Coleção CLE. v. 18).

\_\_\_\_\_. Auto-organização e perspectivismo: algum acréscimo à Ciência Cognitiva? In: GONZALEZ, M. E. Q. BROENS, M.C. (Org.). *Encontro com as ciências cognitivas*. Marília: Unesp-Marília-Publicações, 1998. p. 3-14.

\_\_\_\_\_. *Informação e conhecimento comum: uma análise sistêmica dos processos criativos auto-organizados*. 2004. Tese (Livre Docência), Faculdade de Filosofia e Ciências – Universidade Estadual Paulista, 2004.

HAKEN, H. *Synergetics*. Berlin: Springer Verlag, 1983.

\_\_\_\_\_. WUNDERLIN, A. Synergetics and its paradigm of self-organization in biological systems. In: WHITING, H.T.A. et al. (Ed.) *The natural-physical approach to movement control*. Amsterdam: VU University Press, 1990.

HARTLEY, R. V.L.L. Transmission of information. *Bell System Tech. J.*, v.7, p.535-63, 1928.

HASELAGER, W.F.G. *Cognitive science and folk psychology: the right frame of mind*. London: Sage, 1997.

\_\_\_\_\_. GONZALEZ, M.E.Q. Causalidade circular e causação mental: uma saída para a oposição internalismo versus externalismo? *Manuscrito*, v.25, n.1, p. 217-238, 2002.

PEREIRA JR, A.; GONZALEZ, M.E.Q. Informação, auto-organização e linguagem. In: ÉVORA, F. *Espaço e tempo*. Campinas: Unicamp, 1995. p. 255-290. (Coleção CLE, v.15).

SCHAEFFER, R. Informação e naturalismo esclarecido: o realismo informacional. In: GONZALEZ, M. E. Q.; DEL-MASSO, M. C. S.; PIQUEIRA, J. R. C. (Org.). *Encontro com as ciências cognitivas*. Marília: UNESP-Marília-Publicações; São Paulo: Cultura Acadêmica, 2001. v. 3, p. 281-298.

\_\_\_\_\_. Da identidade biológica à identidade pessoal. In: BROENS, M.C; MILIDONI, C.B. *Sujeito e identidade pessoal: estudos de filosofia da mente*. Marília: Unesp-Marília-Publicações, São Paulo: Cultura Acadêmica, 2004. p. 57-94.

MILIDONI, C. B., GONZALEZ, M. E. Q., BROENS, M. C. Informação e percepção no modelo de mente do “projeto” freudiano. In: SOUZA, G. M.; D’OTTAVIANO, I. L. GONZALEZ, M.E.Q. (Ed.) *Auto-organização: estudos interdisciplinares*. Campinas: Unicamp, 2004. (Coleção CLE. v. 38).

SHANNON, C.; WEAVER, W. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1998. (primeira edição: 1949).

STONIER, T. *Information and the internal structure of the universe*. Londres: Springer-Verlag, 1990.

\_\_\_\_\_. *Information and meaning: an evolutionary perspective*. London: Spring, 1997.

SZILARD, L. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Zeits. Physik*, v.53, p.840-856. English translation: On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, v.9, p.301-310, 1964. Both versions are reprinted in *The Collected Works of Leo Szilard:Scientific*

*Papers*, Bernard T. Feld and Gertrud Weiss Szilard (Ed.). Cambridge, MA: MIT Press, 1972.

WHITEHEAD, A. N. *The concept of nature*. Michigan: University of Michigan Press, 1957.

WIENER, N. *Cybernetics*. Cambridge, MA: MIT Press, 1961.

\_\_\_\_\_. *The human use of human beings: cybernetics and society*. London: Sphere Books LTD, 1968.

ZEIMAN, J. Significado filosófico da noção de informação. In: ROYAUMONT, C. *O conceito de informação na ciência contemporânea*. Rio de Janeiro: Paz e Terra, 1970.



# Aspectos da teoria de controle na cibernética

Ademar Ferreira<sup>1</sup>

## Introdução

A frase que melhor resume o pensamento de Norbert Wiener relativamente às ambições científicas de um grupo de cientistas nos Estados Unidos em toda uma década (1943 – 1953) talvez seja:

Assim, há pelo menos quatro anos, o grupo de cientistas à volta do Dr. Rosenblueth e de mim mesmo já estava cômico da unidade essencial do conjunto de problemas centrados na comunicação, no controle e na mecânica estatística, seja na máquina ou no tecido vivo (WIENER, 1970).

A motivação deste trabalho é a tentativa de explicitar qual teria sido o papel dos conceitos e da teoria de controle na concepção da unidade dos problemas a que se refere Wiener, não só naquela década, mas também nos desdobramentos mais recentes de sua influência.

A julgar pelo nome escolhido, *cibernética*, para designar o domínio que se pretendia comum aos três campos técnicos e científicos mencionados, deveríamos concluir pela influência marcante do controle. De fato, o termo cibernética vem do vocábulo grego para timoneiro ou piloto (WIENER, 1970), e tratava-se de um reconhecimento ao trabalho de James Clerk Maxwell com o artigo *On Governors* (1868), que marca o início da teoria de controle automático.

Entretanto, se considerarmos os rumos tomados pelo movimento, originando o Cognitivismo e a Inteligência Artificial já na década seguinte, seríamos forçados a reconhecer que a influência da teoria de controle, enquanto pertencente a Sistemas Dinâmicos, foi pequena. Mais recentemente, porém, como que entrando pela porta dos fundos do que sobrou da cibernética, ressurge um papel, agora considerável, dessas disciplinas, na ciência cognitiva e na concepção de sistemas inteligentes artificiais, conforme apontaremos.

Este trabalho explora aspectos da trajetória das disciplinas de controle e sistemas dinâmicos, na construção de idéias da primeira cibernética, e de alguns

---

<sup>1</sup> Professor Associado, Escola Politécnica da USP.



dos seus desdobramentos mais recentes, que estão alterando definitivamente nossas concepções do sistema nervoso central, da mente e do próprio fenômeno da vida. Assim, pretendemos analisar, ainda que resumidamente, algumas das possibilidades que se apresentaram aos conceitos e teoria de controle, de representar um papel mais significativo no projeto de edificar uma ciência da mente, tal como poderia ter sido idealizado pelos primeiros cibernéticos. Estaremos também colocando em perspectiva algumas idéias e temas que, num segundo momento, contribuíram para o desenvolvimento da neurociência e da inteligência artificial. O trabalho não tem por objetivo um estudo crítico aprofundado da cibernética, que pode ser encontrado em Dupuy (1996).

### **Um pouco da prática e teoria de controle**

Realimentação é uma noção central nas construções teóricas de Wiener para a cibernética, e por isso é interessante esclarecer a sua origem e evolução. Trata-se de um conceito conhecido há muito tempo, em economia política, representando a idéia de auto-regulação (BENNETT, 1979). A palavra realimentação (feedback), entretanto, foi usada pela primeira vez somente em 1920, por engenheiros de comunicação. Na mesma época, W. B. Cannon utilizava o termo homeostase para designar mecanismos regulatórios no corpo.

Para introduzir o conceito de controle automático, que é inseparável da idéia de realimentação, vamos falar de dois inventos de James Watt, um construtor de equipamentos de Glasgow. O primeiro, em 1765, foi a introdução, na máquina a vapor, de um condensador separado, evitando o resfriamento do cilindro a cada ciclo, e a conseqüente perda de calor latente. Esta invenção permitia aumentar o rendimento de conversão da energia térmica em energia mecânica, economizando  $\frac{1}{4}$  do combustível antes utilizado, tornando a nova máquina o verdadeiro motor da Revolução Industrial. Watt ainda introduziu diversos melhoramentos no equipamento, como a obtenção do movimento circular. Mas foi a adaptação do regulador centrífugo, para manter constante a velocidade do eixo de rotação, o invento cujo significado mudaria definitivamente a forma de utilização das máquinas em geral, qualquer que fosse a sua natureza. É preciso, entretanto, considerar que as duas invenções aqui enfatizadas são de naturezas muito diferentes. Na primeira, a melhoria do funcionamento é obtida por modificação da *estrutura* da máquina a vapor. Já a

segunda, aplicada de forma “externa” ao equipamento, é obtida pelo uso de *informação* do processo representado pela máquina (no caso, a conversão da energia) para tornar sua operação mais próxima de um padrão desejado, aqui, velocidade de rotação constante, independentemente da carga acoplada ao eixo. Trata-se da primeira aplicação nas máquinas do princípio de *realimentação* e, para nossos propósitos, vamos considerá-la como origem prática dos chamados sistemas de controle automático. Uma descrição simples do funcionamento do regulador encontra-se em Franklin et al. (1994).

Generalizando os preceitos acima, podemos dizer que um sistema de controle tem por objetivo manter constante uma ou mais variáveis de um processo, ou então fazê-las variar de acordo com um padrão desejado. Na descrição de tais sistemas, distinguem-se o processo a ser controlado, também chamado de ‘planta’, e o controlador, cuja função é produzir o sinal ou *ação de controle*, que, aplicada à planta, conduzirá a variável controlada ao padrão desejado. A saída do conjunto controlador-planta é realimentada para a entrada, onde é comparada com o sinal de referência ou padrão de comportamento, produzindo o sinal de erro, que por sua vez, será a entrada do controlador.

O conjunto da máquina a vapor mais o regulador (ou controlador) de Watt satisfaz a conceituação de sistema de controle, e dá conta de uma condição de *desempenho*, então considerada satisfatória. De fato, durante os próximos cem anos, o sistema da máquina a vapor controlada apresentou comportamento aceitável. No século seguinte, entretanto, o matemático e astrônomo G. B. Airy descobriu que tal sistema poderia apresentar *instabilidade* de operação, e iniciou o estudo do controle à realimentação por meio de equações diferenciais. Tais estudos tornaram-se mais sistemáticos com o trabalho de James Clerk Maxwell, outro escossês, de Edinburgh, que então escrevia o seu *Tratado sobre Eletricidade e Magnetismo*. Maxwell, cujo interesse no controlador nada tinha a ver com a sua utilidade prática (MAYR, 1971) apresentou um modelo de equações diferenciais do regulador de Watt, no artigo “On Governors”, publicado em 1868, e obteve uma condição de estabilidade, em função dos coeficientes da ‘equação característica’ correspondente à equação diferencial linearizada. Assim, um século após a solução prática da questão do desempenho na máquina a vapor, iniciava-se o estudo teórico da noção de *estabilidade*, completando-se o reconhecimento do par de conceitos, cuja consideração haveria de marcar permanentemente o estudo dos sistemas de controle. Por outro

lado, ficou também claro que tal disciplina deveria revestir-se de caráter matemático, além do aspecto aplicativo.

A partir dos anos 1930, o princípio da realimentação foi aplicado, por meio de grandezas elétricas, aos amplificadores de telefonia de longa distância, para compensação de perdas de transmissão. A teoria do amplificador à realimentação negativa foi descrita por H. S. Black em artigos publicados em 1934 (BENNETT, 1993), mas a inovação exigiu vários anos para ser aceita plenamente.

A facilidade de manipulação dos sinais elétricos e o desenvolvimento de transdutores, que convertem grandezas físicas diversas para elétricas, vão acelerar a aplicação dos sistemas de controle aos diversos segmentos industriais e proporcionar um grande avanço da teoria correspondente. Nesse período, os sinais elétricos e os computadores utilizados eram analógicos, isto é, baseados em grandezas de variação contínua em intensidade e no tempo. A realimentação era aplicada através de um único “laço” ou “malha” de controle, e os sistemas eram do tipo “SISO” (single-input, single-output). Diversas técnicas de análise assistiam os engenheiros em seus projetos de controladores. Esse arsenal de teoria e prática, conhecido por “controle clássico” é o “estado-da-arte” na época da primeira cibernética.

### **A influência de Wiener na cibernética**

Trabalhando em matemática aplicada no MIT, Norbert Wiener teve ocasião de estudar diversos problemas de controle automático, relacionados ao esforço de guerra. Uma dessas questões era sobre o disparo de projéteis de modo a atingir uma aeronave em voo. A solução (WIENER, 1970) corresponderia a prever uma trajetória a partir de dados sobre posições passadas, tarefa a ser realizada por um dispositivo eletromecânico. Wiener vai dizer-se envolvido no estudo de um sistema “destinado a usurpar uma função especificamente humana”, a de previsão do futuro. É também preocupação sua o desempenho do operador humano, a ponto de sugerir a incorporação, na máquina, de um modelo matemático das características do operador. Em casos como esses, na máquina e no homem, trata-se do desempenho de uma atividade voluntária, isto é, destinada a realizar um propósito. Wiener, e seu colaborador H. S. Bigelow, associam a realização de atividade voluntária do sistema nervoso central a



processos circulares, nos quais intervém o efeito de realimentação negativa, com base na diferença entre o realizado e o propósito, como forma de assegurar que este seja atingido. Um exemplo citado é a ação de levantar uma lapiseira. Consta-se também, por confirmação com o Dr. Rosenblueth, que uma realimentação excessiva ocasiona a condição patológica de tremor em pacientes com lesão no cerebelo. Toda essa evidência em torno do conceito de realimentação é divulgada através de um artigo assinado pelos três pesquisadores (ROSENBLUETH et al., 1943). Essa publicação, na realidade, reveste-se de um caráter programático: trata-se de uma ampla proposta de trabalho multidisciplinar sobre o tema que haviam delineado (WIENER, 1970).

Comparada às explicações da neurofisiologia corrente, a nova abordagem representa uma inovação. Antes, a resposta do sistema nervoso central (SNC) sobre os músculos era diretamente ligada ao estímulo dos sentidos, ou seja, em conformidade com as hipóteses do behaviorismo radical. Agora, a relação estímulo-resposta é considerada como ocorrendo através de um processo circular, no qual o estímulo é modificado por uma mensagem realimentada do resultado sobre o ambiente. Trata-se ainda de um ponto de vista externo, mas correspondendo a um behaviorismo atenuado. A originalidade está em que o sistema nervoso central é agora descrito em termos de um todo integrado, e não como simples comando aos músculos a partir de um sinal dos sentidos. O que Wiener faz é encapsular o comportamento voluntário do SNC como um sistema à realimentação, suficiente para descrever sua atuação para um observador externo. Considerada em termos da teoria de controle, esta caracterização corresponderia a uma análise elementar, na qual o controlador, ou seja, o SNC é visto como “caixa-preta”, com uma realimentação negativa para a entrada, fechando o laço.

Em Cibernética (WIENER, 1970), são discutidos exemplos mais complexos de controle de máquinas, e também do corpo humano (p. 144). A título de curiosidade, e para ilustrar uma percepção antecipatória de Wiener, apresentamos um caso que exemplifica a melhor maneira para se dirigir um automóvel numa estrada coberta de gelo. Para controlar o carro neste caso, deve-se “imprimir ao volante uma sucessão de pequenos e rápidos impulsos”, não suficientes para derrapar, mas adequados para “informar ao nosso sentido cinestésico se o carro está em perigo de derrapagem”. Ora, isto dir-se-ia que é uma antecipação, em 1948, da noção de *controle dual*, introduzida por Feldbaum

em 1960, segundo a qual, em situações de incerteza, o sinal de controle deve servir ao duplo papel de identificar e controlar (FELDBAUM, 1973). Wiener, entretanto, apenas salienta o aspecto da realimentação e classifica este esquema de “controle por realimentação informativa”.

Quando se trata de controlar um sistema artificial, seja uma máquina ou processo industrial, o uso de realimentação possibilita o projeto de controle sem dispor-se de um bom modelo do sistema a controlar. Por outro lado, para projetar a realimentação, é necessário um modelo da planta (ZAMES, 1996). Wiener, no caso da atividade voluntária no homem, está apenas descrevendo o controle à realimentação do sistema nervoso central. Mas o faz de um ponto de vista externo, sob influência behaviorista, ignorando qualquer mentalismo e estrutura funcional neurofisiológica. Deve-se ter em conta que o controlador aqui é o sistema nervoso central e, entender o seu funcionamento, significaria conhecer a dinâmica dos processos neuronais envolvidos.

Ora, a primeira cibernética não demonstrou ensejo de investir nessa direção, embora fossem disponíveis alguns recursos teóricos que poderiam subsidiar algumas tentativas de interpretações possíveis. Quanto ao paralelismo entre os tratamentos dados à máquina e ao vivente relativo à categoria de finalidade, ou teleologia, Dupuy (1996) o classifica como um exemplo de “isomorfismo entre domínios diferentes”, que permite à ciência avançar notavelmente no tocante à extensão.

### **A influência decisiva de McCulloch na cibernética**

Em muitos sentidos, Warren McCulloch, um neurofisiologista, mas também grande conhecedor de filosofia, lógica, psiquiatria, entre outros domínios, foi o mais influente cientista na cibernética. Muitas de suas contribuições contaram com o apoio de Walter Pitts em lógica matemática; os dois escreveram um artigo (McCULLOCH, 1943) que juntamente com o primeiro (ROSENBLUETH et al., 1943) foram decisivos no lançamento da cibernética. Nesse trabalho, intitulado *A logical calculus of ideas immanent in nervous activity*, o sistema nervoso central é representado, em princípio, como uma máquina lógica equivalente à de Turing (DUPUY, 1996), constituída de redes de unidades elementares, os neurônios. Estes são modelizados como *portas lógicas* de tipo “tudo ou nada”, isto é, com apenas dois níveis lógicos,



“verdadeiro” ou “falso”, ou “1” ou “0”, de tal forma que “a atividade de qualquer neurônio pode ser representada como uma proposição”. Além disso, “relações fisiológicas existentes entre atividades nervosas correspondem, naturalmente, a relações entre as proposições” (McCULLOCH, 1989). Tais redes de neurônios, que podem comportar realimentações, foram depois classificadas como autômatos finitos, podendo calcular classes de funções aritméticas pertencentes à classe de funções recursivas, equivalentes à classe de funções calculáveis por máquina de Turing (DUPUY, 1996). De qualquer forma, o modelo de McCulloch e Pitts foi revolucionário, e serviu de inspiração para John von Neumann (von NEUMANN, 1961) conceber e construir a máquina que seria o computador digital de nossos dias.

Deve-se observar que as redes lógicas de McCulloch e Pitts não levam em conta a constituição neurofisiológica intrínseca dos neurônios. Jerome Lettvin, que trabalhou com ambos, assim justifica a consideração: “Em certo sentido há propriedades de tais sistemas de conexões (de neurônios) que são mais ou menos independentes da natureza intrínseca dos elementos não-lineares utilizados, sejam portas lógicas ou neurônio” (prefácio a McCULLOCH, 1988), o que está de acordo com a finalidade computacional prevista para tais redes.

De alguma forma, o modelo de McCulloch e Pitts do sistema nervoso central, relativamente ao de Wiener, apresenta maior nível de detalhamento, uma vez que o considera- constituído por unidades elementares, cada uma com o mesmo comportamento, definindo dois estados. Entretanto, haveria um paralelismo entre os dois modelos. Enquanto o objetivo de Wiener, segundo Papert, na introdução ao livro de McCulloch (1988), seria o de estabelecer os princípios gerais para mecanismos que corporificariam o conceito de *propósito*, o modelo de McCulloch, extraído da funcionalidade neurofisiológica do sistema nervoso central seus mecanismos materiais e lógicos, realiza uma corporificação da própria mente (a partir de DUPUY, 1996).

Assim, os objetivos dos dois artigos poderiam ser vistos como complementares relativamente à mente, o de Wiener estabelecendo princípios para obter comportamentos externos, e o de McCulloch, princípios para obter comportamentos “internos à mente”.

## **Algumas consequências da Cibernética**

O modelo da mente como computação lógica, resultante da primeira cibernética, teve o papel de possibilitar a materialização de uma máquina computacional de Turing, dando origem ao computador digital. Já a ciência da mente que a cibernética originou deixou muito a desejar, por caracterizar a cognição como computação sobre representações simbólicas (VARELA et al., 1991). Seu ambicioso programa de inteligência artificial resultante ainda está por realizar.

Se, do ponto de vista da máquina de calcular, o cálculo proposicional por meio da lógica booleana era o caminho certo, tal não deveria ter sido motivo para desprezar totalmente o modelo analógico dos sistemas dinâmicos. Afinal, o próprio McCulloch, nas palavras de Lettvin, “sabia muito bem que os neurônios individuais não são portas lógicas, mas antes que cada um é um processador analógico complexo de dados analógicos” (McCULLOCH, 1988). A complexidade dos neurônios era conhecida meio de resultados experimentais de neurofisiologia. A obtenção de um modelo representativo do funcionamento neuronal certamente demandaria conhecimentos de sistemas dinâmicos não lineares.

A consideração de tais sistemas, todavia, não se ajustava bem mesmo no caso de controle automático. Tais estudos tiveram origem com Poincaré, no final do século XIX, e prosseguiram no século XX com Van der Pol e outros, com estudos no plano de fase. O trabalho de Lyapunov sobre a estabilidade de sistemas dinâmicos foi publicado em 1893, na Rússia, e traduzido ao francês em 1907, e ao inglês somente em 1949, com reprodução em (LYAPUNOV, 1992), e só mais recentemente começou a ser usado em controle. Os métodos utilizados pelos engenheiros de controle privilegiavam o estudo no *domínio de frequência*, para o qual tem sido desenvolvidas várias técnicas lineares de análise e projeto. O estudo das equações não lineares no *domínio do tempo* deveria esperar os anos 1960, com a introdução dos métodos de espaço de estados.

Apesar de toda essa dificuldade, Hodgkin e Huxley, trabalhando na contra-mão da cibernética, obtiveram, em 1952, as equações dos canais iônicos em neurônios, dando origem à neurociência computacional (HODGKIN, 1952).

Ao ler-se o prefácio da segunda edição de Cibernética, muito se esclarece sobre a frase de Wiener que inicia o presente texto. Wiener aí

reconhece que o estágio da teoria de controle no advento da cibernética, por tratar basicamente de sistemas lineares invariantes no tempo, era inadequado para representar os complexos fenômenos que se verificavam, quer na biologia, quer nas máquinas. Wiener chega a comparar a extensão da teoria linear aos problemas mais complexos àquela dos últimos estágios do sistema ptolomaico da astronomia, em que a prática da correção sobre correção não conseguiu evitar seu desabamento. Para o estudo de sistemas não lineares, naturais ou artificiais, uma nova abordagem se tornava necessária, e Wiener propõe iniciar tal estudo com o seu livro *Nonlinear Problems in Random Theory* (WIENER, 1958). Mas a grande complexidade de sua teoria, de tipo entrada-saída de representação geral multivariável para sistemas não lineares, exigia quantidades astronômicas de adições e multiplicações, o que inviabilizava o seu uso prático (ZAMES, 1996). Devido a incertezas nos modelos, recomendava-se, no lugar de representar com generalidade e calcular com exatidão, a análise de aspectos mais qualitativos dos sistemas de controle, tais como robustez das soluções, estabilidade, existência de oscilações (ZAMES, 1996).

Assim, por constatação própria, e a partir das confissões de Wiener, chegamos à conclusão de que as expectativas dos primeiros cibernéticos, quanto às possibilidades da teoria de controle disponível em 1943, para construir uma ciência da mente, eram um pouco exageradas e um tanto ingênuas. Por outro lado, sendo o trabalho proposto por Wiener, Bigelow e Rosenblueth de caráter programático, não se verificou disposição posterior dos cibernéticos em pelo menos seguir uma linha paralela de pesquisas, privilegiando o estudo do neurônio como sistema analógico. Poderiam ter retirado inspiração da história da teoria de controle, construída com intenso trabalho diligente e, às vezes, de gênio, características que também não faltavam entre os cibernéticos. Mas os esforços desses pesquisadores deram origem ao computador digital, o que não é pouco, visto que este possibilitou o próprio desenvolvimento do connexionismo.

Ultimamente, a teoria de controle e sistemas dinâmicos vem representando importante contribuição nos campos da neurociência, ciência cognitiva e robótica móvel, realizando, ainda que muito depois, as expectativas de Norbert Wiener.

## Referências

- BENNETT, S. *A history of control engineering 1800-1930*. Stevenage: Peter Peregrinus Ltd., 1979.
- \_\_\_\_\_. *A history of control engineering 1930-1955*. Stevenage: Peter Peregrinus Ltd., 1993.
- DUPUY, J. P. *Nas origens das ciências cognitivas*. Traduzido por R. L. Ferreira. São Paulo: Edunesp, 1996.
- FELDBAUM, A. *Principes théoriques des systèmes asservis optimaux*. Traduzido por V. Polonski. Moscou: Editions Mir, 1973.
- FRANKLIN, G. F.; POWELL, G. D.; EMAMI-NAEINI, A. *Feedback control of dynamic systems*. Reading – MA: Addison-Wesley, 1994.
- HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, v.117, p.500-544, 1952.
- LYAPUNOV, A. M. *The general problems of the stability of motion*. Traduzido por A. T. Fuller. London: Taylor & Francis, 1992.
- MAXWELL, J. C. On governors. *Proc. Roy. Soc.*, v. 16, p.249-283, 1868.
- MAYR, O. Maxwell and the origins of cybernetics. *Isis*, v. 62, part. 4, n. 214, p. 425-444, 1971.
- McCULLOCH, W.; PITTS, W. A logical calculus of the Ideas Immanent in nervous activity. *Bulletin of the Mathematical Biophysics*, v.5, 1943.
- McCULLOCH, W. *Embodiments of mind*. Cambridge – MA: The MIT Press, 1988.
- ROSENBLUETH, A.; WIENER, N.; BIGELOW, J. Behavior, Purpose and Teleology. *Philosophy of Science*, v. 10, n. 1, 1943.
- VARELA, F. J.; THOMPSON, E.; ROSCH, E. *The embodied mind: cognitive science and human experience*. Cambridge — MA: The MIT Press, 1991.
- VON NEUMANN, J. The general and logical theory of automata. In: TAUB, A. H. (Ed). *John von Neumann collected works*. Oxford: Pergamon Press, 1961.
- WIENER, N. *Cibernética*. Traduzido por G. K. Ghinzberg. São Paulo: USP; Polígono, 1970. Edição original 1948.

WIENER, N. *Nonlinear problems in random theory*. New York: The Technology Press of MIT; John Wiley & Sons, 1958.

ZAMES, G. Input-output feedback stability and robustness, 1959-85. *Control Systems*, v. 16, n. 3, 1996.





## Índice remissivo

*Acaso*, 169, 171, 172, 174, 176 a 178, 183 a 188, 191.

*Anscombe*, 62, 64, 68, 72.

*Astroglia*, 144, 146, 147.

*Auto-organização*, iii, iv, v, 65, 75, 103, 116, 118, 119, 133, 137, 148, 181 a 183, 192, 193, 210, 216, 218, 219.

*Auto-organização, conceitos (ou definições)*, 121 a 129, 131, 132.

*Auto-organização, H. Von Foerster*, 124, 127, 132, 133.

*Auto-organização, W. R. Ashby*, 123 a 127, 132.

*Auto-organização e informação*, 125 a 131.

*Auto-organização, H. Haken*, 128, a 131, 133, 197, 216, 218.

*Auto-similaridade*, 146, 149.

*Background*, 29, 84, 85, 87.

*Behaviorismo*, 28, 105, a 107, 109, 115, 116, 225.

*Cérebro*, ii, iv, 14, 21, 25, 27, 37, 43, 44, 50, 59, 69, 71, 72, 82, 83, 85, 88, 93 a 95, 97, 98, 135, 137, 139, 143, 145 a 149, 167.

*Cibernética*, vi, 124, 127, 175, 176, 199, 201, 221, 222, 224 a 226, 228, 229, 230.

*Coerência*, 126, 128, 137, 138, 144.

*Cognição*, ii, iii, iv, 34, 36, 42, 44, 47, 87, 88, 106, 107, 109, 115, 116, 127, 132, 135, 138, 145, 148, 151, 167, 187, 192, 210, 215, 228.

*Compatibilidade espacial*, iv, 151, 163.

*Compatibilidade estímulo-resposta*, iv, 153, 157, 158, 162, 163.

*Complementaridade*, 64, 67, 70, 72.

*Computação quântica*, iv, 145 a 147, 149.

*Condições de fronteira*, 41.

*Conexionismo*, 106, 108, 115, 196, 229.

*Consciência*, ii, iii, iv, 6, 8, 12 a 16, 21, 22, 27, 29, 30, 37, 48, 49, 61, 82 a 93, 96 a 102, 122, 136 a 139, 142, 144, 145, 147 a 149, 151, 168, 169, 185 a 189.

*Continuum*, v, 186, 187, 192.

*Co-referencialidade*, 66, 67, 70.

*Córtex*, 95, 119, 136 a 138, 142, 144.

*Densidade pós-sináptica*, 140, 147.

*Desempenho*, 99, 108, 158, 223, 224.

*Domínios de descrição*, 54, 59, 60, 65, 69, 71.

*Efeito Simon*, 156, 157.

*Escalas*, iv, 135, 137, 149.

*Espículas*, 144, 146, 147.

*Estabilidade*, 123, 128, 223, 228, 229.

*Evolução*, iii, 30, 32, 64, 75, 91.

*Externalismo*, i, 5, 31, 83, 84, 86, 87, 219.

*Fiscalismo*, 5 a 11, 13, 15 a 18, 20, 28, 32, 48, 51, 63, 64, 68, 71.

*Hemicampo visual*, 154, 155, 160 a 162.

*Idealismo objetivo*, 188.

*Identidade sem redução*, 60, 71.

*Imagem*, 22, 23, 40, 41, 50 a 54, 58, 66, 71, 102, 153, 159.

*Indeterminismo*, 186.

*Informação*, ii, iii, iv, v, 16, 103, 107, 121, 124, 125, 127 a 131, 135, 136, 138 a 141, 143 a 149, 163, 195 a 220, 223.

*Informação molecular*, 139, 146, 147.

*Informação quântica*, 139, 148.

*Inteligência artificial*, iv, v, 16, 75, 80, 89, 105, 121, 167, 178, 196, 215, 221, 222, 228.

*Intencionalidade*, ii, 6, 13, 23, 26, 31, 61, 83 a 85, 110, 175.

*Internalismo*, i, 81, 86, 87, 219.

*Íon de cálcio*, 137, 143 a 147, 149.

*Laser*, 128, 146, 149.

*Lateralidade da mão*, 157.

*Lei da mente*, v, 174, 181, 183, 185 a 192.

*Máquina de Turing*, 227.

*Materialismo eliminativo*, 8, 9, 25, 26, 50, 59, 71.

*Metaciência*, ii, 42, 43, 50 a 52, 71.

*Metáfora visual*, 44, 54, 58, 63, 72.

*Mito so dado*, 61, 62, 73.

*Modelo de Hooker*, 52, 54, 71.

*Movimento imaginado*, 157.

*Naturalismo biológico*, ii, 14, 16, 22, 29, 32, 33, 81, 82, 86.

*Neurônio*, 44, 49, 109, 137 a 145, 226 a 229.

*NMDA*, 141, 142, 144, 145, 152.

*Perspectivismo*, ii, 75, 213, 218.

*Pluralismo descritivo*, 64, 54, 71.

*Portas lógicas*, 226 a 228.

*Postura da mão*, iv, 153, 158 a 162.

*Postura de projeto*, 60.

*Postura intencional*, 19, 60, 61, 209, 212.

*Postura prona*, 157, 159 a 163.

*Postura supina*, 157, 159 a 163.

*Posturas descritivas*, 59, 62, 68, 69.

*Problema da consciência*, 48.

*Processador analógico*, 228.

*Projeto de pesquisa transdisciplinar*, 47.

*Propriedade emergente*, ii, 14, 47, 50, 68 a 71.

*Realimentação*, 222 a 226.

**Reconhecimento pré-atencional,**  
157

**Redução,** ii, 7, 9, 12 a 15, 18, 24,  
25, 39 a 43, 45, 47, 50 a 55, 58, 60,  
61, 67, 68, 70 a 72, 85, 86, 198,  
200, 201.

**Reduccionismo,** i, ii, 7, 9, 11, 14,  
17, 22, 39, 40, 42, 43, 45, 52, 62,  
66, 67, 69, 72.

**Regras de correspondência,** 41.

**Relação análoga,** 41.

**Relevância dimensional,** 154, 155.

**Representação mental,** i, iii, 87,  
110 a 112, 116, 154, 214, 215.

**Representacionismo,** iii, 65, 105 a  
107, 109, 112, 114.

**Robótica,** 116, 123, 196, 229.

**Rotação mental,** 156, 157.

**Sinapse,** 141, 143, 144.

**Sistema de controle,** 223, 224, 229.

**Sistemas inteligentes,** 221.

**Sobreposição dimensional,** 154, a  
156, 163.

**Suposições limitadoras,** 40, 41, 52,  
53.

**Teleologia,** 226.

**Teleológico,** 171, 176, 185, 188.

**Teoria de controle,** vi, 221, 222, 225,  
229.

**Transdução de sinais,** 135, 141, 142.

**Variação,** 188, 224.







ISBN 857139520-9



9 788571 395206

 **CNPq**  
Conselho Nacional de Desenvolvimento Científico e Tecnológico

 **FAPESP**

*Fundunesp*

Fundação para o Desenvolvimento da USP