

*Anupof*



Pragmatismo, Semiótica,  
Filosofia da Mente e  
Filosofia da Neurociência

*Arthur Araújo*

*Jorge L. Viesenteiner*

*Ludovic Soutif*

*Mariana Cláudia Broens*

*Sofia Inês A. Stein*

*(Orgs.)*



# **Pragmatismo, Semiótica, Filosofia da Mente e Filosofia da Neurociência**

Arthur Araújo

Jorge L. Viesenteiner

Ludovic Soutif

Mariana Cláudia Broens

Sofia Inês A. Stein

(Orgs.)



# **ANPOF - Associação Nacional de Pós-Graduação em Filosofia**

## **Diretoria 2019-2020**

Adriano Correia Silva (UFG)  
Antônio Edmilson Paschoal (UFPR)  
Suzana de Castro (UFRJ)  
Franciele Bete Petry (UFSC)  
Patrícia Del Nero Velasco (UFABC)  
Agnaldo Portugal (UNB)  
Luiz Felipe Sahd (UFC)  
Vilmar Debona (UFMS)  
Jorge Viesenteiner (UFES)  
Eder Soares Santos (UEL)

## **Diretoria 2017-2018**

Adriano Correia Silva (UFG)  
Antônio Edmilson Paschoal (UFPR)  
Suzana de Castro (UFRJ)  
Agnaldo Portugal (UNB)  
Noéli Ramme (UERJ)  
Luiz Felipe Sahd (UFC)  
Cintia Vieira da Silva (UFOP)  
Monica Layola Stival (UFSCAR)  
Jorge Viesenteiner (UFES)  
Eder Soares Santos (UEL)

## **Diretoria 2015-2016**

Marcelo Carvalho (UNIFESP)  
Adriano N. Brito (UNISINOS)  
Alberto Ribeiro Gonçalves de Barros (USP)  
Antônio Carlos dos Santos (UFS)  
André da Silva Porto (UFG)  
Ernani Pinheiro Chaves (UFPA)  
Maria Isabel de Magalhães Papaterra Limongi (UPFR)  
Marcelo Pimenta Marques (UFMG)  
Edgar da Rocha Marques (UERJ)  
Lia Levy (UFRGS)

**Diretoria 2013-2014**

Marcelo Carvalho (UNIFESP)  
Adriano N. Brito (UNISINOS)  
Ethel Rocha (UFRJ)  
Gabriel Pancera (UFMG)  
Hélder Carvalho (UFPI)  
Lia Levy (UFRGS)  
Érico Andrade (UFPE)  
Delamar V. Dutra (UFSC)

**Diretoria 2011-2012**

Vinicius de Figueiredo (UFPR)  
Edgar da Rocha Marques (UFRJ)  
Telma de Souza Birchall (UFMG)  
Bento Prado de Almeida Neto (UFSCAR)  
Maria Aparecida de Paiva Montenegro (UFC)  
Darlei Dall'Agnol (UFSC)  
Daniel Omar Perez (PUC/PR)  
Marcelo de Carvalho (UNIFESP)

**Produção**

Antonio Florentino Neto

**Editor da coleção ANPOF XVIII Encontro**

Jorge Luiz Viesenteiner

**Diagramação e produção gráfica**

Editora Phi

**Capa**

Adriano de Andrade

## **Comitê Científico: Coordenadoras e Coordenadores de GTs e de Programas de Pós-graduação**

Admar Almeida da Costa (UFRRJ)  
Adriano Correia Silva (UFG)  
Affonso Henrique V. da Costa (UFRRJ)  
Agemir Bavaresco (PUCRS)  
Aldo Dinucci (UFS)  
Alessandro B. Duarte (UFRRJ)  
Alessandro Rodrigues Pimenta (UFT)  
Alfredo Storck (UFRGS)  
Amaro de Oliveira Fleck (UFMG)  
Ana Rieger Schmidt (UFRGS)  
André Cressoni (UFG)  
André Leclerc (UnB)  
Antonio Carlos dos Santos (UFS)  
Antonio Edmilson Paschoal (UFPR)  
Antonio Glaudenir Brasil Maia (UVA)  
Araceli Rosich Soares Velloso (UFG)  
Arthur Araújo (UFES)  
Bartolomeu Leite da Silva (UFPB)  
Bento Prado Neto (UFSCAR)  
Breno Ricardo (UFMT)  
Cecilia Cintra C. de Macedo (UNIFESP)  
Celso Braidá (UFSC)  
Cesar Augusto Battisti (UNIOESE)  
Christian Hamm (UFMS)  
Christian Lindberg (UFS)  
Cicero Cunha Bezerra (UFS)  
Clademir Luis Araldi (UFPEL)  
Claudemir Roque Tossato (UNIFESP)  
Claudinei Freitas da Silva (UNIOESTE)  
Cláudio R. C. Leivas (UFPEL)  
Clóvis Brondani (UFFS)  
Cristiane N. Abbud Ayoub (UFABC)  
Cristiano Perius (UEM)  
Cristina Foroni (UFPR)  
Cristina Viana Meireles (UFAL)  
Daniel Omar Perez (UNICAMP)  
Daniel Pansarelli (UFABC)  
Daniel Peres Coutinho (UFBA)

Dirce Eleonora Nigro Solis (UERJ)  
Eder Soares Santos (UEL)  
Eduardo Aníbal Pellejero (UFRN)  
Emanuel Â. da Rocha Fragoso (UECE)  
Enoque Feitosa Sobreira Filho (UFPB)  
Ester M. Dreher Heuser (UNIOESTE)  
Evaldo Becker (UFS)  
Evaldo Sampaio (UnB/Metafísica)  
Fátima Évora (UNICAMP)  
Fernando Meireles M. Henriques (UFAL)  
Filipe Campello (UFPE)  
Flamarion Caldeira Ramos (UFABC)  
Floriano Jonas Cesar (USJT)  
Franciele Bete Petry (UFSC)  
Francisco Valdério (UEMA)  
Georgia Amitrano (UFU)  
Gisele Amaral (UFRN)  
Guido Imaguire (UFRJ)  
Gustavo Silvano Batista (UFPI)  
Helder Buenos A. de Carvalho (UFPI)  
Henrique Cairus (UFRJ)  
Hugo F. de Araújo (UFC)  
Jacira de Freitas (UNIFESP)  
Jadir Antunes (UNIOESTE)  
Jelson Oliveira (PUCPR)  
João Carlos Salles (UFBA)  
Jorge Alberto Molina (UERGS)  
José Lourenço (UFSM)  
Júlia Sichieri Moura (UFSC)  
Juvenal Savian Filho (UNIFESP)  
Leonardo Alves Vieira (UFMG)  
Lívia Guimarães (UFMG)  
Luciano Carlos Utteiche (UNIOESTE)  
Luciano Donizetti (UFJF)  
Ludovic Soutif (PUCRJ)  
Luís César G. Oliva (USP)  
Luiz Carlos Bombassaro (UFRGS)  
Luiz Rohden (UNISINOS)  
Manoel Vasconcellos (UFPEL)  
Marcela F. de Oliveira (PUCRJ)  
Marcelo Esteban Coniglio (UNICAMP)

Márcia Zebina Araújo da Silva (UFG)  
Márcio Custódio (UNICAMP)  
Marco Antonio Azevedo (UNISINOS)  
Marcos H. da Silva Rosa (UERJ)  
Maria Cecília Pedreira de Almeida (UnB)  
Maria Cristina de Távora Sparano (UFPI)  
Maria Cristina Müller (UEL)  
Marina Velasco (UFRJ/PPGLM)  
Mariana Cláudia Broens (UNESP)  
Mariana de Toledo Barbosa (UFF)  
Mário Nogueira de Oliveira (UFOP)  
Mauro Castelo Branco de Moura (UFBA)  
Max R. Vicentini (UEM)  
Michela Bordignon (UFABC)  
Milton Meira do Nascimento (USP)  
Nathalie Bressiani (UFABC)  
Nilo César B. Silva (UFCA)  
Nilo Ribeiro (FAJE)  
Patrícia Coradim Sita (UEM)  
Patrícia Kauark (UFMG)  
Patrick Pessoa (UFF)  
Paulo Afonso de Araújo (UFJF)  
Pedro Duarte de Andrade (PUCRJ)  
Pedro Leão da Costa Neto (UTP)  
Pedro Paulo da Costa Corôa (UFPA)  
Peter Pál Pélbart (PUCSP)  
Rafael de Almeida Padial (UNICAMP)  
Renato Moscateli (UFG)  
Ricardo Bazilio Dalla Vecchia (UFG)  
Ricardo Pereira de Melo (UFMS)  
Roberto Horácio de Sá Pereira (UFRJ)  
Roberto Wu (UFSC)  
Rodrigo Guimarães Nunes (PUCRJ)  
Rodrigo Ribeiro Alves Neto (UNIRIO)  
Samir Haddad (UNIRIO)  
Sandro M. Moura de Sena (UFPE)  
Sertório de A. Silva Neto (UFU)  
Silvana de Souza Ramos (USP)  
Sofia Inês A. Stein (UNISINOS)  
Sônia Campaner (PUCSP)  
Tadeu Verza (UFMG)



Tiegue Vieira Rodrigues (UFSM)  
Viviane M. Pereira (UECE)  
Vivianne de Castilho Moreira (UFPR)  
Waldomiro José da Silva Filho (UFBA)

Dados Internacionais de Catalogação na Publicação (CIP)  
Bibliotecária Juliana Farias Motta CRB7/5880

P898

Pragmatismo, semiótica, filosofia da mente e filosofia da neurociência/  
Organização Arthur Araújo...et al -- São Paulo: ANPOF, 2019.

108 p.

ISBN: 978-85-88072-88-6

Outros autores: Jorge L. Viesenteiner, Ludovic Soutif, Mariana  
Cláudia Broens, Sofia Inês A. Stein  
Filosofia da mente.2.Neurociência.3. Semiótica.4. Pragmatismo. I. Araújo,  
Arthur...et al.II. Título

CDD 128.2

Índice para catálogo sistemático:

1. Filosofia da mente
2. Neurociência
3. Semiótica
4. Pragmatismo



## **Apresentação da Coleção do XVIII Encontro Nacional de Filosofia da ANPOF**

O XVIII Encontro Nacional da ANPOF foi realizado em outubro de 2018 na Universidade Federal do Espírito Santo (UFES), em Vitória/ES, e contou com mais de 2 mil participantes com suas respectivas apresentações de pesquisa, tanto nos Grupos de Trabalho da ANPOF quanto em Sessões Temáticas. Em acréscimo, o evento também incluiu conjuntamente o IV Encontro Nacional ANPOF Ensino Médio, sob coordenação do Prof. Dr. Christian Lindberg (UFS), cujos esforços não apenas amplia, mas também inclui os debates e pesquisas vinculados à área do Ensino de Filosofia tanto de professores vinculados ao Ensino de Filosofia quanto também de professores e estudantes do Mestrado Profissional em Filosofia, o PROF-FILO.

A ANPOF publica desde 2013 os trabalhos apresentados sob a forma de livro, com o intuito não apenas de tornar públicas as pesquisas de estudantes e professores, mas também de fomentar o debate filosófico da área, especialmente por ser uma ocasião de congregar uma significativa presença de colegas do Brasil inteiro, interconectando pesquisas e regiões que nem sempre estão em contato. Assim, a Coleção ANPOF sintetiza o estado da pesquisa filosófica naquele determinado momento, reunindo pesquisas apresentadas em Grupos de Trabalho e Sessões Temáticas. O total de textos submetidos, avaliados e aprovados à publicação na atual Coleção ANPOF do XVIII Encontro conta com mais de 650 artigos da comunidade em geral.

É importante registrar nesta “Apresentação” a dinâmica utilizada no processo de organização dos 22 volumes que são agora publicados, cuja concepção geral consistiu em estruturar o processo da maneira mais amplamente colegiada possível, envolvendo no processo de avaliação dos textos submetidos todas as coordenações dos Grupos de Trabalho e dos Programas de Pós-graduação (PPGs) em Filosofia, bem como uma comissão de avaliação específica para os trabalhos que não foram avaliados por algumas coordenações de PPGs. Em termos práticos, o processo seguiu três etapas: 1. cada pesquisador(a) teve um período para submissão dos seus trabalhos; 2. Período de avaliação, adequação e reavaliação dos textos por parte das coordenações de GTs e PPGs; 3. Editoração dos textos aprovados pelas coordenações de GT e PPGs.

Nessa atual edição da Coleção ANPOF, figuraram na co-organização dos volumes não apenas as coordenações de GTs, mas também de PPGs que estiveram diretamente envolvidos no processo, na medida em que ambas as coordenações realizaram as atividades de avaliação e seleção dos textos desde as inscrições ao evento, até avaliação final dos textos submetidos à publicação, exercendo os mesmos papéis na estruturação das atividades. Nessa medida, a Coleção ANPOF conta com o envolvimento quase integral das coordenações, exprimindo justamente a concepção colegiada na organização – seja diretamente na organização dos volumes, seja sob a forma de comitê científico – de modo que os envolvidos figuram igualmente como co-organizadores(as) da Coleção, cujo ganho é, sem dúvida, em transparência e em engajamento com as atividades. O trabalho de organização da Coleção, portanto, seria impossível sem o envolvimento das coordenações.

Reiteramos nossos os agradecimentos pelos esforços da comunidade acadêmica, tanto no sentido da publicação das pesquisas em filosofia que são realizadas atualmente no Brasil, quanto pela conjugação de esforços para que, apesar do gigantesco trabalho, realizarmos da maneira mais colegiada possível nossas atividades.

Boa leitura!

Diretoria ANPOF

## Sumário

Apresentação do GT Filosofia da Neurociência	12
Dores e prazeres de nossa natureza social <i>Daniel Uptmoor Pauly</i>	14
Neurofilosofia dos sonhos e dos sonhos lúcidos: Contribuições para a filosofia da mente <i>Glesicelly Herminia Ferreira</i>	25
Um exame da plausibilidade científica do libertismo neural <i>Guilherme Teixeira Martins Schettini</i>	34
Frege and the problem of the informativity of logic <i>João Vitor Schmidt</i>	42
O cérebro como agente e a mente consciente como guia de ação <i>Jonas Gonçalves Coelho</i>	51
A memética de Dennett está em harmonia com a sua noção de postura intencional? <i>Juliana de Orione Arraes Fagundes</i>	60
Bratman como atualização de Davidson <i>Lucas Ollyver</i>	69
Democracia cooperativa ou socialismo consciente? A proposta de John Dewey de sociedade democrática <i>Maria A. Lima Piai</i>	80
Causalidade, classes naturais e Interpretantes Finais. Uma abordagem semiótica para a causação final <i>Priscila Borges</i> <i>Juliana Rocha Franco</i>	85
Concepções fregeanas e russellianas da noção de número: Uma comparação <i>Rodrigo Sabadin Ferreira</i>	94
Consciência e fenômenos mentais inconscientes nas visões de David Rosenthal e John Searle <i>Tárik de Athayde Prata</i>	99

## Apresentação do GT Filosofia da Neurociência

### Filosofia e Neurociências

Ingressamos em um século no qual a filosofia divide a cena com múltiplas outras ciências que aparentemente estão investigando problemas conhecidos como filosóficos. Este é o caso das ciências cognitivas e das neurociências, que, enquanto ciências multidisciplinares, também interrogam acerca, por exemplo, da relação entre o corpo e a mente, questão eminentemente filosófica. Não sabemos ao certo como a disciplina da Filosofia ficará lotada em termos acadêmicos dentro das universidades, porém uma possibilidade será participar dessas ciências.

Após pelo menos trinta anos de atividades das ciências cognitivas não estritamente computacionalistas dentro das universidades, percebemos que algumas atividades de pesquisa das ciências cognitivas geram novas questões que, em sua primeira aparência, pertencem à filosofia. Uma dessas questões tem relação com o conceito de representação. As discussões das décadas de 1980 e 1990 sobre a necessidade de pressupormos que a mente seja primordialmente representativa de estados de coisas externos, de que tenha como uma de suas principais tarefas representar a realidade —discussões tanto no âmbito sempre renovado da epistemologia e da filosofia da ciência, porém, também, no espaço das teorias da cognição corporificada, que se estabeleceram em oposição ao computacionalismo ligado à inteligência artificial— resultaram, por caminhos diversos, em novas teorias da mente humana —como a teoria do cérebro bayesiano— extremamente interessantes e frutíferas, que conseguem explicar ações humanas de forma muito fidedigna.

Da corrente antirepresentacionista nas ciências cognitivas participaram pensadores oriundos de muitos departamentos e formações. Entre eles, contamos com o engenheiro Rodney Brooks (2002; 1991), que não apenas liderou a construção de robôs “inteligentes”, que “aprendem” a agir de formas variadas, não-programadas, em circunstâncias diversas. Ao contrário do esperado pelo computacionalismo, o programa desses robôs não é ultra-complexo: é simples. É sua simplicidade que capacita o robô a agir de forma flexível, a alcançar resultados que não estão pré-programados. A ideia fundamental é a de que o importante é reagir a *inputs* da forma que em humanos seria chamada de “quase-reflexa”, que não precisa de muitos cálculos, muitas representações. Disto se deduz que, por analogia, a inteligência e a plasticidade humanas poderiam estar baseadas não na capacidade de representar realisticamente o entorno, mas na capacidade de reagir corretamente a estímulos específicos gerados pelo entorno.

A possibilidade de criar robôs que, por analogia, levem a uma reavaliação das teorias acerca da mente humana é o resultado a) de um longo processo tecnológico; b) do desenvolvimento de máquinas cada vez mais capazes em termos computacionais — capazes de estabelecer trilhões de relações lógicas em poucos milissegundos— e c) de teorias filosóficas, antropológicas, psicológicas e biológicas acerca da inteligência em geral —acerca do que seriam e em que consistiriam as funções lógicas, cognitivas e perceptivas que encontramos em organismos vivos e como essas estariam relacionadas às funções de instrumentos mecânicos e eletrônicos.

Examinando a literatura atual na área das ciências cognitivas e das neurociências sobre a possibilidade de identificação de objetos mentais ou neurais —estados mentais, conceitos e representações—, é possível perceber o quanto o discurso filosófico sobre estados mentais pode auxiliar cientistas a cumprirem a sua tarefa de pesquisa. A filosofia pode ajudar neurocientistas a compreender em que consiste, por exemplo, a *diferença entre o discurso de primeira e terceira pessoa* de um ponto de vista semântico e pragmático.

Outra tarefa eminentemente filosófica a ser cumprida seria pensar, mesmo que de forma incompleta, como seria possível uma *unificação das leis* —que estabelecesse leis tanto para os fenômenos chamados mentais ou psíquicos quanto para os fenômenos chamados de físicos ou materiais—, não apenas para encontrar leis psicofísicas, mas leis que estabelecessem regularidades entre fenômenos hoje considerados inexplicáveis. Isso contribuirá, certamente, para avançarmos na compreensão de fenômenos complexos que as neurociências —incluindo as assim chamadas neurociências sociais—, hoje estão enfrentando.

Desde a década de 1990, temos no Brasil grupos de trabalho em Filosofia da Mente e inclusive uma Sociedade de Ciências Cognitivas. Porém, nos últimos anos, com o incremento da Filosofia Experimental (X-Phil) e de inusitados instrumentos neurocientíficos, novos métodos, questões e conceitos emergiram no âmbito científico e acadêmico que merecem nossa atenção.

Em 2014, liderado por Gabriel Mograbi, e com apoio de pesquisadores, como Adriano Naves de Brito, Alfredo Pereira Jr., Beatriz Sorrentino Marques, Carlos Eduardo Batista de Souza, Carlos Eduardo de Sousa Lyra, Gustavo Leal Toledo, João Fernandes Teixeira, Jonas Gonçalves Coelho, Leonardo Ferreira Almada, Marco Aurélio Souza Alves, Noel Struchiner, Nythamar de Oliveira, Osvaldo Pessoa Jr., Pedro Dolabela Chaves, Renato César Cardoso e Sofia Stein, foi fundado o GT em Filosofia da Neurociência, no XVI Encontro Nacional da ANPOF, em Águas de Lindóia.

No XVII Encontro Nacional da ANPOF (2016), em Aracajú, além da sessão do GT em Filosofia da Neurociência, foi organizado pelo GT o Minicurso Neurofilosofia: Percepção, Valor e Decisão, ministrado por Sofia Stein, Adriano Brito e Gabriel Mograbi. De 2016 a 2018, Sofia Stein assumiu a coordenação do GT e organizou a sessão do GT no XVIII Encontro Nacional da ANPOF, sediado em Vitória. Alguns dos trabalhos selecionados e apresentados naquela ocasião estão publicados neste volume: “Dores e Prazeres de nossa Natureza Social”, de Daniel Uptmoor Pauly, “Um Exame da Plausibilidade Científica do Libertismo Neural”, de Guilherme Teixeira Martins Schettini e “Neurofilosofia dos Sonhos e dos Sonhos Lúcidos: contribuições para a filosofia da mente”, de Glescikelly Hermínia Ferreira. Boa leitura.

Porto Alegre, dezembro de 2019

Sofia Inês Albornoz Stein (CNPq/Unisinos)

siastein@gmail.com

## **Dores e prazeres de nossa natureza social**

Daniel Uptmoor Pauly<sup>1</sup>

### **Relação entre processos físicos localizados com a conexão social e o prazer que relações interpessoais podem trazer**

O desenvolvimento de conexões sociais é um aspecto do cérebro humano que contribui para entender a mente social. Áreas que são destacadas dessa rede são: o dACC - que está relacionado com a sensação desagradável de dor; e o estriado ventral - que está relacionado com o sistema prazeroso de recompensa. Provavelmente essas duas funções são as primeiras adaptações neurais relacionadas com a questão social que surgiram na história evolutiva dos mamíferos. É nosso intuito que a compreensão dos mecanismos biológicos envolvidos no tecido das conexões sociais possa lançar novas luzes acerca da natureza da personalidade e comportamento humano. Essa efervescente área de interesse científico guarda profundas implicações filosóficas, devido às análises que fornecem para a interpretação da mente social. Nesse artigo revisamos a literatura que mostra circuitos neurais que estimulam conexão social e produzem processos psicológicos, tais como a dor da rejeição e o prazer de ter uma boa reputação.

### **Desafios do ambiente selecionaram espécimes com boas estratégias de grupo**

É provável que as habilidades sociais foram fundamentais quando o *Homo sapiens* saiu da África para povoar o Mundo. É importante nos perguntarmos como as habilidades sociais surgiram na nossa espécie, uma dessas habilidades é a dor social.

Ao longo das suas pesquisa com resultados bem documentados e replicáveis, Eisenberger<sup>2</sup> em conjunto com Lieberman<sup>3</sup> coletaram evidências com o cuidado de reduzir as possíveis fontes de erro experimental e desenvolveram uma teoria sobre a mente social. Ambos apostam no potencial de funcional Magnetic Ressonance Image (fMRI) para prever quais regiões do cérebro são ativadas quando o cérebro realiza determinados processos psicológicos. Enfatiza o período de rápida expansão da pesquisa que correlaciona processos psicológicos e regiões neurais pelo uso dessa técnica (Ochsner & Lieberman, 2001). Ele defende que, durante nossa história evolucionária, os circuitos neurais no ser humano promovem a vida em sociedade.

---

1 Doutorando em Filosofia; Universidade do Vale do Rio dos Sinos (UNISINOS)

2 Naomi I. Eisenberger, nascida em São Francisco, é psicóloga, professora e diretora do Laboratório Neurociência Social e Afetiva (<https://sanlab.psych.ucla.edu>). Premiada com o APA prêmio por distintas contribuições na jovem carreira científica para psicologia. É reconhecida por sua pesquisa nas bases neurais da dor e conexão social. Eisenberger e Lieberman são casados.

3 Matthew D. Lieberman formou-se em psicologia e filosofia no Rutgers College. É doutor em psicologia social na Universidade de Harvard. É diretor do laboratório de neurociências cognitivas sociais ([www.scn.ucla.edu](http://www.scn.ucla.edu)). Algumas de suas áreas de interesse são a neurociência cognitiva social e as bases neurais da rejeição social e dor física.

A biologia evolutiva do desenvolvimento compara os processos desenvolvimentais de diferentes organismos para inferir relações de ancestralidade. A ideia por trás dessa abordagem pode ser simplificada da seguinte forma: os embriões de diferentes espécies apresentam similaridades e os estágios pelos quais passam guardam uma espécie de memória do seu passado evolutivo (Richardson e Keuck, 2002, p. 501). O desenvolvimento do ser humano, após o nascimento, repete estágios de amadurecimento na mesma ordem cronológica de mudanças biológicas ocorridas nos ancestrais dos quais descendemos. Não me surpreenderia que essa analogia traçada de forma preliminar possa ser apresentada de maneira mais elegante e satisfatória, com mais e melhores evidências apresentadas.

Nossa mente social têm conexões que asseguram que os seres humanos devem criar vínculos de forma natural e disseminada. Os sistemas que têm a maior contribuição para tal efeito são: a conexão social - dividido em uma sedutora rede de recompensa e uma desagradável rede de dor; a “mentalização” - sobre o que os outros estão pensando; e a “harmonização” - o que permite a influência de crenças e valores do grupo em nossas próprias concepções.

Dois circuitos neurais consolidados através de ajuste evolutivo introduziram funções sociais fundamentais no cérebro de nosso ancestral comum. Isso ocorreu quando os mamíferos, ao se separarem de outros vertebrados, ao longo de gerações, sofreram mutações que os tornaram capazes de sentir dores e prazeres sociais. Essa mudança vinculou profundamente nosso bem-estar com a conexão social. O bebê manifesta essa necessidade básica de se manter conectado. Mesmo com o amadurecimento a vontade de participar da vida em comunidade continua presente. O desejo primeiro de estar com os cuidadores é necessário para garantir sua amamentação e proteção. Esse desejo se modifica com o crescimento (surgem novas relações), mas a necessidade pelo contato social não muda. (LIEBERMAN, 2013, p. 11).

A próxima mudança marcante que foi incrementada no cérebro social da linhagem da qual descendemos ocorreu durante o surgimento da ordem dos primatas. Um conjunto de mutações permitiu o surgimento do potencial de compreender as ações e pensamentos daqueles ao seu redor<sup>4</sup>. Como efeito, houve um aumento considerável das probabilidades estatísticas de cada primata se manter conectado e interagir estrategicamente com outros. Ao criar hipóteses acerca do que se passa na mente de seus pares, em determinados contextos, a criança humana se aventura com maiores chances sucesso no complexo mundo das interações sociais. Essa função mental recebe o nome leitura da mente. é o processo de entender estados psicológicos tais como objetivos e percepções dos outros (HERRMANN et al., 2007, p. 1361), de acordo com as evidências que nossos sentidos fornecem para analisá-los.

Apenas muito recentemente (em termos de história evolutiva biológica) é que ocorreu o último processo de mutação pertinente ao cérebro social. O sentido de self - a introspecção sobre nossa própria personalidade e caráter - começa a tomar forma na pré-adolescência e se solidifica na adolescência, guardando ainda certa flexibilidade nos períodos subsequentes. Segundo as teorias tradicionais de desenvolvimento psicológico é durante esse período de desenvolvimento que acontece a maior parte das mudanças de personalidade<sup>5</sup>. A harmonização garante que um conjunto de valores seja amplamente

4 Entretanto, sem o devido tempo de treinamento no convívio de outros da mesma espécie, durante um período crítico, esse potencial se enfraquece permanentemente.

5 Esse não é o consenso, entretanto. Psicólogos como Helson e Caspi se sustentam em estudos de coorte da história recente e do século XX para identificar que a maior parte das mudanças de personalidade ocorrem na etapa de vida



compartilhado pela comunidade (LIEBERMAN, 2013, p. 201). Na pré-adolescência e adolescência existe uma atenção do jovem voltada para se tornar altamente socializado por aqueles em seus círculos de convívio. Essa função só foi identificada na espécie humana, é a mais recente do ponto evolutivo, e é a última mudança drástica no cérebro social durante o desenvolvimento.

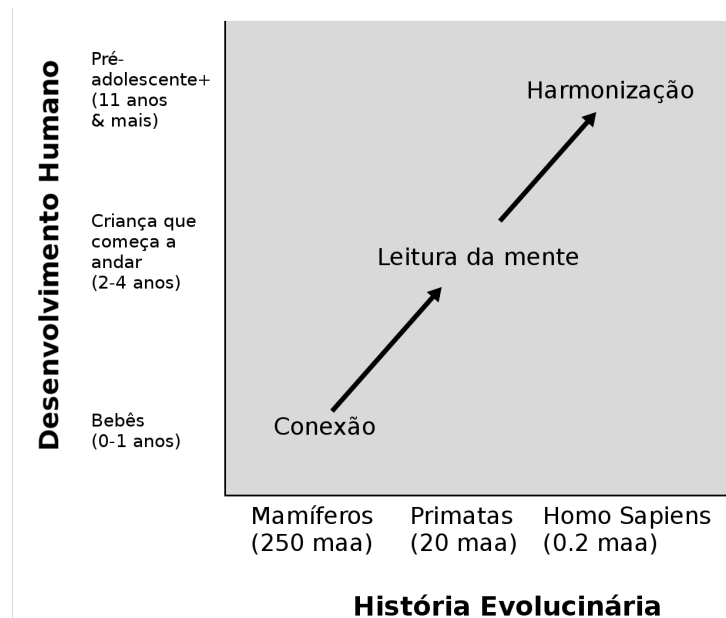


Figura 1. Emergência das adaptações sociais ao longo do desenvolvimento (maa = milhões de anos atrás). (LIEBERMAN, 2013, p. 11, tradução nossa).

Aristóteles chegou a uma conclusão interessante após observar casos particulares<sup>6</sup>: 1) Todos homens observados são sociais; 2) Logo, “O homem é por natureza um animal social.” (ARISTÓTELES, 1992, p. 23). O interessante é que hoje temos uma explicação um pouco mais apurada da razão biológica para sermos animais sociais. O estudo de Lieberman aponta, com uma margem de erro aceitável, que o primeiro circuito neural a receber função de socialização, no nosso ancestral comum, é a conexão social. Ela é responsável pela sensação de dores e prazeres sociais. Localizá-la anatômica (no espaço) e evolutivamente (no tempo) é um passo importante demais para ser ignorado pela filosofia.

### A dor social como alerta para o isolamento.

Eisenberger et al., inspirados no paradigma de Kipling Williams, chamado Cyberball, realizaram um estudo através de fMRI do cérebro durante uma situação de exclusão. É um jogo virtual de troca de passes entre participantes do estudo e outros dois “jogadores” (que foram previamente programados, enquanto o jogador real pensa que os outros dois também são reais). Quando os participantes do estudo param de receber a bola dos seus parceiros, o scanner detecta no cérebro desses pesquisados ativação da mesma região que é ativada pela dor física. Os participantes do estudo relatam sentimentos de desconforto

do jovem adulto.

6 A melhor suposição na qual posso pensar é que o método que o permitiu chegar a sua conclusão tenha sido a indução. O texto não elucida a fonte da qual obteve a informação apresentada, o que não era incomum nos trabalhos acadêmicos da época. Em outra de suas obras, “A Política”, também usa o termo “animal social”, que por vezes é traduzido como “animal político” ou “animal cívico” - essas duas traduções alternativas são mais distantes do sentido no texto original, segundo o tradutor Mário da Gama Kury.

durante os testes. Essa dor é causada pela rejeição social sentida durante o Cyberball, não existe estímulo de dor física algum.

Em suma, um padrão de ativações muito similar ao encontrado em estudo de dor física apareceu durante a exclusão social, fornecendo evidências que a experiência e regulação de dor social e física compartilham uma base neuroanatômica comum. Atividade no dACC, já anteriormente ligado à experiência incômoda da dor, era associada com aumento de desconforto após exclusão social. Além disso, atividade no Córtex Pré-Frontal Ventral direito, anteriormente ligado com regulação do desconforto da dor, se mostrou associado com diminuição do desconforto após exclusão social.<sup>7</sup> (EISENBERGER et al., 2003, p. 291, tradução nossa).

Existe um sistema neural de alarme para dor física e dor social. As sinapses oriundas de diversos sistemas neurais relacionados com dor convergem para o dACC. Ele é responsável pela sensação desagradável da dor física e social (Eisenberger & Lieberman, 2004, pp. 294 e 295). Essa superposição é um fato contraintuitivo. Não é incomum que alguém, ao tomar ciência dela, mostre-se cético e pergunte: afinal, o que eu digo para alguém que foi rejeitado? Indico-lhe um remédio para dor de cabeça? A menção a essa relação estreita entre dor física e dor social põe em dúvida o senso-comum. Entretanto, contrariando os pressentimentos que possam ter suscitado, a tese resistiu ao teste empírico.

Foi feito o seguinte experimento durante 21 dias o primeiro grupo recebeu duas pílulas de 500 mg de acetaminofeno (fórmula ativa do Tylenol com propriedades analgésicas e antipiréticas) depois de acordar e a mesma dose antes de dormir, enquanto o grupo de controle recebeu a mesma quantidade de placebo de mesma aparência - naturalmente, sem que ambos soubessem se estavam tomando o medicamento autêntico, ou não. Concluiu-se que o primeiro grupo apresentou uma diminuição do relato de sentimento de dor social significativa ao longo do tempo, enquanto não houve alteração no grupo de controle. (DeWall et al., 2010, pp. 932 e 933).

Os mecanismos de dor física e social coincidem de tal forma que estratégias para reduzir dor física se mostram eficazes para reduzir dor social. Assim como estratégias para reduzir dor social também contribuem para reduzir dor física, olhar para a foto de uma pessoa amada - um parceiro romântico - atenuou os relatos de dor, assim como a atividade neural relacionada a dor (Eisenberger et al., 2011, p. 4). Para elaborar estratégias para atacar o problema da dor é preciso saber o máximo a seu respeito. Agora que sabemos que existe essa superposição podemos aplicar as estratégias que usamos para lidar com dor física para lidar com a dor social e vice-versa. Também devemos guardar como alerta de que não devemos menosprezar o sentimento de rejeição social, ele tem o mesmo potencial incapacitante quanto qualquer outra dor.

Através de uma meta-análise com o software Neurosynth<sup>8</sup> se estabeleceu a mais provável função principal do Córtex Cingulado Anterior dorsal (dACC). Esse esforço se

7 In summary, a pattern of activations very similar to those found in studies of physical pain emerged during social exclusion, providing evidence that the experience and regulation of social and physical pain share a common neuroanatomical basis. Activity in dorsal ACC, previously linked to the experience of pain distress, was associated with increased distress after social exclusion. Furthermore, activity in RVPFC, previously linked to the regulation of pain distress, was associated with diminished distress after social exclusion.

8 Software aberto cujo projeto está disponível para uso acadêmico em <https://github.com/neurosynth/neurosynth/>. O site, criado e mantido pelo desenvolvedor Tal Yarkoni, introduz o programa ao público interessado e fornece acesso às imagens e meta-análises de fMRI. Atualmente há 14371 estudos na base de dados Neurosynth de imagens tridimensionais que é continuamente alimentado por pesquisadores que utilizam técnicas de fMRI.

justifica pelo fato de que outros estudos neurocientíficos afirmam que o dACC contribui para outros processos psicológicos tais como os processos: executivo, de monitoramento de conflito e de saliência. Ao fazer análises estatísticas de quais os processos psicológicos que aparecem com mais frequência quando a região dACC é ativada, Lieberman e Eisenberger concluem: “Baseados na evidência disponível, a explicação mais clara da função do dACC é que ele é seletivamente envolvido em processos relacionados com dor.”<sup>9</sup> (LIEBERMAN & EISENBERGER, 2015, p. 15255).

Essa abordagem se diferencia da tradicionalmente aplicada em estudos de fMRI. A montagem experimental geralmente busca desencadear determinado processo psicológico e observar qual é a região que consome maior quantidade de oxigênio durante o processo psicológico. A meta-análise desse estudo de 2015 explorou dois sentidos de abordagem.

Na abordagem tradicional, observando-se a região ativada com experimentos que usam os processos psicológicos (*forward inference*<sup>10</sup>) de dor, executivos, de conflito e de saliência, os estudos tratando todos os quatro processos psicológicos produziram ao menos alguma atividade confiável do dACC. Na abordagem de sentido inverso à abordagem anterior, i.e., em que se observa quais os processos psicológicos presentes quando a região do dACC é ativada (*reverse inference*) o processo psicológico que aparece com maior frequência é relacionado com dor e os outros três processos que também ativam o dACC têm uma relevância muito inferior.

O intuito de estudos de fMRI é estabelecer a relação entre processos psicológicos e o local no cérebro que foi respectivamente ativado para produzir tal efeito. A explicação apresentada no artigo que justifica porque *forward inference* não é a abordagem mais adequada, nesse contexto cabe ressaltar que: “Apenas porque um processo psicológico A produz atividade confiável na região X não significa que a atividade na região X em um novo conjunto de dados indica que o processo psicológico A foi invocado.”<sup>11</sup> (LIEBERMAN & EISENBERGER, 2015, p. 15250, tradução nossa).

O artigo esclarece que a teoria que melhor explica os resultados (disponíveis no banco de dados Neurosynth) é que a atividade da região dACC está envolvida com o processo psicológico referente à dor no indivíduo.

O sistema de alarme da dor física, que já tinha uma função no cérebro de nossos ancestrais répteis, possivelmente passou a ser usado também de outra forma no cérebro do ancestral mamífero.

[...] é possível que o sistema de vínculo social, o sistema que nos mantém próximos de outros, pode ter subido nos ombros do sistema pré-existente de dor, emprestando o sinal de dor para alertar e prevenir o perigo da separação social. (Nelson & Panksepp, 1998; Panksepp, 1998).<sup>12</sup> (EISENBERGER e LIEBERMAN, 2005, p. 111, tradução nossa).

9 Based on available evidence, the clearest account of dACC function is that it is selectively involved in pain-related processes.

10 Em 30 de julho de 2018 o site neurosynth foi atualizado, os métodos analíticos nomeados através dos termos *forward inference* e *reverse inference* agora foram rebatizados, respectivamente, como *uniformity test map* e *association test map*. Entretanto, mantivemos a nomenclatura antiga para manter a coerência com o artigo de 2015 de Lieberman & Eisenberger.

11 Just because psychological process A reliably produces activity in region X does not mean that activity in region X in a new dataset indicates that psychological process A was invoked.

12 [...] it is possible that the social attachment system, the system that keeps us near close others, may have piggy-backed onto the pre-existing pain system, borrowing the pain signal to signify and prevent the danger of social separation (Nelson & Panksepp, 1998; Panksepp, 1998).

A tese é que a rede neural responsável pela dor física já estava sendo empregada, porém nos mamíferos essa rede ganhou uma função alternativa: a dor social. Gould<sup>13</sup> foi um defensor de que muitas das funções mais elevadas do cérebro humano são consequências colaterais não intencionais. Claramente o mecanismo de exaptação do tipo que ele batizou como “franklin” é compatível com a forma como foi descrito o possível surgimento da rede responsável pela dor social. “Franklins não são ‘coisas lá fora’ reais, mas sem emprego. Franklins são funções potenciais alternativas de objetos agora sendo usados de outra forma.”<sup>14</sup>(GOULD, 2002, p. 1278, tradução nossa). Uma questão polêmica levantada é se isso de fato corrobora com a tese tão valiosa para Gould que as funções elevadas do cérebro humano (como a sociabilidade humana) são uma consequência evolutiva colateral, ou se uma visão mais determinística da evolução da sociedade e ação humanas devem ser respeitadas. Lieberman menciona o trabalho de Gould, mas, até onde nossa pesquisa nos permite afirmar, não se propõe a mexer nesse vespeiro. O intuito de incluirmos esse autor é apenas fazer menção a uma discussão que está relacionada com o tema apresentado e que merece mais atenção em outro momento.

### **O prazer que conexões sociais podem trazer.**

De maneira análoga com o empréstimo da região que alerta para dor física para cumprir também a função de dor social, as recompensas sociais também compartilham a rede de recompensas relacionadas com prazeres físicos.

[...]recompensas sociais ativam a mesma rede de recompensas que comidas e bebidas desejáveis. Ter uma boa reputação, ser tratado de forma justa e ser cooperativo ativa o estriado ventral. Surpreendentemente, fazer doações para a caridade ativa a rede de recompensas mais do que recebendo a mesma soma de dinheiro para si mesmo.<sup>15</sup> (LIEBERMAN & EISENBERGER, 2009, pp.890, 891).

Essa também é uma associação contraintuitiva, a sensação agradável de ser elogiado não parece ter muito em comum com a deliciosa experiência de comer chocolate. Novamente o cético poderia perguntar: se eu quero sentir o prazer de comer um sorvete e estou de dieta, basta que eu doe dinheiro para alguém em necessidade que o meu cérebro vai me proporcionar a sensação agradável que estou buscando<sup>16</sup>? Quando necessidades sociais são satisfeitas, o cérebro responde de forma muito similar à forma que responde para outras recompensas que são mais tangíveis (Tabibnia et. al.. 2008, Izuma, Saito, & Sadato, 2008, Moll, Krueger, Zahn, Pardini, de Oliveira-souza, & Grafnam, 2006).

As constatações neurocientíficas sobre o estriado ventral parecem corroborar com uma teoria clássica do campo da ética: “[...]são nossas atividades conformes à excelência que nos levam à felicidade.” (ARISTÓTELES, 1992, p. 29). Fazer uma doação para caridade é uma ação cuja excelência poucos questionam. Saber que após a tomada de decisão de doar para caridade temos aumento da atividade no estriado ventral comprova que essa ação promove prazer. Afinal, o estriado ventral está associado com a recompensa na forma da sensação de prazer.

13 Paleontólogo e biólogo evolucionista nascido em Nova Iorque. Que trabalhou nas instituições Harvard University, American Museum of Natural History, New York University.

14 Franklins are not actual but unemployed ‘things out there.’ Franklins are alternative potential functions of objects now being used in another way.

15 [...]social rewards activate the same reward network as desirable foods and drinks. Having a good reputation, being treated fairly, and being cooperative all activate the ventral striatum. Strikingly, making charitable donations activates the reward network more than receiving the same sum of money for oneself.

16 Se as pessoas levassem essa linha de raciocínio a sério com maior frequência, haveria mais pessoas em forma, menos desigualdades e mais pessoas felizes. O preço de ignorar a natureza da mente social pode ser mais alto do que se imagina.

Isso põe em xeque conceito de “altruísmo desinteressado”, pois, sob esse aspecto, o altruísmo em si provoca prazer. Como é possível não carregar o próprio interesse quando a ação gera o próprio prazer como resultado? Boas ações são da nossa natureza, pois as regiões neurais com funções de produção de prazer físico foram provavelmente emprestadas pelos circuitos neurais da conexão social para também serem ativadas em contextos tais como quando se ajuda o próximo. Dalai Lama parecia já prever os dados que a neurociência recolheu, sabiamente aconselhando-nos que: pensar no bem-estar dos outros é a forma mais inteligente de egoísmo.

### **Cérebros interagem fora da cuba?**

O corpo humano vivo é inseparável do cérebro. Essa é uma questão quase livre de debate; quando o cérebro é retirado do corpo de um animal vivo, ambos perdem as condições necessárias para recuperar seu estado original<sup>17</sup>. Alterações e lesões no cérebro alteram profundamente a forma do nosso pensamento. A dissociação que fazemos ao classificar corpo e cérebro com termos distintos é feita por questões anatômicas, não por questões dos processos psicológicos. “[... a] parceria cérebro-corpo interage com o ambiente como um conjunto.” (Damásio, 2012, p. 107). Só faz sentido analisá-los separadamente depois de dissecados. Enquanto organismos vivos, corpo e cérebro são duas metades de um só indivíduo.

O experimento do cérebro na cuba é comentado por Alva Noë, que não foi o criador do experimento<sup>18</sup>, mas um forte combatente da ideia computacionalista de que a essência de alguém poderia ser reduzida simplesmente aos conjuntos de células nervosas dentro da cabeça. Resumidamente seu ponto principal é que a linha que separa o crânio do mundo externo pode ser cruzada (NOË, 2009, p. 49). Para entender o que nós somos precisamos pensar nas interações entre cérebro, o corpo e o mundo. Olhar só para o cérebro não é suficiente.

Com a intenção de fazer uma provocação ao leitor, fazemos alusão à imagem do cérebro fora da cuba. De forma alguma se pretende afirmar que o cérebro possa viver fora do corpo. Do ponto de vista da ciência exata não temos os recursos necessários para fazer os testes, e do ponto de vista ético é um enorme problema. A ideia que queremos apresentar é a seguinte: o cérebro tem redes neurais que o estimulam em direção ao ambiente externo, principalmente às outras pessoas do ambiente. Em uma metáfora: nosso cérebro é circuitado para sair de sua cuba e experimentar o contato com o que está fora. Existem muitos benefícios que o advento da mente social permitiu e estamos apenas arranhando a superfície desse tema.

Nós devemos sempre alertar para o fato de que seres humanos pertencem a uma deriva evolucionária para podermos explicar predisposições e intenções nas interações cérebro-corpo-ambiente.<sup>19</sup> (veja Maturana & Varela, Clark, e Millikan). (Stein, 2012, p. 170, tradução nossa).

17 O advento de ventilação traqueal de pressão positiva na década de 50 e ressuscitação cardiopulmonar na década seguinte levou ao conceito de morte cerebral. Muito recentemente (28 de março de 2018) Nenad Sestan fez menção a um estudo polêmico (do qual participou e que ainda não publicou) no qual as células dos cérebros de 100 a 200 porcos oriundos de abatedouros foram mantidas vivas por 36 horas com o cérebro separado do corpo. Mantinha-se o cérebro em “funcionamento” com o auxílio de bombas, aquecedores e bolsas de sangue artificial em temperatura corporal. (Regalado, 2018).

18 A criação desse experimento mental é atribuída ao filósofo Gilbert Harman.

19 We should always stress the fact that human beings belong to an evolutionary drift in order to be able to explain predispositions and intentions in brain- body-environment interactions (see Maturana & Varela, 18 Clark, 17 and Millikan 19).



Do pouco de evolução biológica tratada neste ensaio, podemos ao menos perceber que foi vantajoso para os nossos ancestrais acrescentar funções sociais para os sistemas neurais que antes só eram ativados por dores e prazeres físicos. Essa é uma das mudanças importantes que aumentaram a capacidade de formarmos grupos maiores e mais coesos. Nos primatas existe uma relação diretamente proporcional da razão **massa do cérebro prevista/tamanho do animal** (chamada **Quociente de Encefalização**) e a **quantidade de indivíduos** que o grupo suporta (**n**). O aumento do volume do neocórtex em relação ao volume corporal possibilitou que novas oportunidades surgissem, principalmente no campo social. Um habitat ecológico que requer um grupo de tamanho maior só é ocupado por uma espécie de primatas em que a razão peso do neocórtex/peso corporal é suficiente. O tamanho dessa área do cérebro dos primatas limita o número de relações sociais possíveis. (DUNBAR, 1992, p. 486).

### **Discussão.**

A teoria do cérebro social diz respeito a nossa qualidade de vida. Entender nossas motivações para criar vínculos sociais nos permite traçar melhores estratégias para melhorar a qualidade de vida. A forma como a sociedade se estabelece ou se constitui sofre forte influência dos conhecimentos aos quais ganhamos acesso. É de suma importância o estudo das funções ou conexões sociais para que possamos colher resultados na forma de revitalizar os vínculos que são tão necessários para nossa espécie. O destaque para os circuitos cerebrais que promovem a conexão social é uma contribuição acadêmica importante.

A própria pesquisa científica não seria possível se não fosse o prazer que sentimos relativo à conquista de boa reputação e a aversão ao ostracismo. Nosso trabalho encontra-se numa circularidade, já que a sua causa e o seu fim perpassam pela questão da conexão social. É difícil estimar a extensão da influência que a conexão social tem sobre nós.

Um de nossos objetivos é a simples reunião de parte das informações disponíveis a respeito da conexão social para contribuir com uma descrição limitada e aproximada das leis da natureza humana que nos inclinam para a vida em sociedade. Tal esforço multidisciplinar tem como meta disseminar resultados que são de interesse da comunidade. O isolamento das pessoas não pode ser ignorado, de tal forma que é necessário pensar estabelecer uma teoria da mente social que abarque as pesquisas recentes sobre dor social. Ações públicas que promovam os contatos sociais em nossa agenda política carecem de um melhor entendimento neurológico do ser-humano. Pesquisa se faz necessária para auxiliar as pessoas que tomam decisões a desenvolver novas soluções para enfrentar o desafio de aproximar de fato as pessoas. Chamar a atenção para o debate envolvendo as dores e os prazeres sociais nos mostra os perigos do individualismo exacerbado.

A natureza do ser humano é estar em sociedade. A tese hobbesiana - de que o estado da natureza humana é o miserável conflito de todos contra todos - não é corroborada pelos estudos aqui citados. Ao olhar para as redes de dores e prazeres sociais surge a convicção de que fazer parte da sociedade é algo que fazemos de acordo com nossos desejos, medos e necessidades e não necessariamente de acordo com um cálculo racional.

As imagens de fMRI reacquecem questões que já habitavam a intuição de pensadores clássicos. Hoje não precisamos nos contentar tão somente em saber apenas como o ser humano se comporta, podemos nos aventurar com mais propriedade na questão subjacente

do porquê para o ser humano agir de determinada forma. Gradualmente estamos ganhando acesso à informação sobre quais áreas do cérebro são ativadas em determinados contextos, ou para alcançar determinados propósitos. As regiões dACC e estriado ventral que estão relacionadas respectivamente com as sensações de dor e prazer sociais que, segundo a interpretação presente na bibliografia estudada até aqui, podem ser entendidos como uma parte central dos circuitos neurais de conexão social.



**Referências bibliográficas:**

- ARISTÓTELES. *Ética a Nicômacos*. Trad. de Mário da Gama Kury. Editora Universidade de Brasília. 3ª edição. 238p. 1992.
- CASPI, Avshalom; ROBERTS, Brent W.; SHINER, Rebecca L. Personality development: Stability and change. *Annu. Rev. Psychol.*, v. 56, p. 453–484, 2005.
- DAMÁSIO, António. *O erro de Descartes: emoção, razão e o cérebro humano*. [s.l.] Editora Companhia das Letras, 2012.
- DEWALL, C. Nathan; MACDONALD, Geoff; WEBSTER, Gregory D.; et al. Acetaminophen reduces social pain: Behavioral and neural evidence. *Psychological science*, v. 21, n. 7, p. 931–937, 2010.
- DUNBAR, Robin IM. Neocortex size as a constraint on group size in primates. *Journal of human evolution*, v. 22, n. 6, p. 469–493, 1992.
- EISENBERGER, Naomi I.; LIEBERMAN, Matthew D.; WILLIAMS, Kipling D. Does rejection hurt? An fMRI study of social exclusion. *Science*, v. 302, n. 5643, p. 290–292, 2003.
- EISENBERGER, Naomi I.; LIEBERMAN, Matthew D. Why rejection hurts: a common neural alarm system for physical and social pain. *Trends in cognitive sciences*, v. 8, n. 7, p. 294–300, 2004.
- EISENBERGER, Naomi I.; LIEBERMAN, Matthew D. Why it hurts to be left out: The neurocognitive overlap between physical and social pain. *The social outcast: Ostracism, social exclusion, rejection, and bullying*, v. 109, p. 130, 2005.
- EISENBERGER, Naomi I.; MASTER, Sarah L.; INAGAKI, Tristen K.; et al. Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proceedings of the National Academy of Sciences*, v. 108, n. 28, p. 11721–11726, 2011.
- GOULD, Stephen Jay. *The structure of evolutionary theory*. [s.l.]: Harvard University Press, 2002.
- HELSON, Ravenna; KWAN, Virginia SY; JOHN, Oliver P.; et al. The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of research in personality*, v. 36, n. 4, p. 287–306, 2002.
- HERRMANN, Esther; CALL, Josep; HERNÁNDEZ-LLOREDA, María Victoria; et al. Humans have evolved specialized skills of social cognition: The cultural intelligence hypothesis. *science*, v. 317, n. 5843, p. 1360–1366, 2007.
- IZUMA, Keise; SAITO, Daisuke N.; SADATO, Norihiro. Processing of social and monetary rewards in the human striatum. *Neuron*, v. 58, n. 2, p. 284–294, 2008.
- KRIPKE, Saul A. Naming and necessity. In: *Semantics of natural language*. [s.l.]: Springer, 1972, p. 253–355.
- LEWIS, D. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy* 50: 24958.[aDR](1980) Mad pain and Martian pain. *Readings in the philosophy of psychology*, v. 1, 1972.
- LIEBERMAN, Matthew D.; EISENBERGER, Naomi I. Pains and pleasures of social life. *Science*, v. 323, n. 5916, p. 890–891, 2009.
- LIEBERMAN, Matthew D.; EISENBERGER, Naomi I. The dorsal anterior cingulate cortex is selective for pain: Results from large-scale reverse inference. *Proceedings of the National Academy of Sciences*, v. 112, n. 49, p. 15250–15255, 2015.
- MOLL, Jorge; KRUEGER, Frank; ZAHN, Roland; et al. Human fronto–mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy*

- of Sciences, v. 103, n. 42, p. 15623–15628, 2006.
- NOË, Alva. *Out of our heads*. [s.l.]: New York: Hill and Wang, 2009.
- OCHSNER, Kevin N.; LIEBERMAN, Matthew D. The emergence of social cognitive neuroscience. *American Psychologist*, v. 56, n. 9, p. 717, 2001.
- REGALADO, A. Researchers are keeping pig brains alive outside the body. *MIT Technological Review*. Disponível em: <<https://www.technologyreview.com/s/611007/researchers-are-keeping-pig-brains-alive-outside-the-body/>>. Acesso em: 31 de julho de 2018,
- RICHARDSON, Michael K.; KEUCK, Gerhard. Haeckel's ABC of evolution and development. *Biological Reviews*, v. 77, n. 4, p. 495–528, 2002.
- SMART, John JC. Sensations and brain processes. *The Philosophical Review*, v. 68, n. 2, p. 141–156, 1959.
- STEIN, Sofia Inês Albornoz. Social minds. *Trends in psychiatry and psychotherapy*, v. 34, n. 4, p. 167–170, 2012.
- TABIBNIA, Golnaz; SATPUTE, Ajay B.; LIEBERMAN, Matthew D. The sunny side of fairness: preference for fairness activates reward circuitry (and disregarding unfairness activates self-control circuitry). *Psychological Science*, v. 19, n. 4, p. 339–347, 2008.
- WILLIAMS, Kipling D.; CHEUNG, Christopher KT; CHOI, Wilma. Cyberostracism: effects of being ignored over the Internet. *Journal of personality and social psychology*, v. 79, n. 5, p. 748, 2000.

# Neurofilosofia dos sonhos e dos sonhos lúcidos: Contribuições para a filosofia da mente

Glescikelly Herminia Ferreira<sup>1</sup>

## O que são sonhos lúcidos

O termo “sonho lúcido”(SL) foi cunhado pelo psiquiatra holandês Frederick van Eeden, em 1913, para definir um estado onírico de plena consciência, o que permite ao sonhador a capacidade de atuar voluntariamente durante o sonho, como explicou:

[...] o tipo de sonho que chamei “sonhos lúcidos” parece-me o mais interessante e digno do estudo e observação mais cuidadosa. [...] Nesses sonhos a reintegração das funções psíquicas é tão completa que a pessoa que está dormindo lembra da vida diária e de sua própria condição, atinge um estado de percepção consciente perfeito, e consegue dirigir sua atenção e tentar executar diferentes atos voluntariamente. Assim mesmo o sono continua profundo e revigorante, como posso afirmar com confiança. (VAN EEDEN, 1913, p. 446)

Van Eeden escreveu seus sonhos por mais de uma década e registrou mais de trezentos sonhos lúcidos, publicando seus resultados em 1913. Porém, um estudo laboratorial controlado do fenômeno só seria empreendido setenta anos depois, com o psicofisiologista Stephen LaBerge.

Sobre o fenômeno, diz LaBerge (1990, p. 109):

Sonhadores “lúcidos” [...] relatam que são capazes de lembrar livremente as circunstâncias da vida acordada, de pensar claramente, e de agir deliberadamente após reflexão, tudo isso ao experienciar um mundo de sonhos que parece vividamente real.

A percepção da importância do fenômeno onírico para estudos da consciência fez com que vários filósofos atuais se debrucem sobre o tema. A filósofa Jennifer Windt e o filósofo Thomas Metzinger desenvolveram uma teoria para os sonhos lúcidos, segundo a qual:

Em particular, ela prevê que a lucidez resulta de uma estabilização semelhante à vigília da perspectiva de primeira pessoa, permitindo que os sonhadores não somente vivenciem

---

<sup>1</sup> Graduação em Filosofia (UFPE), Especialização em Neurociência Aplicada (UFPE), Mestranda em Filosofia, Depto. de Filosofia da Universidade Federal de Pernambuco.

a si mesmos como o centro do mundo onírico, mas também formem um modelo consciente estável [...] da sua relação epistêmica ou de agência para com o mundo do sonho, de maneira a perceberem que eles estão apenas sonhando. (WINDT, 2015, §4)

Sonhar lúcido é, portanto, estar consciente durante o sonho; é perceber no próprio sonho que se está, de fato, sonhando, diferentemente do que acontece no sonho comum, em que o sujeito julga que o fenômeno vivenciado corresponde à realidade externa, e apenas ao acordar percebe que estava sonhando. Dessa forma, durante o SL, o sonhador tem consciência de que o mundo percebido naquele momento é o onírico, e que seu corpo é o corpo sonhado, o que pode torná-lo hábil a tomar decisões e direcionar ações, sendo portanto uma exceção à regra de que o sonho é necessariamente uma experiência em que se tem pouco julgamento racional (MOTA-ROLIM; ARAUJO, 2013).

### Sonhos lúcidos na história da filosofia

O tema da “lucidez onírica”, embora não referido nestes termos, está presente na filosofia desde a Grécia Antiga. Aristóteles foi o primeiro filósofo que escreveu sobre SL. No tratado “Sobre os sonhos” encontrado no *Parva Naturalia* (2012, p. 113), o estagirita escreveu: “de fato, frequentemente quando alguém está adormecido, alguma coisa em sua alma o informa que a aparência diante dele é um sonho”. Ele também afirmava que os sonhos poderiam apontar problemas de saúde e auxiliar no comportamento da vida desperta – o que também é concebido atualmente.

Na Idade Média, Agostinho levantou o problema da responsabilidade moral durante os sonhos e discutiu a possibilidade da consciência onírica, pois temia ser julgado também por suas ações nos sonhos. Embora tenha concluído que esta consciência não era possível e, portanto, não poderia ser julgado por seus sonhos, é considerado dele o primeiro relato de um sonho lúcido no Ocidente, datado de 415, advindo do seu amigo Genádio de Cártago. Para Tomás de Aquino, é possível em alguns momentos julgar e discernir durante o sonho, pois o bom senso é parcialmente liberado, como podemos ver no seguinte trecho da *Summa Theologica* (c. 1270):

no fim do sono, com os homens sensatos e com quem é dotado de imaginação forte [...] não só a imaginação retém sua liberdade, como o bom senso é parcialmente libertado; de modo que, quando um homem está dormindo, às vezes consegue julgar se o que está vendo é um sonho e discernir, por assim dizer, as coisas das respectivas imagens (AQUINO, apud LABERGE, 1998, p. 13).

Isto significa que há, para estes filósofos, um estado de autorreflexão durante os sonhos, que podemos apontar como a distinta experiência do sonho lúcido.

Na modernidade, Descartes trouxe a experiência onírica para incentivar o ceticismo quanto às experiências, pois quando sonhamos acreditamos estar em uma situação de vigília, ou seja, não conseguimos através dos nossos sentidos distinguir entre o estado de vigília e o estado de sonho. No entanto, em meados de 1610, relatou sonhos nos quais ele sabia que estava sonhando, e ainda conseguia solucionar problemas da vigília enquanto sonhava (cf. FRANZ, 2011, p. 164). Ainda no século XVII, Pierre Gassendi escreveu que tivera algumas experiências de lucidez onírica a partir da constatação de que se encontrava com pessoas já falecidas (cf. LABERGE, 1988, pp. 13-14) – o que é um bom sinal para alcançar a lucidez durante o sonho.

Saltando para o século XX, em 1959, o filósofo wittgensteiniano Norman Malcolm, desenvolveu três argumentos para afirmar a impossibilidade da experiência onírica: (1) Não há como saber se um indivíduo realmente sonhou apenas através de seus relatos. (2) Existe uma contradição entre os termos “sonho” e “sono”, pois enquanto dormimos nos encontramos num estado no qual há falta de consciência, e esta é necessária para a experiência. (3) Sem comunicação e sem a capacidade de fazer julgamentos durante o sono, não há como saber se a pessoa sonha enquanto sonha, ou apenas elabora seu sonho em seu relato ao acordar. Para Malcolm, ainda, uma ciência dos sonhos não seria relevante e promissora. Os três pontos podem ser revistos e negados através dos Sonhos Lúcidos, que como já vimos é definido pelo estado no qual a pessoa está autoconsciente durante o sonho.

Em 1976, Daniel Dennett, escreveu o texto “Are dreams experiences?”, no qual desenvolve a ideia de que sonhos não são experimentados realmente, posto que para que haja experiência é necessário estar consciente – o que se assemelha ao argumento de Malcolm. Dennett (1976, p. 161) afirma que os SLs são a ilusão de um sonho dentro de outro sonho. Sobre este texto ele recebeu como resposta o texto “Oneiric experiences” da filósofa Kathleen Emmett, no qual ela desenvolve uma argumentação favorável à experiência onírica e aponta os SLs como uma possível objeção à posição de Dennett. Emmett (1978, p. 449) termina seu artigo destacando: “A perspectiva de se desenvolverem métodos nos quais o sujeito possa ser ensinado a sonhar lucidamente e relatar sobre este processo enquanto este estiver ocorrendo não é inconcebível.” De fato, esta técnica foi desenvolvida alguns anos depois, conforme veremos adiante.

### **Estudos científicos dos sonhos lúcidos**

Em 1981, o psicofisiologista americano Stephen LaBerge conseguiu registrar em tempo real a experiência do SL, por meio de comunicação com movimentos oculares, realizados conscientemente pelo sujeito adormecido. LaBerge aponta que o sono REM (do inglês *Rapid Eye Movement*, ou movimento rápido dos olhos) está ligado às funções de memória e aprendizado. O sonho REM é considerado um produto da evolução do sono, pois de início, na evolução biológica, existia apenas o sono de ondas lentas. Já se sabia que pássaros e mamíferos têm sono REM (RIBEIRO & MOTA-ROLIM, 2013), mas pesquisa recente o verificou em lagartos, indicando que o sono REM está presente em boa parte dos amniotas (SHEIN-IDELSON et al., 2016). É no sono REM que ocorre o período de maior ocorrência do SL. A pesquisa de LaBerge sobre SLs consistiu no registro do sono através de medidas da atividade elétrica cerebral (EEG), dos movimentos oculares (EOG) e do tônus muscular (EMG). Foi combinado entre pesquisadores e sonhadores cinco movimentos oculares para a esquerda e para a direita (medidos por EOG), caso os sujeitos conseguissem ter um SL. O sinal objetivo através do movimento dos olhos foi escolhido por causa da “atonia muscular” (ou paralisia) do nosso corpo ao entrar neste estágio do sono. Durante o sono REM, os únicos músculos que não entram em atonia são os oculares, além dos reflexos. LaBerge e colaboradores conseguiram registrar esse sinal ocular pré-combinado dos sujeitos que tiveram um SL – o que foi verificado tanto pela análise empírica durante o procedimento como também pela coincidência dos relatos realizados por eles. A partir deste estudo de LaBerge et al. (1981), podemos dizer que os SLs ultrapassam a barreira entre o mundo onírico e o mundo desperto, quando há comunicação real entre eles, o que desqualifica o terceiro argumento de Malcolm, de que os sonhos não são experiências pela incapacidade de se comunicar a ocorrência do fenômeno enquanto



ele acontece. Os estudos de LaBerge foram replicados por diversos pesquisadores e sua técnica é utilizada como modo padrão de verificação da ocorrência de SL.

Em 2012, Dresler et al. conduziram uma pesquisa para tentar identificar os correlatos neurais da lucidez onírica a partir do contraste entre SL e Sonho Não Lúcido, durante o sono REM, com diagnóstico por imagem (fMRI) e eletroencefalograma (EEG). Esta pesquisa envolveu quatro sujeitos e encontrou áreas corticais significativamente ativadas durante o SL, em comparação ao Sonho Não Lúcido (NSL). Essas áreas são os córtices pré-frontal e occipito-temporal, os lóbulos parietais, o cuneus e o pré-cuneus. São áreas consideradas responsáveis por elevadas capacidades cognitivas, volição, memória e autoconsciência, o que deve explicar as ações autônomas e a rememoração de fatos da vida desperta e de outros estados, e também a análise de sentimentos para resolução de problemas – que são típicas dos SLs, mas atípicas nos NSLs.

Muita pesquisa neurocientífica tem sido realizada sobre sonhos e SLs. No Brasil, destacam-se as pesquisas sobre sonho e memória de Sidarta Ribeiro e seu grupo, da Universidade Federal do Rio Grande do Norte (UFRN). Em 2010, Ribeiro mostrou como as memórias se consolidam durante o sono; como ocorre o processo de migração destas memórias do hipocampo ao córtex; e como os sonhos auxiliam nesse processo, pois articulam memórias complexas. Pesquisas como estas mostram que a ciência dos sonhos pode ser bem-sucedida e é altamente relevante para entender aspectos importantes da vida humana.

Em 2012, uma pesquisa pioneira sobre SLs no Brasil foi completada no Instituto do Cérebro da UFRN pelo neurocientista Sérgio A. Mota-Rolim, orientada por Sidarta Ribeiro. Este trabalho de doutorado compreendeu dois estudos: primeiramente foram investigados os aspectos epidemiológicos (com  $n = 3430$  sujeitos) do SL e em seguida investigaram-se suas bases neurais. A conclusão sobre o estudo epidemiológico é a de que o SL é uma experiência que pode ser dita comum, mas que não ocorre com frequência, e é difícil de controlar (MOTA-ROLIM, 2012, p. 36). A segunda parte da pesquisa investigou as bases neurais do SL e os resultados apontaram diferenças neurofisiológicas para com o estado REM normal e a vigília. Foi encontrada maior potência do ritmo alfa na região occipital e do ritmo gama na região têmporo-parietal direita. Esta atividade de alfa sugere que o SL pode ser mesmo um estado intermediário entre o REM e a vigília, pois um aumento da potência do ritmo alfa corresponderia a micro-despertares, o que se apresenta como facilitador de comunicação entre o cérebro e o mundo externo. O ritmo gama nas regiões têmporo-parietal direita e frontal, áreas relacionadas à imagética corporal e à autoconsciência, faz com que o pesquisador proponha essa atividade como o “mecanismo neurobiológico subjacente ao SL” (MOTA-ROLIM, 2012, p. ii).

Para LaBerge, o SL é o produto de milhões de anos de evolução biológica. Em 2013, foi publicado o texto “Bases biológicas da atividade onírica”, escrito por Sidarta Ribeiro & Mota-Rolim. Nele, os neurocientistas fazem uma explanação sobre a evolução do sono e do sonho, considerada como fundamental para uma teoria dos sonhos, juntamente com uma fenomenologia, e com a distinção funcional dos estados encontrados no sono e no sonho. Segundo eles, o sono evoluiu concomitantemente ao aparecimento de intensas conexões entre o tálamo e o córtex; durante o sono não REM estas interações tálamo-corticais originam uma grande atividade neuronal e lentas oscilações elétricas, o que configura na falta de conexão sensorial externa (RIBEIRO & MOTA-ROLIM, 2013, p. 204). Já

durante o REM essa ativação tálamo-cortical “contribui para os aspectos perceptuais da consciência onírica” (RIBEIRO & MOTA-ROLIM, 2013, p. 207).

Todos estes estudos mostram a realidade do fenômeno dos sonhos lúcidos, pois conseguem apontar sua ocorrência, áreas de ativação no encéfalo e quais as consequências disto. Estes dados fornecem material para a formulação de teorias que auxiliam na compreensão do que é o fenômeno da consciência, da autoconsciência, da volição e do julgamento racional. Com isto, veremos na seguinte seção como a filosofia tem compreendido o SL e como isso tem sido aproveitado na filosofia da mente.

### **Sonhos lúcidos e as possíveis contribuições para a filosofia da mente**

No âmbito da filosofia o tema dos sonhos tradicionalmente não era muito desenvolvido, mas isso começou a mudar no início dos anos 2000. Hoje temos contribuições significativas de filósofos sobre o assunto. Antti Revonsuo, Jennifer Windt, Evan Thompson e Thomas Metzinger estão entre os nomes da filosofia atual que entendem a importância de estudos sobre sonhos para compreender a consciência. Eles concordam que funções complexas da consciência, como a capacidade autorreflexiva, as memórias de trabalho e autobiográfica, e a volição estão presentes nos sonhos. Por questão de tempo, trataremos nesta apresentação apenas de Revonsuo, Windt e Metzinger.

O filósofo finlandês Antti Revonsuo (2000) afirma que o sonho consciente, típico do sono REM, é um mecanismo evolutivo e simula ações da vigília. Em sua teoria dos sonhos como simulação de ameaça (*threat simulation theory*), defende que este mecanismo foi selecionado por sua capacidade de simular eventos ameaçadores, ensaiando os mecanismos cognitivos necessários para a percepção eficiente de ameaças e a capacidade de evitá-los, aumentando assim a probabilidade de sucesso reprodutivo do animal (o sono REM tendo surgido, como vimos, entre os amniotas). Dessa maneira, nega que os sonhos sejam um mero fruto de ativação cortical aleatória, como alguns cientistas já defenderam, já que os sonhos se organizam de maneira semelhante a como o mundo se apresenta para a consciência na vigília. Além disso, alguns tipos de experiência são mais representadas do que outras, por exemplo, sonhos com elementos negativos e desagradáveis, como medo e raiva.

Com relação aos SLs, Revonsuo tem defendido o seu uso terapêutico para combater pesadelos recorrentes – sonhos que ocorrem sempre da mesma maneira, às vezes com variações mínimas, e que, normalmente, estão ligados a algum evento traumático (ver GAVIE & REVONSUO, 2010), pesadelos estes que podem se repetir por décadas, atuando como revivescência do trauma. Nesse sentido, segue autores como Spormarker et al. (2003), Zadra & Pihl (1997) e Hobson (2009). Em livro recente (REVONSUO, 2017, pp. 117-119), faz uma apresentação didática dos SLs, salientando a pesquisa de Dresler et al. (2012), descrita acima. Revonsuo parte de uma teoria da consciência como fenômeno biológico, encefálico e localizável, podendo ser investigado através dos dados de neuroimagem e de polissonografia durante os sonhos. Isso o leva a salientar a importância de unir filosofia e ciências neurais para avançar na compreensão da consciência, sendo o sonho primordial para este avanço.

Algumas perguntas serão aprofundadas na continuação da pesquisa, com relação a esta discussão. Pode-se dizer que os sonhos lúcidos ativam a faculdade da auto-consciência, tão discutida na Filosofia? Que contribuição trazem os estudos neurocientíficos mencionados? Qual seria a função dos SLs na teoria evolucionista dos sonhos como simulação de ameaças?



A filósofa Jennifer Windt (2011) concorda com a importância desses estudos, mas não considera que os sonhos sejam o caminho privilegiado na busca do correlato neural da consciência, sendo importante o contraste entre diferentes estados de consciência, incluindo os oníricos. Com seu ex-orientador Thomas Metzinger, publicou um importante artigo de revisão filosófica em 2007. Com relação aos correlatos neurais da consciência onírica, apontam para pesquisa de Hobson et al. (2000), e salientam a importância das “ondas PGO” no sono REM, que são geradas na ponte (P), conduzidos para o núcleo geniculado lateral do tálamo (G) e depois para o córtex occipital (O). Já os SLs reativam o córtex pré-frontal dorsolateral, que, normalmente, está desativado no sono REM. Esta região cortical está associada na vigília a habilidades executivas como a auto-observação, planejamento e tomada de decisão (HOBSON et al., 2000, p. 837; WINDT & METZINGER, 2007, p. 198).

Em seu estudo filosófico, Windt & Metzinger (2007) utilizam uma análise do estado de consciência desenvolvida anteriormente por este, baseado no “auto-modelo fenomênico” (*phenomenal self-model*, PSM), ou seja, o modelo de realidade que o sujeito tem em diferentes estados de consciência, como a vigília, sono REM, sonho lúcido e alucinações. Os autores dedicam algumas páginas deste estudo (2007, pp. 201-2) para analisar a fenomenologia dos sonhos, baseando-se em diversos autores. Mas afóra os aspectos fenomenológicos que caracterizam os estados de consciência, propõem três atributos representacionais: globalidade (ativação de um modelo global da realidade), apresentacionalidade (integração deste modelo dentro em uma janela virtual de presença) e transparência (o modelo de realidade não é reconhecido como um mero modelo pelo sujeito) (WINDT & METZINGER, 2007, p. 195). Concordam com Revonsuo que tanto o sonho quanto a vigília envolvem uma “realidade virtual”, e trabalham com a distinção entre uma ativação *online* (como a vigília) e uma *off-line* (como os sonhos). A partir desses e outros conceitos, analisam então as diferenças entre os estados de vigília, sonho não lúcido e sonho lúcido.

A definição de van Eeden (1913) de SL envolve três pontos, segundo os autores (2007, pp. 218-19): (i) reconhecimento completo de que se trata de um sonho e de que há uma vida em vigília; (ii) a habilidade de agir voluntariamente; (iii) o fato de que as sensações corporais ficam bloqueadas (estados parcialmente “desincorporados”). Em outro momento, porém, desdobram essas características dos SL em cinco pontos principais (2007, p. 211): (1) O sonhador lúcido sabe que está sonhando. (2) Possui clareza intelectual completa, que pode até ultrapassar a da vigília. (3) Os cinco sentidos funcionam como no estado de vigília (ao contrário dos sonhos usuais). (4) Tem memória dos seus estados passados, tanto na vigília quanto nos sonhos. (5) A propriedade de agência é realizada. Ressaltam também que os déficits cognitivos usuais dos sonhos são bastante diminuídos nos SLs, como desorientação, raciocínio confabulatório e amnésia.

Como exemplo do uso das categorias de Metzinger em sua análise, defendem que o SL viola a restrição de *transparência*, pois o sujeito está ciente que se trata de um modelo virtual. Dizem então que os SLs são a única classe de estados fenomênicos que é “globalmente opaca”; não há lucidez global na vigília porque isso traria desvantagem adaptativa (2007, pp. 213-16).

Em 2013, a filósofa participou de um estudo juntamente com a psicóloga Ursula Voss, o psiquiatra Allan Hobson e colaboradores, no qual apresentam os resultados da escala LuCID, desenvolvida para medir o estado de consciência onírica no SL:

[...] tanto os sonhos lúcidos quanto os não-lúcidos são condições importantes de contraste para teorias da consciência acordada, fornecendo *insights* valiosos na estrutura da experiência consciente e seus correlatos neurais durante o sono (VOSS et al., 2013, p. 9).

Conclui-se que o achado de padrões distintos de atividade cerebral, a partir do contraste SL/NSL, mostra o SL como um estado que compreende as duas consciências (primária e secundária), o que proporciona a este estado os fenômenos de autorreflexão, memória (de trabalho e autobiográfica), pensamento racional, planejamento e controle comportamental. Esta caracterização serve de aporte para o desenvolvimento de pesquisas laboratoriais tanto para a compreensão fenomenológica dos SLs como para a identificação das diferentes regiões cerebrais envolvidas.

Sobre isso concordam os neurocientistas Mota-Rolim & Araujo (2013), de maneira ampliada, propondo que a investigação seja através do contraste dos correlatos neurais do SL com outros estados, que são os seguintes. (1) *Fisiológicos*: estados acordado atento ou desatento (conhecido como devaneio ou *daydreaming*), sonolência, sono superficial e profundo, sono REM, falso despertar e sonho (lúcido ou não). (2) *Patológicos*: estados que se associam às doenças, como na psicose, no coma, experiência de quase morte, sonambulismo e paralisia do sono. E os (3) *alterados*: estados que são induzidos por (a) técnicas cognitivo-comportamentais (hipnose e meditação), (b) substâncias psicoativas (LSD e ecstasy), ou (c) mistura das anteriores (rituais com ayahuasca).

Como vimos, nas últimas décadas, intensificou-se o interesse em estudar a natureza da consciência com base na neurociência. Da perspectiva filosófica, tais abordagens são ditas “materialistas”, na medida em consideram que a consciência é produzida pelo sistema nervoso humano ou pelo seu corpo, e que na morte do corpo a consciência individual desaparece. Isso se opõe às correntes “dualistas”, que reservam para a mente alguma independência com relação ao corpo, incluindo a sobrevivência (mesmo que parcial) após a morte. O materialismo pode ser dividido em correntes mais reducionistas e correntes mais emergentistas. Na primeira classe, incluem-se correntes que defendem a identidade mente-cérebro, ou seja, processos mentais como sonhos seriam idênticos a certos processos cerebrais. No materialismo emergentista, considera-se que a mente, mesmo sendo um produto do corpo, adquire uma certa independência em relação ao corpo, ou no sentido de não ser redutível a este (emergentismo forte), ou no sentido de tornar-se o agente causal dos processos de tomada de decisão consciente (emergentismo com “causação descendente”, defendido pelo neurocientista Roger Sperry).

Este trabalho quis mostrar os desenvolvimentos das pesquisas neurocientífica e filosófica sobre sonhos, em especial o Sonho Lúcido, com o intuito de apontar como esse fenômeno pode contribuir para a elucidação do problema mente-corpo. Desta maneira, entendemos que a retomada da pesquisa sobre o fenômeno onírico em filosofia contribuirá para solucionar o problema da natureza da consciência. Além da questão sobre a consciência fenomênica e a natureza dos qualia, presente em todos os sonhos, vimos como os sonhos lúcidos parecem abarcar também a nossa capacidade de autorreflexão e de agência por vontade própria, ativando áreas corticais envolvidas também na autorreflexão em estado de vigília. De nossa parte, a continuidade da pesquisa envolverá um aprofundamento das abordagens filosóficas aos sonhos lúcidos apresentadas neste trabalho, como as de Revonsuo, Windt e Metzinger, entre outras.

**Referências bibliográficas:**

- ARISTÓTELES. *Parva Naturalia*. São Paulo: EDIPRO, 2012.
- DENNETT, Daniel C. Are dreams experiences? *Philosophical Review*, v. 85, p. 151-171, 1976.
- DRESLER, M. et al. Neural correlates of dream lucidity obtained from contrasting lucid versus non-lucid REM sleep: A combined EEG/fMRI case study. *Sleep*, v. 35, p. 1017-1020, 2012.
- EMMETT, Kathleen. Oneiric experiences. *Philosophical Studies*, v. 34, p. 445-450, 1978.
- FRANZ, Marie-Louise von. *Sonhos: um estudo dos sonhos de Jung, Descartes, Sócrates e outras figuras históricas*. Petrópolis: Vozes, 2011.
- GAVIE, Josefin; REVONSUO, Antti. Commentary on “The neurobiology of consciousness: Lucid dreaming wakes up” by J. Allan Hobson. *International Journal of Dream Research*, v. 3, p. 13-15, 2010.
- HOBSON, J. Allan. The neurobiology of consciousness: lucid dreaming wakes up. *International Journal of Dream Research*, v. 2, p. 41-44, 2009.
- HOBSON, J. Allan; PACE-SCHOTT, Edward F.; STICKGOLD, Robert. Dreaming and the brain. Toward a cognitive neuroscience of conscious states. *Behavioral and Brain Sciences*, v. 23, p. 793-842, 2000.
- LABERGE, Stephen P. Lucid dreaming in Western literature. In: J. Gackenbach; S. LaBerge (Eds.) *Conscious mind, sleeping brain: Perspectives on lucid dreaming*. New York: Plenum, p. 11-26, 1988.
- \_\_\_\_\_. Lucid dreaming: Psychophysiological studies of consciousness during REM sleep. In: R.R. Bootzen, J.F. Kihlstrom; D.L. Schacter (Eds.), *Sleep and Cognition*. Washington: APA Press, p. 109-126, 1990. Disponível em: <http://www.lucidity.com/SleepAndCognition.html> >. Acesso em: 08 jul. 2015.
- \_\_\_\_\_. et al. Lucid dreaming verified by volitional communication during REM sleep. *Perceptual and Motor Skills*, v. 52, p. 727-732, 1981.
- MOTA-ROLIM, Sérgio A. *Aspectos epidemiológicos cognitivo-comportamentais e neurofisiológicos do Sonho Lúcido*. Tese de doutorado em Psicobiologia, Universidade Federal do Rio Grande do Norte, Natal, 2012.
- MOTA-ROLIM, Sérgio A.; ARAUJO, John F. Neurobiology and clinical implications of lucid dreaming. *Medical Hypotheses*, v. 81, p. 751-756, 2013.
- REVONSUO, Antti. The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioural and Brain Sciences*, v. 23, p. 877-901, 2000.
- \_\_\_\_\_. *Foundations of consciousness*. New York: Routledge, 2017.
- RIBEIRO, Sidarta; MOTA-ROLIM, Sérgio A. Bases biológicas da atividade onírica. In: Ribeiro, L. (org.). *O sono e seus transtornos*. Rio de Janeiro: Atheneu, p. 201-227, 2013.
- SHEIN-IDELSON, M. et al. Slow waves, sharp waves, ripples, and REM in sleeping dragons. *Science*, v. 352, p. 590-595, 2016.
- SPOORMAKER, Victor I.; VAN DEN BOUT, Jan; MEIJER, Eli J.G. Lucid dreaming treatment for nightmares: a series of cases. *Dreaming*, v. 13, p. 181-86, 2003.
- VAN EEDEN, Frederic. A study of dreams. *Proceedings of the Society for Psychical Research*, v. 26, p. 431-461, 1913.
- VOSS, Ursula. et al. Measuring consciousness in dreams: The lucidity and consciousness in dreams scale. *Consciousness and Cognition*, v. 22, p. 8-21, 2013.
- WINDT, Jennifer M. Dreams and dreaming. *Stanford Encyclopedia of Philosophy*. Disponível

em: < <https://plato.stanford.edu/entries/dreams-dreaming/> > . Acesso em: 29 fev. 2017.

WINDT, Jennifer M.; METZINGER, Thomas. The philosophy of dreaming and self-consciousness: what happens to the experiential subject during the dream state? In: Barrett, D. & McNamara, P. (orgs.). *The new science of dreaming*. v. 3: Cultural and theoretical perspectives. Westport, CT: Praeger, p. 193-247.

ZADRA, Antonio L.; PIHL, Robert O. Lucid dreaming as a treatment for recurrent nightmares. *Psychotherapy and Psychosomatics*, v. 66, p. 50-55, 1997.

# Um exame da plausibilidade científica do libertismo neural

Guilherme Teixeira Martins Schettini<sup>1</sup>

## Apresentação: o libertismo neural

Chamemos de *libertismo neural* (*neural libertarianism*) a posição filosófica que afirma que pelo menos algumas das decisões humanas são, ao mesmo tempo, (a) indeterminadas e (b) não randômicas (*nonrandomness*). Por decisões indeterminadas, entendamos aquelas que não se sucedem necessariamente de nenhum evento cerebral prévio, ou de eventos no ambiente em volta do corpo; por decisões não randômicas, entendamos aquelas que encerram as noções de autoria e controle da mente humana.

Agora, chamemos de *decisões divididas* aquelas que são tomadas na seguinte circunstância: (a) o indivíduo possui razões igualmente fortes para duas ou mais opções, e se sente dividido e incapaz de determinar qual razão deve prevalecer; e (b) o indivíduo efetua uma escolha sem resolver esse dilema, e tem a experiência de “arbitrariamente escolher”, ou de escolher sem a “convicção de ter feito a coisa certa”.

Se uma decisão dividida for, de fato, completamente indeterminada no momento da escolha (i.e., se as probabilidades de escolha de cada uma das opções forem, de fato, as mesmas), então ela será não randômica, e o libertismo neural será verdadeiro<sup>2</sup>. Uma argumentação robusta para o que acaba de ser afirmado não pode ser efetuada no curto espaço deste artigo. Na sua ausência, apelemos para uma intuição.

Imaginemos uma situação na qual estivéssemos em um restaurante e quiséssemos fazer um pedido para a sobremesa. Se a nossa mente se defrontasse com razões igualmente fortes para ordenar, digamos, uma salada de frutas e uma torta de chocolate, e, uma vez aparecido o garçom, sem que tivéssemos desempatado a questão, ordenássemos a torta de chocolate, teríamos efetuado uma escolha dividida, segundo a definição proposta acima.

Seria essa decisão completamente indeterminada? A esse respeito, o mais honesto é afirmar: não sabemos. O fato da nossa mente ter ficado dividida durante todo o processo de escolha é um indicativo de que nada em particular determinou ou constrangeu a nossa decisão: havia causas múltiplas, que se equilibraram em força. Em uma linguagem materialista, isso seria equivalente a dizer que nenhum evento físico-neural necessitou a nossa escolha. Não se pode, entretanto, descartar a hipótese de que algum fator oculto (um processo encefálico inconsciente) tenha atuado na última hora e nos feito decidir num sentido ou noutro.

1 Doutorando (2018-) em Filosofia pela USP;

2 Se o termo “randômico” for utilizado como sinônimo de “indeterminismo”, como é usual na Física, essa sentença será evidentemente falsa. Mas vale lembrar que, na teoria apresentada, “randômico” significa simplesmente aquilo que é desprovido de agência e controle – e “não randômico”, a coisa oposta.



De toda forma, é apenas no primeiro cenário que os elementos que caracterizam a autoria e o controle da mente humana estariam plenamente presentes. Se uma decisão dividida é completamente indeterminada no momento de escolha, *quem escolhe e controla* a decisão é a mente humana, não havendo rival para ela. Se uma decisão dividida, a despeito das aparências, não é completamente indeterminada no momento de escolha, a autoria e o controle da mente humana sobre a decisão precisam, no mínimo, ser compartilhados com a razão oculta que desempatou a questão. Essa razão provavelmente terá seu correlato no nível físico-neural.

Esse exemplo do restaurante é bom porque conduz à intuição correta de que efetuamos decisões divididas algumas vezes ao dia (o que pode significar, a depender delas serem completamente indeterminadas, que exercemos plenamente o livre-arbítrio algumas vezes ao dia). Mas as decisões divididas podem estar associadas a temas bem menos banais, como a escolha de aceitar ou não uma boa oferta de emprego em uma cidade desinteressante.

Os conceitos e exemplos mencionados até aqui são do filósofo estadunidense Mark Balaguer (BALAGUER, 2009). Por trás do termo libertismo neural, está a ideia de que o ser humano possui livre-arbítrio em função da estrutura do seu cérebro. A indeterminação de certos eventos neurais, como os que acompanham as decisões divididas, teria a sua origem na indeterminação dos eventos quânticos, mais basilares, de que são constituídos. O indeterminismo seria relevante para a não aleatoriedade, realçando-a.

Não está claro, entretanto, se aquela ideia poderia ser confirmada ou falseada experimentalmente<sup>3</sup>. Por conseguinte, a tese de que, nas decisões divididas, as probabilidades de escolha de cada opção são exatamente as mesmas, pelo que essas decisões seriam completamente indeterminadas no momento de escolha, também é uma questão em aberto.

Se a oferta de um veredicto sobre o libertismo neural está fora de nosso alcance, é ao menos possível levar a cabo um exame da plausibilidade científica dessa posição, comparando-a com posições concorrentes e com aquilo que pode ser legitimamente derivado dos experimentos já realizados na neurociência com implicações no tema do livre-arbítrio. Neste trabalho, faremos isso de uma maneira superficial.

### Posições concorrentes

Por óbvio, o libertismo neural não é a única posição filosófica sobre o tema do livre-arbítrio. Em geral, ele é uma posição pouco prestigiada pelos filósofos, quando não absolutamente desconhecida. Vejamos, então, sumariamente o que afirmam as principais posições concorrentes<sup>4</sup>, a começar pelas de cunho determinista.

Pode-se chamar de *determinismo universal* (BALAGUER, 2009, seção 2) a posição filosófica que afirma que todos os eventos da natureza são necessariamente causados por eventos físicos anteriores, de modo que, pelo menos em princípio, seria possível derivar os primeiros dos últimos (e das leis da física). A objeção óbvia que pode ser dirigida a essa posição (que, em realidade, é uma versão radical de determinismo) diz respeito à natureza dos eventos de nível micro. As interpretações ortodoxas da Mecânica Quântica,

3 Na discussão sobre a flambagem (que examina uma “decisão dividida” na Física), levada a cabo em (GATTI e PES-SOA, 2012), seção 7, há duas posições filosóficas gerais que não podem ser testadas.

4 Por “principais posições concorrentes”, entendemos as posições que são logicamente possíveis a partir de variações nas teses basilares do libertismo neural (o indeterminismo e a não aleatoriedade). Isso não significa, entretanto, que elas estejam bem estabelecidas na literatura, nem que, em caso afirmativo, respondam pelo nome que lhe daremos neste artigo (inspirados em Mark Balaguer). Tampouco é nossa preocupação associá-las a um ou outro filósofo.

5 Embora o nosso objetivo neste artigo não seja defender o libertismo neural, apresentaremos objeções às posições concorrentes.

em particular a teoria da complementaridade de Niels Bohr, são indeterministas. Por outro lado, há uma interpretação determinista, desenvolvida por David Bohm em 1952. Hoje, é uma questão em aberto se tudo na natureza é necessariamente causado por eventos físicos prévios ou não.

O *determinismo de nível macro* (BALAGUER, 2009, seção 3), se assim se pode dizer, é a posição filosófica que afirma a determinação de todos os eventos da natureza, com a exceção dos quânticos. Embora aparentemente mais precavido que o determinismo universal, o determinismo de nível macro é claramente falso: uma medição quântica é um processo macroscópico, e ela é tão indeterminista (de acordo com as interpretações ortodoxas) quanto o processo quântico subjacente. Pode-se dizer, além disso, que essa variante de determinismo faz uso de uma generalização indevida: embora alguns eventos de nível macro pareçam de fato determinados, não é absolutamente claro que todos sejam assim. O lançamento de dados, as partidas de futebol e algumas decisões humanas, por exemplo, são eventos de nível macro aparentemente randômicos.

É claro que os deterministas de nível macro podem “contar uma história” a respeito de como esses eventos aparentemente randômicos são, em realidade, determinados. Mas isso seria insuficiente: o que se espera, aqui, é um argumento mais robusto do que uma mera narrativa de como as coisas poderiam ser.

O *determinismo virtual* (BALAGUER, 2009, seção 3) é a posição filosófica que afirma que, *para todos os efeitos*, os eventos da natureza são causalmente determinados. Não se trata, em verdade, de uma posição determinista: ela reconhece a existência das indeterminações no nível quântico. Ocorre que, segundo essa posição, todas essas indeterminações seriam irrelevantes para o que se observa no nível macro. Isso se explicaria pela Lei dos Grandes Números (LGN), da teoria da probabilidade: o fato dos eventos quânticos existirem em grande quantidade na natureza ofuscaria a indeterminação associada a cada um, pois o somatório dos efeitos de um número muito grande de eventos indeterminados tende a convergir para um efeito esperado. Mas aqui também cabe o contraexemplo da medição quântica – um evento de nível macro que não é sequer virtualmente determinado.

Esgotando o rol das possíveis posições deterministas, mencione-se, por fim, o *determinismo neural* (BALAGUER, 2009, seção 4). Ele afirma que, independente do que se passa no nível quântico, todas as ações humanas são causalmente necessitadas por eventos físico-neurais prévios. Trata-se, possivelmente, da posição predileta dos neurocientistas. Deve-se observar, entretanto, que a neurociência trata uma série de fenômenos de forma probabilística. Na área da codificação neuronal, por exemplo, avalia-se a *probabilidade* de um neurônio emitir um spike. Isso não implica, decerto, a verdade da tese indeterminista. Mas tampouco constitui uma motivação para a crença no determinismo neural.

### **Objeções ao libertismo neural**

Entre as posições incompatibilistas, vale dizer, aquelas que afirmam que há uma incompatibilidade entre livre-arbítrio e determinismo, destacam-se três grandes grupos: as posições *não causais* (que, como sugere o nome, afirmam que as decisões humanas não são causadas por nada; elas simplesmente acontecem), as de *causação por eventos* (que afirmam que as decisões humanas são causadas, mas não necessitadas, por eventos cerebrais prévios), de que é exemplo o libertismo neural, e as de *causação por agente* (que afirmam que as decisões humanas são causadas por um agente mental ou substância não física).



Se as variantes não causais têm dificuldades em lidar com o fato de que algumas decisões humanas (talvez a maior parte delas) se dão por certas razões perfeitamente identificáveis - o que seria equivalente a dizer que se dão por certas *causas* -, as teorias de causalção por eventos oferecem, no mais das vezes, concepções fragilizadas de *agência* e *controle*, centrais para o livre-arbítrio, e as teorias de causalção por agente têm a desvantagem de apelar a uma entidade pouco afeita ao vocabulário científico: o agente mental.

Nenhuma dessas objeções, entretanto, parece mais forte do que a chamada *luck objection* (BALAGUER, 2004, p. 388), velha conhecida das posições indeterministas. O libertismo neural também é instado a enfrentá-la. *Grosso modo*, a objeção afirma que o indeterminismo e a não aleatoriedade são incompatíveis, pois aquele geraria a aleatoriedade. A intuição aqui é simples: se as decisões humanas não são efeitos necessários de certos eventos físico-neurais (nem de nenhuma outra coisa), então elas simplesmente acontecem, vale dizer, acontecem *randomicamente*.

Mas a *luck objection*, se não é de todo falaciosa (pois atinge em cheio posições indeterministas mais radicais, como as não causais), tampouco é decisiva. A combinação de indeterminismo com não aleatoriedade é perfeitamente possível (e plausível) quando se lança mão de determinados arcabouços teóricos provenientes do debate mente-cérebro. Para o caso específico do libertismo neural, podem-se vislumbrar ao menos duas saídas.

A primeira seria o apelo à teoria da identidade mente-cérebro do tipo *token*, altamente prestigiada pelos filósofos da mente. Segundo essa teoria, as decisões são, ao mesmo tempo, eventos mentais e cerebrais (a rigor, a teoria *token* afirma que cada evento mental particular – uma decisão dividida, por exemplo – é equivalente a um evento físico-neural particular, que ocorre concomitantemente no cérebro). Enquanto eventos cerebrais, elas são constituídas (como todo o universo, aliás) por objetos quânticos<sup>6</sup>, que se comportam de maneira indeterminada, de acordo com as interpretações mais aceitas. Enquanto eventos mentais, as decisões são (e isso é conhecido por experiência) atos conscientes, intencionais e propositais. É o caso de se atribuir, portanto, autoria e controle à mente humana, o que implica a não aleatoriedade.

O fato de, na prática, a teoria da identidade mente-cérebro não ser neutra em relação aos dois polos sob análise, o mental e o cerebral (na verdade, a maior parte dos teóricos da identidade é reducionista, i.e., defende a substituição do vocabulário mentalista pelo vocabulário fisicalista), pode causar embaraço a quem recorre a ela para dosar as noções de indeterminismo (fortalecida pelo último) e não aleatoriedade (fortalecida pelo primeiro).

Seria mais prudente, assim, que o libertismo neural procurasse uma saída alternativa para a *luck objection*. Essa saída poderia ser facilitada, por exemplo, por uma teoria da complementaridade mente-cérebro que estabelecesse como princípios básicos a neutralidade (“as noções de mente e cérebro são igualmente relevantes para o livre-arbítrio”), o antirrealismo (“um determinado evento não é, em si mesmo, nem mental nem cerebral. O pesquisador deve tratá-lo, a cada vez, de uma ou outra maneira, a depender do aparato conceitual utilizado”) e a ênfase epistemológica (“o conhecimento atual sobre o binômio mente-cérebro, calcado nos experimentos da neurociência, tem primazia sobre

6 Estes objetos quânticos não seriam os neurônios. Neurônios têm tamanho variado (medem de 1 micron ( $10^{-6}$ m) a 1 metro). Objetos quânticos são da ordem do nanoscópico ( $10^{-9}$ m).

todas as conjecturas ontológicas, até mesmo as mais interessantes e aparentemente afeitas à ciência”). Tratar-se-ia, evidentemente, de uma adaptação da teoria homônima que Niels Bohr desenvolveu para a Mecânica Quântica. Algo nesse sentido já foi tentado pela dupla Paul Oppenheim e Nathan Brody (OPPENHEIM, BRODY, 1969).

### **Alguns resultados da neurociência**

Adicionalmente, há uma série de trabalhos em psicologia experimental e neurociência cujos resultados minariam a possibilidade da existência do livre-arbítrio, desacreditando posições como o libertismo neural. Fazemos um rápido exame desses trabalhos.

Benjamin Libet e colaboradores (LIBET et al., 1983), por exemplo, demonstraram que, em situações de escolha simples, a consciência da vontade é precedida por atividade estritamente cerebral, chamada de *potencial de prontidão* (a flexão consciente do punho direito, como consta de seu célebre experimento, seria precedida, em média, em 350ms pela vontade de movê-lo, e esta, por sua vez, seria precedida, em média, em 150ms pelo potencial de prontidão). Libet concluiu, entretanto, que o livre-arbítrio poderia continuar existindo como poder de veto (LIBET, 1999).

Daniel Wegner e Thalia Wheatley (WEGNER, WHEATLEY, 1999), a partir da análise de um conjunto de casos da psiquiatria e de experimentos da neurociência, desenvolveram a teoria da causação mental aparente, calcada nos princípios da prioridade, exclusividade e consistência. A sua ideia era associar a experiência da vontade a uma ilusão causal. Para tanto, evocaram o problema da terceira variável na análise causal, que pode ser assim enunciado: “dadas duas variáveis, A e B, nunca se pode estar certo de que A é a causa de B, pois é sempre possível que haja uma terceira variável, C, que seja a causa de ambas”. Mas desprover a mente de qualquer poder causal é tratá-la como um epifenômeno. Neste caso, não haveria razão biológica para a evolução da consciência.

Alice Isen e Paula Levin (ISEN, LEVIN, 1972) apontaram para a influência de fatores situacionais (e inconscientes), aparentemente desimportantes, na tomada das decisões humanas. Por exemplo, a dupla julgou ter demonstrado que escolhas altruístas estão relacionadas à sensação subjetiva de “sentir-se bem”. Variáveis ocultas como essa seriam centrais nas ações humanas cotidianas. O problema com essa posição, se o que se quer é tomá-la como um argumento contra o libertismo neural, está no fato de ela não dizer nada sobre as decisões conscientes, em particular as decisões divididas, com as quais propriamente se compromete aquela posição filosófica.

Finalmente, o neurocientista brasileiro Joaquim Brasil-Neto, junto a colaboradores (BRASIL-NETO et al., 1992), realizando experimentos com estimulação magnética na área motora do cérebro, conseguiu manipular decisões simples dos indivíduos pesquisados, sem que isso acarretasse a perda da percepção de vontade. O fato do livre-arbítrio ser ilusório em situações artificiais desse tipo, entretanto, não implica que ele seja também ilusório nas situações da vida normal.

### **Avaliação**

O exame realizado neste trabalho aponta para os seguintes resultados: (i) a verdade do libertismo neural depende da verdade da tese “superempírica” segundo a qual as decisões divididas são completamente indeterminadas no momento de escolha; (ii) as posições concorrentes, sobretudo as de cunho determinista, não parecem estar em posição

de vantagem em relação ao libertismo neural; (iii) as principais objeções dirigidas contra o libertismo neural podem, com algum engenho, ser respondidas; (iv) os experimentos em neurociência e psicologia experimental analisados não constituem argumentos decisivos contra o livre-arbítrio.

### **Questões suscitadas pela apresentação**

A seguir, exponho, de forma adaptada, as questões que foram suscitadas pela minha apresentação no XVIII Encontro, e as respostas que direcionei a cada uma.

Questão 1) De acordo com o libertismo neural, o indeterminismo e a não aleatoriedade não só andam juntos, como um implica o outro. Como e por que isso se daria?

Aqui, a intuição é facilitada apelando-se a cenários contrafactuais. Num universo absolutamente determinado, não haveria espaço para o livre-arbítrio. À medida que o grau de determinação diminui, aumenta o grau de liberdade. Num cenário extremo, sem nenhuma determinação, como o das decisões divididas, o grau de liberdade seria o máximo (não haveria nada, além do livre-arbítrio, a determinar a escolha do indivíduo).

É claro que a indeterminação pura e simples não implica o livre-arbítrio. Se assim fosse, as decisões randômicas deveriam ser consideradas livres. Nos casos em que as decisões são conscientes, entretanto, é forçoso reconhecer que quanto mais próximas da indeterminação, mais realçado será o livre-arbítrio, e quanto mais próximas da determinação, mais o livre-arbítrio será degradado.

Questão 2) Você acredita que há necessidade de se pensar o livre-arbítrio em termos gradualistas?

Sem dúvida. Parece-me que a questão interessante sobre o livre-arbítrio não é se ele existe ou não (a sua não existência seria a maior das trapaças de nossa mente), mas em que grau ele existe.

Particularmente, acho que todas as decisões conscientes são livres, e dificilmente uma decisão inconsciente é livre (pois é difícil lhe atribuir autoria e controle). Dentre as decisões conscientes, as mais difíceis são as mais livres (pois é maior a sua indeterminação), e as mais fáceis são as menos livres (pois é menor a sua indeterminação).

Em certo sentido, o livre-arbítrio está associado à imprevisibilidade. Se consigo prever com 100% de sucesso todas as decisões conscientes de um colega sobre um determinado assunto, posso concluir que ele age segundo razões de acesso público e, portanto, não exerce plenamente seu livre-arbítrio. Se não prevejo nenhuma de suas decisões conscientes, é porque não há nenhuma razão pública para que elas ocorram; elas simplesmente decorrem do seu livre-arbítrio.

Questão 3) Há experimentos neurocientíficos que conseguem prever, com razoável exatidão, qual dos dedos indicadores será movido pelo indivíduo pesquisado alguns milissegundos antes da realização do movimento. Se for possível prever com absoluta exatidão coisas desse tipo, inclusive os resultados das “decisões divididas”, isso significará que o livre-arbítrio não existe?

Sim. Nesse caso, não haveria livre-arbítrio, e a tese “superempírica” de que as decisões divididas são completamente indeterminadas no momento de escolha seria falsa. Isso ainda não foi demonstrado, entretanto.

Questão 4) Por que você usa o termo “plausibilidade” em vez de “compatibilidade”? Parece-me que, na melhor das hipóteses, você conseguiu demonstrar a compatibilidade do libertismo neural com a neurociência moderna, não a sua plausibilidade...

De fato, há uma diferença entre os termos “plausibilidade” e “compatibilidade”. Independente do ser humano, uma coisa pode ser compatível em relação a outra; mas algo só é plausível ou não para o ser humano. Em suma, a compatibilidade é uma noção ontológica, e a plausibilidade é uma noção epistêmica.

Neste trabalho, eu não quis avaliar se o libertismo neural é verdadeiro ou falso, pois acho que a resposta para isso não está ao nosso alcance.

Tampouco quis saber se o libertismo neural é compatível com o que se pode legitimamente derivar dos experimentos em neurociência. Isso me parece que ele é.

A minha proposta foi avaliar a sua plausibilidade, i.e., se as razões para crer nessa posição são tão boas ou melhores do que as razões para crer nas demais. E a conclusão a que cheguei foi esta: o libertismo neural não está em desvantagem em relação a nenhuma das posições concorrentes, pelo que ele pode ser considerado plausível.

**Referências bibliográficas:**

BALAGUER, Mark. A coherent, naturalistic, and plausible formulation of libertarian free will, *Nous*, v. 38, p. 379-406, 2004.

\_\_\_\_\_. Why there are no good arguments for any interesting version of determinism. *Synthese*, v. 168, p. 1-21, 2009.

BRASIL-NETO, Joaquim Pereira et al. Focal transcranial magnetic stimulation and response bias in a forced-choice task. *Journal of Neurology, Neurosurgery, and Psychiatry*, v. 55, p. 964-966, 1992.

GATTI, Fábio Garcia; PESSOA JR, Osvaldo. O debate entre as interpretações reducionista e emergentista da física, In: CELESTINO SILVA, Cibele e SALVATICO, Luis. (eds.). *Filosofia e história da ciência no Cone Sul*; seleção de trabalhos do 7º encontro AFHIC, p. 93-101, 2012.

ISEN, Alice; LEVIN, Paula. Effect of feeling good on helping: cookies and kindness. *Journal of Personality and Social Psychology*, v.21, p. 384-388, 1972.

LIBET, Benjamin et al. Time of conscious intention to act in relation to onset of cerebral activity (readiness potential); the unconscious initiation of a freely voluntary act. *Brain*, v.106, p. 623-642, 1983.

\_\_\_\_\_. Do we have free will?. *Journal of Consciousness Studies*, n. 8-9, p. 47-57, 1999.

OPPENHEIM, Paul; BRODY, Nathan. Application of Bohr's principle of complementarity to the mind-body problem. *Journal of Philosophy*, v.66, p. 97-113, 1969.

WEGNER, Daniel; WHEATLEY, Thalia. Apparent mental causation: sources of the experience of will. *American Psychologist*, v.54, p. 480-492, 1999.

# Frege and the problem of the informativity of logic

João Vitor Schmidt<sup>1</sup>

## 1. Introduction

Logic is a deductive form of reasoning: An argument is logically valid if the conclusion is supported solely by the premises. If there's more in the conclusion, then we lose the validity. In the other hand, it seems clear that we in fact gain knowledge when we do logic, *i.e.* we accept the informativity of the conclusion. Hintikka, (1970) calls this the Scandal of deduction: the fact that, although being essentially tautological, no satisfactory account has been given for the highly informational capabilities of logical truths.

The standard answer, by the Vienna Circle (influenced by Wittgenstein), assumed logic and mathematics to be tautological, thus concluding that no information could possible be carried out through deductions. Any information that we actually gain in theorem-proving is only of psychological relevance. Later, in Bar-Hillel and Carnap (1964) work on the measure of semantic information, the notion of a semantic content of a sentence was interpreted as the worlds that exclude it. Thus, tautologies were assumed to have the minimum amount of information, which keeps the Scandal of Deduction intact, whereas contradictions have the maximum amount, a problem known as The Bar-Hillel-Carnap Paradox.

There is a close connection between the informational capabilities of a logic and its decidability, the existence of mechanical procedures of decision whether a proposition is true. If there is such procedure, not knowing that some proposition logically implies other is a psychological feature, since its explained by the 'lack of our resources' to mechanically check it in advance. But we know that First-order logic is semidecidable and that Second-order logic is undecidable. Even propositional logic, which is in principle decidable (given that truth-tables are mechanical checking procedures), is not feasible in practice: the computational complexity grows quickly.

Here, I focus on Frege's version of the problem and intended solution. Frege defended that logic is informative, and *a fortiori* mathematics as well. The core idea is that Frege's logic is a second-order system, and therefore, is undecidable. Thus, it is informative if its proofs are at least sound. I claim that Frege's logic is bounded with pragmatic features that helps the creative side of deductions which are consistent with its analyticity and soundness. My steps are: first, I'll highlight Frege's remarks on the informativity of logic. Second, I'll show some aspects of Frege's logical notation that have normative or

<sup>1</sup> Universidade Estadual de Campinas - UNICAMP, CNPq



pragmatic features, *i.e.*, assertions, definitions and decompositions. Finally, I'll show how these features play significant role in proofs and analyticity.

## 2. Frege's remarks on the informativity of Logic

Frege's remarks are made in the context of two points, discussed specially in (FREGE, 1967; 1953; 1979): The fruitful role of definitions, and the defense of logic as a language (as opposed to a mere formal calculus). Frege's version of the scandal is given in kantian terms: how can analytic judgements be informative? Kant's famous answer: they can't! This was due to Kant's thesis that concepts are sets of characteristics (other concepts), and that analytic judgements are justified by conceptual containment: " $A$  is  $B$ " is analytically true just in case  $B$  is contained in, or is a sub-concept of  $A$ . Thus, Kant's conclusion was that (General) Logic, by being formal, have no content, and convey no information. Frege's goal: "to put an end to the widespread contempt for analytic judgements and to the legend of the sterility of pure logic" (FREGE, 1953, §17).

Frege's famous denial comes in (FREGE, 1953, §88):

But the more fruitful type of definition is a matter of drawing boundary lines that were not previously given at all. What we shall be able to infer from it, cannot be inspected in advance; here we are not simply taking out of the box again what we have just put into it. The conclusions we draw from it extend our knowledge, and ought therefore, on Kant's view, to be regarded as synthetic; and yet they can be proved by purely logical means, and are thus analytic. The truth is that they are contained in the definitions, but as plants are contained in their seeds, not as beams are contained in a house.

Moreover, Frege argues that fruitful definitions are those in which "[...] totally new boundary lines are drawn", where "[...] we use old concepts to construct new ones, but in so doing we combine the old ones together in a variety of ways by means of the signs for generality, negation and the conditional. (FREGE, 1979, p.34).

The other point where we find Frege's remarks on the informativity of logic is in his defense about the contentful side of logic. Against the algebrists of logic like Boole, Frege says that: "Right from the start I had in mind the expression of a content. What I am striving after is a *lingua characterica* in the first instance for mathematics, not a *calculus* restricted to pure logic" (FREGE, 1979, p.12). On this, Van Heijenoort (1967) famously distinguished the universalists of logic (those who see logic as a language) and the algebrists (those who see logic as a calculus), identifying Frege with the former. In this scenario, Frege's logic is a quantification theory for first and second-order and an universal language with an all-inclusive subject matter: a fully-interpreted language, in Shapiro (1991) words, or a model of 'one-world doctrine' in Hintikka's (1997). Thus, Frege sees logic as contentful, with an unique and fixed interpretation.

### 3. Assertion

Given that, for Frege, logic is a language, it must properly express contents that are true. For this matter, assertions are important. In the *Begriffsschrift* (FREGE, 1967), Frege presented a bidimensional logical notation, concerned with the expression of contents and their logical relations. Among its primitives, there was a sign for the expression of a content,

$\neg A$ ,

known as the content-stroke. It's meant to be read as "the circumstance of  $A$ ". Later in the *Grundgesetze der Arithmetik* (FREGE, 2013), following the sense-reference distinction and the introduction of truth-values as objects, the content-stroke was abandoned in exchange for the horizontal: a total function ranging from objects to truth-values, where it reads "the truth-value of  $A$ ". In both cases, when a content prefixed by the content-stroke or the horizontal is asserted or judged, Frege uses a vertical line on the left, as

$\vdash A$ ,

to indicate that the content, or the truth-value, of  $A$  is affirmed or recognized as denoting the True. This is known as the judgement-stroke, the grandfather of the modern consequence relation symbol. The main reason for adopting such sign for assertion or judgement is to sharply separate the assertion from the asserted content. In the *Grundgesetze*, where Frege have already distinguished the sense (*sinn*) from the reference (*bedeutung*) of a sign, the judgement-stroke serves as a way of connecting both entities of a given expression, linking a sense, the thought of a sentence, with a reference, the True. In the classic paper *Über sinn und bedeutung* (FREGE, 1984, p.164-165), Frege writes that "Judgements can be regarded as advances from a thought to a truth-value", and in the *Grundgesetze*, he defines: "I distinguish the judgement from the thought in such a way that I understand by a judgement the acknowledgement of the truth of a thought" (FREGE, 2013, §5).

In the natural language, the confusion between assertion and what is asserted is due to the fact that assertion is encapsulated in the form of an assertoric sentence. In the later article *Der Gedanke* (FREGE, 1984, p.356n), Frege alerts to this: "It seems to me that thought and judgement have not hitherto been adequately distinguished. Perhaps language is misleading. For we have no particular bit of assertoric sentences which corresponds to assertion; that something is being asserted is implicit rather in the assertoric form".

What is present in the assertoric form that marks that an assertion is being made is the presence of the assertoric force. What distinguishes the sentences "Snow is white" from "Is snow white?" is that the first is uttered as an assertion, hence with assertoric force, while the second is uttered as a question, hence with interrogative force. We can even have the same sentence uttered without the corresponding assertoric force, like when an actor utters something in a play. In this case, he does not aim at the True, therefore, no assertion is made.

As Frege intends, a more perfect language must be able to express a thought without asserting it, therefore, separating the sentence from the assertoric force. That's why such a language needs a judgement-stroke sign. The important point to be made with the presence of the assertoric force, which is marked by the judgement-stroke, is that assertion is an act, and because of that, logic has a distinctive normative role. Frege declares that "From the laws of truth there follow prescriptions about asserting, thinking, judging, inferring. And we may very well speak of laws of thought in this way too." (FREGE, 1984, p.351). Logic is concerned with the laws of truth, and from such laws, and the generality they possess, we derive prescriptive norms on how to think in accordance with such laws. In the unpublished

paper *Logik*, Frege notes the following:

Like ethics, logic can also be called a normative science. How must I think in order to reach the goal, truth? We expect logic to give us the answer to this question, but we do not demand of it that it should go into what is peculiar to each branch of knowledge and its subject-matter. On the contrary, the task we assign logic is only that of saying what holds with the utmost generality for all thinking, whatever its subject-matter. We must assume that the rules for our thinking and for our holding something to be true are prescribed by the laws of truth. The former are given along with the latter. Consequently we can also say, logic is the science of the most general laws of truth. (FREGE, 1979, p.128)

How then, is this expressible in the concept-script language? Through the presence of judgements, hence, the assertoric force of the judgement-stroke. The assertoric force is what is needed to make use of logical laws, or the so-called “laws of truth” in judging. This point is made by Taschek (2008), according to which Frege’s emphasis in the assertoric force is to highlight the normative status of logic. Greimann (2014) develops an alternative reconstruction, arguing that the assertoric force determines the important passage of the level of senses (thoughts) to the level of references (truth-values), something not achievable by the simple predication of truth. Thus, the judgement of a thought as true is not a description of its truth (a predication).

Frege’s judgement-stroke is a sign different from every other in the language. It is not a function and does not have a reference. He points that out in the paper *Funktion und Begriff*: “The assertion sign cannot be used to construct a functional expression; for it does not serve, in conjunction with other signs, to designate an object. ‘ $2+3=5$ ’ does not designate anything; it asserts something” (FREGE, 1984, p.149n). Hence, assertion has a different role in Frege’s logic, that is not a merely descriptive one, but one that is, nonetheless, part of the language.

In this case, we ought to read the expression ‘ $2+3=5$ ’ as an illocutionary act, “that ‘ $2 + 3 = 5$ ’ names the True”, but this reading only makes justice if something is present: the aforementioned assertoric force. Finally, as Greimann (2014) also highlights, this characterization of logic is still a characterization in terms of “truth”, but one that requires the act of assertion as necessary to connect both levels of a concept-script expression (thoughts with truth-values). In sum, logic has a normative role in providing how we ought to think, by offering rules, or laws, on how to rightfully aim at the truth in asserting.

Yet, another reason why logic is in need of a judgement stroke sign is to fill out a defective aspect of language, either natural or formal: the fact that the truth-predicate cannot be defined within language itself without some circularity, as Frege pointed out in *Der Gedanke*. The lesson here is the following: the word “true” represents both the goal of logic and also his major defect. It’s an ineffective predicate, non-definable within language, but one that still marks what logic is about. Only through assertoric force, that is, by taking an illocutionary act into logic, one can unveil what truth is about, by making tacit, *viz.* normatively, its general laws, as it is required in thinking. Thus, Frege is apparently giving a pragmatic solution to a descriptive problem. Assertion has an important role for logic in Frege’s eye precisely because it is not a descriptive (functional) symbol.

With such symbol in the language, axioms are those self-evident asserted truths, e.g.  $A \rightarrow (B \rightarrow A)$ . Rules of inference are prescriptions for assertability: From the assertion of  $A$ , that is  $A$ , and the assertion that  $A$  implies  $B$ , that is  $(A \rightarrow B)$ , we are entitled to assert  $B$ ,

that is, *B*. Finally, Frege sees an inference as a sequence of asserted truths (FREGE, 1980, p.79) that follows from prescriptive laws and other asserted truths. Therefore, the presence of assertion is essential in marking logic as informative for Frege in a twofold way. First, It is essential for marking logic as contentful, giving that only true thoughts can properly inform something. Following the kantian tradition, Frege calls judgements as informative, meaning that they are the “carriers” of information, since it’s the connection between the realm of senses with the realm of references that provides discoveries, precisely by asserting propositions as true. Pedriali argues in this direction as well, stating that assertion is what separate logic from a mere “[...] play with symbols” (PEDRIALI, 2017, p.8). Moreover, he argues, it is by taking assertion into account that logic can differentiate “[...] between sentences that make genuine claims about a scientific subject-matter and those that concern fictional domains” (PEDRIALI, 2017, p.8-9). Second, the assertion sign, and the logic that goes with it, allows deductions to be sound, meaning that asserting something as true is as objective and independent as one desire, avoiding any psychological worries into logic. All this have the practice of logic in mind, by avoiding gaps in proofs, and by separating the assertoric force from the asserted content.

#### 4. Stipulative and Fruitful definitions

The other important notion that marks logic as informative, in Frege’s philosophy, is definitions. Frege has a stipulative notion of definitions, offered by the sign

$$(\varphi=\psi),$$

which stipulate that  $\varphi$  and  $\psi$  have the same content, in (FREGE, 1967), or sense and reference in (FREGE, 2013).  $\psi$  is taken as a new sign, while  $\varphi$  is some expression of the language. These definitions are non-creative. All that follows from it, follows also without it (FREGE, 1967, §24). Their introduction is motivated by pragmatic reasons: to help one perform a proof:

Our sole purpose in introducing such definitions is to bring about an extrinsic simplification by stipulating an abbreviation. They serve besides to emphasize a particular combination of signs in the multitude of possible ones, so that our faculty of representation can get a firmer grasp of it. (FREGE, 1967, §24)].

This is an important point. Even though every sign in the concept-script must objectively have a sense and reference, the manipulation of such language still requires some active part from who manipulates them. Hence, even if definitions have no logical role in proofs, it still have some utility. In an unpublished paper, Frege writes the following passage:

To be without logical significance is still by no means to be without psychological significance. When we examine what actually goes on in our mind when we are doing intellectual work, we find that it is by no means always the case that a thought is present to our consciousness which is clear in all its parts. [...] Our minds are simply not comprehensive enough. We often need to use a sign with which we associate a very complex sense. Such a sign seems, so to speak, a receptacle for the sense, so that we can carry it with us, while being always aware that we can open this receptacle should we have need of what it contains. It follows from this that a thought, as I understand the word, is in no way to be

identified with a content of my consciousness. If therefore we need such signs — signs in which, as it were, we conceal a very complex sense as in a receptacle — we also need definitions so that we can cram this sense into the receptacle and also take it out again. So if from a logical point of view definitions are at bottom quite inessential, they are nevertheless of great importance for thinking as this actually takes place in human beings. (FREGE, 1979, p.209)

Therefore, stipulative definitions have no pure logical use, but only a pragmatic role in the part of the human mind who performs such proofs. It is not required, logic-wise, to define a sign more simply in order to a theorem *to be provable* about it. But it is required if this same theorem is *to be proved* by anyone who manipulates the concept-script language. This shows that stipulative definitions are pragmatically motivated.

The question remains on how can such definitions be fruitful, as we mentioned. Frege famous ancestral definition in part III of (FREGE, 1967), which is an stipulation, is an example of fruitful definition. What are fruitful definitions? Simply put, are definitions with ampliative consequences that are shown in the context of a proof:

[...] definitions must show their worth by their fruitfulness: it must be possible to use them for constructing proofs. (FREGE, 1953, p.xxi)

Definitions show their worth by proving fruitful. (FREGE, 1953, §70)

For fruitfulness is the acid test of concepts, and scientific workshops the true field of study for logic. (Frege, 1979, p.33)

[...] the usefulness of such formulae only fully emerges when they are used in working out inferences, and we can only fully appreciate their value in this regard with practice. (Frege, 1979, p.27)

So how can they be shown fruitful? Enters decompositions. Against Kant's way of forming concepts out of constituents, Frege “[...] think[s] of a concept as having arisen by decomposition from a judgeable content” (FREGE, 1980, p.101). For example, from “ $2^4 = 16$ ”, one can decompose (or dissect) to get new functions: “ $2^x = 16$ ”, which is “logarithm of 16 to the base 2”, or “ $x^4 = 16$ ” which reads “4th root of 16” (FREGE, 1967, §9). This is crucial for stipulations: given the definition of the Ancestral, from  $R^*(a, b)$  (that “ $b$  follows  $a$  in the  $R$ -series”), one decompose to achieve  $R^*(a, x)$  (“ $a$ 's descendent”). This is also crucial in proving that the definition  $R^*$  defines a strict or total order (theorems 98 and 133).

Boolos (1985) showed that decompositions are equivalent to comprehension principles for second-order logic. In this same sense, Landini (2012, p.136) argues that “Decompositions is [...] all that is required for informativity”. But since comprehension principles are defined in second-order logic, there is no decision procedure to “acquire” such functions/concepts mechanically. So, (agreeing with Landini), such decompositions are performed outside the language. But in Frege's case (I add), decompositions can be motivated pragmatically. As Frege says:

The distinction [between functions and arguments] have nothing to do with the conceptual content; it comes about only because we view the expression in a particular way (FREGE, 1967, §9)<sup>2</sup>

<sup>2</sup> The emphasis is mine.



There is nothing in the language itself to specify which decomposition is to be performed, or which function is to be acquired. Frege hints that by speaking about the way we *view*, or the way we *imagine*: “If we *imagine* that an expression can thus be altered [...]” (FREGE, 1967, §9)<sup>3</sup>. This implies that, at some level, there’s a creative side in deductions relying in the agent side, that is, the one who performs the deductions. Dummett (1991, p.42) argues in the same way:

Deductive reasoning [...] has a creative component, involving the apprehension of patterns within the thoughts expressed [...]. Since it has this creative component, a knowledge of the premises of an inferential step does not entail a knowledge of the conclusion [...] and so deductive reasoning can yield new knowledge. Since the relevant patterns need to be discerned, such reasoning is fruitful; but, since they are there to be discerned, its validity is not called in question.

Such patterns, as Dummett calls it, are those functions acquired by decompositions. Adding this to stipulative definitions, we can conclude that they helps one grasping functions and proving theorems *about* the definitions that are not just contained in them.

## 5. Proofs and Analyticity

I finally link such features with the validity (or the analyticity) of the proofs. I’ve already mentioned one version of analyticity, the classical formulation due to Kant. Kant would hardly accept Frege’s proofs as analytical, since decomposition and comprehension could be taken as intuitive features: what does license us to accept and acknowledge the obvious existence of such functions or concepts declared by such principles, if not by some form of intuition?<sup>4</sup> Even if some form of intuition is a necessary condition for visualizing the existence of such functions (the ones acquired by decomposition), it is not clear how the logical validity of a proof featuring it would be affected, at least in the fregean scheme of things. If we do accept the universalist reading of Frege’s logic, then the realm of functions (or likewise, concepts) is an objective fact. They are things that we *grasp*, never construct: “[...] a concept is something objective: we do not form it, nor does it form itself in us, but we seek to grasp it, and in the end we hope to have grasped it, though we may mistakenly have been looking for something where there was nothing” (FREGE, 1984, p.133). Therefore, if every step in a proof is correctly judged as true, the instantiation of an axiom or theorem with some acquired functional expression is justified by the general character of such laws, simply because, if the law holds generally, it holds for a given function.

Frege declared that analyticity is not a feature of contents, as in Kant, but justification:

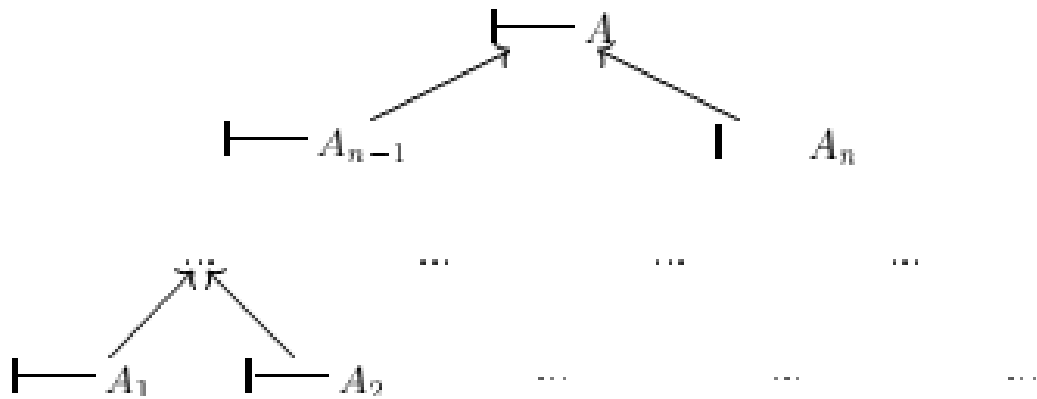
The problem becomes, in fact, that of finding the proof of the proposition, and of following it up right back to the primitive truths. If, in carrying out this process, we come only on general logical laws and on definitions, then the truth is an analytic one. (FREGE, 1953, §3)

A proof in Frege’s system is a deductive chain with an strict order. Each point is an asserted proposition according to the rules of logic:

<sup>3</sup> Both emphasis are mine.

<sup>4</sup> Boolos (1985) was the one to mentioned such kantian objection.





Each pair of arrows denotes a valid deduction, via *modus ponens*, from a pair of asserted truths. If the base of the deduction is composed only of axioms or definitions, then  $A$  is an analytic judgement. In many points, decompositions are allowed simply because, in Frege’s universalist conception of logic, the generality of a logical law allows one to instantiate for any function; *e.g.*,  $(x)f(x) \rightarrow f(a)$  holds for any function  $f$ . Thus, decomposing from a definition preserves analyticity.

## 6. Conclusion

As we saw, in Frege’s logic, validity is explained in terms of assertability, not in terms of truth-preservation. A proof is logically valid, or analytically valid, if it follows from axioms (the general logical laws) and definitions by deductive rules of logic. These rules, axioms and definitions include normative and pragmatic features: axioms and logical truths are *asserted* truths, definitions are *declared* as such and decompositions are visualized and performed outside the language as a way of *grasping* things.

Consider Theorem 98 in (FREGE, 1967) as an example:  $[R^*(x,y) \ \& \ R^*(y,z)] \rightarrow R^*(x,z)$ . This is a consequence of the way  $R^*$  is defined plus the way it is combined (with decompositions) with logical laws and other asserted truths. It is, nevertheless, an ampliative analytic judgement, since theorem 98 is proved with the aid of something not reducible to a mechanical procedure (decompositions). In all this, we can see how Frege’s logic is designed for the human practice, it is a language thought to help one carrying a proof over, with pragmatic components. This language avoids psychological or intuitive reasoning, simply because the assertability condition guarantees the soundness of the deduction. Finally, the proofs are consistent with the undecidability of second-order logic. Thus, it seems to connect informativity with deductive reasoning, answering the scandal of deduction. As Frege have declared it (although in a slightly different context): “[...] our thinking as a whole can never be coped with by a machine or replaced by purely mechanical activity” (FREGE, 1979, p.35)

**Referências bibliográficas:**

- BOOLOS, G. Reading the Begriffsschrift. *Mind*, v. 94, n. 375, p. 331–344, 1985.
- CARNAP, R.; BAR-HILLEL, Y. *An outline of a theory of semantic information*. In: BAR-HILLEL, Y. (Ed.). *Language and information : selected essays on their theory and application*. Reading: Addison-Wesley Publishing, 1964.
- DUMMETT, M. *Frege's Philosophy of Mathematics*. London: Duckworth, 1991.
- FREGE, G. *The Foundations of Arithmetic*. A logico-mathematical enquiry into the concept of number. Oxford: Blackwell, 1953. Translated by J. L. Austin.
- \_\_\_\_\_. Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In: HEIJENOORT, J. V. (Ed.). *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*. Cambridge: Harvard University Press, 1967. Translated by Stefan Bauer-Mengelberg.
- \_\_\_\_\_. *Posthumous Writings*. Oxford: Basil Blackwell, 1979. Edited by Hans Hermes, Friedrich Kambartel and Friedrich Kalbach. Translated by Peter Long and Roger White.
- \_\_\_\_\_. *Philosophical and Mathematical Correspondence*. Chicago: Basil Blackwell, 1980. Translated by Hans Kaal.
- \_\_\_\_\_. *Collected Papers on Mathematics, Logic and Philosophy*. New York: Basil Blackwell, 1984.
- \_\_\_\_\_. *Basic Laws of Arithmetic*. Oxford: Oxford University Press, 2013.
- GREIMANN, D. Frege on truth, assertoric force and the essence of logic. *History and Philosophy of Logic*, Taylor & Francis, v. 35, n. 3, p. 272–288, 2014.
- HINTIKKA, J. Information, deduction, and the a priori. *Noûs*, v. 4, n. 2, p. 135–152, 1970.
- HINTIKKA, J. Is truth ineffable? In: HINTIKKA, J. (Ed.). *Lingua Universalis vs. Calculus Ratiocinator: An ultimate presupposition of twentieth-century philosophy*. Dordrecht: Kluwer Academic Publishers, 1997.
- HEIJENOORT, J. V. Logic as calculus and logic as language. *Synthese*, Springer, v. 17, n. 3, p. 324–330, 1967.
- LANDINI, G. *Frege's Notation: what they are and how they mean*. Hampshire, England: Palgrave Macmillan, 2012.
- PEDRIALI, W. B. The logical significance of assertion: Frege on the essence of logic. *Journal for the History of Analytical Philosophy*, v. 5, n. 8, p. 1–22, 2017.
- SHAPIRO, S. *Foundations without Foundationalism: a case for second-order logic*. first. Oxford: Oxford University Press, 1991.
- TASCHEK, W. W. Truth, assertion, and the horizontal: Frege on 'the essence of logic'. *Mind*, v. 117, n. 466, p. 375–401, 2008.

# O cérebro como agente e a mente consciente como guia de ação

Jonas Gonçalves Coelho<sup>1</sup>

## 1. Introdução

Tenho argumentado, a partir dos exemplos da consciência visual e da consciência de dor, que o cérebro<sup>2</sup> é o agente e que a mente consciente, por ele produzida em sua interação com o corpo no qual está incorporado e com o ambiente externo ao corpo no qual está situado, é seu guia de ação. Nesse sentido, penso que o agente não é uma pessoa cartesiana, ou seja, uma substância consciente distinta e separável do cérebro, do qual, não se sabe como, recebe estímulos, transformando-os em conteúdos conscientes e sobre o qual, também não se sabe como, age e, a partir dele, com as outras partes do corpo e, através deste, com o ambiente externo ao corpo. Ao contrário, considero que a pessoa, com suas sensações, desejos, emoções, crenças, intenções, memórias, imaginações etc. *conscientes* é produzida pelo cérebro a partir de sua interação com seu corpo e com o ambiente externo ao corpo, sendo aquela utilizada pelo cérebro como guia para suas ações conscientes.

Para desenvolver a hipótese de que o cérebro é o agente com vistas a oferecer um esboço explicativo de como ele próprio implementa suas ações, utilizando a mente consciente como guia, privilegiarei no presente texto o modo como alguns filósofos e neurocientistas abordam a fenomenologia da consciência relacionada à ação e, a partir daí, a estrutura e funcionalidade cerebral que a suporta. Procurarei, portanto, primeiramente, apresentar os termos gerais de uma abordagem fenomenológica do *sense of agency*, que aqui traduzirei como “consciência de agir”, a seguir, um esboço de uma abordagem neurofisiológica relacionada à primeira e, por fim, indicar os caminhos para uma abordagem fenomenológica e neurofisiológica mais ampla, a qual permita uma compreensão mais abrangente da relação entre cérebro e mente consciente, sendo o primeiro a causa imediata da segunda e a segunda uma propriedade e guia de ação.

## 2. Consciência de agir: fenomenologia

Sukhvinder Obhi, no artigo de 2011, *Sense of agency and intentional binding in joint action*, argumenta que uma compreensão completa da consciência de agir e dos fatores individuais e contextuais que a afetam é uma das principais tarefas para uma neurociência que se proponha a “iluminar os motivadores da experiência humana e as bases dos sentimentos subjetivos” (OBHI, 2011, p. 655). O autor destaca, neste mesmo

1 Universidade Estadual Paulista – UNESP

2 Utilizo o termo “cérebro” pelo fato de este ser consagrado na literatura filosófica, mas, a rigor, estou me referindo ao “encéfalo”, o qual inclui o tronco cerebral e o cerebelo.

artigo, o fato de que, embora a “experiência de agir ou de autoria da ação” seja “universal e extremamente forte”, ela não é algo sobre o qual as pessoas rotineiramente param para pensar, a não ser em situações nas quais as coisas não acontecem conforme o esperado (OBHI, 2011, p. 655). Por exemplo, “quando aciono o interruptor para acender uma lâmpada, apenas ‘sei’ que a luz acendeu por causa de minha ação – que eu sou o autor da ação e de seu conseqüente efeito”, apenas pensando mais detidamente sobre causa e efeito quando “algo inesperado acontece, como, por exemplo, a luz não acende” (OBHI, 2011, p. 655). Contrariando a atitude natural de desatenção em relação à ação, e indo além da atitude de apenas tentar entender o motivo de uma ação não ter produzido o resultado esperado, alguns filósofos apresentam uma proposta diferente, como Tim Bayne, que, no artigo de 2008, *The phenomenology of agency*, propõe desenvolver uma “fenomenologia do agir”, ou explicitar “como é ser um agente”, o que significa apontar “os estados fenomenais associados ao agir” (BAYNE, 2008, p. 182).

Uma abordagem fenomenológica do agir faz sentido na medida em que, como diz Shaun Gallagher, no artigo de 2007, *The natural philosophy of agency*, a noção de agir está intrinsecamente ligada à consciência de que se é um agente: “o agir depende da consciência de agir do agente” (GALLAGHER, 2007, p. 347). Se eu causo algo e não tenho consciência de que sou eu que o está causando, eu não sou verdadeiramente um agente, embora seja causa do que está acontecendo, estando, assim, numa condição semelhante à de um furacão ao causar uma pane num sistema elétrico: “Um furacão pode causar uma pane no sistema elétrico, mas não atribuiríamos ação ao furacão, considerando-se o uso normal do termo” (GALLAGHER, 2007, p. 347). Eu poderia também, através de meu corpo, causar um acontecimento sem que eu me sinta como um agente, ou seja, sem que eu tenha consciência de agir, embora, diferentemente do furacão, eu tenha a consciência de posse do meu corpo e de seu papel causal desempenhado: “É possível dizer que eu estou me movendo e, conseqüentemente, que este é *meu* movimento e, deste modo, ter uma consciência de posse, em casos onde não há consciência de agir em relação ao movimento, como, por exemplo, nos movimentos reflexos ou involuntários” (GALLAGHER, 2007, p. 347).

Assim como no caso dos movimentos involuntários ou reflexos acima citados, atribuíveis a processos físicos independentes de nossa consciência de agir, poderíamos também não ter consciência de agir em situações nas quais, embora tenhamos consciência de posse de nosso corpo, atribuímos as ações não às nossas próprias intenções, mas às intenções de um outro ser intencional que supostamente age sobre nós, considerando-o, assim, como o verdadeiro agente. É o caso, por exemplo, de certos delírios esquizofrênicos, nos quais, segundo Graham e Stephens, citados por Gallagher, “os sujeitos consideram que seus pensamentos [ou ações] são inexplicáveis em termos de crenças sobre seus estados intencionais” (GALLAGHER, 2007, p. 348). Daí Nicole David, em seu artigo de 2012, *New frontiers in the neuroscience of the sense of agency*, definir a consciência de agir como a “consciência de que eu sou o autor de minhas próprias ações” e a consciência de que sou eu “que está causando ou controlando um movimento ou mudança no mundo externo” (DAVID, 2012, p. 1). Nesse mesmo sentido, Patrick Haggard, no artigo de revisão publicado em 2017, *Sense of agency in the human brain*, define consciência de agir como “o sentimento de fazer algo acontecer” como “a experiência de controlar os próprios atos motores e, através deles, o curso dos eventos externos” (HAGGARD, 2017, p. 2). Da mesma forma, Davide Crivelli e Michela Balconi, num artigo também recente,

publicado em 2017, *The agent brain: a review of non-invasive brain stimulation studies on sensing agency*, definem consciência de agir como “a consciência de que sou eu que está causando uma ação, seja ela física ou mental, e causando seus efeitos” (CRIVELLI; BALCONI, 2017, p. 2).

Gallagher considera que, embora, como ilustrado acima, a consciência de agir e a consciência de posse do corpo não sejam noções coextensivas, no caso das ações incorporadas, os movimentos corporais “contribuem para a consciência de agir, a qual seria muito diferente sem o sentimento de natureza incorporada da ação” (GALLAGHER, 2007, p. 347). Mas Gallagher, que propõe uma “abordagem de múltiplo aspecto da consciência de agir”, também defende que esta não se reduz aos processos incorporados, visto envolver outros aspectos, tais como uma intencionalidade altamente pragmática, ou seja, nossas atenções e ações são dirigidas “*no mundo*, no projeto ou tarefa no qual estamos engajados” (GALLAGHER, 2007, p. 354). Assim, “uma forma de *feedback* intencional, o qual não é o *feedback* aferente de nossos movimentos corporais, mas uma consciência perceptiva de que minha ação está produzindo um efeito, deve contribuir para a consciência de agir” (GALLAGHER, 2007, p. 354). Daí a complexidade da consciência de agir, ainda que apenas considerada num nível mais básico: “ela é complexa porque é o produto de vários elementos: sinais eferentes, *feedback* sensorial (aferente) e *feedback* intencional (perceptivo)” (GALLAGHER, 2007, p. 355).

Ao apresentar e justificar os componentes fenomenais constitutivos da consciência de agir, Davide Crivelli e Michela Balconi, no artigo já citado de 2017, começam por dizer que “a autoatribuição de um papel primário de agente em relação a nossas ações (desde suas manifestações mais simples como atos motores dirigidos a fins às manifestações mais complexas como os raciocínios morais e julgamentos reflexivos concernentes a responsabilidades)” está intimamente relacionada à consciência do “poder causal que podemos exercer sobre o mundo externo” (CRIVELLI; BALCONI, 2017, p. 2). Daí as noções de intenção, desejo e finalidade, por estarem intimamente relacionadas às motivações e resultados de nossas ações, serem inseparáveis de nossa consciência de agir, assim como a básica experiência de nosso corpo como sendo nosso e, ao mesmo tempo, instrumento de nossa ação no mundo.

O agir é, então, um fenômeno complexo influenciado pela intenção, finalidades e desejos, mas também fundado em sinais somatosensoriais e fluxo de informação sensorial aferente. A integração de diferentes níveis é necessária para o surgimento de uma experiência completa. Imagem corporal e estrutura intencional emergem para gerar agentes unificados capazes de perceber, influenciar e exercer poder causal sobre os outros e sobre o ambiente, entendido como um sistema complexo incluindo tanto os objetos com os quais agimos e os sujeitos com os quais interagimos. (CRIVELLI; BALCONI, 2017, p. 2)

Tendo apresentado alguns dentre os principais componentes da fenomenologia da consciência de agir, destacados pela literatura filosófica, passo, a seguir, a tratar de seus correlatos encefálicos estruturais e funcionais, privilegiados na literatura neurocientífica, buscando, a partir daí, entender como o cérebro exerce a ação tomando sua mente consciente como guia.



### 3. Consciência de agir: neurobiologia subjacente

O pressuposto que assumo neste texto é o de que, a partir da descrição da experiência consciente é possível, através do monitoramento e manipulação da atividade cerebral, identificar as suas bases neurobiológicas – embora muitos estudos neurocientíficos utilizem, para esse fim, imagens de ressonância magnética funcional (fMRI), em relação à consciência de agir ter privilegiado técnicas de manipulação direta do cérebro, como é o caso das técnicas não-invasivas de manipulação cerebral (“the non-invasive brain stimulation techniques (NIBS)”). Dentre essas técnicas, que segundo Davide Crivelli e Michela Balconi, podem fornecer “*insights* sobre o papel causal específico de diferentes estruturas corticais” (CRIVELLI; BALCONI, 2017, p. 1), destacam-se os “métodos eletrofisiológicos e de neuroimagem”, os quais “permitem qualificar e quantificar atividades neurais que se sucedem durante tarefas implícitas ou explícitas, e compará-las com experiências subjetivas simultâneas, performances cognitivas ou comportamentais” (CRIVELLI; BALCONI, 2017, p. 2).

Essas técnicas, em especial a estimulação magnética transcraniana (“transcranial magnetic stimulation (TMS)”) e a estimulação elétrica transcraniana (“transcranial electrical stimulation (tES)”)³, possibilitam não apenas identificar, por um processo associativo, os correlatos cerebrais da experiência consciente, isto é, “estabelecer os correlatos anatômico-funcionais das funções e processos investigados por meio de associação” (CRIVELLI; BALCONI, 2017, p. 2). Elas também permitem interferir sobre a atividade cerebral, produzindo experiências específicas: “métodos de estimulação não-invasiva podem ser listados entre as técnicas de interference ou causais, as quais têm a vantagem de estabelecer a causação neural e o papel efetivo de estruturas neurais que suportam ou modulam um processo ou função específica” (CRIVELLI; BALCONI, 2017, p. 2). Daí ser possível “fornecer uma compreensão melhorada da base de nossa habilidade *versus* inabilidade de agir propriamente em contextos sociais complexos” (CRIVELLI; BALCONI, 2017, p. 1) e “de nossa habilidade *versus* inabilidade de sentir e julgar a autoria de nosso comportamento e de seus resultados” (CRIVELLI; BALCONI, p. 2).

A partir, portanto, de vários estudos envolvendo técnicas imagéticas e técnicas de manipulação não-invasiva de estruturas cerebrais, têm sido destacados alguns correlatos neurais mediadores dos diferentes aspectos do agir citados na seção anterior, ou seja, intenções, finalidades, desejos, sinais somatosensoriais, fluxo de informação sensorial aferencial. Acredita-se que os correlatos neurais identificados “espelham a complexidade do fenômeno” (CRIVELLI; BALCONI, 2017, p. 2), visto estarem “associados a atributos específicos da experiência de agir, ou a passos específicos do processo que conduz à consciência do agir” (CRIVELLI; BALCONI, 2017, p. 2). De fato, o que os estudos identificam são áreas específicas constitutivas de circuitos neurais mais amplos, dentre elas, o córtex pré-frontal dorsolateral (dlPFC), o córtex cingulado (CC), as áreas motoras suplementar e pré-suplementar (SMA e pre-SMA), o córtex parietal posterior (PPC) e suas regiões inferiores e, por fim, o cerebelo (CRIVELLI; BALCONI, 2017, p. 2). Considerarei, resumidamente, a seguir, o modo como essas áreas têm sido relacionadas à consciência de agir.

Ao córtex parietal confere-se “um papel crítico na integração e comparação do fluxo de informação sensório-motora para atribuir corretamente os movimentos ao agir”

3 Mais informações a respeito do uso e funcionamento dessas técnicas podem ser encontradas no mesmo artigo na página 3.



(CRIVELLI; BALCONI, 2017, p. 4). Por exemplo, a estimulação magnética transcraniana *offline* repetitiva de baixa frequência do córtex parietal posterior esquerdo, “um protocolo de estimulação que se acredita ter efeitos inibitórios na atividade das porções estimuladas dos tecidos corticais”, prejudica “a avaliação da assincronia entre um movimento e seu *feedback* visual, especificamente quando o movimento é ativo e voluntário” (CRIVELLI; BALCONI, 2017, p. 4), o que implica que a avaliação correta da relação entre um movimento voluntário e seu *feedback* visual é um componente importante da consciência de agir. Já a estimulação magnética transcraniana *online* repetitiva de alta frequência do córtex parietal inferior direito produz um resultado diferente, o qual talvez interfira negativamente na consciência de agir, ou seja, ela “induz pacientes saudáveis a uma percepção equivocada de seu papel de agente durante movimentos autocontrolados experienciando-os como sendo perturbados externamente, mesmo não sendo esse o caso” (CRIVELLI; BALCONI, 2017, p. 4).

Ainda em relação ao papel do córtex parietal na consciência de agir, o uso de estimulação magnética transcraniana não-repetida, ou seja, de pulso único, tem permitido mostrar que “quando a atividade do lobo parietal inferior direito é alterada temporariamente e focalmente, participantes saudáveis tendem a atribuir, erroneamente, o agir a fontes externas, a despeito do fato de os *feedbacks* visuais perceptíveis serem ou não consistentes com as ações executadas” (CRIVELLI; BALCONI, 2017, p. 4). Tem sido também assinalada a “importância das conexões parietal/pré-motora para resolver os conflitos sensório-motores e finalizar a atribuição do agir”, ou seja, o uso do protocolo de estimulação por pulso pareado permitiu observar a “conectividade parietal-motor inibitória durante o repouso e durante condições ilusórias induzindo tanto à consciência de agir quanto à de posse [do corpo]” (CRIVELLI; BALCONI, 2017, p. 4). Esta mesma técnica “tem sido recentemente utilizada para explorar o papel da área parietal inferior esquerda nos componentes prospectivos da consciência de agência, ligados à seleção da ação e processos de programação independentemente dos efeitos da ação” (CRIVELLI; BALCONI, 2017, p. 4). Nesse sentido, tem sido “mostrado que a perturbação da atividade na região parietal inferior esquerda, no momento da seleção e execução da ação, reduz o controle percebido dos participantes sobre os efeitos subsequentes da ação” (CRIVELLI; BALCONI, 2017, p. 4).

Como dito na seção anterior, a consciência de agir também envolve a consciência de posse do próprio corpo, ainda que não se confunda com ela e a área parietal desempenha também aí um papel importante. Por exemplo, estudos sobre a integração visual-tátil-propioceptiva mostram o envolvimento das áreas parietal inferior e visual mais alta na “correta autoatribuição de posse do corpo” (CRIVELLI; BALCONI, 2017, p. 5). Curiosamente, “o paradigma experimental mais utilizado para investigar os correlatos fenomenológicos e neurais da posse do corpo é a ilusão da mão de borracha”<sup>4</sup> (CRIVELLI; BALCONI, 2017, p. 5). A técnica de estimulação magnética permite “fortalecer a experiência ilusória pela aplicação de interferência inibitória sobre a área corpórea extra-estriada”, área esta “envolvida com a integração multimodal definidora da experiência de posse do corpo” (CRIVELLI; BALCONI, 2017, p. 4). Já a estimulação inibitória repetitiva de baixa frequência aplicada

4 O experimento consiste em colocar as duas mãos sobre uma mesa com uma divisória, posicionando uma mão de borracha paralelamente a uma das mãos verdadeiras, deixando a outra mão verdadeira fora do campo de visão, escondida do outro lado da divisória. A seguir, as duas mãos visíveis, uma mão real e a outra de borracha são estimuladas sincronicamente, por exemplo, por toques ou massagens repetitivos contínuos. A partir de um determinado momento, o participante começa a sentir o estímulo na mão de borracha como sente na mão real, como se a primeira também fosse parte de seu corpo, ou seja, ele incorpora a mão de borracha a partir do *feedback* visual do estímulo.

ao lobo parietal inferior parece enfraquecer essa ilusão de posse do corpo no experimento da mão de borracha, ou seja, reduzindo a “relocação da mão real para a mão falsa ainda que apenas imediatamente após a indução da ilusão da mão de borracha” (CRIVELLI; BALCONI, 2017, p. 4). O papel dessas áreas temporais inferiores, em particular, da junção parietal temporal teria sido também corroborado pela estimulação magnética transcraniana de pulso único que, sendo aplicada, produziu “uma redução da incorporação da falsa mão, indiretamente medida pela redução da deriva propioceptiva para a mão real”. A partir daí, os autores concluem que os “processos mediados pelas áreas parietais inferiores parecem ser críticos para a definição dos limites de nossos corpos e para distingui-los de objetos externos e de outros agentes” (CRIVELLI; BALCONI, 2017, p. 5).

Se, como se viu pelo esboço apresentado, o córtex parietal tem sido associado à percepção da integração sensório-motora, como ver-se-á a seguir, o córtex pré-frontal tem sido associado aos processos de controle intencional da ação, ambos constituintes indispensáveis da consciência de agir. Os poucos estudos existentes envolvendo “a conexão causal entre o controle da ação pré-frontal, os processos de intencionalização e o agir” têm utilizado “técnicas de estimulação elétrica para modular a atividade cortical e a conexão intencional (“intentional binding”) com o objetivo de medir implicitamente as mudanças na consciência de agir” (CRIVELLI; BALCONI, 2017, p. 4)<sup>5</sup>. Os autores esclarecem que a conexão intencional está relacionada à “redução do intervalo de tempo percebido entre uma ação voluntária e sua consequência sensorial externa quando sentimos ser o agente primário dessa ação” (CRIVELLI; BALCONI, 2017, p. 4). Por exemplo, pode-se aumentar a consciência de agir aplicando uma estimulação anodal, ou seja, uma estimulação elétrica por corrente contínua, no córtex pré-frontal dorsolateral (dlPFC), em situações nas quais as “pessoas são livres para agir como preferirem”, o que resulta num aumento do “efeito da conexão intencional” (CRIVELLI; BALCONI, 2017, p. 4).

Tanto a estimulação anodal quanto a estimulação catodal – “um protocolo de estimulação para excitabilidade e responsividade cortical mais baixa” – da área pré-motora (pre-SMA) também “conduz a uma redução relevante do efeito de conexão intencional sugerindo a contribuição da rede frontal/pré-frontal medial para a consciência e controle da ação voluntária” (CRIVELLI; BALCONI, 2017, p. 4). A respeito da escolha da área pré-motora, Annachiara Cavazzana e colaboradores, no artigo de 2017, *Neural underpinnings of the ‘agent brain’: new evidence from transcranial direct current stimulation*, justifica-se por sua contribuição relevante para as funções executivas da consciência de agir visto requerer “(i) um plano para realizar uma ação dirigida a fins e (ii) a identificação das consequências de nosso comportamento no mundo externo, evitando e inibindo comportamentos errôneos e não adaptativos” (CAVAZZANA; PENOLAZZI; BEGLIOMINI; BISIACCHI, 2017, p. 1892). Considera-se que esta área pertence ao córtex pré-frontal, e não ao córtex motor, tendo “conexões extensivas” com outras regiões pré-frontais, parecendo “contribuir crucialmente para nosso sentimento de ser os agentes de nossas próprias ações”, o que seria corroborado pelo fato de que “a inibição desta área conduz a um ‘enfraquecimento’ da conexão intencional, o que pode ser interpretado como uma diminuição da consciência de controle ou agir” (CAVAZZANA; PENOLAZZI; BEGLIOMINI; BISIACCHI, 2017, p. 1892).

5 “Quando um indivíduo implementa uma ação que é seguida por um efeito tal como um tom auditivo, ele percebe uma compressão do intervalo temporal entre a ação e o efeito. Desde que isso ocorra apenas para ações intencionais, esta compressão é nomeada ‘conexão intencional’ e tem sido considerada como uma medida implícita do agir” (OBHI, 2011, p. 655).

Tendo resumidamente apresentado alguns dentre os inúmeros componentes neurobiológicos envolvidos com aspectos essenciais da consciência de agir, concluirei o presente texto tecendo algumas considerações sobre a hipótese de que o cérebro é o agente e a mente consciente seu guia de ação.

#### 4. Considerações Finais

Apresentei, resumidamente, nas seções anteriores, uma abordagem fenomenológica e, articulada com ela, uma abordagem neurobiológica da consciência de agir. Certamente, um estudo mais abrangente da consciência de agir deve incluir outros elementos constitutivos da fenomenologia do agir, tais como os *processos cognitivos* - crenças, memória, imaginação, atenção etc. - os *processos volitivos* - desejo/motivação - e os *processos afetivos/emocionais* - prazer, desprazer, medo, alegria, tristeza, amor, ódio etc. -, assim como os circuitos/áreas neurais que os instanciam. Em que pese a incompletude de ambas as abordagens da consciência de agir, a fenomenológica e a neurofisiológica, acredito que o que foi resumidamente apresentado é um ponto de partida estratégico para o desenvolvimento e a fundamentação da hipótese de que o cérebro é o agente e a mente consciente seu guia de ação.

Como foi exemplificado na seção anterior, existem circuitos/áreas neurais específicos responsáveis por formas específicas de estados/processos mentais conscientes e por sua interação com outros circuitos neurais responsáveis por outras formas de estados/processos mentais conscientes, abrangendo os aspectos cognitivos, emocionais, volitivos e motores da ação humana consciente. Assume-se, assim, que a cadeia associativa consciente identificada pelo método fenomenológico é espelhada no nível cerebral num processo de influência específica e mútua, ou seja, os circuitos/sistemas neurais não apenas interagem entre si, mas interagem também com os estados/processos mentais conscientes que instanciam. Como acontece essa interação parece ser o problema filosófico crucial visto que não se trata apenas de explicar a interação entre os circuitos/áreas cerebrais, o que seria um problema exclusivamente científico. O problema torna-se difícil, na medida em que se postula que a interação entre os circuitos/áreas cerebrais é mediada e determinada pelas propriedades mentais conscientes produzidas pelo cérebro. Estamos aqui diante do problema da causação mental, ou seja, diante do problema de acomodar a mente consciente no mundo físico, atribuindo a ela um papel causal.

Penso que esse problema é incontornável apenas quando se substancializa, num sentido cartesiano, a mente consciente, considerando-a como uma substância imaterial, distinta e separável do cérebro, a qual age causalmente de forma descendente sobre o cérebro<sup>6</sup>. Penso que não é apropriado perguntar como a mente consciente, sendo uma propriedade não-física do cérebro, e não uma substância não-física cartesiana, age causalmente, causação descendente, sobre o cérebro. Se se aceita que a mente consciente é uma propriedade do cérebro, ainda que não-física, a pergunta deveria ser formulada nos seguintes termos: como o cérebro, utilizando sua mente consciente, age causalmente sobre si mesmo, desencadeando a partir daí uma ação sobre o corpo e, a partir deste, sobre o ambiente externo ao corpo. O fato de ser o cérebro o agente não destitui a relevância causal da mente consciente, não a transforma num epifenômeno do cérebro, visto que sem a mente consciente o cérebro não seria, estruturalmente e funcionalmente, o que ele é. Assumo a tese de que a mente consciente é um fenótipo criado pela evolução biológica,

6 Apresento e justifico essa ideia nos textos de minha autoria indicados nas referências bibliográficas.

uma propriedade do cérebro, resultante da interação deste com o corpo e o ambiente externo ao corpo, utilizada pelo cérebro como *guia* que lhe permite uma ação adaptativa no mundo.

Estou aceitando a ideia, a ser detalhada em textos futuros, por ser tratar de uma pesquisa ainda em andamento, de que o cérebro é constituído por circuitos/áreas neurais responsáveis pelo processo de tomada de decisão consciente e que esta acontece a partir de outras informações conscientes também instanciadas pelo cérebro em sua interação com o corpo e com o ambiente externo ao corpo. Essas informações conscientes utilizadas pelos circuitos/áreas cerebrais responsáveis pelo processo de tomada de decisão consciente, ou seja, os guias de ação do cérebro, não são apenas aquelas envolvidas com a consciência de agir privilegiados nas seções anteriores, a saber, intenção, finalidade, sinais somatosensoriais e fluxo de informação sensorial aferente. São também cruciais as informações conscientes constitutivas das crenças, memória, imaginação, desejo, atenção, prazer, desprazer, medo, alegria, tristeza, amor, ódio etc. – com seus conteúdos específicos, assim como os circuitos/áreas neurais que os suportam. Afinal, uma abordagem fenomenológica, ainda que não sistemática, indica que não se pode separar de nossos processos de tomada de decisões consciente, ponto de partida de nossas ações conscientes, os nossos desejos, crenças, sensações, lembranças, emoções, afetos etc. Já a abordagem neurobiológica, ainda que não exaustiva, indica que não se pode separar do cérebro esses processos de tomada de decisão consciente de outras formas e conteúdos conscientes, incluindo a consciência de agir. É o cérebro que as instancia e as utiliza como guias para agir sobre o corpo no qual está incorporado e, através do corpo, sobre o ambiente externo ao corpo no qual estão situados.

**Referências bibliográficas:**

- BAYNE, T. The phenomenology of agency. *Revista Philosophy Compass*, v.3, n.1, p. 182-202, 2008.
- CAVAZZANA, A. et al. Neural underpinnings of the ‘agent brain’: new evidence from transcranial direct current stimulation. *Revista European Journal of Neuroscience*. v. 42, p. 1889-1894, 2015.
- COELHO, J. A double face view on mind-brain relationship: the problem of mental causation. *Revista TRANS/FORM/AÇÃO*, v.40, n.3, p. 197-220, jul./set., 2017.
- \_\_\_\_\_. *Abordagem dupla face da relação mente consciente e cérebro: A visão como exemplo paradigmático*. In: TOLEDO, G.L.; GOUVEA, R.A.S; ALVES, M.A.S. Debates Contemporâneos em Filosofia da Mente. São Paulo: FiloCzar, 2018.
- CRIVELLI, D.; BALCONI, M. The agent brain: a review of non-invasive brain stimulation studies on sensing agency. *Revista Frontiers in Behavioral Neuroscience*, v. 11, p. 1-6, 2017.
- DAVID, N. New frontiers in the neuroscience of the sense of agency. *Revista Frontiers in Human Neuroscience*, v. 6, p. 1-5, 2012.
- GALLAGHER, S. The natural philosophy of agency. *Revista Philosophy Compass*, v. 2, n.2, p. 347-357, 2007.
- HAGGARD, P. Sense of agency in the human brain. *Revista Nature Reviews Neuroscience*, v.18, p. 196–207, 2017.
- OBHI, S.S. Sense of agency and intentional binding in joint action. *Revista Experimental Brain Research*, v.211, p. 655-662, 2011.

## **A memética de Dennett está em harmonia com a sua noção de postura intencional?**

Juliana de Orione Arraes Fagundes<sup>1</sup>

Este trabalho procura investigar se essas duas ideias defendidas por Dennett - a memética e a noção de postura intencional - são realmente compatíveis. A noção de postura intencional é bastante central para sua filosofia da mente, perpassa todo o seu trabalho desde a década de 1970 até atualmente. Já a memética é usada por ele para trazer uma compreensão da mente humana como algo singular, diferente das demais mentes, por ser uma mente culturalmente constituída<sup>2</sup>. Ao mesmo tempo, a memética seria compatível com uma perspectiva evolutiva do ser humano, por funcionar de modo abstratamente darwinista. No texto que se segue, a noção de memética será apresentada, em seguida as principais objeções levantadas contra ela. Procurar-se-á mostrar que uma dessas objeções não pode ser apropriadamente respondida, principalmente por quem aceita o mental como holístico. Porém, o holismo do mental está muito atrelado à noção de postura intencional. Disso decorre a incompatibilidade entre essas duas ideias tão centrais ao pensamento dennettiano.

A “memética” consiste em uma tentativa de compreensão da evolução cultural a partir de um modelo tomado da evolução biológica, mas com o reconhecimento de que são processos distintos. A ideia de aplicação de um modelo abstrato tomado da teoria da evolução darwiniana a outros processos foi levantada por Campbell (1960). Esse modelo pressuporia apenas a existência de três condições: um mecanismo produtor de variação, um processo de seleção e um mecanismo de reprodução das variantes selecionadas. As variação devem ser cegas no sentido em que, quando ocorrem pela primeira vez, não são produzidas para resolver os problemas ambientais, mas de modo independente.

Assim, dentre diferentes variantes possíveis, não deve haver maior propensão de uma ocorrer do que as outras. Além disso, uma variante não tende a corrigir um possível fracasso de sua antecessora frente ao processo seletivo. Elas enfrentam cegamente o processo de seleção, podendo ser mantidas por ele ou não. As que são mantidas poderão se reproduzir, mantendo-se no processo. Se as três condições forem preenchidas, entrará em funcionamento um mecanismo análogo a da evolução biológico, independente do que sejam essas variantes. Essa ideia foi aplicada não apenas à cultura, mas também ao crescimento do conhecimento científico e à criatividade humana, por exemplo (POPPER, 1992).

---

1 Departamento de Filosofia e Ciências Humanas, Universidade Estadual do Sudoeste da Bahia.

2 Embora a mente humana tenha sido o primeiro tipo de mente com essa característica cultural, isso não significa, para Dennett, que deva necessariamente ser a única. Em princípio, nada impede que existam outras mentes culturais.



Em 1976, termo “*meme*” foi cunhado pelo biólogo Richard Dawkins para se referir de modo assumidamente especulativo aos replicadores culturais. O termo possui uma relação de parença - semântica e sonora - com “gene”, a unidade de evolução genética. Dawkins propõe que, no caso humano, estamos sujeitos a dois tipos de processos evolutivos paralelos: a evolução cultural e a evolução genética. Embora ambos sejam processos evolutivos, são diferentes em muitos aspectos: as unidades replicadoras são diferentes, os ambientes são diferentes e os processos envolvidos na replicação também. Esses dois processos funcionariam independentemente, embora um acabaria por conduzir ao outro em alguns momentos<sup>3</sup>. O ponto de vista dos *memes* é importante para a memética. Além disso, a noção de “variante” ou de “replicador” é central para que se possa construir um modelo a partir do darwinismo.

De acordo com Dawkins (1976), o processo de evolução pode ser explicado a partir do ponto de vista do gene. Nesse sentido, muitas características que pareceriam desfavoráveis para um organismo, denominado “máquina de genes” pelo autor, poderiam ser explicadas por aumentarem o sucesso dos genes em três aspectos: fidelidade, longevidade e fecundidade. Similarmente, o ponto de vista do *meme* seria o mais apropriado para se explicar a evolução da cultura. Em especial, esse ponto de vista nos permitiria compreender porque alguns itens culturais parecem tão perniciosos para a preservação da própria vida e ainda assim se proliferam entre as pessoas, ganhando adeptos. Podemos aqui dar exemplos que vão desde o racismo até o hábito de consumir alimentos industrializados. Embora nos desfavoreçam enquanto indivíduos ou enquanto espécie, diminuindo nossa variabilidade genética ou nos provocando doenças letais e impedindo a propagação de nossos genes, seriam explicados por terem o poder de aumentar o sucesso reprodutivo dos *memes*. Nós, humanos, seríamos máquinas de genes e *memes*. Dessa forma, para Blackmore (2001) nosso cérebro foi construído a partir de um processo de co-evolução entre genes e *memes*.

Essa ideia é abraçada por Dennett desde a década de 1990. Em um primeiro momento, Dennett aplica um algoritmo abstrato bem simples ao domínio da cultura. Segundo o autor, dados esses três elementos, o algoritmo evolutivo começa a funcionar: replicação, variação e seleção. Conforme Dennett (1997), há diferentes tipos de mentes que se acumulam ao longo da história evolutiva, começando por mentes muito simples e rudimentares e culminando com mentes altamente complexas e sujeitas a dois processos evolutivos paralelos. De todo modo, o autor propõe que todos esses tipos de mentes funcionam por meio de processos abstratamente darwinistas. Certo tipo de mente se diferencia das demais por ser constituída de uma infestação de *memes* que lutam primariamente pela sua existência. O exemplo mais típico seria a mente humana, produto não apenas da evolução biológica como da evolução cultural. Se há outros exemplos, é um assunto a se discutir, mas essa possibilidade não está fechada em princípio.

Porém, há uma crítica à memética que leva a um ceticismo quanto à possibilidade de seu desenvolvimento: diante do caráter holista da cultura e do mental, a dificuldade está em se isolar e identificar os replicadores em um processo de evolução cultural compreendido conforme uma compreensão abstratamente darwinista, pois isso exigiria a existência de um replicador discreto e fiel.

---

3 Essa é a metáfora da coleira, provinda da sociobiologia, para quem a cultura sempre estará presa à coleira da evolução biológica. Na memética, há uma alternância de quem conduz a coleira. Às vezes os *memes* estão no comando, outras, os genes, mas genes e *memes* frequentemente caminham lado a lado.

### Críticas à memética

A memética enfrentou muitas críticas. Conforme Toledo (2017), as principais foram o *problema da unidade* e o *problema ontológico*. O problema da unidade reside no fato de que o *meme*, ao contrário do gene, não é uma unidade discreta. Nesse sentido, não poderia ser considerado um replicador. O problema ontológico é a crítica segundo a qual não é possível encontrar um substrato físico para os *memes*. Para Toledo, ambos são pseudoproblemas fundamentados na ignorância do funcionamento dos processos biológicos.

Em relação à questão denominada “problema ontológico”, de fato não é possível encontrar um substrato físico para os *memes*, mas isso não é um problema real, pois os *memes* são compreendidos como itens culturais e a cultura não é algo sobre o que se possa estabelecer uma localização precisa no tempo e no espaço. Quer dizer, a cultura se modifica ao longo do tempo e do espaço, mas não sou capaz estabelecer seu lugar de modo preciso assim como consigo estabelecer o lugar do livro de Dennett que está agora sobre minha escrivaninha. A cultura constitui o ambiente impalpável das relações sociais humanas, permeia nossas vidas intersubjetivas, ainda que não seja possível determinar seu substrato físico. Isso não impede a compreensão da ideia de que a cultura evolui. Portanto, o problema ontológico não se constitui um problema real. Nem tudo precisa ser localizado em um ponto determinado no espaço e tempo.

Contudo, frente ao problema da unidade, a memética parece estar diante de dificuldades maiores. Como separar a cultura em partes? Os itens culturais dependem uns dos outros, ao mesmo tempo em que essa dependência não se pode estabelecer de modo preciso. Quer dizer, na grande maioria dos casos, não é possível determinar algo como “o item cultural A depende dos itens culturais B e C, mas não do D”. Há uma rede de relações que é difusa, extremamente dinâmica e complexa. De que modo se tornaria possível identificar o replicador? Toledo lida com a questão adotando uma postura agnóstica em relação aos *memes*: “[...] o mais importante é entender que não há oposição entre o holismo ontológico e o reducionismo metodológico” (2017, p. 126). Para ele, vale manter a memética, mesmo diante de uma postura holista em relação à cultura, isso permitiria uma análise mais simples da realidade, uma vez que permitira a unificação de dois processos - a evolução biológica e a evolução cultural - sob o mesmo algoritmo abstrato.

A solução de Toledo não resolve o problema. Ainda que adotemos uma postura anti-realista ou agnóstica em relação aos *memes*, tomando-os como termos teóricos por razões instrumentais e práticas, o problema da unidade não desaparece. Fica o problema prático de identificar o replicador na cultura. É certo que a cultura é composta de uma série de informações aprendidas por imitação. A questão é que não é possível separar essas informações umas das outras para identificar onde está o replicador. Richerson e Boyd (2005) lidam com essa questão por meio da substituição do termo “*meme*” por “variantes culturais”, o que também não resolve o problema.

Uma abordagem darwinista da cultura não implica que a cultura tenha que ser divisível em minúsculos pedacinhos que sejam fielmente replicados assim como os genes. Melhor que isso, as melhores evidências sugerem que as variantes culturais são apenas frouxamente análogas dos genes. A transmissão cultural frequentemente não envolve replicação de alta fidelidade; nem são as variantes culturais sempre minúsculos fragmentos de informação. Não obstante, a evolução cultural é fundamentalmente darwinista em sua estrutura básica” (RICHERSON; BOYD, 2005, p. 60, tradução nossa).

O problema é que para que se estabeleça um processo evolutivo, é preciso haver replicação *fidedigna* de algo. A substituição do termo “*meme*” por variante cultural, embora seja uma forma de reconhecer que o problema existe, não traz uma solução. A cultura não forma complexos fechados nem partes que possam ser identificadas separadamente. Aparentemente, nada há nela que possa ser um replicador fidedigno.

### **Palavras como exemplo de *memes***

Para lidar com essa crítica de que os *memes* não podem ter uma transmissão discreta e fidedigna no ambiente da cultura, Dennett oferece as *palavras* como exemplo de *meme* que, segundo ele, podem ter uma transmissão discreta e fiel. “As palavras, eu argumentarei, são os melhores exemplos de *memes*, itens culturalmente transmitidos que evoluem por replicação diferencial - ou seja, por seleção natural” (DENNETT, 2017, p. 176). Nessa perspectiva, as palavras teriam algo de parecido com os genes, ou os genes com as palavras. Há algo nelas que se presta a uma transmissão fidedigna e discreta: sua grafia ou sua expressão sonora.

A linguagem complexa e articulada, na concepção de Dennett, é um meio muito fértil para a proliferação de *memes*, embora se deva observar que nem todos os *memes* sejam linguísticos. Por exemplo, talvez a técnica das rendeiras seja uma atividade em grande medida não linguística mas culturalmente transmitida, embora aqui seja difícil delimitar os aspectos linguísticos dos não linguísticos. De todo modo, o autor argumenta que é a linguagem que permite o desenvolvimento de uma cultura complexa e ricamente acumulativa. As palavras, segundo ele, são o melhor exemplo de *meme*.

As palavras certamente desempenham um papel central e ineliminável na nossa evolução cultural explosiva, e informar-se sobre a evolução das palavras servirá como uma entrada viável a questões desafiadoras sobre a evolução cultural e seu papel na modelagem de nossas mentes (DENNETT, 2017, p. 179, tradução nossa).

Conforme o autor, alguns escritores, como o próprio Dawkins (2004, *apud* DENNETT, 2017, p. 225) tentaram relacionar *memes* com frases, e as palavras seriam comparadas a trechos de *memes*. Isso foi pensado para fazer um paralelo com os genes, considerando que cada gene é composto por uma sequência de ácidos nucleicos e cada frase é composta de uma sequência de palavras. O problema, nesse caso, é que as frases não se replicam, ao passo que os genes, sim. Por exemplo embora os diferentes exemplares de um mesmo título de livro contenham as mesmas frases, em cada título de livro diferente as frases são completamente diferentes. O tempo todo, de cada boca, saem frases diferentes ao longo da vida de alguém, da boca de cada ser humano, também, as frases proferidas são distintas. Por isso, as frases não podem ser consideradas replicadores.

As palavras, segundo Dennett, contam com algo que lhes permite a replicação fidedigna e discreta: a digitalização por meio dos fonemas. “Fonemas conseguem a digitalização do meio auditivo para o discurso. Cada língua falada tem um alfabeto auditivo com variedades finitas, os fonemas com os quais as palavras são compostas” (DENNETT, 2017, p. 199). Por meio dos fonemas, casos particulares de palavras são criados, os quais são identificados e categorizados, ainda que cada um desses casos particulares de um mesmo tipo de palavra seja muito diferente dos outros. O que permite isso são as características da cognição humana e, em especial, dos seus sistemas vocal e auditivo. As réplicas de uma

palavra, segundo o autor, não são réplicas físicas, obviamente, mas sim réplicas virtuais. Cada falante de uma língua recorre a uma série de normas implícitas por meio das quais ajusta inconscientemente suas percepções auditivas e suas emissões vocais. É isso que permite a transmissão de alta fidelidade.

Conforme Dennett, há tipos de palavras e os seus diversos casos particulares<sup>4</sup>. Um caso particular da palavra “gato” ocorre em quem ouve a palavra. A palavra escrita no papel é outro caso, assim como a palavra dita por diferentes pessoas e por pessoas de diferentes sotaques. Embora cada caso particular da palavra seja completamente diferente dos outros em termos físicos, eles não deixam de ser casos particulares da mesma palavra. Dessa forma, o tipo de palavra teria uma replicação fidedigna, independente das múltiplas variações nos casos particulares.

### **Palavras: uma solução artificial**

Porém, a solução de Dennett a essa questão da fidedignidade e da identificação do *meme*, embora engenhosa, parece ser muito artificial. Na prática, nós compreendemos os casos particulares e diversos tipos de palavras dentro do contexto, não apenas o contexto da frase, mas o contexto da comunicação como um todo. Por exemplo, talvez a palavra “gato” possa soar como “gado” quando dita por alguém, mas o interlocutor sequer percebe isso, pois “gado” não faria sentido dentro daquele contexto. Então, nesse caso, o que promove a remissão do interlocutor do caso particular da palavra para o tipo geral é o contexto. Em que medida se poderia, de fato, separar a palavra como um replicador discreto, já que ela não faz sentido fora de um contexto de comunicação?

Assim como ocorre com os demais itens culturais, as palavras aparecem sempre em relações complexas com outras palavras, dentro de frases e contextos de diálogo, relações entre os sujeitos em comunicação e desses sujeitos com o mundo circundante, seja esse mundo físico ou social. Nesse sentido, compreender a palavra como exemplo paradigmático de *meme* leva a uma supersimplificação da cultura, da linguagem e das relações de comunicação. Embora a simplicidade possa ser uma virtude científica, a supersimplificação não é!

Explicando um pouco o problema da supersimplificação, as explicações mais simples e abrangentes, em geral, são preferíveis. Quer dizer, se duas hipóteses científicas forem igualmente explicativas, a tendência é que se opte pela mais simples. Por exemplo, se todo o meu condomínio ficar sem energia elétrica, seria melhor a explicação de que houve um problema geral no sistema elétrico do condomínio do que a explicação de que cada unidade residencial tenha tido um problema específico em seu sistema elétrico. E quanto à supersimplificação?

Nesse caso, o problema é que, ao tentar simplificar as coisas muito rápido e exageradamente, atropela-se justamente o que é mais importante. Em outras palavras, a supersimplificação deixa de lado os aspectos centrais daquilo que se quer explicar em nome da simplicidade, tornando a explicação excessivamente superficial. Aqui, é possível fazer uma analogia entre a supersimplificação e o que Dennett (1998, p. 85) chama de “reducionismo ganancioso” [“greedy”]. Ele critica a atitude exagerada de se tentar reduzir todas as explicações, inclusive as explicações sobre a cultura e a mente ao nível da física, pois isso deixa de capturar os aspectos mais importantes do que precisa ser explicado e acaba por gerar uma série de dualismos e paradoxos. Por outro lado, a atitude supersimplificadora

4 Trata-se, aqui, da distinção “type”/ “token”.

da memética, abraçada pelo próprio Dennett, parece cair nas mesmas desvantagens do reducionismo ganancioso que ele critica.

Além disso, há outro ponto que precisa ser considerado: Ainda que se considere a palavra como unidade constituinte da linguagem, ela constitui apenas um tipo de linguagem e a linguagem não é o todo da cultura. Por exemplo, a palavra não constitui a linguagem musical ou gráfica, ou a linguagem cinematográfica. Também não é capaz de dar conta dos aspectos não-linguísticos da cultura, como técnicas para construção de artefatos e instrumentos, danças, culinária, temperos, hierarquias dentro de grupos, um jeito de olhar, um gesto, um corte de cabelo. No mais, os diversos tipos de linguagens também se relacionam aos aspectos não linguísticos da cultura.

Assim, ainda que a palavra cumprisse o papel de unidade discreta - embora aparentemente não cumpra - isso não seria capaz de solucionar o problema da unidade, pois como poderíamos encontrar a unidade quando algo do jeito do olhar de um professor passa a ser copiado inconscientemente por um aluno, por exemplo?

### **A postura intencional**

Desde a década de 1980, Dennett defende o que chama de “postura intencional”. É uma ideia que perpassa toda a sua obra, podendo ser considerada um conceito central e integrador das ideias do autor. É possível encontrar nessa noção uma herança de Quine, seu professor. De acordo com Dennett, nós temos a necessidade de atribuir intencionalidade a determinados sistemas para que possamos compreender o que está acontecendo. Por exemplo, só podemos manter um diálogo com alguém por meio da atribuição a essa pessoa de uma série de estados mentais, incluindo atribuições de crenças sobre o significado das palavras. Essa atribuição de intencionalidade, contudo, não é algo que façamos linguisticamente, ao contrário, é uma habilidade natural da nossa cognição, nós aplicamos o tempo todo, sem que percebamos. Ocorre que, uma vez que a comunicação é bem sucedida, isso acontece porque nós também obtivemos sucesso em nossas atribuições. Ou seja, muito daquilo que nós atribuímos ao nosso interlocutor de fato constitui a vida mental daquele sujeito.

Dois pontos podem ser levados em consideração aqui. O primeiro é que isso pressupõe um alto grau de coincidência entre os estados mentais de um sujeito e os de seu interlocutor. O outro ponto é que as divergências não podem ser identificadas uma por uma. Os estados mentais participam de redes complexas e só podem ser compreendidos por meio de uma compreensão do seu papel dentro dessa rede de eventos mentais. Ainda que as palavras pareçam uma entidade de reprodução fiel da forma como Dennett as apresentou, quando observamos que ela tem um significado que só é dado pela sua participação em uma rede complexa de estados mentais que podem ser linguísticos ou não, fica mais difícil compreendê-las como replicadores fiéis e discretos. Parece que é preciso fazer uma escolha entre a memética e o holismo do mental.

Ao explicar nossos próprios estados mentais, adotamos uma postura parecida com a que é adotada quando tentamos explicar os estados mentais alheios: construímos uma explicação linguística para dar conta de uma série de eventos ocorridos internamente. Dessa forma, a noção de subjetividade é construída de forma holista, a mente é holista, como resultado da própria adoção da postura intencional.

Vale aqui mencionar que, na concepção de Dennett, a mente humana é caracteristicamente cultural e está sujeita ao processo de evolução memética. Ora, mas se



é holista, não se podem separar estados mentais discretos. Nesse sentido, a memética, ao pressupor entidades que se reproduzem de modo discreto e fiel, não entra em harmonia com a postura intencional. Após convencidos que o mental é holista, torna-se difícil aceitar a noção de *meme* e torna-se difícil também conciliar as duas ideias. A mente é holista ou ela é constituída por uma infestação de *memes*?

### Um argumento sobre a história do darwinismo

Às vezes, os defensores da memética usam um argumento que apela para a história do darwinismo. Diz Blackmore (1999, p. 56, tradução nossa), “[n]ós podemos percorrer um longo caminho com os princípios gerais da seleção memética sem entender os mecanismos dos quais ela depende.” O ponto aqui tem a ver com uma analogia com o darwinismo. Darwin conseguiu desenvolver os princípios gerais da Teoria da Evolução sem que conhecesse os genes, portanto, não sabia como eles eram armazenados e replicados. De acordo com esse argumento, analogamente, seria possível desenvolver bastante a memética, ainda que não se tenha uma noção clara de quais sejam as entidades replicantes da cultura, nem como elas são armazenadas e replicadas.

Porém, esse argumento que faz uma analogia entre a história futura da memética e a história passada do darwinismo traz consigo uma aposta muito arriscada. Primeiro, deve-se considerar que, na própria história da biologia, o darwinismo se tornou amplamente aceito após a chamada Síntese Evolutiva Moderna, que une o darwinismo e a genética mendeliana. A Síntese ocorreu em meados do último século, quase cem anos depois da primeira publicação da “Origem das Espécies” de Darwin e dos trabalhos de genética de Mendel, que vieram um pouco depois. Ou seja, embora as duas teorias já tivessem sido desenvolvidas separadamente, foi a união delas que provocou toda uma mudança de olhar sobre o mundo orgânico. Uma compreensão mais detalhada dos mecanismos genéticos acabou fundamentando e fortalecendo o darwinismo a partir da nova síntese.

Assim, apostar que acontecerá o mesmo tipo de coisa com a memética, que haverá uma teoria sobre os *memes* e que ela será sintetizada com os mecanismos darwinistas para fortalecer essa suposta nova disciplina é arriscar alto demais. Provavelmente não existem razões fortes o suficiente para se fazer tal aposta. Os únicos argumentos favoráveis seriam a simplicidade e abrangência explicativa, mas esses argumentos se tornam fracos diante dessas circunstâncias.

Deve-se notar que o argumento de que a memética pode ser desenvolvida ainda que se desconheçam os seus mecanismos fundamentais possui contornos falaciosos. Incorre-se, aqui, na famosa falácia de apelo à ignorância. “O *argumento ad ignorantiam* é o uso de um argumento que soa convincente aos outros porque estes ignoram a fraqueza do argumento e dos fatos que se contrapõem a ele.” (JOSEPH, 2014, p. 243). Quer dizer, do fato de que ignoramos os mecanismos de armazenamento e transmissão de uma suposta evolução memética não podemos concluir nada, nem que a memética seja uma teoria, nem que não seja. A ignorância não pode servir de base para uma conclusão cheia de conteúdo de promessas. Diante disso, o argumento que apela para a história da biologia não resiste.

### Holismo ou memética?

O argumento principal exposto aqui é de que a memética não está em harmonia com todo o sistema de atribuições de estados mentais defendido por Dennett como fundamento de sua filosofia da mente: denominado “postura intencional”. A adoção



da postura intencional pressupõe o holismo do mental. Este, por sua vez, não permite a fragmentação dos estados mentais, não torna possível isolar estados mentais uns dos outros. Isso se alastra pelo campo da linguagem. Não podemos compreender a linguagem sem atribuímos estados mentais intencionais aos outros. Mas também não podemos fazer a atribuição de estados mentais isolados. Cada palavra tem que ser entendida dentro de seu contexto, e isso é um exemplo da dificuldade para se isolarem os *memes*. A linguagem constitui uma parte da cultura, mas é provável que o problema afete as outras partes da cultura. Para Dennett, a mente humana é caracteristicamente cultural. Deve-se notar que, assim como o mental é holista, a cultura também é. Não podemos separá-la em fragmentos. Diante disso, parece que é necessário fazer uma escolha entre a memética e o holismo do mental. Não é possível abarcar as duas posições ao mesmo tempo.

**Referências bibliográficas:**

- BLACKMORE, Susan. *The meme machine*. Oxford: Oxford University Press, 1999.
- \_\_\_\_\_. Evolution and Memes: The Human Brain as a Selective Imitation Device. *Cybernetics and systems*, v.32, n.01, p.225-55, 2001.
- CAMPBELL, Donald. Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological review*, v. 67, n. 6, p.380-400, 1960.
- DAWKINS, Richard. *The selfish gene*. Oxford: Oxford University Press, 1976.
- \_\_\_\_\_. *The ancestor's tale: A pilgrimage to the dawn of evolution*. Boston: Houghton Mifflin, 2004.
- DENNETT, Daniel. *Tipos de mentes: rumo a uma compreensão da consciência*. Rio de Janeiro: Rocco, 1997. (Tradução de Alexandre Tort; revisão técnica de Marcus Pinto).
- \_\_\_\_\_. *A perigosa idéia de Darwin: a evolução e os significados da vida*. Rio de Janeiro: Rocco, 1998. (Publicado originalmente em 1995. Tradução de Talita M. Rodrigues).
- \_\_\_\_\_. *From bacteria to Bach and back. The evolution of minds*. New York: W. W. Norton & Company, 2017.
- JOSEPH, M. *O Trivium: As artes liberais da lógica, da gramática e da retórica*. São Paulo: Realizações Editora, 2014.
- POPPER, Karl. A Evolução e a árvore do conhecimento, In: \_\_\_\_\_. *Conhecimento objetivo*. Belo Horizonte: Itatiaia, 1992, p. 234-56. (Publicado originalmente em 1972. Tradução de Milton Machado.)
- RICHERSON, Peter J.; BOYD, Robert. *Not by genes alone: How culture transformed human evolution*. Chicago: The University of Chicago Press, 2005.

## Bratman como atualização de Davidson

Lucas Ollyver<sup>1</sup>

Vamos começar então com o conceito de ação: “A grande maioria dos filósofos concordariam em dizer que ações são eventos, mudanças de certo tipo iniciadas por uma agente.” (LECLERC, 2014, p. 71) A doutrina mais aceita sobre teoria da ação é do filósofo americano Donald Davidson. Para ele as ações são eventos. Eventos são particulares espaço temporais, que têm uma característica particular: elas podem ser descritas intencionalmente, porque tem pelo menos uma intenção em sua história causal.<sup>2</sup> Desde então muitos tem concordado com Davidson de que existe uma estrita ligação causal entre uma razão e uma ação. Esta relação é responsável por explicar o motivo da ação acontecer, também chamada de racionalização: “quando a relação entre uma razão e a ação quando a razão explica a ação por dar as razões do agente por ter feito algo.” (DAVIDSON, 1980, p. 3)

### Racionalização

Uma razão racionaliza uma ação apenas se nos leva a ver algo que o agente viu, ou pensou ter visto, em sua ação - alguma característica, consequência, ou aspecto da ação do agente queria, desejava, valorizava, sustentava com apreço, pensava obedecer, benéfico, obrigatório ou agradável. (DAVIDSON, 1980, p. 3)

Se eu ligo o interruptor com a intenção de iluminar a sala é porque eu tenho o desejo de iluminar o ambiente, e a crença de que isso irá me ajudar a encontrar as chaves de casa, algum tipo de objeto perdido ou simplesmente que não irei tropeçar dentro da sala. Se eu procuro um copo para tomar água é porque eu tenho o desejo pelo alívio que a água pode me trazer e a crença de que se eu tomar uma determinada quantidade de água, não irei mais sentir sede.

Raciocinando desta maneira é sempre possível encontrar um desejo e/ou uma crença que preencha a minha necessidade de racionalização de uma determinada ação. Nessa direção está o trabalho de Donald Davidson. Inclusive, o primeiro exemplo é dele.

“Sempre que alguém faz algo por uma razão, portanto, ele pode ser caracterizado como (a) ter algum tipo de pro atitude em relação a ações de um determinado tipo, e (b) acreditando (ou sabendo, percebendo, lembrando) que sua ação é desse tipo. Sob (a) devem ser incluídos desejos, vontades, impulsos, sussurros, e uma grande variedade de visões morais, princípios estéticos, preconceitos econômicos, convenções sociais, objetivos públicos e privados e valores na medida em

---

1 Professor IFPE, doutorando UFPE

2 Esta descrição intencional tem um caráter mentalista e de primeira pessoa.

que estes podem ser interpretados como atitudes de um Agente dirigido para ações de um determinado tipo.” (DAVIDSON, 1980, p. 3)

Davidson acredita que sempre que agimos há subjacente a nossa ação uma crença e o desejo que a faz tomar forma. Como somos seres racionais, nossas ações herdam a nossa racionalidade, sempre devendo responder à pergunta do por que agimos como agimos? Para ele, exibir a crença e/ou desejo que motiva a minha ação é exibir a razão pela qual agimos. Este processo de racionalização sempre pode ser exigido para que não sejamos acusados de irracionalidade.

### **Mantendo o senso comum**

A tentativa dele é manter a estrutura de senso comum para explicar as ações. Sempre que alguém age, perguntamos o porquê de aquilo ter ocorrido daquela maneira e não de outra. Na maior parte dos casos temos uma razão como resposta. Acreditamos que o agente saberá, por introspecção, qual razão o levou a agir, e por meio da exposição da razão ele pode apresentar uma justificativa plausível para sua ação. Esta depende do olhar do agente sobre o que foi feito e confia na autoridade do agente para explicar a ação realizada. Utilizando-se de um discurso mentalista de primeira pessoa. É deste modo que o Davidson apresenta esta intuição: «O que quer que alguém faça por uma razão, portanto, pode ser caracterizada como:

- (a) como tendo algum tipo de pro atitude para ações de um certo tipo, e
- (b) acreditando (ou sabendo, percebendo, notando, lembrando) que sua ação é daquele tipo” (DAVIDSON, 1980, p. 3-4)

### **Aplicando a estrutura**

Davidson define pro atitude como: uma atitude proposicional (crença-desejo) que nos move a uma ação, considerada por ele como sua causa, aqui também está incluído o motivo da ação. Para ele tudo aquilo que é realizado tem uma crença ou desejo como seu ponto de partida. Modelo posteriormente conhecido como crença-desejo. Este processo se sofisticava ao longo do texto. Ele assume que quando alguém faz algo intencionalmente podemos retornar sempre a pro atitude (crença-desejo) que o levou a agir, fazendo deste par (crença-desejo) a razão primária da ação, constituindo a razão pela qual alguém agiu. Reformulando sua formação inicial do que é uma racionalização para dar maior poder de explicação, nos seguintes termos:

1. Para entender como uma razão de qualquer tipo racionaliza uma ação é necessário e suficiente que nós possamos entender, pelo menos em linhas gerais, como se constrói uma razão primária.

2. A razão primária de uma ação é sua causa.” (DAVIDSON, 1980, p. 4)

Voltar ao exemplo anterior. Se mudo o interruptor de posição, ligo a luz e ilumino a sala, sem saber, também alerto um ladrão para o fato de que eu estou em casa. O que eu fiz? Davidson diz que somente uma coisa foi realizada, da qual temos as quatro descrições possíveis, a saber: mudar o interruptor de posição, ligar a luz, iluminar a sala e alertar o ladrão. No entanto, minha razão primária para fazê-la é meu desejo de ligar a luz. Para evitar esses problemas Davidson propõe mais um refinamento.

### Condição necessária para razões primárias

C1. R é uma razão primária pela qual um agente realizou a ação A sobre a descrição d somente se R consiste de uma pro atitude do agente para a ação com certa propriedade, e uma crença do agente que A, sob a descrição d, tem aquela propriedade. (DAVIDSON, 1980, p.5)

C2. Uma razão primária de uma ação é sua causa. (DAVIDSON, 1980, p.12)

Uma ligação direta e necessária entre razões e ações é posta a partir deste argumento. Conectando a primeira (razão) como causa direta da segunda (ação), no contexto de uma ação na qual há um propósito de sua realização e o agente está ciente do que faz, ou seja, uma ação intencional. Para tal:

“Davidson visa mostrar que quando essa relação entre razões e ações se verifica – isto é, quando a razão de facto explica a ação – essa explicação é uma forma de explicação causal. Isto quer dizer que as razões explicam as ações na medida em que são as suas causas, ou melhor, é precisamente porque a razão é a causa da ação que esta pode ser explicada explicitando aquela.” (CADILHA, p. 138)

Assim, Davidson acredita que o que distingue uma ação é o fato de poder ser descrita de um ponto de vista mentalista ou intencional, recorrendo a razão primária que a causou, reduzindo a ação a crenças e desejos. É o fato de ser causada por razões que distingue uma ação de outros eventos no mundo. Uma ação consistirá, então, numa descrição intencional de um evento. Possuindo também como traço marcante que “justificação e explicação de uma ação estão frequentemente de mãos dadas, com frequência indicamos a razão primária para uma ação por afirmar que, se verdadeira, deve também verificar, sustentar ou suportar a crença ou atitude relevante do agente.” (DAVIDSON, 1980, p.8)

Há para ele uma força incrível em sua argumentação, conseguindo manter a estrutura da psicologia de senso comum e explicar o que é uma ação, como ela acontece e demonstrar que há uma relação de causação mental, na qual uma razão ou crença é o motor/causa da ação. Deste modo, no contexto de uma ação intencional, a razão exerce uma determinação causal na ação e serve de justificativa para a ação a partir do conceito de racionalização. Assim, Davidson acredita que é possível saber qual é a intenção da ação. “Saber a razão primária pela qual alguém age como age é saber a intenção com a qual a ação foi feita.” (DAVIDSON, 1980, p.7)

### Análise metodológica do modelo de Davidson

Destrinchando o processo de análise que o Davidson realiza no seu modelo crença-desejo, em geral, podemos dizer que procede nos seguintes passos:

A prioridade metodológica da intenção sobre a ação.

Se é possível imputar intencionalidade a uma ação é necessário que ela esteja em uma relação apropriada com a crença e o desejo do agente. Se alguém limpa a piscina de sua casa pela manhã é porque tem o desejo de relaxar e a crença de que poderá passar o dia inteiro de folga. Realizando a ação intencional de limpar a piscina.

2) A teoria desejo-crença da intenção na ação

Consiste em entender a ação intencional e a ação feita com uma intenção em termos das crenças e desejos do agente, e ações como estando em relações apropriadas a estas crenças e desejos. A estrutura apresentada por ele é que existem dois tipos de estados mentais importantes para a ação: os que representam o papel das crenças e os que representa o papel dos desejos. Entendidos aqui da forma mais ampla possível. A proposta de Davidson é pensar as intenções nas ações presentes, passadas ou futuras da mesma maneira. “A ideia é que uma vez que tenhamos uma explicação adequada do que é agir intencionalmente e agir com uma certa intenção nós podemos esperar ter a disposição todos os materiais essenciais requisitados para um tratamento satisfatório das intenções direcionadas ao futuro.” (BRATMAN, 1987, pos.175)

### 3) A estratégia da extensão

Este é o passo mais simples. Já que temos a estrutura para explicar a intenção na ação direcionadas ao tempo presente e ações no passado, não devemos nos incomodar com uma nova estrutura para as ações direcionadas ao futuro. Aplicamos a estrutura que já possuímos.

### 4) A redução das intenções direcionadas ao futuro para desejos e crenças adequados.

Esta etapa, assume que a estrutura esboçada para a análise da ação a partir do background crença-desejo é suficiente para a caracterização tanto da mente quanto da ação. A intenção não representa nenhum papel significativo neste processo.<sup>3</sup> Chamo esse processo metodológico de sanfona da crença e do desejo. Visto que ele consiste no processo de análise numa redução e posterior ampliação exatamente como uma sanfona. Ganhando fôlego ao assumir uma classe restrita de ações, as já concluídas, estabelecendo uma estrutura explicativa e posteriormente expandindo a sua análise as intenções direcionadas ao futuro. Assumindo que elas possuem exatamente as mesmas características das ações voltadas ao presente.

Embora a postura de Davidson tenha sido predominante por um tempo, Michael Bratman é um dos filósofos que discordam do tratamento dado a intenção. Ele concorda com o Wilson e Shpall que “o conceito de ‘intenção’ tem várias nuances conceituais que não são simples ou fáceis de delinear. Uma delas, talvez a mais importante é sua relação com o futuro, agir intencionalmente e agir com uma certa intenção.” (WILSON; SHPALL, 2012) E é exatamente sobre as intenções direcionadas ao futuro que Bratman irá trabalhar.

A primeira e maior diferença entre a posição filosófica de Davidson e Bratman é metodológica. O primeiro procura analisar a ação através de um processo de racionalização, posterior a execução da ação e buscar os porquês de sua realização. O segundo procura a intenção voltada ao futuro. Observa a intenção antes da sua realização enquanto ação, seu processo formação e o papel que esta exerce para a sua efetivação.

O objetivo é mostrar como Bratman compreende o papel da intenção na ação. Ele assume que planos fazem parte de nossa natureza: “O fato é que nós somos agentes planejadores.” (BRATMAN, 1987, pos. 85) E esta é uma atividade diária que agentes limitados e racionais como nós realizamos, a todo tempo “configuramos planos que são na maior parte das vezes parciais e então os preenchemos com o que julgamos necessário com o passar do tempo.” (BRATMAN, 1987, pos. 193)

<sup>3</sup> No artigo *Intending*, Davidson muda um pouco de opinião, mas ainda não é suficiente para caracterizar a intenção como um elemento no mesmo nível das crenças e dos desejos. Dizendo que a intenção tem um papel importante na agência, mas não possui o mesmo papel causal das crenças e desejos.



### Digressão sobre a racionalidade limitada

Bratman assume textualmente ser influenciado por Herbert Simon,<sup>4</sup> (SIMON, 1983) e que todo o seu texto é uma contribuição ao conceito de racionalidade limitada proposto pelo trabalho de Simon. Pensar e escrever acerca do ser humano como limitado em suas capacidades é hoje uma obviedade, mas na primeira metade do século XX uma corrente de pensamento chamada teoria racional da escolha atribuía à razão humana, papel central na avaliação das opções e deliberações. Eles afirmavam que tudo isso poderia, a grosso modo, ser reduzido a uma função de utilidade subjetiva que após preenchida suas variáveis, as deliberações seriam feitas sem maiores esforços.

Simon, por outro lado, acompanhando o desenvolvimento da psicologia experimental, desenvolve outra perspectiva acerca desta questão. Ele a chama de teoria comportamental, cunhando o conceito de racionalidade limitada em oposição a racionalidade plena para designar as limitações que possuímos estruturalmente enquanto agentes humanos. “Para os seres humanos, o processo de deliberação sobre o que fazer é uma atividade que consome recursos. Um agente que age como se sua ação fosse sem custo descobre que teria, muitas vezes, o mesmo sucesso do que se tivesse simplesmente, adivinhado, ou confiado no resultado da deliberação no momento em que ele não tinha nada melhor do que deliberar sobre as informações imperfeitas que possuía.” (VARGAS; YAFFE, 2014, p. 2)

Esta observação motivou Simon a elaborar a teoria da racionalidade limitada que reconhece a nossa racionalidade como imperfeita em muitas maneiras diferentes. Nossa lista de limitações é extensa. Mas vamos a alguns exemplos para ilustrar esse ponto. Foi realizada uma pesquisa com oito juízes de condicional em Israel, que passaram vários dias revisando pedidos de condicional. Os pesquisadores registraram quanto eles levavam para tomar cada decisão, de conceder ou não a condicional, e o período dos intervalos para refeição, embora à primeira vista possa parecer estranha a relação entre as decisões dos juízes e o horário da alimentação dos mesmos ficará clara ao final do exemplo. “Os autores fizeram um gráfico da proporção de pedidos aprovados em relação ao tempo desde a última refeição. A proporção conhece picos após cada refeição, quando cerca de 65% dos pedidos são concedidos. Durante as duas horas, mais ou menos, até a refeição seguinte dos juízes, a taxa de aprovação cai regularmente, até chegar perto de zero pouco antes da refeição.” (KAHNEMAN, 2011, pos. 937)

Após analisar os dados, os pesquisadores chegaram a uma conclusão indesejada por todos: sob o efeito de fome e cansaço é mais fácil que os juízes tomem a decisão padrão (em média 35% dos pedidos são negados e essa é considerada a decisão padrão para tais casos). Eles se tornam mais eficientes (rápidos em seu julgamento), no entanto, com prejuízo aqueles que estão sendo julgados. Partindo do panorama da racionalidade limitada, Bratman criou o que ele chama de Teoria de planejamento da intenção, que é “uma descrição dos mecanismos psicológicos através dos quais os agentes racionais agem em conformidade com os objetivos que pretendem alcançar, e tentam, ao menos diminuir, o impacto de nossas limitações cognitivas.” (BRATMAN, 1987, p. 215)

O trabalho dele vai na direção de descrever o processo deliberativo a partir de uma perspectiva funcionalista. No curso de investigação, ele depara-se com várias características da intenção. O argumento central é que o modelo da teoria crença-desejo não é capaz de lidar com o fenômeno do planejamento, pois o mesmo não leva em consideração o

4 Reason in Human Affairs é um dos seus trabalhos mais importantes.

raciocínio prático como algo que impulsiona a formação da intenção, e não explica como os agentes podem construir planos e permanecer firmes com eles. Sob essa perspectiva, planos são intenções formalizadas, que nos ajudam a restringir nossas opções futuras, facilitando nosso trabalho deliberativo e economizando nossos limitados recursos.

Um exemplo de como isso funciona é o plano do jantar. Se eu decido pela manhã que vou a um restaurante jantar, com a intenção de tomar uma sopa. Ao chegar lá não é necessário perder tempo nem ficar perdido entre as diversas opções que existem no menu. Vou direto as opções de sopas me atendo ao plano original e escolhendo dentre as que estão disponíveis, a que irei tomar. Isto faz com que nós não tenhamos que reconsiderar todas as possibilidades quando há uma mudança no contexto, a não ser que naquele dia, por algum motivo especial, não haja sopa no menu. Mesmo que as crenças e os desejos sejam capazes de dar razões para a ação, elas não são suficientes para explicar a estabilidade e eficiência do raciocínio prático humano, mesmo com suas limitações.

### **Relevância direta das intenções**

Este raciocínio nos traz de volta o ponto de independência e relevância da intenção como atuante direta nas ações, deixando de ser um elemento de relevância indireta (como era no modelo crença-desejo) para se tornar um elemento forte e direto no curso do que fazemos. Por isto se dá a importância da construção de planos, tendo em vista que nossas capacidades são limitadas, eles são necessariamente incompletos, apresentando estruturas gerais que preenchemos conforme a necessidade e o passar do tempo. Além disso eles tem um caráter extremamente prático, facilitando a nossa coordenação intrapessoal e interpessoal, esta última determina aquilo que podemos ou não esperar dos demais agentes, dado a estabilidade de suas ações que traduzem o comprometimento com suas intenções.

As intenções apresentam também uma forma hierárquica, a intenção mais geral, prevalece sobre os menos gerais, que geralmente estão sujeitas a mudanças para preservar as intenções gerais. As menos gerais são flexíveis e sensíveis a mudanças de contextos, as intenções gerais só mudam a partir de mudanças nas crenças. Essas informações nos levam a concordar com Bratman que a intenção tem um papel diferente das crenças e dos desejos, merecendo uma análise especial.

### **O trilema**

Antes de avançarmos é importante perceber o motivo da preocupação de Bratman com as intenções direcionadas ao futuro. Por que deveríamos nos incomodar de estabelecer hoje uma intenção para uma ação futura? Temos que enfrentar três objeções à noção de intenção direcionada ao futuro:

A ação voltada ao futuro é uma ação que ocorre à distância. Algo como uma mão fantasma invisível que atuaria no futuro a partir da nossa intenção.

Uma intenção formada agora estaria passível de mudança, dado que obviamente, o contexto no qual ela deve se concretizar provavelmente será diferente do que tínhamos em mente ao formar essa intenção.

Seria uma perda de tempo se preocupar com o futuro desta maneira.<sup>5</sup>

---

5 Estas objeções podem ser encontradas no capítulo 2 de IPPR.

## Planos

A intenção direcionada ao futuro é pensada na forma de um plano.<sup>6</sup> Os planos são estados mentais envolvendo um tipo apropriado de comprometimento com a ação: Eu tenho um plano A somente se é verdadeiro que planejo A. Se e somente se o agente está disposto a tentar executar o plano, mesmo sem ter a certeza, crença ou convicção de que será bem sucedido. As principais características dos planos são:

1) Parcialidade - Planos são parciais, essencialmente incompletos, preenchidos com o passar do tempo. Esta característica é um reflexo da nossa limitação cognitiva e escassez de recursos.

2) Estrutura hierárquica - Planos em geral são sobre fins e incorporam planos sobre meios e passos preliminares. Da mesma forma que intenções gerais incorporam intenções específicas. Esta característica deriva do ponto anterior.

Dado a sua estrutura e organização eles apresentam uma vantagem pragmática, auxiliando na coordenação entre nós mesmos e com os outros. Pensar tudo antecipadamente é humanamente impossível. Lidar com pequenas partes e realizar tarefas simples é muito mais fácil e está em acordo com nossas capacidades limitadas e deliberação e processamento de informação. Respeitando a estrutura limitada de nossas mentes para podermos tirar mais proveito dessa forma de organização. Deste modo, nos tornamos agentes mais confiáveis, aumentando e melhorando a avaliação que os outros tem de nós enquanto parceiros de trabalho. A parcialidade dos planos também permite que outras pessoas possam nos auxiliar neste processo de preenchimento dos passos preliminares, dividindo o trabalho e tornando-o executável. Todas as atividades complexas, em geral, seguem este tipo de estrutura.

## Estrutura dos planos anteriores

Nossos planos não começam do zero, estamos dentro de uma estrutura psíquica que possui inúmeros planos anteriores, todos eles em maior ou menor grau interligados entre si. Devido a coerência meio-fim, as intenções anteriores põem problemas para a deliberação. Nos fazendo escolher entre meios possivelmente conflitantes ou alternativos. Direcionando nossa energia e fluxo de ações para as suas execuções, estas demandas de consistência restringem o tipo de intenção que pode ser introduzida no raciocínio.

Os planos e intenções anteriores põe problemas para a deliberação, estabelecendo padrões de relevância para as opções que serão consideradas na deliberação, soma-se a isto as restrições de soluções para estes problemas, temos assim um filtro de admissibilidade que filtra as opções relevantes. Deste modo planos e intenções anteriores ajudam a tornar o processo deliberativo mais tratável para seres com capacidades limitadas como nós. Pois ao Diminuir as variáveis envolvidas em nossas escolhas ganhamos tempo para outras atividades. Por exemplo, um cardápio com opções variadas num restaurante faz o cliente perder muito mais tempo pensando no que comer do que realmente aproveitando o ambiente ou a companhia.

Além disso os planos e intenções anteriores provém uma estrutura de background contra a qual podemos pesar as razões das crenças e desejos, seja a favor, seja contra. Ajudam a determinar quais opções são relevantes e admissíveis.

<sup>6</sup> Como vimos anteriormente faz parte da nossa natureza ser um agente que planeja e estar envolvido nessa rede de planos.

## **Demandas sobre planos**

### 1) Restrição de consistência

Os planos precisam ser internamente consistentes, devem estar de acordo com minhas crenças, não existindo contradições entre o que eu tenho a intenção de fazer e aquilo que acredito ser possível. Não quer dizer que o agente deva acreditar que será bem-sucedido, o único requerimento aqui é que ele não acredite que não será impossível de realizar tal intenção. A relação entre o plano e a crença deve ser como segue: “Ele deveria ser bem-sucedido num mundo em que minhas crenças são verdadeiras.”

### 2) Coerência meio-fim

Requer que meus planos sejam preenchidos com especificações tão detalhadas, da minha perspectiva, quanto seja necessário para sua execução com sucesso. Este tipo de detalhamento é num estágio inicial ainda bastante geral, estando de acordo com a incompletude apresentada acima. Neste momento são os primeiros passos que devem ser tomados em direção a execução do plano, é necessário também que neste processo todos os passos estejam de acordo entre si e não sendo incompatíveis, sob risco de colocar o sucesso da execução do planejamento em risco.

### Intenção e Comprometimento

No entanto, nada do que vimos até agora funciona sem comprometimento com o plano (formalização da intenção). Existe uma diferença fundamental entre a intenção e o desejo quando analisamos o aspecto volitivo, a motivação para a ação, envolvido em cada um dos casos. A volição do desejo é potencialmente controladora de conduta, o agente pode ter o desejo de tomar um milk shake ao invés de almoçar e não fazer isto. Já quando ele estabelece uma intenção e a formaliza através de um plano, o caso é diferente. A intenção irá controlar a conduta do agente para que seu plano seja bem-sucedido. Para tanto, o agente se vale de outra característica dos planos, a inércia.

O agente age em função de uma intenção anterior, mas poderia ter realizado uma ação diferente da que fez, se tivesse parado para reconsiderar aquela intenção. Bratman assume que a racionalidade prática é uma questão de satisfazer desejos racionais. Para ele, embora a intenção anterior não nos forneça razões, como as crenças e desejos fornecem. “Os planos e intenções anteriores fornecem uma estrutura de fundo dentro da qual muitas das deliberações tomam lugar.” Estas deliberações dão forma sobre o que é racional, a fazer, quando estamos tomando uma decisão. (BRATMAN, 1992, p. 4)

### **Inércia ou estabilidade**

A opção padrão dos planos consiste em:

- 1) Retenção da intenção anterior e não reconsideração;
- 2) Disposição para raciocinar de uma intenção anterior para outra (frequentemente dos meios ou passos preliminares direcionados ao fim desejado)
- 3) Restrição de outras intenções (filtro de admissibilidade)

A intenção tem o papel de influenciar nossa conduta e controlar as opções que teremos, criando um filtro de admissibilidade para as ações e um comprometimento com a ação que irá realizar essa intenção. A partir de tais constatações, Bratman afirma que a intenção não pode ser redutível nem a crença nem ao desejo, as “intenções são estados mentais no mesmo nível de crenças e desejos. (BRATMAN, 1987, pos. 487) Devido às suas

características particulares, as intenções dão suporte às ações futuras e controlam a nossa conduta, elas não podem ser contraditórias entre si, isso levaria o agente a ser acusado de irracionalidade, assim como a dificuldade prática de que compromisso assumir.

### **Os três papéis e as três disposições**

#### **Papéis**

Intenções normalmente põe problemas para o agente: O agente precisa determinar uma maneira de realizar a intenção.

Intenções promovem um “filtro de admissibilidade” para a adoção de outras intenções. Enquanto desejos podem ser inconsistentes, por exemplo, o desejo de comer uma sobremesa e emagrecer. Os agentes normalmente não adotam intenções que eles acreditam que são conflitantes com suas intenções direcionadas ao presente ou ao futuro.

Os agentes perseguem o sucesso de suas das tentativas de realização das intenções. Não somente estão preocupados com o sucesso do que fazem, como estão dispostos a replanejar se a tentativa anterior não foi bem-sucedida. (COHEN; LEVESQUE, 1990, p. 213-261)

#### **Disposições**

Nos identificamos três tipos de disposições associadas com a intenção. Elas são: controle de conduta - e não meramente um potencial influenciador de conduta - pro atitudes resistem a reconsideração, tendo a característica da inércia; e representam papéis característicos como inputs para o raciocínio prático e outras intenções. (BRATMAN, 1987, pos. 487)

Elas facilitam a orientação do raciocínio, e a coordenação, seja intrapessoal, seja interpessoal, restringindo outras possibilidades de raciocínio e outros cursos de ação. Desta maneira, as intenções dão suporte a expectativa da ação que o agente se propõe a executar, fazendo com que as demais pessoas esperem que tal ação seja concretizada, inclusive planejando em conjunto com ele ou na expectativa da concretização de sua intenção.

### **Lidando com o trilema**

Voltando a questão apresentada acima, em que consiste a preocupação com ações futuras, e por que deveríamos nos incomodar com isso? Ter uma Intenção de algo é estar de algum modo comprometido com a realização de algo. Mas não se deve entender esse comprometimento como um passo ou uma preliminar do que irá ser pretendido fazer. É somente com uma parte das condutas futuras que eu estou agora comprometido a realização. Até mesmo porque, como foi dito, os planos são parciais e não é possível estar no momento em que deliberamos em favor de uma intenção estar comprometido com todos os detalhes não existentes de um plano. “Intenções são, por assim dizer, os tijolos de construção de tais planos; e planos são intenções em larga escala.” (BRATMAN, 1987, pos. 205) Com o passar do tempo o agente deve se comprometer com as demais ações para a realização completa daquilo que foi planejado.

## **Conclusão**

Nosso objetivo aqui foi demonstrar como se dá o processo de autonomia do conceito de intenção a partir de Michael Bratman. Pôr este conceito no mesmo patamar das crenças e desejos na explicação de ações intencionais, criando o modelo crença-desejo-intenção (DBI em inglês). Isto foi feito através da perspectiva das intenções direcionadas ao futuro como paradigma para pensar as ações intencionais, da percepção que estas ações estão incorporadas em planos e que existem regras de regem este jogo e suas relações com as crenças e desejos que formam o background no qual pesamos nossas intenções antes de formalizá-las em planos. Por fim, a finalidade do trabalho de Bratman é atualizar a forma de pensar o que é uma intenção dentro do escopo da pesquisa de Herbert Simon, através do conceito de racionalidade limitada e atualizar o trabalho do Davidson.



**Referências bibliográficas:**

BRATMAN, M. *Intention, Plans and Practical Reasoning*. Harvard University Press, 1987.

\_\_\_\_\_. *Planning and Stability of Intention*, 1992. Disponível em: < <https://link.springer.com/article/10.1007/BF00261286>>

CADILHA, Susana. A teoria da ação de Donald Davidson e o problema da causação mental. Disponível em: < <http://ler.letras.up.pt/uploads/ficheiros/9973.pdf>>

COHEN, Philip; LEVESQUE, Hector J. Intention is a Choice with Commitment. *artificial intelligence* 42 (1990) 213-261

DAVIDSON, Donald. *Action, Reasons and Causes*. In: *Essays on Actions and Events*, Oxford, Oxford University Press, 1980.

DAVIDSON, Donald. *Agency*. In: *Essays on Actions and Events*, Oxford, Oxford University Press, 1980.

LECLERC, A. *Questões Fundamentais da Teoria da Ação*. In: João Carlos Brum. (Org.). *Manual de Ética. Questões de Ética Teórica e Aplicada*. 1ed. Petrópolis: Vozes, 2014, p. 71-90.

SIMON, Herbert. *Reason in Human Affairs*. Stanford University Press, 1983.

VARGAS, Manuel; YAFFE, Gideon. *Rational and Social Agency*. Oxford University Press, 2014.

WILSON, George and SHPALL, Samuel, "Action", *The Stanford Encyclopedia of Philosophy* (Summer 2012 Edition), Edward N. Zalta (ed.), disponível em: <http://plato.stanford.edu/entries/action/>

# **Democracia cooperativa ou socialismo consciente? A proposta de John Dewey de sociedade democrática**

Maria A. Lima Piai<sup>1</sup>

## **Introdução**

A proposta de Dewey de sociedade democrática é de uma sociedade mais igualitária possível, sem separação de classe, existindo o maior número de empreendimentos e experiências das quais todos possam participar sem dualismo educacional, sem uma educação para senhores e uma educação para escravo, mas uma educação para a igualdade, para a democracia. O homem, para Dewey, é um ser social, desta forma, a educação é uma necessidade, pois não há uma natureza fixa, mas plástica. Daí a necessidade de que cada sujeito compreenda sua importância e papel social, assim como a importância e o papel social de cada membro da sociedade para o desenvolvimento dessa mesma sociedade. A liberdade que no liberalismo econômico voltou-se para o individualismo e a propriedade privada são criticados por Dewey que propõe o intercâmbio e fruição cultural e o estímulo intelectual. O ideal democrático deweyano envolve maior participação dos interesses, cooperação e maior confiança na reciprocidade desses interesses, regulamentação e direção social, mudança nos hábitos sociais e consciência dessas mudanças e da necessidade de uma contínua readaptação desses hábitos diante de novas situações criadas pela comunicação e intercâmbio das experiências. Debater a temática democrática significa consolidar cada vez mais a ação democrática e uma postura de dialogada diante dos acontecimentos políticos atuais que permeiam o Brasil em mais um momento de enfraquecimento das bases democráticas, portanto, de enfraquecimento do princípio interno de autoridade democrática. Numa sociedade com aspirações democráticas o processo educativo estaria de acordo com esse ideal democrático e as experiências escolares seriam experiências democráticas. Dewey afirma que uma sociedade “deve procurar fazer que as oportunidades intelectuais sejam acessíveis a todos os indivíduos, com iguais facilidades para os mesmos [...] assim, a democracia é mais do que uma forma do governo, é uma forma de vida associada, de experiência conjunta e mutuamente comunicada” (DEWEY, 1979, p 94).

## **Educação e Democracia**

Uma sociedade ou comunidade se constitui por interesses comuns, mas esses interesses comuns nem sempre visam o bem de toda sociedade, mas geralmente de um grupo. A educação varia de acordo com a qualidade da vida do grupo social e toda “educação ministrada por um grupo tende a socializar seus membros, mas a qualidade

---

1 Doutoranda em Educação- usp, professora de filosofia da educação – UNESPAR

e o valor da socialização dependem dos hábitos e aspirações do grupo” (1979, p. 88-89). Assim o fator moral não está dissociado das pretensões educacionais e todo esse contexto evidencia, para Dewey, a necessidade de fazer julgamento aos diferentes modos de vida social. No entanto, Dewey indaga sobre os critérios adequados para fazer esse julgamento e salienta a necessidade de evitar dois extremos ao estabelecer esses critérios: o de imaginar uma sociedade ideal sem referência na realidade tornando possível alcançar aquilo que foi idealizado e, por outro lado, não se fixar na realidade a ponto de somente reproduzi-la sem novas perspectivas de progresso. “O problema consiste em extrair os traços desejáveis das formas de vida social existente e emprega-los para criticar os traços indesejáveis e sugerir melhorias” (DEWEY, 1979, p. 89).

Quais seriam as características dos grupos sociais? Seriam a interação e a reciprocidade cooperativa. Pois “em qualquer grupo social, mesmo em malhas de ladrões, encontramos algum interesse comum e, além dele, certa porção de interação e reciprocidade cooperativa com outros grupos” (DEWEY, 1979, p. 89). Mas até que pontos os interesses comuns são conscientemente compartilhados e até que ponto as relações com outras formas de associação são livres e intensas? As considerações de Dewey são: a) numa quadrilha de malfeitores: conclui-se que a educação é parcial e falseada; b) numa família: existem muitos interesses conscientemente comunicados e compartilhados e pontos de contato com outras modalidades de associação ou instituições: conclui-se uma educação mais completa e verdadeira. Cada grupo oferece um tipo de vida, portanto um tipo de educação.

Os valores comuns se fortificam e multiplicam conforme a igualdade de oportunidade vai atingindo todos os membros de uma sociedade. Há uma necessidade de um intercâmbio entre o conhecimento, entre as diversas formas de pensar, de produção social. Há uma necessidade de comunicação das experiências.

A sociedade democrática proposta por Dewey é de uma sociedade que busca a igualdade entre seus membros, possibilitando maior número de empreendimentos e experiências das quais todos possam participar sem dualismo educacional, sem uma educação para senhores e outra para servos, mas uma educação para a igualdade, para a democracia. No processo educacional dual mesmo a classe supostamente privilegiada perde, pois cria uma cultura fechada em si mesma, sem fruição, superespecializada e muito mais artificial que humana. O homem enquanto ser social necessita de uma educação, pois para Dewey esse não possui uma natureza fixa, mas plástica.

### **Liberdade e democracia**

Dewey compreende a democracia como uma experiência de participação voluntária e cooperativa que permite resolver as dificuldades apresentadas pela sociedade e “é somente a experiência de participar voluntária e cooperativamente em grupos para resolver problemas e aproveitar oportunidades, que pode apontar para o indivíduo a necessidade de um espaço público democrático” (FRANCO, 2008, p. 18). O indivíduo toma consciência da sua condição de agente político, de sua responsabilidade e da necessidade de cooperação democrática a partir da sua participação nos empreendimentos públicos/comunitários.

É preciso estar atento, segundo Dewey, às ações que desviam as práticas democráticas tais como a “Intolerância, abuso, ofensas pessoais por causa de diferenças de opinião sobre religião ou política ou negócios, bem como por causa de diferenças de raça, cor, riqueza ou nível cultural são uma traição ao modo de vida democrático” (DEWEY, 2008b, p. 135).

Dewey, quando propõe a radicalização da democracia, traz o caráter cooperativo, comunitário e socialista do estilo vida democrática. Para Franco, “Dewey percebeu que toda democracia é local, no sentido de que a democracia é um projeto comunitário. Ele não tinha, como é óbvio, as palavras atuais para descrever o que pensava, mas farejou os conceitos – como se ouvisse ecos do futuro” (2008, p. 20).

De acordo com Dewey o capitalismo manipula a democracia e o liberalismo, e o que nos são apresentados como Estados democráticos são democracias burguesas, na quais o poderio se encontra “em última instância, nas mãos do capitalismo financeiro, a despeito das reivindicações que são feitas por um governo do povo, pelo povo e para o povo” (DEWEY, 2008a, p. 125). Dewey enfatiza que a democracia, como definida por Lincoln: do povo, para o povo e pelo povo, ainda não se alcançou por isso ele está buscando e propondo novas perspectivas em sua filosofia da educação.

A liberdade que no liberalismo econômico voltou-se para o individualismo e a propriedade privada é crítica por Dewey, que propõe o intercâmbio e fruição cultural e o estímulo intelectual. O trabalho e a técnica precisam andar juntos. A ciência desvinculada das questões morais e sociais reduz o homem à maquinaria.

Quando analisadas historicamente, o desenvolvimento dos governos supostamente democráticos, é possível constar a transferência de interesses que outrora eram agrário para, agora, industriais, comerciais e financeiros. Aquilo que chamamos liberalismo era, segundo Dewey a transferência do poder, pois os representantes do poder comercial e industrial legitimou aquelas reivindicações como sendo em prol da liberdade do indivíduo, mas o que estava por trás era o máximo de liberdade econômica, ou seja, de ações econômicas individuais, sem preocupação e sem controle social, ações focadas no interesse daqueles empenhados na produção e comércio” (DEWEY, 2008a, p. 125-126). Dessa forma, o chamado liberalismo econômico se apresentou, num primeiro momento, como um cordeiro dócil ao grupo, mas era um lobo faminto e antissocial.

O liberalismo é visto por Dewey como ultrapassado, pois não atende as necessidades sociais dando poder para um pequeno grupo decidir sobre o modo de ser e pensar de muitos. O liberalismo econômico cria barreiras e reprime a liberdade e o desenvolvimento dos indivíduos e da própria individualidade. Assim, para Dewey existe uma combinação entre os meios democráticos liberais com as finalidades socialmente radicais. Ele enfatiza ainda que “nem a história nem a natureza humana dão motivos para se supor que fins socialmente radicais possam ser atingidos por outros meios que não os meios democráticos liberais” (DEWEY, 2008a, p. 132).

O capitalismo e o liberalismo não oferecem, segundo Dewey, condições para a concretização da democracia, pois ao dar ênfase à competitividade, ao consumismo, ao acúmulo de bens materiais, cria exclusão. A democracia na definição deweyana é uma experiência social e a experiência só é social se for compartilhada e solidarizada, o liberalismo capitalista não tem preocupação social, coletiva ou solidária, mas individualista; bem diferente da proposta liberal deweyana que tem simpatia ao bem comum, ao livre desenvolvimento da capacidade humana que se dá no contexto socialmente democrático.

### **Considerações**

As oportunidades de desenvolvimento não deveriam se restringir a um grupo; as ações democráticas pautam num interesse mútuo e não individualista. A educação tem um

papel crucial com o propósito de evitar essa separação entre ciência/técnica e moralidade/intelectualidade permitindo o equilíbrio entre os interesses sociais e os estímulos intelectuais para que se fortaleçam as relações humanas envolvidas nas atividades industriais.

O ideal democrático deweyano envolve maior participação dos interesses comuns e maior confiança na reciprocidade desses interesses, regulamentação e direção social, mudança nos hábitos sociais e consciência dessas mudanças e da necessidade de uma contínua readaptação desses hábitos diante de novas situações criadas pela comunicação e intercâmbio das experiências. Essas características delimitam uma sociedade democraticamente constituída ou não, pois democracia para Dewey não se limita a uma forma de governo, e sua constituição depende da educação.

Fazer as instituições políticas funcionarem é uma aposta muito pequena na democracia. Apostar na democracia é confiar nas ações e capacidades do “homem comum” de guiar sua vida, é apostar na natureza humana, nas relações livres e mútuas da vida em sociedade e não de relações coercitivas. Para Dewey, a educação democrática é aquela na qual a igualdade de oportunidades é um elemento fundamental, isto é, todos os indivíduos presentes no processo de ensino e aprendizagem devem ter a mesma oportunidade de ensino.

Para que a democracia seja alcançada de forma efetiva seria necessário, de acordo com Dewey, uma grande mudança nas instituições jurídicas, econômicas, sociais e culturais, como por exemplo: as escolas. A escola deveria ser, para esse filósofo, um lugar de experiência democrática, de inteligência social. Alcançar a democracia para Dewey, depende de uma postura social nova, de uma confiança no outro, de dar novos sentidos para a sua ação, para a sua vida. A educação, então se apresenta para Dewey como o elemento capaz de proporcionar o desenvolvimento da democracia, o desenvolvimento para todos numa aposta nas relações sociais e nas pessoas como capaz de autogoverno, como pessoas capazes de direcionar suas próprias vidas no contexto das relações com sociais. A democracia seria um espaço de oportunidade para todos, apresentando-se com uma proposta socialista.

**Referências bibliográficas:**

DEWEY, John. Democracia e Educação: introdução à filosofia da educação. Trad. Godofredo Rangel e Anísio Teixeira. 4. ed. São Paulo: Companhia Editora Nacional, 1979.

\_\_\_\_\_. A democracia é radical (1937). In: DEWEY, J. Democracia Cooperativa: Escritos Políticos Escolhidos de John Dewey (1927-1939). (Orgs. Augusto de Franco e Thamy Pogrebinski) Porto Alegre: Edipucrs, 2008a. p. 125-128.

\_\_\_\_\_. Democracia criativa: a tarefa diante de nós (1939) In: DEWEY, J. Democracia Cooperativa: Escritos Políticos Escolhidos de John Dewey (1927-1939). (Orgs. Augusto de Franco e Thamy Pogrebinski) Porto Alegre: Edipucrs, 2008b. p. 131-138.

FRANCO, Augusto de. Introdução. In: DEWEY, J. Democracia Cooperativa: Escritos Políticos Escolhidos de John Dewey (1927-1939). (Org: Augusto Franco e Thamy Pogrebinski) Porto Alegre: Edipucrs, 2008. p.13-20.



## **Causalidade, classes naturais e Interpretantes Finais. Uma abordagem semiótica para a causação final<sup>1</sup>**

Priscila Borges<sup>2</sup>  
Juliana Rocha Franco<sup>3</sup>

Apresentamos neste artigo uma discussão sobre os tipos de causas finais que guiam o modo de ação dos signos. Tratamos da teoria da causação de Peirce, sua relação com as classes naturais e da possibilidade de relacionar a causação final e a causação eficiente à semiótica. Propomos observar um sistema de classes de signos considerando seu aspecto teleológico, isto é, como um sistema que descreve uma cadeia de possíveis eventos que apresentam uma certa tendência.

A causação final em Peirce é entendida como algo que determina um certo caráter geral para o resultado de um processo semiótico (1902, EP 2:120; CP 1.211). Para Peirce, toda classificação, natural ou artificial, deve estar relacionada a algum propósito. Classificações naturais não envolvem nossos propósitos, mas as causas finais das próprias coisas classificadas.

Delineiam-se duas coisas: a ideia de classe como classificação e a relação entre a classe e a causalidade final. Certas coisas pertencem à mesma classe natural devido a uma dimensão metafísica que unifica as características de uma classe. Tal dimensão corresponderia à causalidade final que é um princípio geral em virtude do qual os membros da classe tendem a se comportar de uma maneira específica.

A causalidade final estaria relacionada ao interpretante final no sistema de classes de signos, pois o interpretante final “é aquilo que deveria finalmente ser decidido como sendo a interpretação verdadeira se considerações sobre o assunto fossem levadas tão longe que uma opinião final fosse alcançada” (PEIRCE, 1909, EP 2:496-7). O interpretante final revelaria, então, a essência geral das coisas.

No sistema de 66 classes de signos Peirce propõe modos para o interpretante final fundamentado nas três categorias fenomenológicas. Signos que apresentam interpretantes finais do modo da primeiridade, secundidade e terceiridade são denominados respectivamente gratificantes, práticos e pragmáticos. Tais nomenclaturas podem ajudar a pensar modos de causação final a partir dos modos dos interpretantes finais.

---

1 Agradecimento à FAPDF pelo financiamento concedido.

2 Professora da UnB

3 Professora da Escola de Design da Universidade Estadual de Minas Gerais (UEMG)

### Causação final e classes naturais

No texto *On Science and Natural Classes*, Peirce (1902, EP2: 115-132) desenvolve um ponto de vista original sobre a causalidade. Para ele, aquilo que define uma classe natural é o fato dos membros dessa classe terem uma causa final comum (1902, EP2: 117). Causas finais naturais podem ser vagas ou desconhecidas do homem, mas são necessárias, especialmente se considerarmos processos evolutivos, que para Peirce, não são fortuitos, mas guiados por uma causa final (1902, EP2: 117). No caso de objetos artificiais, a causa final pode ser associada a um propósito. Mas o propósito é apenas o modo de operação do desejo, que é um geral. Portanto, “o propósito é apenas a forma da causa final que é mais familiar à nossa experiência” (1902, EP2: 120).

Para Peirce, em cada ato de causalidade há um componente eficiente e final. As causas finais são tipos gerais que tendem a realizar-se por processos teleologicamente determinantes de causação eficiente. Dessa forma, as causas finais não são propriamente eventos futuros, mas possibilidades gerais. O aspecto eficiente da causação é que cada evento ou fato é causado por um evento ou fato anterior (a causa eficiente). O aspecto teleológico está presente na ideia de que cada evento ou fato faz parte de uma cadeia de eventos com uma tendência definida. A tendência é determinada pela causa final do processo. Isso implica que cada ato de causação é mediado por uma causa final.

As causas finais são gerais e não podem ser eventos, pois os eventos são sempre individuais: “A causa final é, em todos os casos, um fim definido de um processo que influencia as fases iniciais desse processo para exibir uma tendência em direção a esse fim.”<sup>4</sup> (PAPE, 1993, p. 588). A causa final não é uma coisa existente. Hulswitt (2002, p.78) afirma que pode ser enganoso chamá-la de “causa”, pois este termo - em seu sentido moderno comum - sugere algo concreto, coisa existente, evento ou fato, que tenha uma influência determinante sobre uma outra coisa, evento ou fato. A causa final não é uma coisa concreta, mas um estado final ideal para o qual um processo tende. A causa final determina que tipo de meios (causas eficientes) são adequados para atingir o fim geral (HULSWIT, 2002, p.78).

De acordo com Peirce, as causas finais são tipos gerais que controlam a causação eficiente e determinam que os efeitos trazidos pela causação eficiente são de um certo caráter geral. A teoria das classes naturais de Peirce está intimamente relacionada à sua teoria da causação final. Hulswitt (2002, p. xviii), resume assim a concepção peirceana de classe:

Todas as coisas que pertencem a classes particulares fazem isso em virtude de alguma “essência” e algumas propriedades de classe adicionais. (...) A “essência”, então, é um princípio geral em virtude do qual os membros de uma classe possuem a mesma causa final, a qual está implicada na tendência deles de se comportarem de uma maneira específica.

Se as causas finais não podem ser eventos futuros, então, elas são tipos gerais que poderão ser realizados no futuro. Toda classificação tem referência a uma tendência, a um fim (PEIRCE, 1902, NEM IV: 65).

<sup>4</sup> “The final cause is in all cases a definite end of a process influencing earlier phases of this process to exhibit a tendency towards that end.” (PAPE, 1993, p. 588)

Toda classificação, seja artificial ou natural, é o arranjo de objetos de acordo com ideias. Uma classificação natural é o arranjo deles de acordo com aquelas ideias das quais sua existência resulta. Nenhum mérito maior pode ter um taxonomista do que ter seus olhos abertos para as ideias na natureza; não há mais deplorável cegueira que possa afligi-lo do que não ver que existem ideias na natureza que determinam a existência de objetos<sup>5</sup>. (PEIRCE, 1902, CP 1.231)

Para Peirce, se esta tendência é a tendência que determinou os caracteres de classe dos objetos, é uma classificação natural (1902, NEM IV: 65). O termo natural é utilizado para designar uma classe cuja existência é devida a uma causa final comum e peculiar.

Assim, a condição para causas e os efeitos serem mediados por causas finais é que eles pertencem a uma classe natural. As coisas pertencem a uma mesma classe natural, não por causa de certas qualidades (primeiros), que possibilitaria uma descrição das qualidades de tais objetos, mas conforme uma essência metafísica que é uma causa final (ou terceiro).

### **A causação final a partir das 66 classes de signos**

A semiótica apresenta sistemas de leis gerais criados a partir de experiências particulares para explicar como podemos atingir esses ideais, relacionando os fenômenos com seus fins. Dessa forma, é de suma importância compreender a semiótica a partir de um paradigma processual-pragmático como o estudo dos processos e das práticas de produção de sentido, sobre como os signos se configuram como “eventos produtivos” na produção de outros signos.

Os sistemas de classes de signos traçam caminhos possíveis para que os signos alcancem seus fins (PEIRCE, ca.1897, CP 2.227), por isso são eficientes como método para análises semióticas. Nos textos de Peirce há propostas de sistemas com 3, 10, 28 e 66 classes de signos. O que propomos neste artigo é uma discussão a respeito da causação final a partir das 66 classes de signos. Tal sistema, por traçar o maior número de caminhos possíveis para que os signos alcancem seus fins, apresenta muitos padrões que dizem sobre o modo de funcionamento possível dos signos a partir dos quais podemos discutir diversas questões filosóficas, entre elas a causação final.

A causação final é um tipo de lei, ou princípio geral, que guia a ocorrência individual de ações. No entanto, essa lei não determina precisamente o modo como a ação ocorrerá. Apesar de governada pela mesma lei, as ações podem ocorrer de muitos modos, pois a lei define apenas que o resultado terá um certo caráter geral (PEIRCE, 1902, EP 2:120; CP 1.211). Em termos semióticos, a causação final é uma lei que guia o modo de ação dos signos de forma que os interpretantes gerados pela semiose tenham um certo caráter geral.

O signo, por sua vez, é aquilo que faz a mediação entre objeto e interpretante. O signo é determinado pelo objeto e ele é capaz de determinar um interpretante que estabelece com o objeto o mesmo tipo de relação que o próprio signo tem com o objeto (PEIRCE, 1904, SS:31-2; CP 8.332). A descrição do modo de ação do signo tem pelo menos dois processos de determinação: a determinação do signo pelo objeto e a determinação do interpretante pelo signo. Esses processos de determinação são modos de ação do signo por causação final (SANTAELLA, 2006). Em um processo semiótico ideal, o interpretante

5 All classification, whether artificial or natural, is the arrangement of objects according to ideas. A natural classification is the arrangement of them according to those ideas from which their existence results. No greater merit can a taxonomist have than that of having his eyes open to the ideas in nature; no more deplorable blindness can afflict him than that of not seeing that there are ideas in nature which determine the existence of objects. (PEIRCE, 1902, CP 1.231)

produzido pelo signo corresponderá ao objeto do signo. Tal interpretante revelará o objeto do signo. Uma vez que não temos acesso direto ao objeto do signo, será por meio de um signo que representa seu objeto e determina um interpretante que poderemos ter ideia do objeto do signo.

Essa capacidade do objeto de determinar, que inicia um processo sgnico dentro de uma cadeia de semioses infinitas, pode ser vista como o objetivo da semiose. O que nos levaria a pensar que a causação final está de certo modo na capacidade do objeto de determinação (PAPE, 1993, p. 585). Tal capacidade, assim como a causação final é uma tendência que não depende de suas atualizações. Embora a causação final diga respeito a uma tendência a ser alcançada no futuro, a tendência em alcançar certo estado acaba por influenciar o modo das ações que antecedem esse estado futuro.

Segundo Pape (1993, p. 594) não pode haver um processo sgnico sem que haja uma causa final ativamente envolvida no processo<sup>6</sup>. Podemos perceber na capacidade do objeto de determinar um signo uma tendência quase inicial (quase, pois não é possível alcançar o início da semiose) de um processo sgnico que se relaciona com a ideia de causação final. Contudo, a causação final não poderia se esgotar no processo de determinação do signo pelo objeto, uma vez que é essencial ao funcionamento do signo, principalmente ao signo ideal, que ele produza interpretantes dotados de hábitos ou regras gerais que possam vir a ser novos signos capazes de agir conforme essa lei ou tendência em ocasiões futuras.

O processo de determinação do signo pelo objeto pode ser identificado como sendo a causação final básica ou fundamental para o processo sgnico, mas tal causação precisa se manifestar no processo de geração de interpretantes para que tais interpretantes correspondam também ao objeto do signo. Quando passamos ao conceito de signo mais detalhado, com a subdivisão dos objetos e dos interpretantes, a cadeia de interpretantes pode ser ainda mais rica para percebermos o modo como a causação final aparecerá no processo sgnico.

A divisão dos objetos em objeto dinâmico e imediato, diferencia o objeto real do objeto tal como representado pelo signo. O objeto dinâmico é o objeto real que determina o modo como o signo irá representá-lo. Já o objeto imediato é o modo como o objeto dinâmico foi representado no signo. A relação entre o objeto dinâmico e a metafísica é evidente, uma vez que o objeto dinâmico é o real externo ao signo, que pode ser conhecido por meio do signo que o representa em parte, mas não completamente. O real, para Peirce, é justamente o que independe de suas representações (1870, W2:439-40), não é algo completo a ser descoberto, mas algo incompleto que se completa no futuro. O objeto dinâmico pode ser considerado o aspecto no processo semiótico que indica a relação da semiótica com a metafísica. Os três tipos de objetos dinâmicos também podem ser relacionados com o modo como as categorias fenomenológicas são expressas na metafísica. Abstrativo, concreto e coletivo são os tipos de signos que possuem objetos dinâmicos da natureza da primeira, segunda e terceira categorias fenomenológicas, respectivamente. Acaso, existência e lei são as categorias fenomenológicas expressas na metafísica. Portanto, o acaso se relaciona aos signos abstrativos, a existência aos signos concretivos e as leis aos signos coletivos.

Os **signos coletivos** possuem objetos dinâmicos da natureza da terciridade. Tal objeto mostra que a realidade, expressa na semiótica pelos objetos dinâmicos, tende à generalização e à formação de leis, chamada lei da causação final.

6 "there cannot be a sign-process without there being a final cause actively involved in it." (PAPE, 1993, p.594)

Devemos entender por *causação final* aquele modo de trazer à tona fatos de acordo com os quais uma descrição geral de resultado acaba sendo gerada, sem muita relação com qualquer compulsão particular para que ela seja gerada, muito embora os meios podem ser adaptados para o fim<sup>7</sup>. (PEIRCE, 1902, CP 1.211)

Podemos entender que os signos coletivos possuem como objeto as leis gerais da realidade. Os **signos concretivos**, por sua vez, possuem objetos dinâmicos da natureza da secundidade. Tais objetos representam as leis eficientes, aquelas que mostram como a *causação final* ocorre efetivamente fazendo com que as coisas se conformem à lei. Na *causação eficiente*, o resultado da ação não tem importância para o processo de determinação, ao contrário do que ocorre na *causação final*. Os signos coletivos, que possuem objetos dinâmicos da terceiridade, podem ser relacionados à *causação final*, pois eles mostram que a realidade tende à formação de leis. Já a possibilidade das leis se efetivarem é garantida pelos signos concretivos que com objetos dinâmicos da secundidade expressam a lei da *causação eficiente*. Para completar esse sistema que tende à formação de leis falta algo que impulse a formação de leis.

Os **signos abstrativos** têm esse papel, uma vez que possuem objetos de possibilidade e que podem ser relacionados à lei do acaso na realidade. Tal lei expressa o princípio ontológico gerador de diversidade, capaz de incorporar novos elementos e expressar a capacidade de crescimento da realidade. Os modos como os objetos dinâmicos se apresentam, portanto, referem-se aos modos como esses objetos determinam os signos. Os objetos dinâmicos não são coisas ou formas representadas pelo signo, mas leis, ou causas, que guiam o processo de determinação *sígnico* de três modos, abstrativo, concretivo e coletivo.

Ainda que os objetos dinâmicos no sistema de 66 classes expressem as causas presentes na metafísica de Peirce, os modos como essas causas atuam nos processos *sígnicos* devem aparecer também na determinação de signos e seus interpretantes.

A realidade é incompleta, ou seja, os objetos com os quais lidamos em nossas interpretações da realidade são incompletos porque permitem uma gama de interpretações bastante diferentes e, às vezes, mutuamente exclusivas. Portanto, o que completará a realidade depende de serem verdadeiros interpretantes finais realmente eficazes. Mas essa é apenas outra maneira de dizer que, em qualquer momento futuro, a realidade se tornará mais completa pela criação ou incorporação de novos propósitos.<sup>8</sup> (PAPE, 1993, p. 601)

Veremos a seguir como seguindo a lógica das categorias fenomenológicas determinados tipos de objetos dinâmicos podem gerar um grande e diferente número de interpretantes.

Do objeto dinâmico passamos para a observação dos interpretantes finais e de sua relação com os objetos dinâmicos. Apesar do processo de determinação estar presente em cada uma das passagens de uma tricotomia para a outra, que leva a uma descrição do

7 “We must understand by final causation that mode of bringing facts about according to which a general description of result is made to come about, quite irrespective of any compulsion for it to come about in this or that particular way; although the means may be adapted to the end” (PEIRCE, 1902, CP 1.211).

8 “Reality is incomplete, that is to say, the objects we deal with in our interpretations of reality are incomplete because they allow for a range of rather different and, at times, mutually exclusive interpretations. Therefore, what will complete reality depends on their being really effective final interpretants. But that is just another way to say that at any future time reality will become more complete by the creation or embodiment of new purposes.” (PAPE, 1993, p. 601)



processo sgnico em 10 etapas no sistema de 66 classes, focaremos no interpretante final, pois ele, assim como o objeto dinmico, tem a caracterstica de funcionar como um princpio guia e no como um interpretante efetivo, como  o caso do interpretante dinmico. Alm disso, o interpretante final se refere ao modo como o signo tende a representar seu objeto (PEIRCE, 1906, CP 4.536). Essa capacidade do interpretante final de mostrar como o signo tende a representar no futuro pode estar relacionada com a causao final, principalmente se considerarmos que o processo sgnico  contnuo e evolutivo.

Nas 66 classes de signos os interpretantes finais se apresentam de trs modos: 1) so chamados gratificantes os signos cujo modo de ser do interpretante final  da primeiridade; 2) prticos, os signos com interpretantes finais de secundidade; e 3) pragmticos, os signos com interpretantes finais que seguem a lgica da terceiridade. Os interpretantes finais dizem respeito aquilo que seria finalmente decidido como a verdadeira interpretao, se o assunto fosse considerado a tal ponto que uma opinio final fosse alcanada (PEIRCE, 1909, EP 2:496-7). Veremos a seguir como cada objeto dinmico pode determinar certos tipos de interpretantes finais nas classes de signos. As restrioes nas determinaoes so dadas pela lgica da determinao das categorias fenomenolgicas.

O signo abstrativo, por ter um objeto cujo modo de ser  da primeiridade, s determina aspectos e relaoes sgnicas que tambm seguem o modo da primeiridade. Conseqentemente, h apenas uma classe de signo abstrativo. Nessa classe, temos qualissignos icnicos que geram interpretantes finais gratificantes e cuja relao tridica entre signo, objeto dinmico e interpretante final  instintiva.

Signos concretivos, por terem objetos que tm o modo da secundidade, determinam aspectos e relaoes sgnicas que seguem o modo da primeiridade ou da secundidade. Isto , eles formam classes de signos possveis, qualissignos, e de signos existentes, sinsignos com relaoes icnicas e indiciais. As classes de sinsignos icnicos determinam interpretantes finais de primeiridade, isto , signos gratificantes, enquanto as classes de sinsignos indiciais podem determinar interpretantes finais de primeiridade ou secundidade, isto , signos gratificantes ou prticos.

J as classes de signos coletivos, com objetos de terceiridade, so aquelas nas quais observamos aspectos e relaoes sgnicas de todos os trs modos, primeiridade, secundidade e terceiridade. Alm de observarmos qualissignos e sinsignos precedidos por objetos de terceiridade, podemos observar que os objetos de terceiridade determinam legissignos, isto , uma lei que funciona como signo (PEIRCE, 1903, EP 2:291). As classes que possuem legissignos dizem respeito aos signos do tipo geral que significam por meio de suas instncias (PEIRCE, 1903, EP 2:291). Os legissignos coletivos podem estabelecer com seus objetos dinmicos, relaoes icnicas, indiciais e simblicas. A relao simblica  um tipo de relao que se estabelece por meio de uma lei ou conveno. Peirce (1904, EP 2:323), define o smbolo como sendo essencialmente um propsito, isto , uma representao que procura tornar-se definida, ou produzir um interpretante mais definido que ela mesma.<sup>9</sup> Lembrando que, para Peirce o propsito  a forma da causa final mais familiar  nossa experincia (EP2: 120, 1902). O smbolo requer uma associao de ideias gerais para representar o seu objeto (PEIRCE, 1903, EP 2:294). Os interpretantes finais produzidos por tais smbolos, necessariamente precedidos por objetos dinmicos da terceiridade, podem ser dos trs modos, configurando signos gratificantes, prticos ou pragmticos.

<sup>9</sup> "A symbol is essentially a purpose, that is to say, is a representation that seeks to make itself definite, or seeks to produce an interpretant more definite than itself." (PEIRCE, 1904, EP 2:323)



No sistema de 66 classes de signos, observamos 28 classes de signos que apresentam uma relação simbólica. Tais classes referem-se a signos gerais, isto é, signos que se efetivam em réplicas que dizem respeito a circunstâncias particulares. No entanto, se estamos observando a causação final nas classes de signos, talvez seja mais importante focar nas classes de signos gerais, como as classes de signos simbólicos, do que observar as classes de signos que dizem respeito a signos particulares que nos mostrariam as possibilidades das causas eficientes. Os interpretantes finais de signos simbólicos devem mostrar verdadeiramente seus objetos dinâmicos de terceiridade. Sabemos que tais objetos não são coisas particulares, mas leis gerais, ou causações finais. Isto é, os interpretantes finais mostram como uma mente deverá atuar. No caso de signos simbólicos coletivos, os interpretantes finais indicam o modo como as causações finais que de certo modo deram início à ao processo desses signos podem continuar determinando processos semióticos subsequentes.

O modo **gratificante** do signo, indica um modo de agir futuro dado por uma familiaridade com o objeto, por hábitos. Tais hábitos densamente sedimentados acabam por criar modos de agir que configuram as características das coisas (COLAPIETRO, 2009, p. 352). São modos de agir que coincidem com o processo de determinação iniciado pela causação final. Isto é, os interpretantes seguem a tendência da causação final. Já o modo **prático** do signo, dado pelo interpretante final de secundidade, mostra um modo de agir futuro fundamentado em experiências práticas, como uma reação, sem que haja reflexão sobre elas. Nesse caso, o interpretante final alcançado é fruto da relação entre a tendência apontada pela causação final do objeto de terceiridade e a os signos efetivamente produzidos por essa lei. Isto é, o signo prático levaria em conta os outros signos efetivamente produzidos ou a ação das causas eficientes. Conseqüentemente, ele poderia indicar um modo de agir ligeiramente diferente da causação final inicial, uma vez que modificado por processos sígnicos que ocorreram de fato.

Por fim, o modo pragmático do signo envolve um modo de agir futuro fundamentado em reflexão e capaz de produzir auto-controle (PEIRCE, 1908, EP 2:490). As classes de signos pragmáticos, que possuem interpretantes finais de terceiridade, mostram signos que são submetidos ao pensamento crítico e seu significado é foco de atenção crítica de uma mente investigativa que por meio de um processo auto-controlado é capaz de modificar os hábitos e, conseqüentemente, criar novas causações finais:

Parece-me que a função essencial de um signo é referir-se a relações ineficientes e eficientes - não para pô-las em ação, mas para estabelecer um hábito ou uma regra geral por meio da qual agirão ocasionalmente.<sup>10</sup> (PEIRCE, 1904, SS:31)

Interpretantes finais pragmáticos dizem respeito a regras gerais criadas a partir de um processo reflexivo e auto-controlado.

(...) o pragmaticista não faz com que o summum bonum consista na ação, mas que consista naquele processo de evolução pelo qual o existente cada vez mais incorpore aqueles gerais aos quais estava destinado, sendo o que nos esforçamos para expressar ao chamá-los, a esses últimos, razoáveis [reasonable]. Em seus estágios mais elevados, a evolução toma lugar

10 "It appears to me that the essential function of a sign is to refer inefficient relations efficient, - not to set them into action, but to establish a habit or general rule whereby they will act on occasion." (PEIRCE, 1904, SS:31)

cada vez mais amplamente através do auto-controle, e isso fornece ao pragmaticista uma certa justificativa para fazer com que o teor racional seja geral. (PEIRCE, 1905, CP. 5.433).

O modo de agir pragmático está restrito aos signos simbólicos, isto é, signos gerais que agem a partir de suas réplicas que são signos particulares. As réplicas dos símbolos são tipos especiais de sinsignos indiciais (PEIRCE, 1903, EP2:295-6). Se observarmos os interpretantes finais dos sinsignos indiciais, veremos que eles variam entre interpretantes de primeiridade e secundidade, sem alcançar o modo pragmático do signo.

### Considerações Finais

Identificamos, portanto, nos signos três tendências que podem guiar signos simbólicos. Os signos tenderiam a 1) manter hábitos já estabelecidos pela própria força do hábito, que corresponde ao modo gratificante do signo; 2) modificar ou manter hábitos a partir da relação entre a causa final e as ações dos signos efetivamente produzidos, que corresponde ao modo prático do signo; e 3) modificar ou manter hábitos a partir de uma conduta racional, que corresponde ao modo pragmático do signo. Lembrando que, embora a máxima pragmática de Peirce afirme que o significado de um conceito, ou termo geral, está nas consequências práticas que concebemos a partir do conceito (1875, EP1:132). Peirce não restringe o seu pragmatismo à força das ações particulares. Tais consequências práticas podem estar relacionadas aos modos de conduta racional e não apenas a força bruta das ações.

A possibilidade de modificar a tendência dos signos expressa no modo prático e pragmático dos signos indica uma mudança nos objetos dinâmicos que incorporam essas novas tendências. Além do acaso, que possibilita a diversidade a partir de uma ruptura, a realidade, expressa pelo objeto dinâmico, se modifica por meio do movimento dos signos que possibilitam a criação de novas tendências que passarão a regular o fim de novas semioses.

Os processos sígnicos na teoria de Peirce ocorrem impulsionados por causas finais, isto é, leis gerais, que fazem com que o processo tenda a alcançar um determinado fim. Não há garantia desse fim e na geração de signos efetivos novas regras podem ser criadas, configurando diferentes causas finais. Conhecemos o mundo por meio de signos e os resultados que alcançamos devem tender a mostrar o modo como o mundo se organiza. Isto é, quando dizemos que entendemos um certo fenômeno, estamos afirmando que conseguimos prever o fim para o qual ele tende. Peirce afirma que “Tudo aquilo que for totalmente incomparável a alguma outra coisa é totalmente inexplicável porque a explicação consiste em colocar as coisas sob leis gerais ou sob classes naturais”. (Peirce, 1868, EP1:41; 2010, p. 272).

Se seguirmos métodos confiáveis, como o método científico (Peirce, 1877, EP1:109-141), nosso conhecimento pode de certo modo corresponder à causa final do fenômeno observado. Portanto, os processos de categorização e classificação, fundamentais ao processo de conhecimento, pois agrupam os signos de acordo com certa tendência a alcançar um fim, podem ser entendidos como sendo o resultado de uma certa ordem natural dos objetos dinâmicos.

**Referências bibliográficas:**

- COLAPIETRO, Vincent. Habit, Competence and Purpose: How to make the Grades of Clarity Clearer. *Transactions of The Charles S. Peirce Society*. Vol 45, no. 3, 2009.
- HAWKINS, Stephen B. Desire and Natural Classification: Aristotle and Peirce on Final Cause. *Transactions of The Charles S. Peirce Society*. Vol 43, no. 3, 2007.
- HULSWIT, Menno. From cause to causation – a peircean perspective. Dordrecht, Boston e Londres: Kluwer Academic Publishers, 2002.
- PAPE, Helmut. Final Causality in Peirce's Semiotics and His Classification of the Sciences. *Transactions of the Charles S. Peirce Society*. Vol. 29, No. 4 (Fall, 1993), pp. 581-607.
- PEIRCE, Charles S. *Collected Papers of Charles S. Peirce*, vols. 1-6, Hartshorne, C; Weiss, P. (Eds.); vols. 7-8, Burks, A.W. (Ed.), Cambridge, Mass.: Harvard Univ. Press, 1931-58. (Citado como CP, seguido do volume e página)
- \_\_\_\_\_. *The New Elements of Mathematics*, vols. 1-4, Eisele, C. (Ed.), La Haya: Mouton, 1976. (Citado como NEM, seguido do volume e página)
- \_\_\_\_\_. *Semiotic and Significs: The Correspondence between Charles S. Peirce and Victoria Lady Welby*, Hardwick, C. (Ed.), Bloomington, IN: Indiana University Press, 1977. (Citado como SS, seguido do volume e página)
- \_\_\_\_\_. *The Essential Peirce. Selected Philosophical Writings*, vols. 1-2, Houser, N. et al. (Eds.), Bloomington, IN: Indiana University Press, 1992-98. (Citado como EP, seguido do volume e página)
- \_\_\_\_\_. *Writings of Charles S. Peirce*, vols. 1 to 6. Vol. 1, Max Fisch et al. (Eds.); vol. 2, Moore E.C. et al. (Eds.); vols. 3-5, Kloesel, C. et al. (Eds.); vol. 6, Peirce Edition Project (Eds.). Bloomington: Indiana University Press, 1980- 2000. (Citado como W, seguido do volume e página)
- \_\_\_\_\_. *Semiótica*. São Paulo: Ed. Perspectiva, 2010.
- SANTAELLA, Lucia. *Matrizes da Linguagem e do pensamento. Sonora visual verbal*. São Paulo: Iluminuras, 2001.
- \_\_\_\_\_. Final causality and semiosis. *RS.SI* vol. 26, n. 2-3, 2006. P. 177-188.

## Concepções fregeanas e russellianas da noção de número: Uma comparação

Rodrigo Sabadin Ferreira<sup>1</sup>

Uma das teses centrais da Filosofia da Matemática de Frege e de neo-fregeanos, como Crispin Wright é que números são objetos, ou seja: que números são entidades que podem ser valores de variáveis de primeira ordem e termos numéricos são termos singulares. No caso de Frege, o número de um conceito  $F$  é a extensão do conceito [ $G$ :  $G$  é equinúmero a  $F$ ], sendo extensões de conceitos introduzidas pelo princípio de abstração conhecido como Axioma V. No caso de neo-fregeanos, como Crispin Wright, números são tratados como entidades sui generis cujas condições de identidade são introduzidas e explicadas em termos de um princípio de abstração tal como a equivalência conhecida na literatura como “Princípio de Hume”, geralmente expresso como (WRIGHT, 1983, pp.145-6):

$$(HP) \quad (F)(G)(\#F = \#G \leftrightarrow F \approx G)$$

Sendo a relação  $\approx$  de equinumerosidade, definida usualmente como:

$$(\approx) \quad F \approx G =df (ER)[(x)(Fx \rightarrow (E!y)(Gy \& \mathcal{E} xRy)) \& (x)(Gx \rightarrow (E!y)(Fy \& \mathcal{E} xRy)]$$

A motivação básica para adoção desse princípio é o que o ficou conhecido como Teorema de Frege: a Aritmética de segunda ordem PA2 é interpretável na lógica de segunda ordem + (HP) (BOOLOS & HECK, 1998, pp.334-336).

Um aspecto fundamental da concepção russelliana da noção de número é que números não são objetos. A posição russelliana pode ser caracterizada em termos de duas teses: Números são conjuntos de conjuntos equinúmericos e conjuntos são construções lógicas, ou seja: termos da linguagem que supostamente denotam conjuntos são símbolos incompletos definidos em termos de quantificação de segunda ordem (e superior). Em *Principia Mathematica*, o compromisso ontológico com conjuntos é evitado utilizando quantificação de ordem superior. Isso é possível mediante as definições contextuais da seção \*20, que traduzidas para notação moderna podem ser expressas do seguinte modo (WHITEHEAD & RUSSELL, 1962, p.190):

$$(*20.01) \quad A(\{x : Fx\}) =df (EG)[(x)(Gx \leftrightarrow Fx) \& A(G)]$$

Sendo  $A$  um contexto (fórmula bem formada contendo ocorrências da expressão em questão) qualquer. Essa definição permite que expressões da forma “ $\{x : Fx\}$ ” não sejam tratadas como termos genuínos, mas possam ser eliminados contextualmente: todo contexto  $A$  em que uma expressão como ocorre “ $\{x : Fx\}$ ” pode ser reformulada como

<sup>1</sup> Doutorando Pela Universidade Federal do Rio Grande do Sul.

uma sentença que afirma a existência de um conceito ou, na terminologia russelliana, uma função proposicional, que possui a mesma extensão de  $F$ . Esse modelo eliminativo de definição contextual de expressões para classes/conjuntos funciona de modo análogo àquele da teoria das descrições definidas (WHITEHEAD & RUSSELL, 1962, p.173):

$$(*14.01) \quad A([(ix)Fx]) =df (Ex)[Fx \mathcal{E} (y)(Fy \rightarrow y = x) \& A(x)]$$

O que essa definição contextual permite é tratar enunciados da forma “o tal e tal” como afirmações existenciais, analisando a estrutura dessas proposições como sendo da forma “existe um  $x$ ,  $x$  é tal e tal, e para todo  $y$ , se  $y$  é tal e tal, então  $y$  é igual a  $x$ ”. Em ambos os casos, o que temos é uma eliminação contextual de certas classes de expressões. A fim de comparação com os Neo-Fregeanos, veremos como essa estratégia geral de eliminação poderia ser aplicada no caso de termos numéricos, ignorando certos detalhes da teoria dos tipos de *Principia*. Para isso seguiremos em detalhe a formulação proposta por Kevin Klement em *A Generic Russellian Elimination of Abstract Objects* (KLEMENT, 2015, pp.91-114).

Em primeiro lugar, introduzimos a seguinte definição contextual para eliminar termos do tipo “ $\#F$ ” (KLEMENT, 2015, p.101):

$$(\#F\text{-Elim.}) \quad A(\#F) =df (EG)(F \approx G \mathcal{E} A(G))$$

Isso afirma que o contexto  $A$  é verdadeiro do número que convém a  $F$  se, e somente se o contexto  $A$  é verdadeiro de algum conceito equinúmero a  $F$ . Definimos a relação de identidade entre conceitos usando quantificação de terceira ordem (KLEMENT, 2015, pp.105):

$$(C\text{-Ident.}) \quad F = G =df (M)[M(F) \leftrightarrow M(G)]$$

E com isso podemos obter uma versão de (HP) que não introduz termos do tipo “ $\#F$ ” como termos singulares (KLEMENT, 2015, p.105):

$$(HP\text{-Elim.}): \quad (F)(G)[((EH)(F \approx H \mathcal{E} A(H)) \leftrightarrow (EH)(G \approx H \mathcal{E} A(H))) \leftrightarrow F \approx G]$$

Por fim, introduzimos variáveis especiais para os pseudo-termos do tipo “ $\#F$ ” por meio de definições contextuais que introduzem pseudo-variáveis “ $m$ ”, “ $n$ ”, etc., como *placeholders* de termos da forma “ $\#F$ ” (KLEMENT, 2015, p.105): (Todo  $n$ ):  $(n)A(n) =df (F)A(\#F)$  e (Algum  $n$ ):  $(En)A(n) =df (EF)A(\#F)$ . E com esse aparato podemos introduzir definições similares às Fregeanas das noções aritméticas (KLEMENT, 2015, p.105):

$$(Zero\text{-Elim.}) \quad A(0) =df (EF)[F \approx (\lambda x : x \neq x) \& A(F)]$$

$$(Suc\text{-Elim.}) \quad nSm =df (EF)(Ez)[\#F=n \& m = \#(\lambda x : Fx \ x \neq z)]$$

$$(N\text{-Elim.}) \quad N(l) =df (M)[M(0) \& (m,n)(M(m) \& nSm \rightarrow M(n)) \rightarrow M(l)].$$

Com (HP-Elim.) e essas definições, versões dos postulados de Peano, com exceção do quarto, que garante que todo número tem um sucessor único e, portanto que a série dos números naturais é infinita, podem ser provadas utilizando lógica de terceira ordem (KLEMENT, 2015, p.106; pp.113-114). Isso caracteriza uma diferença fundamental entre as concepções fregeanas e russellianas de número. Em formulações fregeanas da Aritmética enquanto uma teoria lógica, o teorema que afirma a existência de infinitos números naturais segue-se sem a necessidade de hipóteses adicionais usando apenas (HP). No caso de reconstruções russellianas da Aritmética nas quais termos numéricos são compreendidos como símbolos incompletos (de ordem superior), só é possível provar que a série dos números naturais é infinita assumindo alguma espécie de Axioma do Infinito,



que garante que, dada uma cardinalidade finita qualquer, há um conceito que instancie essa cardinalidade (LANDINI, G., 2011, pp.187-191).

De acordo com a concepção Neo-Fregeana, princípios de abstração como (HP) legitimam a introdução de termos numéricos como termos singulares. A concepção russelliana assume como princípio metodológico que termos singulares da gramática de superfície de teorias matemáticas devem ser eliminados em favor de símbolos incompletos sempre que possível. Como Russell afirma em *Logical Atomism* (RUSSELL, 1956, pp.323-343), a reconstrução de uma teoria qualquer em termos de uma teoria lógica deve preservar a *estrutura* da teoria, evitando sempre que possível, o compromisso ontológico com entidades inferidas ou postuladas:

When some set of supposed entities has neat logical properties, it turns out, in great many instances, that the supposed entities can be replaced by purely logical structures composed of entities which have not such neat properties. [...] In that case, in interpreting a body of propositions hitherto believed to be about the supposed entities, we can substitute the logical structures without altering any of the detail of the body of propositions in questions. This is an economy, because entities with neat logical properties are always inferred, and if the propositions in which they occur can be interpreted without making this inference, the ground for the inference fails, and our body of propositions is secured against the need of a doubtful step. The principle may be stated in the form: "Wherever possible, substitute constructions out of known entities for inference to unknown entities". (RUSSELL, 1956, pp.325-6)

Esse princípio, que Russell chamou de sua "Máxima Heurística" e de "[...] uma forma da Navalha de Occam" (RUSSELL, 1956, p.325) tem como duas de suas aplicações mais fundamentais, o tratamento de números como conjuntos de conjuntos equinúmericos e a eliminação contextual de conjuntos nos moldes da seção \*20 de *Principia Mathematica* (RUSSELL, B., 1956, pp.326; WHITEHEAD & RUSSELL, 1962, p.187-191).

Um modo de compreender o ponto de disputa diz respeito à famosa tese da "anterioridade da verdade em relação à referência" defendida por Neo-Fregeanos (KLEMENT, 2013, pp.134-5). O que essa tese afirma é que uma vez que (a) as condições de verdade de um certo conjunto de sentenças contendo termos singulares tenha sido estipulado e (b) estejamos aptos a tomar essas sentenças como verdadeiras de acordo com algum critério pré-estabelecido, podemos concluir (c) esses termos singulares referem a objetos. Encontramos essa posição expressa de modo paradigmático em *Frege's Conception of Numbers as Objects* de Crispin Wright:

Frege requires that there is no possibility that we might discard the preconceptions inbuilt into the syntax of our arithmetical language, and, the scales having dropped from our eyes, as it were, find in reality there are no natural numbers, that our old way of speaking we had not succeeded in referring to anything. Rather, it has to be the case that when it has been established, by the sort of syntactic criteria sketched, that a given class of terms are functioning as singular terms, and when it has been verified that certain appropriate sentences containing them are, by ordinary criteria, true, then it follows that those terms genuinely refer. And, being singular terms, their reference will be to objects. There is no



further, intelligible question whether such terms really have a referencem whether there are really such objects. (WRIGHT, 1983, p.14)

Nossa posição é que o desacordo entre as duas concepções se dá em termos metodológicos em relação ao próprio modo como teorias matemáticas devem ser reconstruídas em termos de uma teoria lógica. De um lado, a metodologia Neo-Fregeana é guiada por duas teses fundamentais: (1). Que há uma prioridade de categorias sintáticas (Nome, predicado, etc.,) em relação à categorias ontológicas (objeto, propriedade, etc.,) e (2) Que há prioridade da verdade em relação à referência. Ambas essas teses são plausíveis uma vez que assumimos, como faz Wright, alguma versão forte do Princípio do Contexto. De fato, como nota Michael Dummett, o que o Neo-Fregeano precisa é da garantia de que as questões pertinentes à referência de certo grupo de termos (como os numerais) é uma questão *interna* à linguagem, justamente o que o princípio do contexto, interpretado de maneira forte, garante .

The context principle, as enunciated in *Grundlagen*, can be interpreted as saying that questions about the meaning (*Bedeutung*) of a term or class of terms are, when legitimate, internal to the language. We know the meaning of a term, say, “the Equator”, when we know the conditions for the truth of any sentence containing it; that is all we need to know, and all we can know. Hence, to determine the meaning of a term, what we have to do, and all we to do, is to fix the senses of sentences in which it occurs. Reference therefore does not consist in a mental association between the term and the object, considered as apprehended by the mind independently of language; nor can it consist, we may add, in the existence of a causal chain leading from the object to an utterance of the term. It follows that any legitimate question about the meaning of a term, that is, about what we should call its reference, must be reducible to a question about the truth or otherwise of some sentence of the language. (DUMMETT, 1991, p.192)

Por outro lado, russellianos que aceitam o tratamento eliminativo dos termos numéricos devem negar que a tese da prioridade e, portanto, o princípio do contexto, possuem a validade irrestrita defendida por Wright. Mas nesse caso, o russelliano nos deve uma explicação de algum critério extra-linguístico que nos permita identificar uma classe de termos como fazendo referência ou não a objetos. Russell, ele próprio, ofereceu como explicação sua epistemologia do Acquaintance. Entretanto, a pressuposição dessa espécie de faculdade intuitiva que garante uma relação direta (não explicada) com entidades abstratas está em forte tensão com a motivação nominalista do eliminativismo russelliano.

Nossa conjectura é que a possibilidade de uma filosofia Russelliana da Aritmética nos moldes indicados depende de uma resolução satisfatória do seguinte problema: *como rejeitar a validade da tese da prioridade e do princípio do contexto para os termos aritméticos?* E, em particular, como rejeitar a tese de prioridade sem recorrer a um critério que invoque, como meio de explicar nossa capacidade de referir a objetos abstratos e caracterizar certos termos como termos singulares, alguma espécie de faculdade que possibilite uma relação cognitiva *direta* com tais entidades?

**Referências bibliográficas:**

- BOOLOS, G. & HECK, R. *Die Grundlagen Der Arithmetik §§82-83*. In: JEFFREY, Richard. *Logic, Logic and Logic*. Londres: Harvard University Press, 1998, pp.315-338.
- DUMMETT, Michael. *Frege: Philosophy of Mathematics*. Cambridge: Harvard University Press, 1991.
- LANDINI, Gregory. *Logicism and The Problem of Infinity: The number of numbers*. In: *Philosophia Mathematica*, vol 3, n.19, 2011, pp.167-212.
- KLEMENT, Kevin. *Neo-Logicism and Russell's Logicism*. *Russell: the journal of Bertrand Russell studies*, n. 32, 2013, pp.127-59.
- \_\_\_\_\_. *A Generic Russellian Elimination of Abstract Objects*. *Philosophia Mathematica*, v. 25, n. 1, 2015, pp.91-115.
- RUSSELL, Bertrand. *Logical Atomism*. In: MARSH, Charles (ed.). *Logic and Knowledge: 1901-1950*. Londres: Allen & Unwin, 1956, pp.323-343.
- WHITEHEAD, Alfred N. & RUSSELL, Bertrand. *Principia Mathematica to \*56*. Cambridge: Cambridge University Press, 1962.
- WRIGHT, Crispin. *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press, 1983.

## Consciência e fenômenos mentais inconscientes nas visões de David Rosenthal e John Searle

Tárik de Athayde Prata<sup>1</sup>

John Searle, um dos mais influentes e conhecidos filósofos contemporâneos, alega repetidamente ser um adversário do pensamento cartesiano. Ele diz que costumava pensar que o cartesianismo estava desacreditado (cf. Searle, 1992, p. xii; Searle, 1997, p. 2; Searle 1992, p. 13-4; Searle, 1997, p. 24-5), defende que ele não é capaz de dar conta do problema da causação mental (cf. Searle, 1999, p. 47; Searle, 2000, p. 51), e chega ao ponto de nomear uma das seções de um de seus livros de “Descartes e outros desastres” (cf. Searle, 2004, p. 13). Entretanto, é curioso notar que ele parece estar em sintonia com uma tese fundamental de Descartes, uma tese amplamente atribuída ao filósofo francês (cf. Rosenthal, 1986, p. 331; Armstrong, 1997, p. 721; Armstrong, 1999, pp. 14-16; Searle, 2004, p. 13; Searle, 2015, p. 202; Marques, 2017, p. 67-68), segundo a qual a *consciência* é algo constitutivo da existência psicológica. Descartes, que considerava o pensamento como uma propriedade essencial da alma, pensada por ele como uma “substância pensante” (*res cogitans*), caracterizou o pensamento como “tudo quanto está de tal modo em nós que somos disso imediatamente conscientes [*conscii simus*]” (Descartes, 1979, p. 169 [AT, VII, p. 160; AT, IX, p. 124]), o que o levava a recusar a possibilidade de fenômenos inconscientes (cf. AT, VII, p. 107; AT, IX, p. 85; AT, VII, p. 232; AT, IX, p. 180; AT, VII, p. 246; AT, IX, p. 190). Em consonância a esse modo de pensar, Searle entende que a consciência é o fenômeno mental mais importante, pois “não temos noção do mental independentemente de nossa noção de consciência” (Searle, 1992, p. 18; Searle, 1997, p. 31). Na visão dele, “todas as outras noções mentais – como intencionalidade, subjetividade, causação mental, inteligência, etc. – só podem ser plenamente compreendidas como *mentais* por meio de suas relações com a consciência” (Searle, 1992, p. 84; Searle, 1997, p. 125-26), de maneira que “o estudo da mente é o estudo da consciência, exatamente no mesmo sentido em que a biologia é o estudo da vida.” (Searle, 1992, p. 227; Searle, 1997, p. 326).

Porém, diferente do autor do *Discurso do método*, Searle não está em posição de, pura e simplesmente, negar a existência de fenômenos inconscientes, pois a influência de Sigmund Freud (cf. Searle, 1992, p. 151; Searle, 1997, p. 218; Gennaro, 1996, p. 6) e das ciências cognitivas (cf. Talvitie, 2009, p. 9; Boag, 2015, p. 240; Searle, 2015, p. 203), tornaram a noção de inconsciente amplamente aceita, e justamente para conciliar a visão cartesiana (da consciência como essencial ao psíquico) com a aceitação do inconsciente ele propõe uma perspectiva *disposicional* dos fenômenos inconscientes. Searle parte do princípio de que (1) há uma intencionalidade *genuína*, não meramente metafórica, e

1 UFPE

que (2) todo estado intencional representa suas condições de satisfação *sob determinados aspectos*, e não outros. Porém, já que fenômenos inconscientes existem na forma de *processos cerebrais*, e já que processos cerebrais, na medida em que são puramente *objetivos*, não possuem uma forma aspectual, a única maneira como o estado inconsciente pode ser pensado é como um estado consciente *possível*, no sentido de que o processo cerebral possui a *capacidade de causar* o estado mental em sua forma consciente, dotado de sua devida forma aspectual (para examinar os argumentos em detalhe, cf. Searle, 1992, p. 156-61; Searle, 1997, p. 224-31; Searle, 2004, p. p. 247-48; Searle, 2015, p. 207-8; Prata, 2017a, p. 208-10).

Porém, essa teoria disposicional do inconsciente entra em profundas dificuldades quando Searle atribui aos fenômenos inconscientes uma *eficácia causal* capaz de produzir determinados comportamentos, pois tal eficácia só é concebível em termos da atuação de propriedades mentais *ocorrentes* do estado mental inconsciente. Tanto no caso de fenômenos sensoriais inconscientes quanto no caso de fenômenos intencionais inconscientes, temos particulares concretos subjacentes (ou seja, determinados processos cerebrais), que exemplificam inúmeras propriedades (cf. Prata, 2017b, p. 394). Segundo os princípios da concepção disposicional, tais fenômenos não exemplificam a propriedade de serem conscientes, permanecendo essa propriedade, por assim dizer “em potência”, e não “em ato”. Todavia, mesmo assim, se esses estados inconscientes estão exercendo papel causal (provocando efeitos no comportamento do indivíduo), então tem que estar presente uma propriedade mental *efetiva* (não disposicional) independente da consciência (pois a consciência está ausente) – cf. Prata, 2017b, p. 397.

Diante das profundas dificuldades da perspectiva disposicional proposta por Searle, o presente trabalho empreenderá uma *comparação* entre essa perspectiva e a perspectiva proposta por um outro influente filósofo contemporâneo, o Professor David Rosenthal, autor cujas ideias exercem influência bastante significativa no debate atual a respeito da consciência. Para comparar as duas perspectivas, será seguido o seguinte percurso: primeiramente, será discutido o próprio conceito de consciência, no sentido de mostrar que as distinções propostas por Rosenthal a respeito da consciência são bem mais elucidativas do que o obscuro conceito articulado por Searle. Em seguida, será discutido o conceito de disposição, no sentido de mostrar que, como coloca Rosenthal, a caracterização de um fenômeno enquanto uma disposição implica que esse fenômeno possua uma base ocorrente, o que abre espaço para a ideia de que entre as propriedades ocorrentes que constituem a base de uma disposição estejam propriedades *mentais*, o que rompe com a ideia de Searle de que propriedades mentais inconscientes só podem existir enquanto disposições. Por fim, serão discutidos alguns argumentos de Rosenthal a favor da visão na qual propriedades intencionais e sensoriais podem existir de modo inconsciente, argumentos que, no meu modo de entender, produzem uma perspectiva sobre o inconsciente que é muito mais proveitosa do que aquela que encontramos nos trabalhos de Searle.

### **O conceito de consciência**

Em diversos escritos publicados ao longo das últimas décadas, Searle tem apresentado mais ou menos o mesmo conceito de consciência, que podemos sintetizar como se segue: a consciência consiste em uma sequência de estados de sensibilidade (“*feeling*” ou “*sentience*”) e ciência (“*awareness*”) que começa quando despertamos e terminam quando adormecemos (desmaiamos ou morremos), sendo que tais estados são essencialmente caracterizados

por um aspecto *qualitativo*, que por sua vez implica uma *subjetividade*, de modo que a consciência se constitui em um *campo unificado* de consciência *qualitativa* e *subjetiva* (cf. Searle, 1992, p. 83; Searle, 1997, p. 124; Searle, 2002, p. 7; Searle, 2010, p. 1-2; Searle, 2002, p. 21; Searle, 2010, p. 26; Searle, 2007, p. 170; Searle, 2015, p. 46-47; Searle, 2013, p. 10344; Prata, 2017a, p. 204-6; Prata, 2017b, p. 397-401; Prata, 2017c, p. 437-42).

Salta aos olhos que esse conceito de consciência envolve remissão a diversos fenômenos distintos, fenômenos cujas relações não são devidamente elucidadas nesta formulação. Já há muito, Ned Block denunciou o conceito de consciência de Searle como *confuso*, porque ele “aponta para coisas demais, muitas consciências diferentes” (Block, 1997, p. 380). Nesse sentido, as distinções conceituais propostas por Rosenthal são muito úteis para esclarecer a *estrutura* do próprio fenômeno da consciência, estrutura que permanece nebulosa no conceito searleano exposto acima. Para Rosenthal, é importante ter clareza que a consciência pode ser pensada como uma propriedade de *criaturas* (coisa que não constitui um problema filosófico, porque podemos caracteriza-la intuitivamente como a propriedade de estar *desperto* e *sensorialmente receptivo*), ou como uma propriedade<sup>2</sup> de *estados mentais* (o que constitui um problema filosófico genuíno, pois não temos uma concepção intuitiva do que essa consciência de estados poderia ser). Além disso, é importante notar que quando um estado mental está consciente, ele pode estar consciente de modo *não introspectivo* (ou seja, espontâneo, pois é apenas a presença desse estado mental no fluxo de consciência do indivíduo) ou de modo *introspectivo* (fruto do esforço deliberado de examinar os próprios estados mentais). Mas Rosenthal procura explicar a consciência de estados mentais (não a de criaturas), tanto em sua forma não introspectiva quanto em sua forma introspectiva, em termos de uma terceira distinção conceitual.

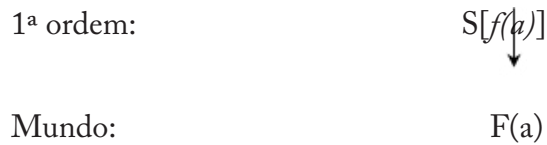
Por um lado, o predicado “x é consciente” pode ser usado para exprimir uma propriedade *não relacional*, no sentido de que, em primeira linha, diz respeito apenas ao seu portador, mas, por outro lado, esse predicado, com a adição da preposição “de”, pode ser usado para exprimir uma propriedade *relacional*, que constitui uma relação psicológica de um sujeito com um estado de coisas, como quando dizemos que “o ministro está consciente de que a inflação está subindo”. Trata-se da distinção entre consciência *intransitiva* e consciência *transitiva*.

O caso é que Rosenthal concebe o estado mental consciente como o estado *do* qual estamos conscientes de nos encontrar, de modo que a explicação do que é a consciência como propriedade (intransitiva) de um estado mental pode ser formulada em termos de nossa consciência (transitiva) desse estado. Por isso, ele se posiciona no campo das chamadas “teorias de ordem superior” a respeito da consciência (cf. Van Gulick, 2012, p. 47), já que a consciência de um estado mental (de primeira ordem) decorre da ação de um *outro* estado mental (de segunda ordem) a respeito do primeiro. Mas ele defende uma forma específica da teoria de ordem superior, na qual o estado de segunda ordem é um *pensamento* assertórico, e não uma percepção.

Para deixar mais claro o que ele tem em vista, podemos recorrer à seguinte descrição. Se pensarmos em um estado de coisas simples, no qual um ente individual (a)

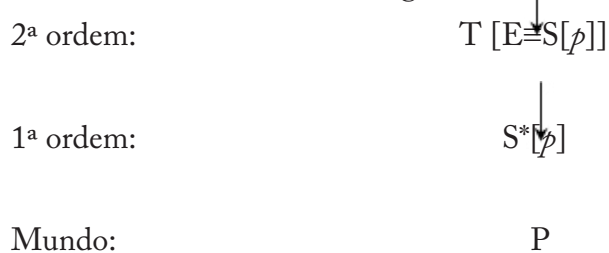
2 Apesar de usarmos o predicado monádico “x é consciente” para descrever a consciência, Rosenthal me esclareceu (em contato por email) que ele não entende que a consciência seja uma *propriedade* de estados mentais. Para ele, a consciência de estados mentais é um assunto do que se manifesta a nós do ponto de vista subjetivo – o que implicaria a possibilidade de *pensarmos* nos encontrar em estados mentais que não existem (uma visão dele que é objeto de críticas – cf. Block, 2011) – e se a consciência é uma questão de *aparências*, ela não pode ser pensada como uma propriedade do próprio estado mental. Para mais detalhes sobre a visão de Rosenthal, cf. Rosenthal, 2011.

exemplifica uma propriedade (F), podemos considerar a situação na qual esse estado de coisas (F(a)) é representado pelo conteúdo ( $f(a)$ ) de um estado mental qualquer (S), seja ele uma percepção, uma crença, um desejo, uma intenção, etc. De acordo com Rosenthal, enquanto esse estado mental permanecer desacompanhado de um pensamento que seja (de um determinado modo) *a respeito* dele, tal estado mental permanecerá inconsciente. Representemos essa situação da seguinte maneira:



Na visão de Rosenthal, um estado mental (S), seja ele de que tipo for, com o seu conteúdo ( $f(a)$ ), se torna consciente quando há um *pensamento* (T) que é a respeito desse estado mental (S). Porém, como aludido acima, o pensamento (T) tem que ser a respeito do estado mental (S) de um modo bem determinado: o pensamento (T) tem que ser a respeito do fato de que o *portador* desse pensamento *se encontra*, ele mesmo, no estado mental (S). Portanto, o pensamento (T), que é de segunda ordem em relação ao estado mental (S), tem que ter um conteúdo proposicional mais complexo do que aquele de primeira ordem.

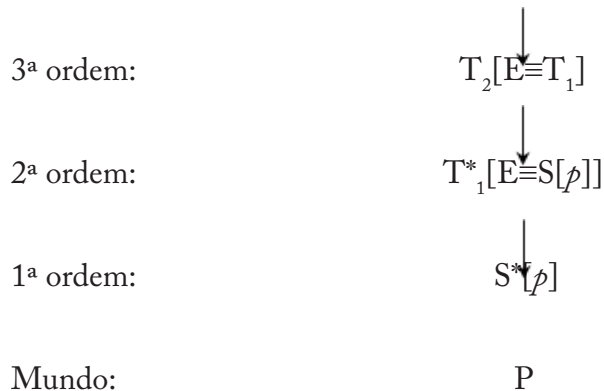
Para fins de simplicidade, representemos o estado de coisas F(a) com a letra “P”, e representemos o conteúdo proposicional  $f(a)$  com a letra “p”. Representemos, em seguida, a referência ao próprio sujeito psicológico, portador do estado mental (S) de primeira ordem, e do pensamento (T) de segunda ordem, referência que é algo como a noção de “eu”, pela letra “E”, e representemos a relação pela qual esse sujeito psicológico *se encontra* no estado mental (S) pelo símbolo “≡”. Já a consciência de um estado mental será representada pelo índice “\*”. A partir dessa notação podemos representar a situação na qual um estado mental de primeira ordem é consciente da seguinte maneira:



Confirme explicitado acima, o índice “\*” indica que o estado mental (S) de primeira ordem está consciente. Mas é importante perceber que esse índice está ausente do pensamento de segunda ordem (T), pois isso significa que esse pensamento é *inconsciente*, o que evita o problema de um regresso ao infinito (cf. Rosenthal, 1986, p. 340), e faz com que o conteúdo de T permaneça, por assim dizer, nas sombras, pois embora o sujeito psicológico E também se encontre no pensamento T, ele não tem consciência disso. Portanto, trata-se, aqui, de um caso de consciência *não introspectiva*, em que o foco da atenção do sujeito está voltado para o estado de coisas no mundo, já que a referência a si mesmo (E) se encontra, como conteúdo, em um pensamento inconsciente.



A consciência introspectiva, por sua vez, entra em cena quando temos um pensamento de *terceira ordem*, que torna consciente o pensamento de segunda ordem (cf. Rosenthal, 1986, p. 337). Podemos representar essa situação da seguinte maneira:



Para fins de simplicidade, o conteúdo do pensamento de segunda ordem  $T_1$  foi omitido em sua representação no conteúdo do pensamento de terceira ordem  $T_2$ . O caso é que uma vez que  $T_1$  se torna consciente, seu conteúdo se torna manifesto ao sujeito psicológico que é portador de todos esses estados mentais ( $S$ ,  $T_1$  e  $T_2$ ), e já que a referência a si mesmo ( $E$ ) se torna manifesta, o pensamento  $T_2$  opera uma *introspecção*, na medida em que passa para o foco da atenção o estado mental no qual o portador se encontra, e não mais o estado de coisas no mundo que a respeito do qual o estado mental de primeira ordem ( $S$ ) é. E considerando que o conteúdo de  $T_2$  já é bastante complexo, fica claro que não é nada fácil que possa surgir um pensamento de quarta ordem  $T_3$  que represente todos esses detalhes, que dirá um pensamento de quinta ordem, e assim por diante.<sup>3</sup>

Sem me comprometer, no presente trabalho, com a avaliação de que a teoria dos pensamentos de ordem superior (*higher-order thoughts*) de Rosenthal é adequada para a explicação da consciência de estado, irei me limitar a defender que as três distinções acima expostas fornecem uma clareza a respeito da consciência que *não* encontramos nos trabalhos de John Searle, pois essas distinções nos permitem discernir aspectos que, de fato, não deveriam ser confundidos.

### O conceito de disposição

Se entendermos uma *disposição* como a capacidade de um ente concreto de causar certos efeitos, mas apenas se forem dadas as circunstâncias adequadas, fica implícito que, enquanto essa capacidade não é exercida, ela tem que existir na forma de alguma *propriedade* que é efetivamente exemplificada por esse ente concreto. Rosenthal dá o exemplo da propriedade de ser inflamável, que tem de existir enquanto uma propriedade química ocorrente (mesmo que disjuntiva) – cf. Rosenthal, 1997, p. 732.

Sendo assim, se considero que um estado mental inconsciente (como uma crença ou um desejo *do* qual não me dou conta) existe na forma de uma disposição, sou obrigado a admitir que, enquanto esse estado mental permanece inconsciente, ele tem que existir na forma de certas propriedades *ocorrentes*. Se, além disso, baseado em indícios convincentes (que abundam na literatura), suponho que um estado mental inconsciente exerce, efetivamente, poder causal sobre o comportamento de um indivíduo, então sou obrigado

<sup>3</sup> Para uma representação alternativa das relações entre estados mentais e pensamentos de ordem superior, cf. Prata, 2017c, p. 439.

a atribuir esse poder causal a alguma(s) dessas propriedades ocorrentes, pois apenas uma propriedade ocorrente pode fazer uma diferença causal. Entretanto, se considero, realmente, tal episódio de causação (a causação de um determinado comportamento por um estado inconsciente) como um episódio de causação *mental*, então sou obrigado a atribuir esse poder a uma propriedade genuinamente *mental*, ainda que, nesse caso, por hipótese, ela permaneça inconsciente.

Mas se é assim (se tenho uma propriedade mental *ocorrente* causando determinado comportamento), então em que sentido o estado mental inconsciente é uma disposição? O único sentido que resta é um sentido muito fraco, a saber: o sentido de que o estado inconsciente possui a *capacidade* não atualizada de gerar uma experiência consciente. Isso é muito fraco, para os objetivos de Searle (ou de qualquer teórico que pensa o inconsciente como uma disposição), porque significa que a *consciência* é que é uma disposição, ao passo que as *propriedades mentais* inconscientes são ocorrentes. Esse cenário rompe com (a) a visão cartesiana, pois admite propriedades mentais inconscientes, mas rompe também com (b) a própria visão disposicional do inconsciente, pois essas propriedades mentais são ocorrentes, e não disposicionais.

O tratamento que Rosenthal dá ao conceito de disposição se mostra, muito mais adequado do que o tratamento oferecido por Searle, pois mesmo que seja o caso que este aceite que a disposição existe na forma de propriedades ocorrentes (posição que não aparece com clareza nos trabalhos de Searle), ele não pode aceitar que entre elas se encontrem propriedades *mentais*.

### **Rosenthal sobre a existência mental inconsciente**

Em diversos escritos, Rosenthal discorre sobre as razões que temos para aceitar a existência de estados mentais inconscientes, rompendo assim com a visão cartesiana. Por um lado, quando percebemos que uma pessoa pensa ou se sente de determinado modo, sem que ela própria esteja ciente de que pensa ou se sente assim, estamos constatando a presença de estados mentais que não são conscientes (cf. Rosenthal, 1986, p. 329). Além disso, é muito difícil negar que as pessoas possuem, por exemplo, crenças e preferências em um número muito maior do que o que elas podem pensar conscientemente em um dado momento (cf. *Ibid.*, p. 334). Mas ele entende que temos razões para concluir que não só estados intencionais (como crenças, desejos e preferências), mas também *sensações*, podem existir de maneira inconsciente. Apesar da ideia intuitiva, que encontramos em pensadores da estatura de Thomas Reid e Saul Kripke (cf. *Ibid.*, p. 347-48; p. 358, nota 19), de que a consciência parece inseparável das qualidades sensoriais, Rosenthal apresenta diversos argumentos para provar que isso não é verdade.

Além de estados mentais como desejos, pensamentos e emoções, e até mesmo sensações inconscientes, Rosenthal enfatiza que temos motivos para admitir que ocorrem *percepções* inconscientes, por exemplo nos casos de percepções subliminares e de visão periférica (cf. Rosenthal, 1997, p. 731). Na verdade, os indícios da existência de estados mentais (intencionais ou sensoriais) inconscientes surgem tanto de contextos de *investigação científica* quanto de *contextos cotidianos*. Estudos empíricos sistemáticos mostram evidências de percepções inconscientes, seja em investigações sobre percepção subliminar e visão periférica, seja no caso de patologias dissociativas como a célebre “visão cega” (cf. Rosenthal, 1993, p. 357; Rosenthal, 2008, p. 833). Por outro lado, algumas circunstâncias do dia a dia nos mostram evidências de sensações inconscientes, como quando uma dor ou

um desconforto moderado se manifesta de modo intermitente à consciência, por exemplo, em virtude de alguma distração (algo que afasta nossa atenção desse desconforto). Em tais casos, não parece razoável dizer que existiram diversas dores descontínuas, mas sim que uma única dor, contínua ao longo do tempo, ora se manifestava ora desaparecia da consciência, coisa que, na visão de Rosenthal, abre espaço para a noção de sensações inconscientes no nosso senso comum (cf. Rosenthal, 1997, p. 732).

### Considerações finais

Ainda que Searle admita que a noção de inconsciente é *indispensável* para a devida compreensão do comportamento humano (cf. Searle, 1992, p. 151; Searle, 1997, p. 217; Searle, 2004, p. 239), a forma como ele concebe os fenômenos mentais inconscientes, na tentativa de acomodá-los à visão cartesiana, é uma forma que *fracassa* quando tenta explicar a eficácia causal que esses fenômenos possuem sobre nosso comportamento, eficácia que, justamente, é o elemento responsável pela importância do inconsciente na compreensão de nossas ações.

A perspectiva oferecida por Rosenthal, de acordo com a qual a consciência é uma propriedade *relacional* e *contingente* dos estados mentais que a possuem é muito mais adequada aos fatos (tal como descritos na seção anterior), e é bem sucedida na explicação da eficácia causal dos estados mentais inconscientes, uma vez que, nessa perspectiva tais estados preservam suas propriedades mentais (sejam intencionais, sejam sensoriais) enquanto eles estão inconscientes, propriedades estas que são o que efetiva uma *causação mental* inconsciente.

Mas, para concluir, é importante mencionar uma outra vantagem da visão de Rosenthal sobre a de Searle: o fato de que aquela, no mínimo, torna plausível uma explicação teórica da consciência de estado, ao passo que esta parece aprisionada em uma concepção muito pobre de explicação (cf. Prata, 2009). No que Searle concebe a consciência como um “campo unificado” do qual os estados mentais singulares são modificações, ele está meramente articulando uma *metáfora* a partir de um fenômeno físico, metáfora que não parece abrir nenhum espaço para qualquer explanação da propriedade de um estado mental ser consciente, apresentando a instanciação dessa propriedade como uma espécie de *fato bruto*.

Isso pode ser percebido ao examinarmos a concepção que Searle defende a respeito da *autoconsciência*, pois ao mesmo tempo em que ele recusa a presença constante de uma autoconsciência, no sentido de uma constante consciência de si mesmo (enquanto indivíduo) ou dos próprios estados mentais, ele aceita um sentido técnico de autoconsciência, a saber: o sentido de que, quando nos encontramos em um estado mental consciente, sempre podemos deslocar nossa atenção para o estado mental em si (cf. Searle, 1992, p. 143; Searle, 1997, p. 206). Entendo que isso significa que todo estado mental é autoconsciente no sentido específico de que ele sempre pode se tornar o centro de nossas atenções. Mas se isso ocorre, então, nos marcos da teoria de Searle, isso só pode ser entendido como consequência do fato de que esse estado mental é uma modificação do campo de consciência. Mas essa teoria não oferece resposta à pergunta sobre *como* a consciência caracteriza esse estado mental. A visão dele é simplesmente que:

Em tais situações, sem introspecção, todo o campo de consciência do sujeito é *perpassado* (de um modo que *não fica claro*) por uma autoconsciência implícita, que pode

ser caracterizada como pré reflexiva (ou não introspectiva, nos termos de Rosenthal). Isso significa que, na visão de Searle, a autoconsciência é uma *característica constitutiva* do próprio campo unificado, mesmo quando não há introspecção (embora sua teoria não explique *como* essa característica é efetivada no campo unificado). (Prata, 2017c, p. 449).

Já na teoria de Rosenthal, a consciência é pensada como uma propriedade *contingente* dos estados mentais que a possuem, ou seja, uma propriedade sem a qual esses estados mentais podem continuar a existir, de maneira que se abre a possibilidade de uma explicação *psicológica* da consciência de estado, ou seja, uma explicação da consciência de estados mentais em termos de *outros* estados mentais. É justamente para *abrir* essa possibilidade que a aceitação do inconsciente serve, pois se o segundo estado mental (que torna consciente o primeiro estado) é um estado inconsciente, então a explanação se torna *informativa*, pois se ele fosse, também, um estado consciente, a explanação se tornaria *circular* (explicaríamos consciência em termos de consciência, o que não nos diria *nada*). Por isso Rosenthal enfatiza que para entender como a consciência pode ocorrer em entidades físicas, precisamos dissolver a força intuitiva do aparente *abismo* entre o físico e a consciência. E, de acordo com ele: “só podemos fazer isso explicando a consciência de estados mentais em termos de estados mentais que não são conscientes.” (Rosenthal, 1986, p. 353). Nesse sentido, Rosenthal me parece estar defendendo uma posição bastante razoável, quando ele afirma: “Reconhecer que pensamentos, sentimentos e sensações frequentemente ocorrem sem serem conscientes é, portanto, o primeiro passo para entender sua natureza.” (1997b, p. 740).

**Referências bibliográficas:**

- Adam, Ch. (Org.); Tannery, P. (Org.). *Ouvres de Descartes – Meditationes de prima philosophia*. Paris: Vrin, 1996. (Vol. VII).
- Adam, Ch. (Org.); Tannery, P. (Org.). *Ouvres de Descartes – Méditations et principes (traduction française)*. Paris: Vrin, 1996. (Vol. IX).
- Armstrong, A. “What is Consciousness?”. In: Block, N. (Org.); Flanagan, O. (Org.); Güzeldere, G. (Org.). *The Nature of Consciousness: Philosophical Debates*. Cambridge (Massachusetts): MIT Press, pp. 721-728, 1997.
- Armstrong, A. *The Mind-Body Problem: An Opinionated Introduction*. Boulder (CO); Oxford (UK): Westview Press, 1999.
- Boag, S. “In defense of unconscious mentality”. In: Boag, S. (Org.); Brakel, L. A. W. (Org.); Talvitie, V. (Org.). *Psychoanalysis and Philosophy of Mind: Unconscious Mentality in the Twety-First Century*. London: Karnac Books, pp. 239-65, 2015.
- Block, N. “On a Confusion About a Function of Consciousness” In: Block, N. (Org.); Flanagan, O. (Org.); Güzeldere, G. (Org.). *The Nature of Consciousness: Philosophical Debates*. Cambridge (Massachusetts): MIT Press, pp. 375-415, 1997.
- Block, N. “The Higher-Order Approach to Consciousness is defunct”. In: *Analysis*, Vol. 71, No 3, pp. 419-31, July 2011.
- Descartes, R. *Discurso do método; Meditações; Objeções e Respostas; As Paixões da Alma; Cartas*. 2ª Edição. São Paulo: Abril Cultural (Os Pensadores), 1979.
- Gennaro, R. *Consciousness and Selfconsciousness: a Defense of the Higher Order Thought Theory of Consciousness*. Amsterdam/Philadelphia: John Benjamins Publishing, 1996.
- Marques, E. “Consciência e autoconsciência em Leibniz”. *Analytica*, Vol. 21 nº 1, pp. 67-83, 2017.
- Prata, T. A. Pode-se explicar a consciência através de processos cerebrais? A argumentação de John Searle contra a concepção de Thomas Nagel. *Kalagatos – Revista de Filosofia*, Vol. 6, Nº 11, p. 137-172, inverno de 2009.
- Prata, T. A. “A concepção disposicional do inconsciente na Filosofia da Mente de John Searle” In: *Revista Reflexões*, Ano 6, Nº 11, pp. 201-16, 2017a.
- Prata, T. A. “Uma crítica à concepção disposicional de Searle sobre os fenômenos mentais inconscientes”. In: Araújo, A. (Org.) [et al.]. *Pragmatismo, filosofia da mente e filosofia da neurociência*. São Paulo: ANPOF, pp. 387-403, 2017b.
- Prata, T. A. “A estrutura da autoconsciência na filosofia da mente de John Searle” *Veritas* (PUCRS), Vol. 62, Nº2, maio-ago., p. 428-452, 2017c.
- Rosenthal, D. “Two Concepts of Consciousness”. In: *Philosophical Studies* 49, pp. 329-59, 1986.
- Rosenthal, D. “State Consciousness and Transitive Consciousness”. In: *Consciousness and Cognition* 2, pp. 355-363, 1993.
- Rosenthal, D. “A Theory of Consciousness”. In: Block, N. (Org.); Flanagan, O. (Org.); Güzeldere, G. (Org.). *The Nature of Consciousness: Philosophical Debates*. Op. cit., pp. 729-753, 1997a.
- Rosenthal, D. “Commentaries”. *Journal of the American Psychoanalytic Association*, 45, pp. 740-746, 1997b.
- Rosenthal, D. “Consciousness and its Function”. *Neuropsychologia*, 46, pp. 829-40, 2008.
- Rosenthal, D. “Exaggerated Reports: Reply to Block” In: *Analysis*, Vol. 71, No 3, pp. 431-37, July 2011.

- Searle, J. R. *The Rediscovery of the Mind*. Cambridge Mass., London: MIT Press, 1992.
- Searle, J. R. *A Redescoberta da Mente*. São Paulo: Martins Fontes, 1997.
- Searle, J. R. *Mind, Language and Society: Doing Philosophy in the Real World*. New York: Basic Books, 1999.
- Searle, J. R. *Mente, linguagem e sociedade: filosofia no mundo real*. Rio de Janeiro: Rocco, 2000.
- Searle, J. R. *Consciousness and Language*. Cambridge (UK): Cambridge University Press, 2002.
- Searle, J. R. *Mind: A Brief Introduction*. Oxford: Oxford University Press, 2004.
- Searle, J. R. *Consciência e linguagem*. São Paulo: Martins Fontes, 2010.
- Searle, J. R. "Theory of Mind and Darwin's Legacy". In: *PNAS*, Vol. 110, Suppl. 2, pp. 10.343-48, 2013.
- Searle, J. R. *Seeing Things as They Are: A Theory of Perception*. Oxford: Oxford University Press, 2015.
- Talvitie, V. *Freudian Unconscious and Cognitive Neuroscience: From Unconscious Fantasies to Neural Algorithms*. London: Karnac Books, 2009.
- Van Gulick, R. "Consciência". *Investigação filosófica*, Vol. E2, Artigo digital 2, 2012.