

Colóquio Internacional de Filosofia da Mente

SÃO JOÃO DEL-REI – MG – BRASIL

ABRIL, 2017



APRESENTAÇÃO

O IX *Colóquio Internacional de Filosofia da Mente* foi realizado entre os dias 26 e 28 de abril de 2017, no Campus Dom Bosco da Universidade Federal de São João del-Rei (UFSJ). O evento reuniu alguns dos principais pesquisadores brasileiros na área, além de três convidados estrangeiros. Foram cerca de 20 apresentações de palestrantes convidados e mais de 40 comunicações, durante 3 dias de intenso debate e confraternização entre todos os participantes e o público presente. Ao contrário de edições anteriores, nessa edição do evento não foi proposto nenhum tema específico, sendo os palestrantes e comunicadores incentivados a apresentarem as pesquisas que estavam realizando no momento. A intenção, que se mostrou bem-sucedida, era estimular a contribuição e o conhecimento mútuo entre os pesquisadores da área no Brasil. Em sua nona edição, o Colóquio deu continuidade a uma já consolidada tradição de inegável sucesso.

Os 15 artigos que compõem esses *Anais* foram selecionados dentre as mais de 40 comunicações realizadas no Colóquio; comunicações essas que foram selecionadas de um total de cerca de 80 submissões. Trata-se, portanto, de uma amostra selecionada e representativa da pesquisa hoje feita no Brasil na área de filosofia da mente e áreas correlatas.

Muitas pessoas contribuíram para a execução do evento, e certamente seríamos injustos com alguns nomes se tentássemos nomear cada um. Vamos registrar apenas nossos mais sinceros agradecimentos aos colegas do GT-Filosofia da Mente, em especial aos professores Maria Eunice Quilici Gonzalez e André Leclerc, que iniciaram essa tradição há quase 20 anos atrás.

Desejamos longa vida ao Colóquio, e bons estudos para aqueles que porventura vieram a consultar essas Atas.

São João del-Rei, abril de 2017

Prof. Dr. Gustavo Leal Toledo (DFIME/UFSJ)

Prof. Dr. Marco Aurélio Sousa Alves (DTECH/UFSJ)

Prof. Dr. Rodrigo Gouvea (DFIME/UFSJ)

SUMÁRIO

1	O conteúdo da representação na teoria teleosemântica	4
	<i>Ana Paula Monteiro dos Reis Emmendorfer</i>	
2	Ensaio sobre a concepção de <i>psykhé</i> em Platão no <i>Cármides</i> e no <i>Fédon</i>	16
	<i>Bárbara de Abreu Freitas</i>	
3	A Hipótese da Mente Estendida: argumentos e objeções	27
	<i>Bruno Coelho</i>	
4	<i>Qualia</i> e estados subjetivos: A defesa de Searle da teoria dos <i>qualia</i> frente ao funcionalismo computacional	43
	<i>Bruno Cunha</i>	
5	Em que sentido há um problema mente-corpo?	53
	<i>Cínthia Roso Oliveira</i>	
6	O desenvolvimento do conceito de intencionalidade: Brentano, Husserl e Searle em diálogo	70
	<i>Danilo da Cunha Pontes</i>	
7	Paul Ricoeur e a investigação sobre o sujeito: considerações sobre o cogito ferido e a mediação da psicanálise	78
	<i>João Bosco Batista</i>	
8	Sobre estados mentais relacionais e <i>qualia</i>: o totalmente subjetivo não pode ser conhecido	93
	<i>Juliana de Orione Arraes Fagundes</i>	
9	Há filosofia da mente no pensamento de Guilherme de Ockham?	108
	<i>Laiza Rodrigues de Souza</i>	
10	O conceito de mente na filosofia indiana: breves tópicos introdutórios	124
	<i>Lincoln T. Zacconi</i>	
11	Filosofia da mente no pensamento pré-platônico: o caso de Parmênides	139
	<i>Nicola Stefano Galgano</i>	
12	A mente e a evolução da ficção: estado atual do problema	156
	<i>Pedro Dolabela Chagas</i>	
13	Se a consciência é indispensável para que computadores sejam considerados artisticamente criativos	172
	<i>Ricardo de Oliveira Toledo</i>	
14	Bergson: o problema mente-cérebro sob uma perspectiva pragmática	188
	<i>Ricardo Pinho Souto</i>	
15	Pampsiquismo em Nietzsche? Apontamentos introdutórios	198
	<i>William Mattioli</i>	

O conteúdo da representação na teoria teleosemântica

The content of representation in teleosemantics theory

Ana Paula Monteiro dos Reis Emmendorfer¹

RESUMO: O objetivo deste artigo é esboçar como Dretske e Millikan lidam com o conteúdo de uma representação. As teorias de Dretske são compreendidas em duas versões: a teoria informacional e a teleosemântica informacional. A primeira teoria diz que se um signo natural S carrega informação I, então S Indica I, caso a probabilidade disso acontecer for igual a 1, o erro consiste quando não há compatibilidade entre o conteúdo informacional e o conteúdo representacional. E a segunda teoria, diz que uma representação R tem a função F de indicar um estado de coisas, caso carregue informação deste estado de coisas, a representação é falsa quando a representação não cumpre a sua função de Indicação. Millikan, por sua vez, naturalizou a intencionalidade. Um dos pilares da sua teleosemântica é a teleofunção. A noção de função define-se por meio do propósito e/ou finalidade de um ser vivo, órgão e/ou organismo. Além disso, essa noção auxilia na explicação da função das representações mentais humanas. Para ela, há um mecanismo cerebral capaz de representar. Millikan faz uso do termo técnico *função normal própria de projeção (ou mapeamento)* quando trata das funções das representações mentais. A *função própria* deve estar de acordo com a história evolutiva e não com suas disposições atuais. Diante disso, ocorre uma falha quando, por uma razão ou outra, a função própria não desempenha o seu papel.

PALAVRAS-CHAVE: representação; conteúdo; indicar; função; normal.

ABSTRACT: The goal of this paper is to sketch how Dretske and Millikan deal with the content of representation. Dretske's theories are understood in two ways: informational theory and informational teleosemantics. The first one says that, if a natural sign S carries I information, then S indicates I, in the case of the probability of this happening being equal to 1, the mistake would occur when there is no compatibility between the informational and representational content. And the second theory argues that a representation R has the function F of indicating a state of affairs, in case it carries the information of this state of affairs, the representation will be fake when the representation does not accomplish its function of Indication. Millikan, in turn, has naturalized the intentionality. One of the cornerstones of his teleosemantics is the telefunction. The notion of function is defined through the purpose and/or finality of a living being, organ and/or organism. Beyond that, this notion helps in the explanation of the human's mental representational functions. According to her, there is a mental mechanism capable of representing. Millikan uses the technical term *normal function proper of projection (or mapping)* when he deals with the function of mental representation. The proper function must be in accordance with the evolutionary history and not to its actual disposition. Then, a fail occurs when, by some reason or another, the proper function does not perform its role.

KEYWORDS: representation; content; Indicate; function; normal.

¹Doutoranda Programa de Pós-graduação em Filosofia da Unisinos (bolsista CAPES), sob a orientação da Prof^a. Dr^a. Sofia Stein. E-mail: anapaullamonteiro@yahoo.com.br.

1. Considerações iniciais

Os naturalistas têm tentado solucionar questões relacionadas ao significado (representações) naturalisticamente. Em vista disso, qualquer ideia acerca do significado e/ou representação que se preze busca especificar o que é verdadeiro. Contudo, é delicado estipular o que é falso. Dessa forma, teorias teleosemânticas buscam resolver este enigma por meio da noção de função. Mesmo assim, os filósofos que defendem tal posição tratam a temática de maneiras distintas e, por conseguinte, estabelecem a ideia de função de modo diverso.

Nesse sentido, teorias teleosemânticas buscam resolver o problema do conteúdo fixado, principalmente, no que tange as representações enganosas. Para este artigo, escolhemos trabalhar com Dretske e Millikan. No entanto, eles têm visões distintas sobre as teorias das representações, seus conteúdos e suas possibilidades de erros. Nessa perspectiva, a teoria de Dretske pode ser dividida em duas partes, isto é, ele possui duas teorias que tratam a mesma temática, sua primeira versão foca na ideia de informação e a segunda, une informação e função. Por conseguinte, as duas tomam como referencial a noção de informação.

Já Millikan julga relevante reformular a ideia de intencionalidade, pois acredita que a intencionalidade não pode ser uma característica exclusiva de objetos mentais, como Brentano afirmava, porque tanto objetos físicos, quanto mentais, possuem um propósito de representar um estado de coisas, embora possam falhar. A partir disso, a filósofa determina uma gama de elementos, os quais são essenciais para a constituição de sua teoria cuja base é a biologia evolucionista, tais como função própria, mecanismo normal, signos-produzidos e signos-consumidos.

A proposta deste artigo, então, é analisar as teorias propostas pelos dois filósofos, de modo que possamos expor quais as suas diferenças, já que seus pontos de vistas parecem opostos. Destarte, consideraremos se é possível determinar qual das teorias propostas parece ser mais razoável.

2. Teoria informacional de Dretske

A teoria da informação da matemática² ou teoria da comunicação vincula uma informação a um estado de coisas, bem como essa informação é transmitida e analisada. É uma teoria estatística, quantitativa, avalia possibilidades e probabilidade. Lida, basicamente, com a correção de erros. “Information theory identifies the amount of information associated with, or generated by, the occurrence of an event (or the realization of a state of affairs) with the reduction in uncertainty, the elimination of possibilities, represented by that event or state of affairs”. (DRETSKE, 1995, p.04).

Dretske coloca a seguinte situação³: há uma seleção para o preenchimento de uma vaga de emprego com oito candidatos na disputa, cuja escolha será feita a partir da análise de algumas características dos candidatos. As características são: 1-homem, 0-mulher, 0-novo, 1-velho, 0-baixo, 1-alto. A partir disso, é gerado uma informação de três bits por meio do código binário. Diante disso, o selecionado é Herman por meio da informação 101 e o seu nome é enviado ao chefe. Assim, 101 representa Herman em que 1 indica homem, 0 para novo e 1 para alto, mas digamos que outra informação seja gerada, 010, que representa Margarete em que 0 indica mulher, 1 para velho e 0 para baixo. Não importa qual das duas informações é escolhida, pois em termos de conteúdo é irrelevante, em uma teoria quantitativa, escolher Herman ou Margarete. Todavia, o filósofo julga incoerente tratar sobre informação e desconsiderar o conteúdo.

Dessa maneira, Dretske toma emprestado essa teoria da informação da matemática para elaborar a sua. Ele questiona acerca das condições para avaliação de um conhecimento correto e, sobretudo, em quais condições podemos dizer que um conhecimento é incorreto e/ou uma representação é deturpada? Em seu artigo *Misrepresentation* (1986), ao levar em conta a noção de conteúdo semântico, ele afirma que a representação pode falhar devido, em parte, a uma pequena parcela do problema geral de significado ou intencionalidade. Sua pretensão, então, é desenvolver uma teoria genuína semântica de informação que possa ser aplicada na epistemologia e na filosofia da mente para melhor compreender a ideia de conteúdo semântico.

The problem is not how, for example, a diagram, *d*, can misrepresentation the world, *w*. For if we have another system, *r*,

²Inspirada no trabalho de Claude Shannon's.

³Esse exemplo foi retirado do primeiro capítulo do livro *Knowledge and Flow of Information*, p.03-39.

already possessed of representational powers, *d* can be used as an expressive extension of *r*, thereby participating in *r*'s representational successes and failures. When this occurs, *d* can come to mean that *w* is F when, in fact, *w* is not F, but *d*'s meaning derives, ultimately, from *r*. (DRETSKE, 1986, p.157).

A comunicação ocupa-se com fontes e não com mensagens em particular, lida com a média da quantidade de informação. Como nos sorteios de moedas em que os sorteios são as fontes, bem como no exemplo anterior em que a seleção de emprego é a fonte, cuja média é os três bits de informação gerados na escolha do candidato. A transmissão de informação, por sua vez, é dependente da inter-relação causal entre a fonte e o receptor. Teixeira, para exemplificar Dretske, usa a seguinte situação, trata-se da relação entre dois sistemas: dois sistemas S e S' possuem uma relação causal em que S é a fonte de informação e S' é o receptor. Nisso, há uma história causal, S transmite informações para S'. Essa transmissão indica que S carrega informação sobre S'.

A relação entre a fonte e o receptor é o ponto de partida para a análise de Dretske. Nesse sentido, baseando-se na distinção de Grice⁴ entre signos naturais e signos não-naturais em que um signo é natural quando P significa P *sse* P for o caso e é não-natural quando P significa P mesmo que P não seja o caso, ele estabelece a sua distinção. Grice considera importante expor que os signos naturais possuem um vínculo factual entre signo e estado de coisas, diferentemente dos signos não-naturais, os quais têm relação com a intenção do comunicador, por essa razão, tomando a comunicação verbal é possível perceber que aquilo que está sendo dito nem sempre corresponde àquilo que se pretende dizer. Grice expõe,

This question about the distinction between natural e non-natural meaning is, I think, what people are getting at when they display an interest in a distinction between “natural” and “conventional” signs. But I think my formulation is better. For some things which can mean _{NN} something are not signs (e.g., words are not), and some are not conventional in any ordinary sense (e.g., certain gestures); while some thing which mean naturally are not signs of what they mean. (GRICE, 1957, p.379).

Para Dretske⁵, conteúdo informacional são signos naturais e conteúdo representacional são signos não-naturais. Considerando os signos naturais (conteúdo informacional), esses só

⁴Grice explica melhor essa distinção no seu texto *Meaning* (1957) – p.377-388.

⁵Em *Misrepresentation* (1986), Dretske coloca signo natural e signo funcional.

podem representar um estado de coisas corretamente se a informação carregada pelo signo for o caso. Quanto à natureza da informação, refere-se a algo dado no mundo, isto é, independe da avaliação e/ou reconhecimento do sujeito cognitivo. As informações independem do fato de termos conhecimento e/ou quais são as nossas crenças a respeito dessas⁶.

Se temos:

Nuvens pretas Indicam chuva.

Fumaça Indica fogo.

Manchas vermelhas Indicam Sarampo.

Em qualquer um desses exemplos acima algo carrega informação sobre algo. X carrega informação sobre Y, X é o receptor da informação, o signo natural, e Y é a fonte, o estado de coisas representado. X carrega informação *sse* Y for o caso cuja probabilidade de Y ser o caso é igual a 1. Mas em que circunstâncias pode acontecer ou acontece uma falha? Na circunstância de Y não ser o caso e/ou a probabilidade condicional ser menor que 1, ou seja, quando o conteúdo informacional não é equivalente ao conteúdo representacional. Exemplificando, manchas vermelhas costumam Indicar sarampo, mas se a pessoa não está com sarampo não será o caso, pois pode ser que esteja com alguma alergia e/ou comeu doces demais.

De qualquer modo, a relação deve ser unívoca entre fonte e receptor para ser verdadeira. O signo natural não pode Indicar o estado de coisas específico e/ou esperado, se não houver a existência daquilo que está sendo Indicado. Em *Misrepresentation*, ele diz: “Natural signs are indicators, more or less reliable indicators, and what they mean is what they indicate to be so”. (DRETSKE, 1986, p.18).

Temos outro exemplo, o da campainha: digamos que a campainha da minha casa foi tocada, mas eu não sei quem a tocou, apenas tenho certeza que não pode ter sido tocada por um animal. Por outro lado, não posso garantir, não tenho a probabilidade igual a 1, que realmente tenha sido uma pessoa, pode ser que tenha ocorrido um curto circuito. O conteúdo representacional é determinado pelo estado de coisas representado a partir do conteúdo informacional carregado pelo signo. No entanto, nem sempre o conteúdo representacional é

⁶ Dretske adota um realismo bruto.

compatível com o conteúdo informacional, por conseguinte, enquanto conteúdos representacionais podem ser falsos, conteúdos informacionais não podem.

Existem conexões causais capazes de estabelecer relações constantes entre signos naturais e as informações carregadas por esses signos as quais tornam o vínculo entre signo e o estado de coisas unívoco. Então, como podem ocorrer erros nas representações? Pela não compatibilidade do conteúdo informacional e conteúdo representacional. Em seu livro *Knowledge & the Flow of Information* (1981), Dretske expõe que o erro de uma representação acontece em uma situação de não aprendizagem.

Destarte, para Dretske (1981) durante um determinado tempo T, um organismo O fica submetido a muitos signos X que carregam a informação Y. A partir disso começa a ser construído uma codificação, uma espécie de mapeamento cujo objetivo é fazer com que o organismo o possa identificar Y. O conteúdo semântico é estabelecido neste processo de aprendizagem ideal, por meio desta codificação. Entretanto, tendo em vista que, há muitos exemplares, no caso, Y e não-Y, então este processo de aprendizagem ideal não funciona no início deste tempo T, pois a adaptação do organismo O para com o reconhecimento de Y é gradativa. As representações falhas ocorrem quando o sistema S troca Y por não-Y, havendo, desta maneira, um processo de não aprendizagem.

2.1. A teleosemântica informacional de Dretske

Em sua segunda teoria, a teleosemântica informacional, Dretske une a noção de informação com a de função. Dessa maneira, enquanto a sua primeira teoria consolida a sua teoria na noção de informação carregada pelo signo, a teleosemântica determina que uma representação R tem a função F de Indicar um estado de coisas, caso carregue informação acerca deste estado de coisas. Assim, há uma conexão entre função e informação. Conforme o filósofo, todos os fatos mentais são representacionais, embora nem todas as representações sejam mentais, porque há representações públicas, como mapas, gráficos, fotografias, etc.

The fundamental Idea is that a system S, represents a property, F, if and only if S has the function of indicating (providing information about) the F of a certain domain of objects. The way S performs its

function (when it performs it) is by occupying different states $s_1, s_2, \dots, s_n$ corresponding to the different determinate values $f_1, f_2, \dots, f_n$ of F . (DRETSKE, 1995, p.02).

Diante disso, a representação de um estado mental ocorre *sse* este estado tem a função de Indicar. Aqui é capturada a normatividade, o que para Dretske (1995) é um elemento inerente à representatividade. Quanto à noção de função, ela é etiológica, vincula-se a ideia de efeito, isto é, todos os itens e/ou dispositivos foram feitos para, em condições normais, executarem uma função específica, uma função própria, como a função do coração é bombear o sangue. Teleosemântica é uma teoria de significado, por consequência, uma teoria de Representação. “Teleosemantics⁷ does this by identifying the meaning representational content of a neurobiological state (organ, mechanism, etc.) with what it has the function of indicating about the world”. (DRETSKE, 2006, p.71).

Quanto ao conteúdo de uma representação, Chediak expõe que o conteúdo de uma representação não informa, não se refere ao objeto, mas às suas propriedades. “O conteúdo de uma representação não informa sobre o objeto que ela visa, mas sobre as propriedades do objeto, propriedades que poderiam ser de outro objeto, pois o objeto varia de contexto a contexto”. (CHEDIAK, 2014, p.10).

E, embora a informação seja relevante para Indicar a função de uma representação, nem toda informação possui uma função, mas toda representação deve possuir uma função. Dessa forma, toda função de uma representação carrega informação equivalente àquilo que ela Indica. Além do mais, para Dretske, há dois tipos de representações: as naturais e as convencionais. As naturais são, basicamente, as experiências, percepções, sensações e os sentimentos, que são desprovidos de conceitos e os estados mentais intencionais, tais como pensamentos, crenças, intenções, etc. E as convencionais, são a linguagem, altímetros, velocímetros, termômetros, etc.

Nesta segunda versão da teoria de Dretske, é relevante expor a relação existente entre Indicação e representação. Assim, enquanto a Indicação sempre será verdadeira, pois não pode ocorrer se não for o caso, a representação pode ser falsa. Como no caso das manchas vermelhas, aqui⁸, elas podem não representar, isto é, podem não ser signo natural de sarampo. Porém, como

⁷De acordo com sua versão (1988, 1995).

⁸Nesta teoria teleosemântica informacional.

já foi colocado, em relação à ideia da natureza da informação, indicador e indicado possuem conexão mesmo que não haja conhecimento de um sujeito cognitivo.

Em suma, quando a representação não desempenha o seu papel, não realiza sua função de Indicação, então a representação falha. Isso porque deve haver uma correspondência entre a representação, como signo natural, e o fato e/ou estado de coisas que ela representa, considerando também as condições normais. Logo, uma representação R é verdadeira quando executa corretamente a sua função F de Indicação, conforme a informação que carrega.

3. A teleosemântica de Millikan

A teleosemântica trata, especialmente, sobre a possibilidade do erro, mais especificamente: Como é possível fixar o conteúdo semântico de uma representação, já que signos intencionais podem ser falsos? Em uma resposta curta, Millikan coloca que todas as coisas podem falhar quando não realizam seus propósitos de representarem os estados de coisas esperados. Com isso, Millikan repaginou a ideia de Intencionalidade⁹, tornando-a natural.

Segundo Millikan, a Intencionalidade é uma propriedade relacional capaz de mediar tanto objetos físicos, quanto objetos mentais. Naturalizar a intencionalidade é considerar que todas as coisas existentes têm como propósito representar estados de coisas. Em vista disso, por vezes, alguns itens e/ou dispositivos podem falhar em suas representações.

A teia teórica de Millikan é grande, além do termo técnico propósito há outros elementos que auxiliam para corroborar a sua teoria baseada na biologia evolutiva. O segundo termo técnico é o de função própria. Função própria significa que algo foi designado e/ou projetado para desempenhar um determinado papel. De acordo com Millikan, ainda, categorias biológicas, psicológicas e/ou linguísticas são categorias de funções próprias, as quais fazem parte de uma categoria biológica.

⁹A confusão começa, ou pelo menos passa, por meio da comum explicação dos escritos de Franz Brentano. Conforme essa interpretação (Chisholm, 1967), Brentano manteve aquilo que distingue fenômeno mental de fenômeno físico, em que o primeiro possui uma propriedade chamada “intencionalidade”. Ele caracterizou intencionalidade de dois modos diferentes. Ele fala de “intencionalidade de um objeto”, “direção para um objeto”, ou “referência para um conteúdo” (MILLIKAN, 2002, p.63). (*Tradução Nossa*).

If language device tokens and mental intentional states (believing that, intending to, hoping that) are members of proper function or “biological” categories, then they are language devices or intentional states not by virtue of their powers but by virtue of what they are supposed to be able to do yet perhaps cannot do. (MILLIKAN, 1985, p.17).

Conforme a história evolutiva de todos os itens e/ou dispositivos, todas as coisas existentes têm uma função própria, independentemente de suas disposições atuais. Aliás, as funções próprias foram adquiridas e preservadas conforme suas histórias evolutivas, por isso em um dado momento podem, por uma circunstância ou outra, deixar de desempenhar essa função. Um estômago, por exemplo, sua função própria é digerir a comida, todavia se houver algum problema que impeça este órgão de desempenhar o seu papel, mesmo assim sua função permanece, pois ele foi feito para isso por meio de um processo evolutivo.

Ademais, ter êxito em realizar a função própria quer dizer que este desempenho foi feito em condições normais, por meio de mecanismos normais. Normal também é um termo técnico e possui um sentido quase normativo, tendo relação com história e uso. Ao longo de seus escritos (1984, 1993, 2002), Millikan utilizou tal termo de vários modos, embora todos queiram dizer o mesmo: condições normais, explicação normal e mecanismos normais. “A ‘normal explanations’ explains the performance of a particular function, telling how it was (typically) historically performed on those (perhaps rare) occasions when it was properly performed”. (Millikan, 1993, p.86).

Nesse sentido, a função própria tem relação com a potencialidade e a história evolutiva. E isso é verificado a partir da FRE, (família reprodutivamente estabelecida), da seguinte forma: no passado, um item foi feito, ganhou uma característica determinada e passou a ser responsável para desempenhar uma determinada função. A função própria é o efeito desse item que têm continuado, ao longo da história evolutiva, por seleção e uma boa adaptação ao ambiente.

A definição de função própria é recursiva e para seu cumprimento, se faz necessário pelo menos duas condições:

- (1) A originou-se de uma reprodução, no caso, é uma cópia ou cópia de uma cópia.
- (2) A originou-se como produto de algum dispositivo a priori.

Contudo, uma vez estabelecida, a função própria não perde o seu posto, embora possa falhar. Millikan (em um dos seus vários exemplos) exemplifica com a máquina de café, sua função própria é fazer café, mas mesmo se for quebrada e/ou ocorra algum outro problema, a máquina permanece sendo uma máquina que faz café.

Para Millikan, o nosso sistema cognitivo nos possibilita representar, e conforme a sua teoria, há dois sistemas: um que produz a representação e outro que consome, interpreta a representação. O seu exemplo mais comum é a dança das abelhas¹⁰. As abelhas fazem uma espécie de dança, em formato de oito, a qual direciona as outras abelhas ao néctar. Essa dança é feita graças a um mapeamento semântico. “É a correspondência entre os movimentos das abelhas e as áreas do mundo onde há néctar (o que Millikan chama de mapeamento), tal como interpretada ou utilizada por outras abelhas, o que constitui o funcionamento correto e, portanto, a representação correta”. (RITCHIE, 2008, p.234).

Nesse sentido, há signos-produzidos e signos- consumidos. Os signos-produzidos, por meio de um isomorfismo, produzem representações que correspondam ao mundo e os signos-consumidos e/ou intérpretes, quando realizam suas funções, por meio de mecanismos normais verificam essa correspondência. Em vista disso, para Millikan (2002), se houver variações no mundo, essas variações devem corresponder aos signos que produzem variações adaptativas nas atividades dos signos-consumidos e/ou intérpretes. O mais relevante, então, são as representações produzidas para algum fim, que tenham algum uso, uma utilidade biológica que contribua para com a sobrevivência, adaptação e permanência das espécies.

A determinação do conteúdo da representação, por sua vez, é dada pelo modo que o sistema fixa este mapeamento. “(...) content is determined by the rules or functions in accordance with which representations need to correspond to world affairs (to what they represent) if the responses of these interpreters are going to produce the effects they are designed to produce by way of normal mechanisms”. (Millikan, 2006, p.100). A preocupação de Millikan concentra-se nas representações que trazem algum tipo de benefício ao sistema, órgão e/ou organismo.

¹⁰A dança das abelhas é o tipo de representação mais simples, primitiva, chamada por Millikan de Pushmi-pullyu.

4. Considerações finais

Como foi visto Dretske e Millikan lidam com a questão do erro nas representações de modos distintos. Dretske foca na noção de Indicação. Na sua primeira teoria trata acerca de uma relação causal entre signos e estados de coisas. E na segunda, coloca que toda representação tem a função de Indicar. No entanto, a sua primeira teoria, especialmente, pode gerar alguns equívocos, tais como: a relação causal pode gerar o problema da disjunção, isto é, o mesmo signo pode Indicar diferentes estados de coisas, como o exemplo das manchas vermelhas, as quais não necessariamente Indicam apenas sarampo.

Ademais, quando ele coloca que o erro nas representações pode ocorrer por meio de um processo de aprendizagem não ideal, não explica como diferenciar aprendizagem ideal da não ideal. Já a visão teleosemântica de Millikan mantém a sua base na ideia de intencionalidade, no caso, o seu significado, refere-se à noção de propósito. Para tanto, redefine o conceito tratado por Brentano e alarga a sua aplicação. Sua ideia, ao considerar a intencionalidade naturalizada, parece ser de ultrapassar os limites da distinção entre físico e mental, cujo objetivo parece ser explicar como é possível lidar com a "falha" e não só com os casos de sucesso.

A função própria, embora um tanto abstrata, tem relação com a propriedade do objeto, sem descaracterizá-lo. Analisar as coisas por meio de suas histórias e evolução equivale a identificá-las com suas origens, embora tenham sido projetadas para um determinado fim, salva as exceções quando essas fracassam. A abrangência dessa teoria, ao compará-la com as teorias de Dretske, parece ser maior devido ao fato de deixar claro o quanto cada coisa, mesmo possuindo um determinado rótulo, não é definida por tal rótulo, mas unicamente por seu propósito.

Referências bibliográficas

CHEDIAK, K. Função Natural e Indicação – A atividade de Representação em Fred Dretske. *Filosofia, Darwinismo e Evolución*. Ed. Alejandro Rosas. Universidad Nacional de Colombia. Bogotá, p. 100 – 119. 2007.

_____. Representação e Contexto: A crítica de Millikan a Dretske. *Dissertation*, Pelotas, p.257-271. Jan.2014, v.40.

- DRETSKE, F. *Knowledge and the Flow of Information*. Cambridge, Massachussets: MIT PRESS, 1981.
- _____. Misrepresentation. In: STICH, S.P; WARFIELD, T.A. (Org.). *Mental Representation: A Reader*. Oxford, Reino Unido e Cambridge, Massachussets: Basil Blackwell, 1994. p.157-173.
- _____. *Naturalizing the Mind*. Cambridge, Massachussets: MIT PRESS, 1995.
- GRICE, H.P. *Meaning*. The Philosophical Review. p.377-388, v.66, jul 1957.
- MILLIKAN, R. *Language, Thought and Other Biological Categories*. Cambridge: MIT Press, 1984.
- _____. *Varieties of Meaning*. Cambridge: MIT Press, 2002.
- _____. *White Queen Psychology and Other Essays*. Cambridge: MIT Press, 1993.
- PAPINEAU, D. *Teleosemantics*. Oxford: Clarendon Press, 2006.
- RITCHIE, J. *Naturalismo*. Tradução de Fábio Creder. Petrópolis: Vozes, 2012.
- SILVA, P. Intencionalidade: Mecanismos e Interação. *Revista Principia*. UFSC. Santa Catarina, p.255-277, v. 14, 2012.
- SOUZA FILHO, S.F. *Seguir Regras e Naturalismo Semântico*. 2013. Dissertação (Mestrado em Filosofia) – Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2013.
- STEIN, S.I.A. Ações Linguísticas Significativas: Relação entre Propósitos e Projeções. *IX SAPPGFILPUCRS*. Porto Alegre, p.1-11, jun. 2012.
- TEIXEIRA, J.F. Knowledge and the Flow of Information (Review). *Transformação*. São Paulo, p.133-140, v.12, 1989.

Ensaio sobre a concepção de *psykhé* em Platão no *Cármides* e no *Fédon*

Essay on the conception of psykhé in Plato in Charmides and Phaedo

Bárbara de Abreu Freitas¹¹

RESUMO: O homem ocidental, desde sua origem mais remota, começa a questionar o que vem a se constituir esse “eu” que nos consolidaria como pessoas. Por mais anacrônico que sejam os termos, “eu”, “subjetividade” e “consciência”, acreditamos que, de certa maneira, quando os primeiros pensadores identificaram a *psykhé* como fonte de vida ou mesmo como a verdadeira parte do homem, eles já estavam refletindo sobre a mesma problemática que a filosofia da mente na contemporaneidade. O primeiro a elaborar de forma mais concisa o que seria a *psykhé* foi Platão, que por meio dos seus diálogos, procurou entender e refletir sobre essa parte do ser humano tão misteriosa e difusa e que continua até hoje sendo alvo de várias interpretações. Platão não elaborou uma única maneira de pensar a *psykhé* e, no decorrer de sua obra apresentou várias perspectivas sobre o mesmo problema. No fluir dessa comunicação buscaremos apresentar duas concepções distintas e radicalmente diferentes que encontramos nos diálogos de Platão. A primeira concepção se apresenta em seu diálogo *Fédon*, diálogo no qual, Platão reflete sobre a imortalidade da alma e estipula umas das principais teses da *psykhé*, sendo está apresentada como oposta ao corpo. Em contrapartida, notamos no diálogo *Cármides*, a tese de um *psykhé* que não mais é diferente do corpo, mas se apresenta como parte dele e assim, como o corpo fica doente, a *psykhé* também ficaria, sendo que uma poderia influenciar a outra.

PALAVRAS-CHAVE: *psykhé*; mente; *Cármides*; *Fédon*; Platão.

ABSTRACT: Since the very beginning, Western people have questioned what, after all, constitutes this ‘Self’ that makes us persons. Though it may sound anachronic, we believe that terms like ‘Self’, ‘subjectivity’, or ‘consciousness’, capture partly what the first thinkers meant by *psykhé*: when they used it to identify the source of life or the true part of man, they were already discussing the same problems now discussed in contemporary philosophy of mind. The first one to elaborate, in a more concise way, the notion of *psykhé* was Plato, who, by means of his dialogues, tried to understand and reflect upon this mysterious and diffuse part of human being which is still today a source of multiple interpretations. Plato did not elaborate a unique approach on *psykhé*. Along his works, he presented various perspectives on the same topic. We will try here to present two radically distinct conceptions that can be found in Plato’s dialogues. The first one appears in *Phaedo*, which is a dialogue in which Plato reflects upon the immortality of the soul and stipulates one of the main thesis of *psykhé*: i.e. it is distinct from the body. On the other hand, in *Charmides*, the favored thesis says that the *psykhé* is not distinct from the body, but a part of it, and when the body is sick, the *psykhé* is equally sick, showing that one can influence the other.

KEYWORDS: *psykhé*; mind; *Charmides*; *Phaedo*; Plato.

¹¹Graduada em Filosofia pela Universidade Federal de São João del-Rei (UFSJ). Email: barbara-fr-ab@hotmail.com.

Desde a antiguidade, o homem se pergunta por questões que até hoje nós fazemos: Quem somos nós? Do que somos constituídos? O que é essa voz que fala dentro de nós? Eu sou isso? E meu corpo, faz parte de quem eu sou ou é apenas um empecilho? Isso que eu chamo de eu, está dentro de mim?

Na contemporaneidade a filosofia da mente continua a se debruçar sobre essas perguntas que até os dias atuais não possuem uma única resposta, se é que um dia terão. Desde o berço do pensamento ocidental, essas perguntas, embora não tão explicitamente, já eram feitas. Por mais anacrônico que sejam os termos, “eu”, “mente” e “consciência”, acreditamos que, de certa maneira, quando os primeiros pensadores identificaram a *psykhé* como fonte de vida ou mesmo como a verdadeira parte do homem, eles já estavam refletindo sobre a mesma problemática que a filosofia da mente na contemporaneidade.

A primeira identificação do que chamamos “mente” fora feita pelos gregos arcaicos com os nomes *thymos*, *θῦμός* (ânimo), *phren*, *φρήν* (mente), *nous*, *νοῦς* (pensamento) e *psykhé*, *ψυχή* (alma). Nas traduções dos poemas *Íliada* e *Odisseia* para as línguas latinas, frequentemente essas palavras foram traduzidas como alma ou espírito. Notamos que, diferentemente dos outros termos, o termo *psykhé* é usado apenas para designar a alma no momento da morte. Assim, a *psykhé*, neste ponto de vista é encarada como vida ou “sopro de vida”, pois ao sair do corpo morto, por meio de ferimentos ou voando pela boca, ia em direção ao *Hades* onde permaneceria para sempre como espectro. Segundo Reale, o termo *psykhé* está diretamente ligado à respiração. *Psykhé* significa soprar, assim, a ideia da morte está ligada ao exalar o último suspiro (REALE, 2002, p. 70). Pode-se deduzir que *psykhé* era diretamente ligada à ideia de *πνεύματα* (pneumata), costumeiramente traduzido como espírito. Todavia, essa palavra advém de *πνέω* (pneu) que significa sopro ou respiração¹².

Todavia, essa visão da *psykhé* fora se modificando e se assemelhando com o conceito que entendemos como alma hoje em dia: uma segunda pessoa, ou em alguns casos, a verdadeira pessoa que habita em nós. Rompendo com a tradição épica da poesia homérica, quais as ações não se separavam da influência dos deuses, Platão começa a formular um intelecto que consegue

¹²Snell, historiador da antiguidade grega nos relata: “*Psykhé*, palavra usada para “alma” no grego mais tardio, nada tem a ver na origem com a alma pensante e sensiente. Em Homero, *psykhé* só é alma enquanto anima o homem, isto é, enquanto o mantém vivo” (SNELL, 2009, p.8).

se distanciar do objeto, e assim, pode construir suas próprias leis e ideais. Nesse plano moral, a construção de uma *psykhé* ativa é o que baseará as ações virtuosas, assunto qual ele discute em vários diálogos.

Até os séculos IV e V a. C., a ideia da *psykhé* como um segundo eu ou como uma parte intrínseca do homem era estranha aos olhos do homem grego, porém, algo mudou no percurso da história¹³. Assim, se fomos transportados para a contemporaneidade de Platão, iríamos notar que o cidadão grego não possui uma única ideia de *psykhé*. Boa parte da população ainda acreditava na formulação homérica da alma apenas como um espírito. Notamos isso presente na comédia de Aristófanes, *As nuvens*, qual ele satiriza a ideia socrática da alma, chamando Sócrates de fantasma. Porém, mesmo no pensamento socrático/platônico, notamos uma mudança de visão ao pensar a *psykhé*.

Todavia, mesmo quebrando com o modo de pensar da época, Platão fora influenciado por concepções anteriores. Não podemos dizer que Platão elaborou uma única forma de pensar a *psykhé*, pois no decorrer de sua obra, ele repensou, sob diversos ângulos, o que viria ser essa parte do homem tão intrínseca e tão misteriosa.

No diálogo Crítón 77 d,e, Platão se refere a *psykhé* como "Essa qualquer coisa que vive em nós". De forma curiosa o tema da alma se encontra presente em quase todos os diálogos de Platão. Seja refletindo em torno do bem ou da justiça, o problema da alma aparece e é fonte de várias especulações.

Mediante a isso, escolhemos delimitar nossa comunicação a dois diálogos, para exemplificar duas concepções de *psykhé* diferentes uma da outra. O primeiro diálogo analisado é o de Cármides ou da Moderação, um diálogo de juventude¹⁴, considerado por muitos pesquisadores

¹³ Segundo Dodds: O fato do termo *psyche* causar um sentimento de estranheza para os cidadãos atenienses do século V a. C., uma coisa é certa: a palavra não possui nenhum sabor de puritanismo, e nem sequer gozava de qualquer *status* metafísico. A alma não era nenhuma prisioneira relutante do corpo, mas sim a vida ou espírito do corpo, sentindo-se perfeitamente à vontade ali. Foi nesse momento que o novo padrão religioso fez sua fatídica contribuição- ao creditar ao homem um "eu" oculto, de origem divina, e, por conseguinte, colocar em desacordo corpo e alma, este padrão introduz em meio à cultura europeia uma nova interpretação da existência humana (DODDS, 2002, p. 143).

¹⁴ Segundo Gregory Vlastos, pode-se separar os diálogos de Platão em 'juventude', 'maturidade' e 'tardios'. A investigação da alma é fundamental para essa filosofia. O período de maturidade começa quando Platão retorna a Atenas (387. A.C.), funda a "academia" e escreve os diálogos *República*, *Fedro*, *Banquete* e *Fédon*. G. (VLASTOS, 1991).

como um diálogo com mais influência socrática. O segundo diálogo é o Fédon ou sobre a imortalidade, um diálogo de maturidade e que, segundo a hipótese de alguns comentadores, foi escrito depois da viagem de Platão ao sul da Itália, na Sicília¹⁵, que segundo estes, foi onde ele teve reais influências órfico/pitagóricas, o que poderia explicar alguns conceitos apresentados no diálogo como da transmigração e da metempsicose da alma.

Assim, ao analisar o conceito de *psykhé* presente nesses diálogos, outros conceitos também se farão presentes, e assim, a temática perpassará pela visão de Platão sobre a morte e suas influências órficas/pitagórica.

Iniciando pela análise do diálogo Cármides, a questão da *psykhé* é introduzida bem no início do diálogo, quando Sócrates ao analisar o seu futuro interlocutor, Cármides, reflete sobre este ter o físico bonito, mas ainda não saber de sua *psykhé*, necessitando antes de admirá-lo, despir-lhe a alma para ver se ele é bom (155 a). Para isso, Sócrates usa de uma desculpa para começar um diálogo com Cármides. Se passando por um médico, Sócrates diz poder curar a dor de cabeça de Cármides, por meio de um encantamento¹⁶ que este aprendera com os médicos trácios, discípulos de Zalmoxis. Segundo os ensinamentos desses médicos, não era possível curar os olhos sem curar a cabeça, da mesma forma, diz Platão:

“não é possível tratar do corpo sem cuidar da alma, sendo essa a causa de desafiarem muitas doenças o tratamento dos médicos helenos, por desconhecerem este o conjunto que importa ser tratado, pois não se pode ir bem à parte, quando vai mal o todo” (156e).

Zalmoxis, segundo alguns mitos fora um homem que morreu e seguiu em rumo à *katabasis* (*kata*- descida, *basis*- ir), ao Hades, onde ficou por três anos e depois retornou. Porém, segundo os relatos de Heródoto, Zalmoxis era um escravo de Pitágoras¹⁷. Podemos então entender que os

¹⁵ Segundo Gorman: “A Sicília e a Itália foram, na Antiguidade, os principais centros do pitagorismo. Foi para a Sicília que Platão se dirigiu a fim de comprar o famoso livro pitagórico de Filolau de Arquitas, um importante político e matemático pitagórico do século IV a. C.” (GORMAN, 1979, p. 15).

¹⁶ Como nos mostra Pinheiros, os encantamentos se mostram importante, pois o questionamento dialético significaria a entrega total e irrestrita da alma do ouvinte, que assim como um remédio, provariam “uma transformação de caráter de tal magnitude a produzir a sensatez e temperança no homem” (PINHEIRO, 2005, p.178).

¹⁷ Segundo Gorman: “É digna de nota a afirmação segundo o qual Pitágoras teria libertado seu escravo Zalmóxis, tornando-o seu amigo. Mais tarde, alguns pitagóricos, tais como Plotino e Porfírio, também deixaram de ter escravos ou servos em nome da independência e da liberdade interior”. (GORMAN, 1979, p. 15)

pitagóricos foram influenciados pelo mito ou mesmo que ambos foram associados por terem uma crença da imortalidade da alma. Contudo, segundo Dodds:

O trácio Zalmoxis na obra de Heródoto, reunia “os melhores cidadãos” para anunciar que a alma humana não era imortal, mas que ele e seus descendentes iriam viver para sempre- aparentemente eles seriam pessoas escolhidas, uma espécie de elite espiritual. Que existia alguma analogia entre Zalomoxis e Pitágoras, é algo que deve ter ocorrido aos colonizadores gregos da Trácia, de quem Heródoto ouviu a estória, tanto assim que fizeram de Zalmoxis um escravo de Pitágoras. Isto é um absurdo, como percebeu Heródoto, pois o verdadeiro Zalmoxis era um daemon, talvez um xamã do passado transformado em herói¹⁸ (DODDS, 2002, p. 148)

O motivo principal pelo qual a alma deve ser tratada junto com o corpo é apresentado logo em seguida, quando Sócrates diz: “é da alma, declarou, que saem todos os males e todos os bens do corpo e da cabeça” (157 a). Por isso que a único jeito de curar a alma é tratando-a com certas fórmulas mágicas, que segundo Sócrates, seriam os belos argumentos demonstrados pela *sofrósina* (σοφροσύνη)¹⁹ traduzida quase sempre por temperança ou prudência. Assim, Sócrates mostra que não se pode tratar da temperança e da saúde de forma separada, pois estas dependem mutualmente uma da outra para a saúde do homem. A Cármides é sugerido, como remédio para sua dor de cabeça, que este se cure intelectualmente(156d). Porém, para ser curado pelos belos argumentos, Cármides precisaria possuir a temperança, assunto o qual se desenrolará por todo o diálogo.

No Cármides, Platão propõe uma conjunção entre a *psykhé* e corpo, de modo que ambos buscassem um equilíbrio²⁰. Apenas por meio dessa unidade, ou seja, pelo equilíbrio entre ambas as partes, que poderia haver a cura do corpo e da alma. Isso não significa que ambos sejam uma só coisa, o “Homem”, mas que ambos são necessário para a cura e para o bem-estar do homem. Segundo Robinson:

Se a linguagem do médico trácio deve ser considerada exata, de algum modo, a relação entre alma e o corpo é obviamente não de simples adição numérica e a combinação seria igual ao “eu” ou ao

¹⁸ Dodds compara a figura xamânica de Zalmoxis a figura De Orfeu, também de origem trácia e dotado de poderes mágicos e curandeiro (DODDS, 2002, p. 150)

¹⁹ Segundo PINHEIRO: “A palavra *sophrosyne* é de difícil tradução. Tem, em seu radical, relação com *saos*, e *phren*, saudável e espírito, mente. *Phren*, no entanto, também designa a parte entre o coração e o abdômen, o diafragma, mostrando forte relação com o mundo concreto dos sentidos. Acreditamos que a palavra “sensatez” traduz bem o termo *sophrosyne* por conter tanto o aspecto sensorial em seu étimo quanto o espiritual em seu sentido. As palavras temperança e prudência parecem perder sua força se tivermos em mente toda discussão da segunda parte do *Cármides* sobre a *sophrosyne* ser ou não conhecimento de si e, por isso, ser ciência das ciências. No entanto, a tradução mais comum para o português é temperança” (PINHEIRO, 2005, p. 175)

²⁰ Essa busca pelo equilíbrio pode ser vista como uma possível semelhança com o pitagorismo, que buscava pelo culto dos números, a conquista do equilíbrio universal.

“homem todo”. A chave para a relação (e para o próprio conceito de alma) parece estar na oração “porque eles desconsideram o todo”. Esse “todo” pode ser razoavelmente referir-se a duas coisas: (a) o corpo todo; (b) o homem todo, isto é, corpo e alma combinados. Se é (a), os médicos gregos estão sendo castigados por desprezar os cursos de fisiologia geral; eles não conseguem compreender que as doenças que afetam órgãos particulares só podem ser adequadamente diagnosticadas à luz de uma apreciação mais geral das operações do corpo humano como um todo. Entretanto, isso dificilmente combina com a frase imediatamente precedente combina com as frases imediatamente precedentes e imediatamente seguintes, nas quais é feita a referências à “alma”. Portanto, a referência é presumivelmente a (b) e os médicos estão sendo repreendidos por serem incapazes de avaliar que muitas doenças são, de fato, (para usar o jargão moderno), psicossomáticas. Se isso é verdade, trata-se de um insight médico que antecede em muito as descobertas supostamente modernas no campo da psicologia, mas a noção de alma que ele parece implicar parece até mais notável em sua agudeza. (ROBINSON, 2010, p. 67,68)

Ao contrário, como veremos no Fédon, não se trata de coisas que sofrem uma constante tensão, mas que precisam um do outro. Podemos dizer, assim, que corpo e *psykhé* são vistos como uma unidade participativa, ou seja, haveria algo entre o corpo e alma que o interligaria e que faria a comunicação. Essa unidade não é entendida matematicamente como um, mas sim como uma conjunção de partes, ou em outras palavras, uma unidade complementar (COLTINHO, 2013, p. 94). Assim, a interação entre ambos geraria um processamento de informações necessário para o bom funcionamento do homem²¹. Na interpretação de Robinson:

À primeira vista, isso parece com a afirmação de que o “homem todo” é a soma aritmética de corpo e alma, mas o paralelismo feito nas palavras finais deixa claro que a posição de Sócrates ainda permanece. A influência da alma sob o corpo é como a da cabeça sobre os olhos, e, ao se buscar a fonte dos problemas, deve-se examinar, num caso, a cabeça, no outro, a alma. Se isso é verdade, a frase “o homem como um todo” (πᾶςὁἄνθρωπος) só pode significar “a pessoa” ou “o homem inteiro”, no sentido que foi visto, isto é, “a alma” (ROBINSON, 2007, p. 43).

Contrário à ideia homérica de que a cura viria por meio dos deuses, Platão no Cármides, propõe uma cura interna por meio da *psykhé* em conjunto com o corpo. Em termos contemporâneos, Platão fala aqui, nesse diálogo, - embora anacronicamente-, em uma “psicossomática”, termo cunhado séculos mais tarde, mas que denotaria o mesmo fator, a saber: o corpo seria influenciado pelos estados mentais²².

²¹Robinson nos atenta para isso, quando diz: “Os dois, corpo e alma, estão intimamente ligados, sendo a destruição de qualquer um, pressuposto da destruição de ambos. Assim como cabeça e corpo se implicam como côncavo e convexo, e o mesmo pode ser dito de corpo e alma. Em síntese, a alma parece ser compreendida aqui como a pessoa total ou o homem todo, do qual o corpo vivo (σῶμα) é uma parte integral” (ROBINSON, 2007, p.42)

²² Ainda segundo esse pesquisador: “O Cármides ainda afirma que é a pessoa total que é a alma e que só a uma pessoa podem ser atribuídas atividades morais e intelectuais de modo significativo. Se essa análise é correta, o médico trácio está dizendo algo que não havia sido verbalizado pelos gregos. Até então, a maior parte dos comentários sobre o corpo e sobre a alma tinham sido francamente dualistas” (ROBINSON, 2010, p. 72).

Por outro lado, vemos a *psykhé* sendo apresentada no diálogo Fédon de um modo sutilmente diferente. Platão agora não fala de uma *psykhé* em conjunto com corpo, mas de uma separação entre eles. A *psykhé* passa a ter um caráter distinto do corpo, correspondendo à verdadeira parte do homem. Porém, a *psykhé* não é delimitada apenas dessa maneira. O conceito de *psykhé* é analisado de diversas formas no diálogo Fédon. Platão e o seu pensar sobre a temática ainda desenrola e reformula sua concepção de *psykhé*, assim, notaremos durante sua exposição uma *psykhé* que é apresentada, principalmente, de três formas: como um princípio cognitivo, como características pessoais e como uma pessoa interna (ROBINSON, 2007, p. 60). Dessas três formas a *psykhé* é apresentada como um “eu” duplo do homem²³, que está em constante tensão com o corpo, mas que sofre interação com ele. A *psykhé* aparece de forma mais veemente como a verdadeira parte do homem, todavia, não podemos falar do homem separado do corpo e da alma. Porém, a essência deste, ou seja, seu verdadeiro eu estaria presente na *psykhé*. Há em Platão uma inovação ao pensar a *psykhé*, que além de sinalizar a ideia de um “eu” do sujeito o apresenta como um ente racional, tal ideal, segundo Dodds, se mostra bem distinta da tradição grega²⁴.

A *psykhé* é apresentada no diálogo como outro “nível”²⁵. O ponto é que, mesmo distintos, corpo e alma sofreriam uma tensão constante, como observamos nas passagens, 64 a, 65 a-c, 93 a. Porém, mesmo havendo trechos nos quais notamos uma exata diferenciação entre ambas, há passagens que destoam; se, por um lado, o corpo é descrito em 82 e, 92 a como uma prisão para alma, como algo prejudicial por ser a fonte dos prazeres e da ignorância (64 e, 65 a, 81 a), e como empecilho na busca pela Verdade (66 a-b), em outra passagem, a *psykhé* é descrita como usando o corpo para operar os seus sentidos, visão e audição (79c). Assim, embora a *psykhé* não seja corpórea, ela pode ser influenciada pelo corpo, e segundo uma possível interpretação dessa

²³ Segundo Goldschmidt, não se deve dizer que é alma de Sócrates, que sobrevive, mas sim Sócrates mesmo. Assim, a “alma imortal é o que é, propriamente, nós mesmos, o que nos permite dizer: eu, o que constitui minha personalidade e nós dá um nome” (GOLDSMIDT, 1970, p. 71)

²⁴ Segundo Dodds: Em escritores áticos do século V a. C., assim como em seus predecessores jônicos, o “eu” designado pela palavra *psyche* é normalmente mais emocional do que racional. Fala-se dela como da sede da coragem, da paixão, da piedade, da ansiedade, do apetite animal. Mas antes de Platão, raramente, ou quase nunca, ela é citada como sede da razão – sua extensão sendo tão ampla quando a do *thumos* homérico” (DODDS, 2002, p. 142)

²⁵ Poderíamos usar para designar a diferença entre corpo e *psykhé* o termo substância ou mesmo fenômeno, porém ambas seriam ainda mais anacrônicas e poderiam gerar uma longa discussão, pois estão impregnadas com os termos empregados na história da filosofia. Assim, o emprego do termo nível apontaria tanto que *psykhé* e corpo são “coisas” distintas como estão em graus diferentes, diferente da matéria, sendo corpo e alma descritos como pertencentes a naturezas diferentes.

passagem, chegar a usá-lo. Porém, mesmo que a *psykhé* necessite do corpo enquanto em vida, o exercício de morte se torna estritamente necessário.

Essa noção do corpo como uma prisão é mostrado no diálogo recorrendo a crença órfica: “A esse respeito há mesmo uma fórmula que usam os adeptos dos ministérios: É uma espécie de prisão o lugar onde nós, homens, vivemos, e é dever não libertar-se a si mesmo nem evadir-se”. (62 b)

Segundo a tradição dos mistérios órficos, o corpo era uma espécie de cárcere da *psykhé*. Estando ela presa a este corpo não deveria tentar se soltar dele por meio do suicídio. A razão pela qual é estrita a proibição é apresentada logo em seguida no diálogo: “Os Deuses são aqueles sob cuja guarda estamos, e nós, homens, somos uma parte da propriedade dos Deuses”. (62 b)

Dessa maneira, mesmo sendo de alegria o momento da morte, pois esta significa a libertação das prisões corporais, esta não deve ser forçada; em contrapartida deve ser com consentimento divino. Notamos que Sócrates encara o mandato de morte como um sinal de que está na hora de partir. Assim, quando de frente com a morte, um filósofo não deve fugir, pois, segundo o diálogo: “Quando uma pessoa se dedica à filosofia no sentido correto do termo, os demais ignoram que sua única ocupação consiste em preparar-se para morrer e em estar morto” (64 a). Deste modo, a relação entre alma e morte se torna clara no filósofo, pois é na morte que a alma se apresenta de forma mais clara.

Compartilhando a crença na imortalidade, Sócrates estava convencido de que apenas na morte se poderia alcançar a Verdade, estando junto da companhia dos deuses. Ao explicar os motivos de sua tranquilidade, mesmo estando prestes a morrer, Sócrates mencionará uma antiga tradição -que pode se referir aos Mistérios de Elêusis- que esperavam ter melhor sorte do que os demais mortos, devido a sua esperança espiritual.

Platão mais uma vez fará referências ao orfismo:

E purificação não vem a ser, precisamente, o que dissemos antes: separar do corpo, quanto possível, a alma, e habituá-la a concentrar-se e recolher-se a si mesma, a afastar-se de todas as partes do corpo e a viver, agora e no futuro, isolada quanto possível e por si mesma, e como que libertada dos grilhões do corpo? (67 d)

Ao referir-se à purificação, Platão remete-se às purificações ou à catarse, através da prece da oferenda praticada pelo orfismo²⁶. A crença órfica considerava tudo que era impuro como algo que afastava os deuses. Para a Antiguidade grega, quando se cometia um crime, o castigo recaía não só sobre o criminoso como sobre seus familiares. Os órficos resolveram o problema da “culpa” de modo inovador: a culpa é sempre de responsabilidade individual e por ela se paga aqui, em vida e em reencarnações posteriores, até que o erro seja redimido.

Para além de um tipo de dualismo, corpo e alma, notamos que o diálogo acena para uma terceira forma: o filósofo. Assim, haveria “níveis” ou “estratos”: a *psykhé*, encarnada ao corpo, o corpo, que muitas vezes tem um caráter negativo e o filósofo, que procura pelos conhecimentos se purificar e distanciar o máximo possível do corpo, tendo em vista a busca pela contemplação das ideias, mesmo enquanto imagens. Essas purificações, descritas em várias partes do diálogo, remetem às purificações órficas, que buscavam pelo conhecimento de certos domínios sobre o corpo e pelo treinamento para a morte (64 a), se prepararem para o pós-morte e assim, para a contemplação da verdade no mundo das ideias. Nesta perspectiva, podemos interpretar o exercício para a morte como uma forma de treinamento para a visão do mundo das ideias. Porém, antes de visualizá-la o filósofo precisaria treinar-se com as imagens das ideias, ou seja, com o mundo material. Assim, o filósofo poderia treinar sua *psykhé* para não sucumbir aos desejos corporais usando de práticas de ascese e de purificações.

Tentamos mostrar nesse artigo as diferentes formas que Platão apresenta a *psykhé* em dois de seus diálogos. A análise comparativa não se deu de forma a testar o conceito de *psykhé* usado nos diálogos, mas apenas em mostrar como Platão ao pensar sobre o problema, analisa de vários ângulos. O problema da *psykhé* estará presente em quase todos os diálogos de Platão. Esse “algo” do homem, que se mostra de grande fascínio de Platão, o impulsiona a tentar descobrir o que de fato é a alma, ou que chamamos de mente, fazendo-o criar, de forma inovadora, um pensar sobre a *psykhé*, sendo ele o precursor de todo modo ocidental de enxergar este problema.

²⁶ Segundo Robinson: “Sócrates conseguiu transformar a ética órfica. Ele aparentemente aceita as ideias de uma existência prévia, de metempsicose e a possibilidade de uma outra vida ou série de vidas depois desta (81 E), e a visão da alma que estas necessariamente implicam pode ser claramente percebida em grande parte do diálogo. Essa visão é essencialmente, uma noção de ‘contrapessoa’ ou ‘duplo de uma pessoa’”(ROBINSON, 2007, p 63)

Notamos que, como no *Cármides* a *psykhé* e corpo assumem um papel de unidade conjuntiva, qual o equilíbrio entre ambos que mostraria essencial para uma boa saúde e para um saber filosófico, no *Fédon*, notamos uma unidade disjuntiva, ou seja, mesmo fazendo parte do homem, corpo e *psykhé* estariam em uma tensão em vida; mesmo necessitando um do outro, a *psykhé* em vez de buscar o equilíbrio com o corpo, buscaria se distanciar dele.

A possível diferença entre os dois diálogos pode ser explicada, como apontamos no início, pela ordem cronológica dos diálogos, e que, segundo alguns intérpretes, haveria uma mudança na perspectiva de Platão depois que este vai para o sul da Itália, na Sicília²⁷ e retorna em 387 a.C., fundando a sua academia. Segundo esses intérpretes, Platão teria tido contato com muitos pitagóricos na viagem, o que explicaria sua mudança ao ver a *psykhé*. Todavia, por mais que exista essa possibilidade de diferenciação entre os dois diálogos, sendo o *Fédon* como mais influência órfica/pitagórica, notamos já no *Cármides* certa alusão ao problema da imortalidade da alma, só que de forma sutil, sendo mais predominante o pensar sobre o equilíbrio entre *psykhé* e corpo enquanto em vida.

Referências Bibliográficas

- COLTINHO, S. L. C. "O mito de Zalmoxis e o sacerdócio de Sócrates: a cura no *Cármides* de Platão". Tese de mestrado, UnB. Brasília, 2013.
- DODDS, E.R. *Os gregos e o irracional*. São Paulo: Escuta, 2002.
- GOLDSCHMIDT, V. *A Religião de Platão*. São Paulo: editora Difusão europeia do livro, 1970.
- GORMAN, P. *Pitágoras: uma vida*. São Paulo: Cultrix/Pensamento, 1979.
- PLATÃO. *A república*. Rio de Janeiro: Globo, 1964.
- _____. *Diálogos: Menon-Banquete, Fedro*. Rio de Janeiro: Edições de Ouro, 1969.
- _____. *Diálogos VI. Crátilo, Cármides, Laques, Ion, Menexeno*. São Paulo: Edipro. 2016.
- _____. *Eutifron; Apologia de Sócrates; Criton*. Lisboa: Verbo, 1972.
- _____. *Fédon*. Belém: Ed. UFPA, 2011.
- _____. *Obras completas*. Madrid: Aguila.
- PINHEIRO, R, M, "Filosofia, Saúde, Alma e Corpo no *Cármides* de Platão". *Princípios*, Natal, vol. 12, nos 17-18, jan./dez. 2005, p.173 -182.

²⁷ Segundo Gorman: "A Sicília e a Itália foram, na Antiguidade, os principais centros do pitagorismo. Foi para a Sicília que Platão se dirigiu a fim de comprar o famoso livro pitagórico de Filolau de Arquitas, um importante político e matemático pitagórico do século IV a. C." (GORMAN, 1979, p. 15).

REALE, G. *Corpo, alma e saúde: o conceito de homem de Homero a Platão*. São Paulo: Paulus, 2002.

ROBINSON, M.T. *As origens da alma: Os gregos e o conceito de alma de Homero a Aristóteles*. São Paulo: Annablumes, 2010.

_____. *A psicologia de Platão*. São Paulo: Loyola, 2007.

SNELL, B. *A cultura grega e as origens do pensamento europeu*. São Paulo: perspectiva, 2009.

VLASTOS, G. *Socrates: Ironist and Moral Philosopher*. Cambridge: Cambridge University Press, 1991.

A Hipótese da Mente Estendida: argumentos e objeções

The Extended Mind Hypothesis: arguments and objections

Bruno Coelho²⁸

RESUMO: A hipótese da mente estendida argumenta que na execução de algumas tarefas, processos cognitivos se estendem para o ambiente. Neste artigo, apresento os argumentos que deram origem à hipótese, assim como algumas críticas. Na primeira seção apresento o externismo ativo, a versão que inicialmente provocou o debate. Na segunda seção, duas críticas contra a hipótese serão detalhadas. Por fim, na terceira seção, argumento que a resolução do debate é uma questão empírica, e apresento algumas evidências em psicologia ecológica que favorecem a existência de sistemas cognitivos estendidos.

PALAVRAS-CHAVE: Mente Estendida; Externismo Ativo; Psicologia Ecológica.

ABSTRACT: The extended-mind hypothesis argues that in performing some tasks, cognitive processes extend into the environment. In this article, I present the arguments that gave rise to the hypothesis, as well as some criticisms. In the first section I present the active externalism, the version that initially provoked the debate. In the second section, I present two criticisms against the hypothesis. Finally, in the third section, I argue that the resolution of the debate is an empirical question, and I present some evidence in ecological psychology that favors the existence of extended cognitive systems.

KEYWORDS: Extended Mind; Active Externalism; Ecological Psychology.

28 Mestre em Filosofia pela Universidade Estadual Paulista (UNESP-Marília).

Introdução

A hipótese da mente estendida foi inicialmente proposta por David Chalmers e Andy Clark no artigo “The Extended Mind” (1998). A ideia básica é que na execução de algumas tarefas, processos e objetos externos podem constituir um processo cognitivo. A tese um tanto incomum parte da ideia que o uso recorrente de determinadas ferramentas nos autorizam a afirmar que elas fazem parte da mente, por vezes. Se considerarmos o cérebro como âmbito da cognição -- pressuposto compartilhado pela maioria dos cientistas cognitivos -- iremos recusar a possibilidade da mente “fora da cabeça”. Independente da estranheza, a hipótese se tornou motivo de controvérsia²⁹. Ainda que nem todos os autores que possuem algum alinhamento à hipótese, penso que há evidências fortes a seu favor.

No entanto, há inúmeras críticas dedicam atenção a hipótese, com intuito de falsificá-la. Supostamente, haveria propriedades genuinamente cognitivas que os processos internos possuiriam que estes externos não. Acredito que este não seja o caso, mas antes de expor as razões para tal, vejamos os argumentos de Clark e Chalmers.

1. Externismo Ativo

A hipótese da mente estendida, afirma que os processos cognitivos não estão contidos somente “na cabeça”, mas dependem substancialmente do uso de artefatos externos para o desenvolvimento de algumas tarefas. As expressões “mente estendida” e “cognição estendida”, “externismo de veículo” são encontradas na literatura, mas como é esperado de filósofos, não equivalem. A versão dos autores chama-se externismo ativo e possui a seguinte estrutura:

- 1 – Todos os componentes do sistema (internos e externos) desempenham um papel causal ativo;
- 2 – Em conjunto eles governam o comportamento do mesmo modo como a cognição usualmente faz;

²⁹ Em uma coletânea dedicada à hipótese (Menary 2010) pode-se encontrar inúmeros artigos, seja em apoio ou contra. Há também nesta coletânea o artigo original que provocou o debate.

3 – Se removermos o componente externo, a competência geral do sistema irá ser reduzida, da mesma forma que ocorreria se removêssemos uma parte do cérebro;

4 – Logo, este tipo de processo acoplado equivale a um processo cognitivo, esteja ou não em sua totalidade na cabeça (Clark 1998, p. 10).

O termo “ativo” é usado em contraponto às outras versões de externismo defendidas por Putnam (1975) e Burge (1979), a qual supostamente são “passivas”³⁰: “Nestes casos, as propriedades externas são passivas. Devido a sua natureza distal, elas não atuam no direcionamento dos processos cognitivos aqui-e-agora.” . (Clark, 1998, p. 8). Outro aspecto presente na definição diz respeito a performance, ou melhor, a redução em caso da retirada de um componente. Alguém realizando um cálculo com uma calculadora tende a ser mais eficiente e menos sujeito a erros do que alguém sem uma. Da mesma forma, ao elaborar um texto, alguém pode utilizar cadernos, notas e notebooks sem os quais o desenvolvimento do texto seria substancialmente comprometido. Se considerarmos a hipótese nestes termos, não haverá muita novidade, pois são apenas descrições diferentes do uso de diversas ferramentas tecnológicas. Contudo, segundo os autores, não se trata de uma disputa verbal, mas sim uma mudança metodológica substancial:

“Ademais, ao ver a cognição como estendida não estamos tomando uma decisão terminológica; isto implica uma mudança na metodologia científica. Com efeito, métodos de explicação que podem ser pensados como apropriado somente para a análise de processos “internos” são agora adaptados para o estudo de processos externos.” (Clark 1998, p. 5).

Compreender ambos ambientes, interno e externo, como constituindo um sistema único requer que, inicialmente, admitamos que as propriedades dos dois âmbitos nos autorizem a considerar ambas as partes como pertencentes a um mesmo sistema, ou seja, uma descrição científica que abarque sistemas desta natureza. Na versão anteriormente exposta dos autores, é a similaridade funcional.

30 O externismo semântico e epistêmico (anti-individualismo) não versam sobre a constituição dos processos cognitivos, mas acerca da dependência do significado e do conhecimento de fatores externos, respectivamente (Leclerc 2014).

Além do uso recorrente dos objetos, há outros dois aspectos que os autores acreditam estarem a favor da hipótese: portabilidade e evolução:

O cérebro (ou o cérebro e o corpo) compõe um pacote de recursos cognitivos básicos e portáteis. Estes recursos podem incorporar ações corporais em processos cognitivos, quando utilizamos os dedos como memória de trabalho ao realizar um cálculo, porém eles não compõem aspectos mais contingentes do ambiente externo, como uma calculadora de bolso. Ainda assim, a contingência da relação não elimina o status cognitivo conferido (Clark, 1998, p. 9)

Nesta acepção, aqueles processos que importam para um defensor do externismo ativo são processo confiáveis. Os mais confiáveis são certamente presentes no cérebro, que evolutivamente compõem o mesmo órgão, mas que resultam de problemas adaptativos distintos. Se utilizarmos algo como um livro ou notebook para memorizar ou lembrar algo, ou como na afirmação de Feynman “pensar no papel” ao realizar uma dedução incrivelmente exigente, este uso precisa ser confiável, no sentido de estar disponível constantemente. Se pensarmos nas falhas recorrentes de memória, ou a facilidade com que informações são esquecidas, percebe-se que uma alta confiabilidade é requerida.

Além da portabilidade, há uma questão adicional, saber se a teoria da seleção natural favorece uma mente que se estende para o ambiente:

“Ademais, pode ser o caso que o cérebro biológico evoluiu e amadureceu de maneira a favorecer a presença do ambiente externo confiável. Certamente, nos parece que a evolução favoreceu capacidades que são especialmente feitas para parasitar o ambiente local com objetivo de reduzir a carga da memória, e mesmo para transformar a natureza dos problemas computacionais”. (Clark, 1998, p. 10)

A noção de ferramentas cognitiva pode no ajudar aqui. Há quem pense na linguagem como uma ferramenta que permitiu a execução de pensamentos com graus de abstração maiores (Dascal 2000). De acordo com o argumento, a capacidade sintática e semântica nos permite elaborar pensamentos mais complexos e conectar uma maior quantidade de itens. Contudo, ainda que reconheçamos a linguagem como ferramenta cognitiva, isto por si só não indica que objetos externos constituem processos cognitivos genuínos. Certamente, os mais eficientes para seu nicho têm maiores chances de permanecer no *pool* genético. E quem explora seu ambiente com intuito de reduzir a carga cognitiva tem uma vantagem. Disso se segue que uma mente que

parasita seu ambiente próximo com intuito de reduzir o trabalho desempenhado é em teoria mais eficiente. Todavia, para um cético, biólogo ou não, este argumento não é por si só suficientemente convincente.

O exemplo da linguagem anteriormente mencionado:

“A linguagem é novamente o exemplo central aqui, assim como são os diversos artefatos físicos e computacionais rotineiramente utilizados como extensões cognitivas por crianças em escolas e por estagiários em inúmeras profissões. Nestes casos o cérebro desenvolve uma forma de se complementar as estruturas externas, e aprende seu papel em um sistema densamente acoplado”. (Clark, 1998, p. 10)

Além destes aspectos como portabilidade e a seleção natural estarem supostamente favorecendo processos cognitivos estendidos há um experimento mental que os autores utilizam na defesa da hipótese.

1.1. Otto e Inga

De todos os passos da argumentação nenhum obteve tanta atenção quanto o experimento mental de Otto e Inga. Por ser central à argumentação penso ser importante a descrição na íntegra:

“Primeiro, considere um caso normal de crença embebida na memória. Inga escuta de um amigo que há uma exibição no Museu de Arte Moderna, e decide ir vê-lo. Ela pensa por um momento e relembra que o museu é na rua 53, e vai ao museu. Parece claro que Inga acredita que o Museu é na Rua 53, e ela acredita nisto mesmo antes de consultar sua memória. Não era previamente uma crença ocorrente, mas a maioria de nossas crenças não são.

Considere agora Otto. Otto sofre de Alzheimer e como inúmeros pacientes com esta doença, necessitam de informações disponíveis no ambiente para seguir com sua vida. Otto carrega um notebook consigo que leva a todos os lugares. Quando ele aprende algo novo, escreve no notebook. Para Otto, o notebook faz o papel desempenhado por uma memória biológica. Hoje, Otto ouve que haverá uma exibição no Museu de Arte Moderna, e decide ir ver a exibição. Ele consulta seu notebook, onde diz que o museu fica na Rua 53, o que o leva a andar até esta Rua 53 para ir ao museu.

Claramente, Otto anda até a rua 53 porque ele deseja ir ao museu e acredita que o museu se encontra na Rua 53. Assim como Inga tem

sua crença mesmo antes de consultar sua memória. Parece razoável afirmar que Otto acredita que o museu se encontra na Rua 53 mesmo antes de consultar seu notebook. Os aspectos relevantes dos casos são totalmente análogos: o notebook exerce a função que a memória exerce para Inga. A informação no notebook funciona da mesma forma que a informação que constitui crenças não-ocorrentes. O que “ocorre é que esta informação está fora da pele”. (Clark, 1998, p. 12)

Há diversos aspectos filosoficamente importantes aqui, como a função memória, o conteúdo informacional e a equivalência funcional entre os dois casos. Para os autores, a diferença estrutural não impede que a relação entre Otto e o notebook seja semelhante à de Inga. Na realidade, o funcionalismo na filosofia da mente define os estados mentais justamente como a reprodução dos padrões causais. O fato de ambos atingirem o objetivo parece ser suficiente para os autores: “Certamente, como as crenças e desejos são caracterizados por seu papel explicativo, os casos de Otto e Inga parecem similares: a dinâmica causal dos dois casos espelham um ao outro precisamente” (Clark, 1998, p. 11). Seguindo o raciocínio dos autores, as diferenças existentes entre os dois casos seriam superficiais se considerarmos que a crença de que o Museu está na rua 53 é mantida tanto por Inga como por Otto. Ao exercer o papel de guiar tanto Inga como Otto ao seu destino final, mesmo que a relação informacional de Otto com o notebook seja menos confiável, inconstante e sujeita a falhas, isto por si só não impede que façamos a equivalência: “As pequenas diferenças entre os casos de Otto e Inga, são diferenças superficiais. Focar nas diferenças é desconsiderá-las o modo como o notebook desempenha o papel que a crença o faz na vida das pessoas.” (Clark, 1998, p. 11).

Um filósofo não muito simpático a hipótese pode argumentar que estas diferenças são crucialmente importantes, e traçar um limite entre os dois casos. Em última instância, o pressuposto largamente difundido do cérebro como órgão central da cognição serve de base tanto para sustentar a tese de todos os aspectos centrais da cognição residem nesse órgão, como para recusar propostas de explicação que não o fazem; nesta linha de defesa, recusar a existência de processos cognitivos estendidos.

Outro aspecto da discussão acerca do experimento é processamento informacional. O conceito de informação é bastante controverso, com inúmeras definições (Floridi, 2011). Por

vezes, utiliza-se processamento de informações e computação como indistintos. Contudo, há que se especificar corretamente casos onde esta equivalência não pode ser feita (Picinnini, 2010)³¹.

1.2. Princípio da Paridade e Causação Recíproca Contínua

Seguindo a argumentação, há uma noção que novamente se sustenta em uma definição funcional, o Princípio da Paridade:

se quando confrontado com alguma tarefa, um sistema do mundo funciona como um processo que, se ocorresse na cabeça, não hesitaremos em reconhecer como um processo cognitivo, então defendemos que esse sistema do mundo é parte do processo cognitivo (Clark 1998, p. 29).

Se um braço mecânico ou *chip* implantado executa funções de maneira similar a uma região neural ou a um membro amputado, então há uma equivalência em algum grau. Pode-se argumentar que estas partes não constituem o sistema cognitivo “central”, ou são auxiliares. Em certo sentido, pode-se admitir um espaço para gradação, talvez quando nos referimos ao uso de novas tecnologias. Quando, por exemplo, um macaco controla o cursor em uma tela de computador, transmitindo sinais elétricos que são interpretados por computador (Nicolelis, 2011).

Além do PP há outra noção. Clark avança outra noção em suporte da hipótese:

A causação recíproca contínua (CRC)) ocorre quando algum sistema *S* simultaneamente afeta e é afetado pela atividade de outro sistema *O*. Internamente, nós podemos confrontar a complexidade causal do cérebro sem que muitas áreas neurais estejam conectadas por relações de *feedback* e *feedforward*. [...] Em uma perspectiva mais abrangente, frequentemente encontramos processos de CRC que transpassam cérebro, corpo e ambiente local. Pense em uma dançarina, cuja orientação corporal continuamente afeta e é afetada pela do seu parceiro, ao qual está continuamente respondendo. (Clark, 2008, p. 24)

A ideia é destacar a interdependência causal entre os elementos: ainda que seja possível identificar elementos internos e externos, a separação deles e o funcionamento só são

31 Os conceitos não são equivalentes por haver casos de processamento de informações que não executam computações, a saber, no processamento de informações analógicas.

compreendidos a partir da integração das diversas partes do sistema. Os dois conceitos utilizados *feedback* e *feedforward*, bastante comuns em biologia. O *feedback* ocorre quando a resposta ou output de um processo traz consigo informações que são enviadas para o restante do sistema. O *feedforward* consiste em repassar determinadas informações para outras partes do sistema.

Este acoplamento causal entre o corpo humano e objetos externos pode ser entendido de duas maneiras: uma influência simétrica ou assimétrica (Menary, 2010). Na influência assimétrica: características do ambiente possuem uma influência causal nos processos internos, mas o mesmo não ocorre inversamente. No caso da influência simétrica, a influência causal entre fatores internos e externos é mútua. Neste sentido, não é somente o uso de um caderno ou *notebook* que permite à pessoa acessar informações que lhe interessam, mas sim o uso recorrente destes artefatos que lhe permite afirmar que seu comportamento futuro será influenciado de maneira significativa por tal uso.

2. A Falácia da Constituição

A objeção conhecida como falácia da constituição argumenta que do fato de algo causar um processo cognitivo, não se segue que constitua um (Adams & Aizawa, 2001). Esta distinção é utilizada em suporte de uma resposta ‘internista’, que busca preservar a ideia do cérebro como âmago da cognição. A distinção teórica não se aplica em todos os âmbitos, primeiramente, porque o próprio conceito de causação é sujeito a um debate filosófico profundo³². A questão que resulta da objeção é: os conceitos de causação e constituição equivalem em uma descrição científica?

Em um artigo dedicado somente a este aspecto do debate, J. Ladyman avalia se a distinção entre causação e constituição é real nas ciências especiais. Após considerações metodológicas acerca dos fenômenos presentes na física, conclui-se que esta distinção não é real (Ladyman 2007). Para ele, no nível mais fundamental da realidade, descrito pela física, a distinção não existe, ou seja, do fato de um evento causar outro implica que os dois fazem parte do mesmo sistema. Apesar de seu argumento ser sustentado por um alinhamento a favor da hipótese da

32 Para as diversas teorias que versam sobre o conceito de causação, ver (Beebe, 2009)

mente estendida, penso ser importante considerar este ponto em mais detalhe. A ideia básica de Ladyman não é que a mente esteja 'fora da cabeça', mas sim que:

Mais especificamente, argumentamos que: (a) a distinção intuitiva entre causa e constituição, para a qual Adams e Aizawa apelam, tende a ser abandonada à medida que as ciências avançam em direção a modelos mais robustos de uma estrutura geral; (b) a noção metafísica de constituição ou composição é uma abstração que não corresponde a nenhuma ideia geral que figure metaforicamente em ciência; (c) a noção de causação, ainda que seja relevante para a ciência, não é aplicável à física fundamental, o que levanta sérias dúvidas sobre seu papel explanatório nas restrições e unificação dos modelos. (Ladyman, 2010, p. 157)

Se considerarmos a filosofia da física, especialidade de Ladyman, certamente sua colocação é válida. Todavia, não penso que transferir esta análise para outros âmbitos seja algo simples de fazer; principalmente, quando se busca descrever processos cognitivos que são modelados utilizando descrições científicas conflitantes.

Transpor análises e noções da física para outras ciências não é algo simples. Se a distinção teórica entre causação e constituição é relevante, ela é relevante para os fenômenos que buscamos explicar, fenômenos ainda mal compreendidos e que não dependem diretamente de uma descrição física subjacente, ou seja, não envolvem necessariamente uma explicação no nível mais fundamental da realidade. Se pensarmos que a explicação mais precisa do mundo é feita em física, isto por si só não garante que distinções teóricas feitas em um âmbito possam ser feitas em outro³³.

Outra razão para Ladyman atentar para este debate é sua posição claramente naturalista, onde argumentos *a priori* tem uma força menor do que aqueles empiricamente informados, pois “considerações metafísicas não devem ter nenhuma participação ao se decidir como modelar a cognição” (Ross & Ladyman, 2010, p 156)³⁴. Novamente, depende de qual aspecto da realidade estamos a falar. A metafísica é uma das áreas centrais da filosofia, e argumentos *a priori* são

33 O realismo estrutural defendido por Ladyman, argumenta que no nível fundamental da realidade não há objetos, mas somente relações sem os relata. Contudo, a ideia de relações sem relata é bastante estranha, além de não ser consensual entre filósofos da física. Esta me parece uma das motivações para ele recusar a distinção.

34 Em seu livro *Every Thing Must Go* (2007), James Ladyman e Don Ross sustentam que a metafísica do modo como é estudada hoje está novamente envolvida em confusões conceituais e falsas explicações, devido a um distanciamento dos resultados mais recentes na pesquisa científica.

bastante comuns. Inicialmente, porque há questões que não podem ser solucionadas por meios empíricos somente e exigem uma análise filosófica substancial, considerações acerca do emprego da linguagem, experimento mentais diversos e quando necessário, informações resultantes de dados empíricos.

As especificidades de cada ciência, noções teóricas aplicáveis aos seus fenômenos específicos, e aspectos metodológicos particulares levam a outro debate, acerca da unidade da ciência. A ideia de uma unidade da ciência é uma questão em aberto. De um lado, autores como Ladyman irão sustentar que apesar de empregarem métodos diferentes, todos os fatos existentes podem ser reduzidos a fatos físicos. Todavia, esta tese está longe de ser consensual, mesmo no nível pragmático. Se observarmos a ciência como é desenvolvida hoje, trabalha-se majoritariamente como modelos simplificados da realidade, justamente porque uma previsão de fenômenos com um grande número de variáveis é, na maior parte das vezes, algo bastante difícil e computacionalmente dispendioso. Neste sentido, pode até ser o caso que se tivéssemos um poder computacional infinito e uma quantidade de tempo ilimitada, ser possível descrever a realidade a partir de uma teoria física fundamental. Contudo, esta tarefa está ancorada em um pressuposto que não é consensual.

3. Objeção da Trivialidade

Alguém que não esteja inclinado a aceitar a hipótese pode responder de duas maneiras: (i) que o externismo ativo não prova o que pretende provar, e que podemos manter a ideia da mente como sendo constituída somente pelo cérebro ou (ii) argumentar que o externismo ativo prova demais, ou seja, não estabelece critérios claros que separa a mente do mundo. Seguindo a segunda rota, Rupert (2004) argumenta que se corre o sério risco de trivializar os processos cognitivos:

Minha estratégia é focar um tipo específico de estado cognitivo, a memória, e a partir daí seguir uma discussão com dois objetivos: argumentar que as porções externas da “memória” estendida diferem largamente da memória interna (os processos de rememoração) e devem ser tratados como tipos diferentes; isto elimina qualquer tentativa de argumentar em favor da hipótese da cognição estendida (HEC) a partir de uma analogia bruta

(nomeadamente, estados cognitivos estendidos são semelhantes aos internos; portanto, possuem a mesma força explicativa e, por conseguinte são estados cognitivos estendidos)” (Rupert, 2004, p. 407)

Ao estender os processos internos, segundo o autor, iremos acabar tornando-os triviais, não diferenciando assim o que é cognitivo do que não é. Esta objeção se assemelha em muito às objeções feitas contra o funcionalismo. Sucintamente, o funcionalismo é a tese de que os estados mentais são individuados pelos padrões causais que apresentam. Inicialmente proposto por Putnam (1967), o funcionalismo se tornou uma teoria em filosofia da mente com muitos adeptos. A razão central por trás da definição era incluir máquinas como possivelmente inteligentes. A analogia com computadores, ou mais especificamente, com a Máquina de Turing, se revelou interessante para os matematicamente inclinados. Porém, abriu espaço para críticas como a feita em Block (1980). A objeção de Block leva esta definição funcional ao limite e diz que se pudéssemos reproduzir estes padrões causais em outra escala, também poderíamos chamar este padrão de um estado mental. Sendo assim, se colocássemos toda a população da China instanciando estes padrões causais, então estaríamos autorizados a afirmar que as pessoas formam um estado mental. Apesar da conclusão absurda, há quem se mantenha inamovível.

Apesar do debate ter sido iniciado e debatido utilizando uma definição funcional, acredito que esta não seja a melhor forma de defender a hipótese, ainda que seja uma. Irei apresentar na próxima seção, o que acredito serem evidências empíricas que sustentam a existências de sistemas cognitivos estendidos, e apresentar uma linha paralela de argumentação que não apresenta as falhas apontadas até agora.

4. Percepção Háptica de Affordances

O trabalho que será objeto de atenção nesta seção é desenvolvido por psicólogo W. Wagman, em especial, acerca da percepção háptica de affordances. Sucintamente, affordances são possibilidades de ação que permitem ao organismo se comportar de uma maneira ou de outra, dependendo da informação obtida do ambiente. Em geral, o nicho de atuação no qual o organismo se encontra define qual affordance está sendo ativada. O exemplo típico da cadeira ilustra bem. Uma cadeira permite a um ser humano a affordance “sentável” enquanto para uma

formiga pode fornecer algo como “subível”. Em psicologia ecológica, o conceito é de extrema importância.

O estudo que irei apontar aqui não é exatamente do agente no mundo, mas de sistemas agente+objeto, ou seja, de como sistemas desta natureza se comportam. Neste sentido, a percepção háptica consiste no estudo dos agentes quando estão utilizando objetos ou alguma espécie de carga ao navegar pelo mundo. O intuito do estudo é compreender como nos comportamos quando estamos com objetos acoplados, por exemplo, carregando um bastão ou uma mochila. De acordo com a proposta aqui defendida, os objetos alteram as capacidades perceptivas do agente, indicando a presença de sistemas cognitivos estendidos.

O primeiro estudo (Wagman 2005) realiza três experimentos com pessoas carregando um objeto em forma de “T” de diversos tamanhos (50-140 cm). Na frente da pessoa há uma abertura de cerca de 90 cm de largura a qual os agentes terão de avaliar se conseguem atravessá-la com os objetos em forma de “T”. No Experimento 1 há duas situações: (i) na primeira a pessoa carrega o objeto próximo ao seu corpo e observa a abertura em sua frente; (ii) na segunda situação a pessoa apenas observa o objeto a sua frente em um balcão e compara com a abertura. Nos dois casos, duas questões são feitas: se eles conseguirão atravessar a abertura com o objeto, e qual a sua confiança na resposta (onde o número 1 corresponde a um grau alto de incerteza e 7 para alto grau de certeza). Após avaliar os diversos indivíduos o psicólogo compara as respostas entre as duas situações.

Neste primeiro experimento, o que se verificou foi que as pessoas conseguem identificar com grande acurácia a largura necessária para quem transpassar a abertura. Além disso, o grau de confiança nas respostas variou de acordo com a largura dos objetos. À medida que a largura do objeto se aproxima da largura da abertura, as pessoas vão ficando mais incertas acerca da possibilidade. Neste sentido, quando a largura do objeto é muito maior que a abertura, há uma alta confiança de que a pessoa+objeto não transpassará a abertura, e a confiança vai diminuindo a medida que o tamanho do objeto também diminui. O segundo ponto a ser destacado neste primeiro experimento, é que as pessoas não precisam tocar nos objetos para chutar corretamente se conseguirão passar ou não pela abertura. Os resultados nas duas situações, vendo os objetos apenas e carregando-o, são bastante próximos, especialmente nos

graus de confiança, definido aqui como a maior proporção entre a largura da abertura e a largura do objeto que é percebida como passável (a proporção é de aproximadamente 1).

No mesmo estudo agora com uma montagem distinta, o qual os psicólogos chamam de Experimento 2, tem-se os mesmos objetos, porém com pesos nos braços. Neste segundo momento, mantém-se a mesma largura do objeto e variam-se os pesos; nesta variação a intenção é saber se a presença dos pesos atrapalha a percepção dos agentes. Novamente, os resultados se mantêm (a proporção de aproximadamente 1); mesmo com pesos alterando o ponto de inércia do sistema pessoa+objeto, há um reconhecimento de qual tamanho permite transpassar a abertura. Na interpretação dos pesquisadores, a variação do ponto de inércia do sistema pessoa+objeto é compensada pela pessoa com intuito de melhorar a percepção da tarefa, ocorre uma adaptação.

Por último, no Experimento 3 busca-se saber se a largura do ombro da pessoa interfere no modo como a pessoa percebe a affordance. Neste caso, em vez de carregar um objeto em forma de “T”, utiliza-se extensões ao lado de cada braço, novamente, com variações na largura dos objetos. De acordo com esta construção, o objeto não está referenciado no centro do corpo, mas encontram-se ao lado. Ao montar a situação desta forma, busca-se saber como irão responder à situação, pessoas com largura dos ombros diferentes. E o resultado, novamente, indica uma capacidade de percepção acurada, apesar da variação na largura dos ombros, as pessoas conseguem inferir com grande acurácia em qual tamanho ocorrerá o limite para transpassar a abertura.

Em um estudo diferente, verificou-se que pessoas carregando mochilas pesadas por longas distâncias ou em superfícies inclinadas tem sua percepção alterada, diferente do que caso não tivessem carregando-as. No artigo de Bhalla e Proffitt (1999) os indivíduos reportam que as distâncias parecem ser maiores e a inclinação maior com mais difícil equilibrar.

Nas situações que um indivíduo carrega um objeto pesado tem-se a alteração das propriedades dinâmicas, isto é, do agente em movimento, e não apenas das propriedades geométricas. A pesquisa nesta linha chama-se de toque dinâmico: a descoberta de propriedades através do toque. Se considerarmos um livro, mesmo sem poder vê-lo, ainda podemos ter informações acerca de sua espessura, peso, largura etc. O livro exerce uma força mecânica nos dedos que conseguimos perceber e avaliar ao tocá-lo. Esta ação exploratória nos permite avaliar

inúmeras outras características: acerca da inclinação do livro em relação ao chão. Esta capacidade de avaliação é permitida pelas características motoras dos músculos, do seu braço que consegue manter o livro erguido, em uma posição de leitura, visível ou não. As ações são possíveis devido aos receptores mecânicos que são estimulados e fornecem diversas informações sobre os objetos (Gibson 1962, 1966; Turvey 1996; Carello e Turvey 2000).

Tanto no estudo que busca compreender as propriedades geométricas percebidas pelo agente como no toque dinâmico, fica claro que há uma clara relação de pertencimento que os processos cognitivos adotam ao carregar objetos ou tocá-los. Neste sentido, a hipótese não precisa necessariamente ser defendida utilizando uma descrição funcional ou computacional, ainda que muitos o façam. Em termos empíricos, há toda uma linha de pesquisa dedicada a compreender os agentes em relação ao seu ambiente, nos fornecendo uma maior compreensão da percepção e da ação.

Considerações Finais

A hipótese da mente estendida foi apresentada neste trabalho, inicialmente, como do externismo ativo. Após detalhar a definição de externismo ativo, os casos de Otto e Inga, a noção de causação recíproca contínua e o alinhamento com o funcionalismo foram detalhados, assim como algumas objeções. O pressuposto básico por trás das críticas é que há algo de especial no cérebro confinado no crânio, uma marca do mental, que processos cognitivos estendidos não possuem. Argumentei que a partir da definição mais comum da hipótese, as críticas têm força. Defendi que a resolução do debate se dá a partir de pesquisas sendo desenvolvidas em psicologia ecológica, mais especificamente, da percepção de affordances. Nesta linha, assume-se que a melhor maneira de compreender os processos cognitivos não traça esse limite de maneira estática, mas assume que dependendo da tarefa, há capacidades perceptivas que se alteram ao acoplar objetos.

Referências bibliográficas

ADAMS, F. & AIZAWA, K. (2001). The bounds of cognition. *Philosophical Psychology*, 14 (1), 43-64. New York, NY: Oxford University Press.

- _____ (2008). *The bounds of cognition*. Malden, MA: Blackwell Publishing.
- ADAMS, F., and K. Aizawa, 2009, "Why the Mind is Still in the Head," in *The Cambridge Handbook of Situated Cognition*, P. Robbins and M. Aydede (eds.), Cambridge University Press, pp.78–95.
- BEER, R. (2000), Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3), 91-99.
- _____ (2003). The dynamics of active categorial perception in an evolved model agent. *Adaptive Behavior*, 11, 209-243.
- BEEBEE, H., Hitchcock, C., and Menzies, P., eds. (2009). *Oxford Handbook of Causation*. Oxford: Oxford University Press.
- BHALLA, M., & Proffitt, D. R. (1999). Visual-motor recalibration in geographical slant perception. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1076-1096.
- BURGE, T. 1979. "Individualism and the Mental." In French, P., Uehling, T., and Wettstein, H., eds., *Midwest Studies in Philosophy* 4. Minneapolis: University of Minnesota Press, 73-121.
- BLOCK, Ned (1980), 'What is Functionalism?'. *Readings in Philosophy of Psychology*, 1. MA: Harvard University Press. pp. 171-184.
- CHEMERO, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.
- CLARK, Andy (1997), *Being There: Putting Brain, Body, and World Together Again*. Cambridge: MIT Press.
- _____. (2006). Language, embodiment and the cognitive niche. *Trends in Cognitive Sciences*, 10(8), 370-374.
- _____, Andy (2008), *Supersizing the Mind*. Oxford: OUP.
- CLARK, Andy and CHALMERS, David (1998), 'The Extended Mind'. *Analysis*, 58, 1, p. 7-19.
- DASCAL, M. (2000). Language as a cognitive technology. *International Journal of Cognition and Technology*, 1(1), 33-59.
- FARIA, Paulo. Externalismo semântico. *Compêndio em Linha de Problemas de Filosofia Analítica*. Editado por João Branquinho e Ricardo Santos, Lisboa: Centro de Filosofia da Universidade de Lisboa, 2013. Disponível em: . Acesso em: 20 maio 2015.
- FODOR, 1975, *The Language of Thought*, Cambridge, MA: Harvard University Press.
- _____, J.A. and , Z.W. Pylyshyn, 1988, "Connectionism and cognitive architecture: A critical analysis," *Cognition*, 28: 3–71.
- GIBSON, J. J.(1997). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing*(pp. 66-82). Hillsdale, NJ.
- _____.(1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.
- LECLERC, A. Externismo e conteúdo mental. *Compêndio em Linha de Problemas de Filosofia Analítica*. Editado por João Branquinho e Ricardo Santos, Lisboa: Centro de Filosofia da Universidade de Lisboa, 2014. Disponível em: . Acesso em: 20 maio 2015.

- LADYMAN, J. , and Ross , D. (2007). *Every Thing Must Go: Metaphysics Naturalised*. Oxford : Oxford University Press .
- LYCAN. W. 2001. "The Case for Phenomenal Externalism." *Philosophical Perspectives* 15, 17-35..
- MENARY, R. (2007). *Cognitive Integration: Mind and Cognition Unbounded* .Basingstoke : Palgrave Macmillan .
- MILNER, D. & Goodale, M. (1995). *The visual brain in action*. Oxford: Oxford University Press.
- PICCININI, G. & Scarantino, A. (2010) 'Information processing, Computation, and Cognition'. *Journal of Biological Physics* 37 (1):1-38.
- PUTNAM, H. (1967). 'The Nature of Mental States'. Reprinted (2008) in *Mind, Language and Reality: Philosophical Papers Vol. 2*. Cambridge: Cambridge University Press.
- ROSS, D., & Ladyman, J. (2010). The alleged coupling-constitution fallacy and the mature sciences. In R. Menary (ed.), *The Extended Mind* (pp. 155-166). Cambridge, MA: The MIT Press.
- RUPERT, Robert (2004), 'Some Problems for the Thesis of Extended Cognition.' *Journal of Philosophy*. 101. pp. 389–428.
- STICH, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, M.A.: MIT Press.
- TURVEY, M. T. (1992). Affordances and prospective control: an outline of the ontology. *Ecological Psychology*, 4, 173-187.
- _____, M. T., & Carello, C. (2011). Obtaining information by dynamic (effortful) touch. *Philosophical Transactions of the Royal Society B* , 366, 3123 – 3132.
- WAGMAN, J. B., & Taylor, K. R. (2005a). Perceiving affordances for aperture crossing for the person-plus-object system. *Ecological Psychology*, 17, 105-130.
- _____, J. B., & Taylor, K. R. (2005b). Perceived arm posture and remote haptic perception of whether an object can be stepped over. *Journal of Motor Behavior*, 37, 339-342.
- WARREN, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 371-383

Bruno Cunha³⁵

RESUMO: A teoria dos *qualia* é uma defesa do estado *sui generis* dos fenômenos mentais ou da consciência. Em acepção geral, os *qualia* representam o caráter subjetivo da experiência mental ou, mais particularmente, uma qualidade fenomenológica subjetiva da consciência impassível de redução. Meu objetivo nessa comunicação é demonstrar como Searle se apoia na teoria dos *qualia* para responder ao funcionalismo computacional de autores como Daniel Dennett e Jerry Fodor, que consideram a consciência, em termos reducionistas e sob o ponto de vista objetivo da terceira pessoa, como um fenômeno que, assim como um programa de computador, é explicável mediante estados funcionais e estímulos de entrada e saída.

PALAVRAS-CHAVE: *qualia*; subjetividade; intencionalidade; programas de computador.

ABSTRACT: Qualia theory is a defense of the *sui generis* state of mental phenomena or consciousness. In a general sense, the *qualia* represent the subjective character of mental experience or, particularly, a subjective phenomenological quality of consciousness which is impassive of any reduction. My purpose in this paper is to demonstrate how Searle bases on qualia theory to respond to the computational functionalism of authors as Daniel Dennett and Jerry Fodor, who consider the consciousness, in reductionist terms and from the objective viewpoint of the third person, as a phenomenon which, like a computer program, can be explained by functional states and input and output stimuli.

KEYWORDS: *qualia*; subjectivity; intentionality; computer programs.

³⁵Professor Adjunto, DTECH / UFSJ. Email: brunocunha@ufsj.edu.br.

Dentre as abordagens contemporâneas da mente, o *naturalismo biológico* de John Searle é conhecido por ter contribuído em grande medida para a redescoberta da consciência. Diante das diversas tendências fisicalistas que, nesse contexto, buscaram reduzir toda espécie de fenômeno ao modo de ser objeto, Searle propõe uma compreensão *ontológica* especial do âmbito do mental na qual a consciência é concebida em termos de estados dotados de características peculiares tais como a intencionalidade e a subjetividade. Apesar de suas críticas ao autor, ao defender a irredutibilidade da consciência, Searle direciona-se por um caminho similar ao que Thomas Nagel já havia antes proposto: a mente não é analisável em termos de nenhum sistema explicativo de estados funcionais e só pode ser compreendida a partir de um ponto de vista muito particular, uma vez que seu conteúdo subjetivo é altamente específico. A consciência, portanto, só pode ser entendida como um fenômeno de primeira pessoa, já que se constitui como uma experiência qualitativa interna específica (*qualia*). Meu objetivo nessa comunicação é, em vista desse problema, demonstrar como Searle se apoia na teoria dos *qualia* para responder ao funcionalismo computacional de autores como Daniel Dennett e Jerry Fodor, que consideram a consciência, em termos reducionistas e sob o ponto de vista objetivo da terceira pessoa, como um fenômeno explicável mediante estados funcionais e estímulos de entrada e saída.

Em seu artigo *O problema mente-corpo*, publicado em 1981, Jerry Fodor apresenta a abordagem *funcionalista* como decorrente dos avanços nos estudos sobre inteligência artificial, teoria computacional, linguística, cibernética e psicologia, situando-a em um âmbito distinto das tradicionais interpretações dualistas e materialistas da mente. Compreendido como uma tentativa filosófica de explicar o nível de abstração aferido a partir do processamento de informação, o funcionalismo interpreta os processos mentais em termos de estados funcionais. Ou seja, compreende a mente como um sistema formal de manipulação de símbolos relacionados mutuamente de modo a executar, em determinadas condições, funções específicas. Assim um estado funcional é definido pela estrita relação causal entre estímulos de entrada (*input*) - estes compreendidos como certos estados causais internos - e estímulos de saída (*output*) - seus efeitos ou comportamentos decorrentes. Ao conceber a mente nesses termos, o funcionalismo vai reconhecer, com efeito, a possibilidade de que sistemas tão distintos como os seres humanos, as máquinas de calcular e os espíritos desencarnados possam, de algum modo, ter estados mentais, já que, segundo Fodor, a psicologia de um sistema não deve depender da matéria da qual ele é

formado, mas tão somente do modo como ela é arranjada (1981, p.124). Já que o estado funcional nunca depende da estrutura física do sistema, pode-se dizer, do mesmo modo, que o caráter do estado mental nunca se refere à estrutura física (1981, p. 130 a).

É possível compreender, a partir disso, mesmo que em linhas gerais, a proposta do funcionalismo computacional e das diversas teorias de inteligência artificial. Uma máquina computacional existe exclusivamente com o propósito de desempenhar funções. Nesse processo, os computadores decodificam certos *inputs*, que são comandos específicos de entrada, transformando-os, por conseguinte, de acordo com um conjunto estabelecido de regras formais, em respostas ou *outputs* que podem ser expressos em certos comportamentos. Sendo assim, pode-se dizer, em analogia, que da mesma forma que o estado funcional de uma máquina computacional pode ser descrito exhaustivamente a partir da relação entre *inputs* e *outputs*, como quando uma máquina de *Coca-Cola* entrega o refrigerante e subsequentemente o troco ao receber duas moedas (Fodor, 2011, p.129 ab), o comportamento mental dos seres humanos pode ser concebido, da mesma maneira, partindo da mesma relação. Mas os *inputs* se apresentam, no caso dos seres humanos, como determinados estímulos sensíveis, sejam visuais, auditivos ou outros, enquanto os *outputs* decorrentes encontram expressão em certos tipos de comportamentos, como quando alguém sente uma dor de cabeça e apresenta um conjunto de reações comportamentais dentre as quais está a atitude de tomar um remédio. Portanto, se os estados mentais são realmente comparáveis a estados funcionais e os papéis funcionais sempre dependem de estímulos de entrada e saída, nada impede que o funcionalista proponha uma analogia estrita entre a mente humana e um software, alegando que “o caráter relacional do mental” está sempre preservado em uma máquina computacional (2011, p.130 a). Segundo as palavras de Fodor (2011, p.129 c), “[u]ma vez que o papel funcional de um estado depende de sua relação com outros estados e também com as entradas e saídas, o caráter relacional do mental fica preservado na versão da máquina-de-Turing do funcionalismo”.

A grossa modo, dentre as vantagens dessa abordagem do mental, de acordo com o próprio manifesto de Fodor, é que ela se adéqua melhor às exigências fisicalistas sem ser reducionista (2011, p.128 a)³⁶. Ao definir as atividades mentais em termos de suas relações e, portanto,

³⁶Como destaca Maslim (2009, p.142), o funcionalismo pode ser reducionista na medida em que elimina os conceitos mentais.

enquanto estado funcional, o funcionalismo não renuncia às exigências do materialismo ao mesmo tempo em que proporciona uma resposta para a complexidade do mental mais elaborada do que a proposta pelo behaviorismo lógico a partir de suas disposições comportamentais (2011, 129 b). Além disso, a hipótese dos estados funcionais parece lidar satisfatoriamente com o problema da causação mente-corpo ao compreender a causação mental com um tipo especial de causação física. Nesse sentido, a dor, por exemplo, não é identificada de antemão com um conteúdo mental causada por um evento físico. Embora o comportamento de dor seja causado por eventos no sistema nervoso, por outro lado, os eventos isolados não podem ser estritamente identificados com a dor. De outro modo, a dor deve ser vista como um estado completo em que se considera a relação completa entre *inputs* e *outputs* (p.128 a). Segundo a interpretação funcionalista, o estar com dor não seria, dessa forma, um evento mental em específico, mas um conjunto de estados funcionais em próxima relação e considerados holisticamente.

Apesar do funcionalismo ser apresentado por Fodor, em seu texto, como uma teoria emergente capaz de capturar as melhores características das abordagens materialistas (2011, p. 129 a) , o autor certamente reconhece a força daquilo que considera a principal objeção diante de sua teoria. Da pergunta sobre o que caracteriza a mente decorre a caracterização dos estados mentais como dotados de conteúdo *qualitativo* e *intencional* (2011, p.130 b). " A razão pela qual o conteúdo qualitativo representa um problema para o funcionalismo é evidente. O funcionalismo está comprometido com a definição dos estados mentais em termos de suas causas e seus efeitos". Não obstante, como Fodor reconhece frente a sua teoria, parece ser possível "que dois estados mentais tenham as mesmas relações causais, diferindo, porém, em seu estado qualitativo". Isto é, se for possível que dois observadores distintos apresentem os mesmos papéis e conexões causais, mas experiências qualitativas internas distintas, ou seja, se puderem ter conteúdos qualitativos distintos da mesma experiência como no exemplo do espectro tivesse invertido no qual a experiência de alguém relacionada ao vermelho, é equivalente à experiência do verde em outro observador, isso seria um verdadeiro problema. "Se isto for possível, então a abordagem funcionalista não funciona para os estados mentais que possuem conteúdo qualitativo" (FODOR, 2011, 130b).

Sobre o conteúdo qualitativo, Thomas Nagel havia escrito, em 1974, um importante artigo cujo impacto resvalou fortemente na comunidade científica. Em sua tentativa de ressaltar a

intratabilidade da consciência, Nagel foi de encontro ao aspecto mais característico e *sui generis* dos fenômenos mentais ou da consciência. Trata-se do caráter subjetivo da experiência mental que, segundo ele, não foi capturado e compreendido adequadamente por quaisquer das recentes análises redutivas do mental. Algo que, segundo ele, escapa, da mesma forma, a toda análise explicativa em termos funcionais e intencionais que poderiam ser atribuídos a autômatos e robôs (2005, p.246). Isto decerto não é negar, de acordo com Nagel, a possibilidade de qualquer explicação funcional de eventos mentais, mas é recusar que tais fenômenos se esgotem nessas explicações. Pois, no que diz respeito à consciência, não é possível excluir os elementos fenomênicos subjetivos da experiência e a razão disso, como explica Nagel, é porque todo fenômeno subjetivo está vinculado necessariamente a um ponto de vista singular e privado (2005, p.248) que dificilmente pode ser enquadrado no paradigma de explicação objetivo de uma teoria física. Para exemplificar essa estrita relação entre subjetividade e ponto de vista e, ao mesmo tempo, a disparidade entre subjetivo e objetivo, Nagel propõe seu famoso experimento mental hipotético descrito na suposição de “como é ser um morcego”. Os morcegos são dotados de um aparato sensorial bastante distinto que os permite reconhecer o mundo externo primariamente através de um sonar. Por meio deste aparato, eles detectam os ecos de seus próprios gritos emitidos em alta frequência e ao correlacionar os impulsos enviados com os ecos subsequentes, eles adquirem informações precisas acerca de distância, tamanho, forma, movimento e textura de objetos e coisas (2005, p.249). Esse tipo de percepção não é, em absoluto, similar a nenhum tipo de percepção ou sensação dos seres humanos. A questão que fica é se podemos realmente ter, mesmo em posse de uma descrição física do fenômeno em questão, uma noção verdadeira de como é estar na pele de um morcego. Parece que o máximo que podemos ter a esse respeito é uma concepção esquemática de como é ser um morcego. Se, por um lado, é possível compreender alguns dos aspectos dessa experiência por analogia, por outro, é preciso admitir que o seu caráter qualitativo subjetivo está além de nossa capacidade de concepção. Visto, portanto, que o conteúdo subjetivo da consciência é altamente específico e que há limites, talvez intransponíveis, em nosso sistema de representação, Nagel esvazia nossas expectativas de alcançar, pelo artifício da linguagem humana, uma *descrição* detalhada da fenomenologia subjetiva dos morcegos ou, até mesmo, dos marcianos, o que não significa necessariamente afirmar que tanto morcegos, quanto marcianos, não devam ter experiência subjetivas tão ricas quanto as nossas (2005, p. 251). Em suma, isso quer dizer, em outras

palavras, que embora seja possível fazer uma descrição em terceira pessoa de certas experiências da consciência, a especificidade da experiência fenomenológica de primeira pessoa suplanta essa possibilidade, mostrando-se, portanto, incompreensível por meio de uma linguagem estritamente objetiva.

O artigo de Nagel apresenta-se, sem dúvida, como uma defesa substancial da teoria dos *qualia*. Embora, em filosofia da mente, o termo *qualia* tenha sido empregado, não de maneira estrita, em vários sentidos diferentes. Aqui compreendemos *qualia*, em acepção geral, como uma qualidade fenomenológica subjetiva da consciência impassível de redução. Uma das abordagens da mente que, fazendo frente ao funcionalismo computacional, tem se ocupado, em especial, com o problema da consciência, tomando como referência a teoria dos *qualia*, é o naturalismo biológico de John Searle. Em seu livro de 1992, *Redescoberta da Mente*, Searle acredita estar indo além de Nagel ao compreender este aspecto da subjetividade não apenas partindo de um ponto de vista epistêmico, em relação ao nosso modo e capacidade de concepção, mas, sobretudo, a partir de um ponto de vista ontológico (1997, p.139). Dessa forma, para Searle, a subjetividade não seria tanto uma questão de adequação de juízos subjetivos aos fatos, mas um modo subjetivo de *existência* atribuído ao próprio fenômeno (1997, p.169). Como consequência adicional, este modo de subjetividade traz consigo formas conscientes de *intencionalidade* que devem ser concebidas sob um ponto de vista bastante especial (p. 1997, 140). Como se pode notar, nessa premissa estão apresentados, sobretudo, os dois aspectos que Jerry Fodor já havia apontado, em seu texto, como objeções reais contra a teoria funcionalista. O argumento de Searle contra o funcionalismo, desenvolvido desde 1984 em seu livro *Mente, Cérebro e Ciência*, no entanto, vai se dirigir especialmente e diretamente a outro adversário, Daniel Dennett. Em um livro publicado em 1997, *Mistério da Consciência*, Searle faz uma manifesta defesa da teoria dos *qualia* frente ao funcionalismo ou, usando os termos do próprio autor, da teoria da IA forte defendida por Dennett em seu livro *Consciência explicada*. O ponto de partida da objeção contra Dennett parte da ideia, já anteposta pelo artigo de Nagel, de que a dor, sentida por exemplo em um beliscão, é uma sensação desagradável associada a um tipo de impressão subjetiva de acesso inteiramente particular. Embora essa acessibilidade tenha certamente consequências epistêmicas, já que nenhum expectador é capaz de *conhecer* a minha sensação, a impressão subjetiva em questão, segundo Searle, tem, ademais, um caráter especialmente ontológico (1998, p.118). Isto é, enquanto um fenômeno de primeira pessoa, ela extrapola o campo da descrição para existir

como um sentimento qualitativo interno específico. Dessa forma, uma teoria da consciência em sentido amplo, segundo Searle, deve ser capaz de explicar não apenas o processo que ocorre em nossas trilhas neuronais em um evento mental específico como por exemplo a dor, mas precisa lidar satisfatoriamente com esses sentimentos e dados subjetivos. O cerne da crítica de Searle a Dennett encontra-se no fato de que o último nega esses dados. Ao defender que os processos mentais ocorrem a partir de uma relação entre inputs de estímulo, estados discriminatórios e disposições reativas, Dennett recusa a existência dos *qualia*, das experiências subjetivas e dos fenômenos de primeira pessoa. (1998, p.119). Ele vai ainda mais longe do que as teorias funcionalistas convencionais ao negar *completamente* os estados mentais. Ao recusar a existência dos dados conscientes, Dennett apresenta uma explicação alternativa, articulando noções como máquinas von Neumann, connexionismo, máquinas virtuais e memes que, para Searle, não é mais do que uma nova versão da Ia forte.

Um computador digital, conhecido como máquina von Neumann, opera realizando uma série de passos realizados em uma pequena fração de segundos. Esses computadores são capazes de operar em paralelo ou em rede, interagindo com outros computadores. Esta ideia é, à grosso modo, o que chamamos de connexionismo. Qualquer computação capaz de ser realizada em uma estrutura de conexão ou arquitetura possui algumas propriedades peculiares, a saber, operam mais rapidamente e podem aprender. A rede pode supostamente aprender porque é capaz de mudar seu comportamento de acordo com certas situações e isso acontece porque é possível obter o emparelhamento correto entre valores de *input* e *output* que são ajustados por um computador principal. Um sistema computacional operando nesses termos - como Dennet supõe - se assemelharia aos cérebros na medida em que as conexões poderiam ser pensadas como algo semelhante aos axônios e dendritos. Como no funcionalismo tradicional, essa computação não seria dependente tanto da estrutura física do sistema, mas de um padrão regular de funções. É nesse sentido que se insere a noção de máquina virtual. Por fim, essas conexões estariam ocupadas com a transmissão de memes, noção que, em analogia aos genes da biologia, é entendida como uma unidade de transmissão cultural ou uma unidade de imitação. Como exemplo de memes, temos melodias, ideias, roupas, modas etc. Considerando esses aspectos, segundo Searle, Dennet compreende, de maneira conclusiva, que a consciência é uma imensa coleção de memes transmitida por meio da operação de uma máquina virtual implementada na arquitetura paralela de um cérebro (1998, p.125).

Dessa forma, de acordo com a interpretação searleana, o estar consciente, na perspectiva de Dennett, é simplesmente a expressão da programação de uma máquina paralela que evolui na natureza. Mas o antimentalismo extremo de Dennett - observa Searle - vai se deparar com sérias dificuldades. Fazendo alusão a um argumento apresentado por Dennis Chalmers, a teoria de Dennett não se mostra capaz de distinguir entre seres humanos e zumbis inconscientes que se comportam tão complexamente quanto os homens. A sua objeção maior, entretanto, baseia-se certamente em seu anterior argumento do quarto chinês. Embora Dennett tenha tentado desqualificar este argumento alegando sua inexequibilidade, Searle acredita que sua força argumentativa se mantém intacta: a sintaxe de um programa não é suficiente para garantir o seu conteúdo semântico. O experimento mental hipotético apresentado, primeiramente, por Searle em *Mente, Cérebro, e Ciência* propõe que imaginemos um programa capacitado a compreender o chinês. Ao ser confrontado com uma questão, o computador consultará sua complexa base de dados, estando habilitado a produzir respostas adequadas. Mas a questão que se coloca por meio do experimento é: pode o computador compreender o chinês como um nativo? Suponhamos agora que nós, mesmo sem compreender uma palavra em chinês, estivéssemos dentro de um quarto no qual há cestos com vários símbolos em chinês. Ao mesmo tempo, temos um manual em nossa língua com instruções precisas que nos permite ordenar os símbolos de determinada maneira. Alguns símbolos novos são introduzidos no quarto (inputs) e somos instruídos, de acordo com algumas regras, a passar alguns símbolos para o exterior do quarto (outputs) em determinada ordem. Sem que saibamos, os símbolos novos introduzidos são perguntas de pessoas que estão fora do quarto enquanto os símbolos mandados para fora, de acordo com as instruções, são respostas para essas perguntas. Suponhamos que, da mesma forma que um programa de computador avançado, esse processo de manipulação de símbolos é realizado perfeitamente por alguém dentro do quarto de tal forma que as respostas enviadas para fora sejam indistinguíveis da de um falante nativo. Em analogia com a pergunta acima, será que podemos concluir a partir disso que o manipulador de símbolo compreende Chinês? (1987, p.40). Tanto quanto os estados mentais, a compreensão de uma língua requer mais do que a manipulação puramente formal de símbolos em termos de sua sintaxe, mas exige a compreensão de seus significados. Nos processos computacionais, a sintaxe não é capaz de produzir a semântica.

Searle aceita, nesse contexto, que ao menos há uma sintaxe relativa ao programa. Ou seja, que há regras governando a manipulação de símbolos. Não obstante, posteriormente, em *Redescoberta da Mente*, Searle radicaliza sua própria perspectiva passando a não aceitar nem mesmo que haja uma sintaxe intrínseca aos eventos físicos operacionais de um computador. Ora, pois semântica e sintaxe - como ele pressupõe - são sempre relativas aos seres humanos que projetam e interpretam esses *inputs* e *outputs*. Isso exclui qualquer capacidade autônoma para mecanismo físicos, pois estados sintáticos não são descobertos dentro da física, eles devem ser vistos, de outro modo, como atribuídos à física (MASLIM, 2009, p.153).

O exemplo de Searle tem como objetivo nos mostrar que, além do aspecto qualitativo da subjetividade, um pensamento verdadeiro sempre possui intencionalidade, ou seja, ele sempre carrega a característica de se referir a algo diferente dele mesmo. Um pensamento, uma crença ou um sentimento, portanto, são muito diferentes das operações computacionais porque estas não significam nada para o sistema que as realiza. Dessa maneira, uma frase de texto proferida por um computador pode adquirir um sentido apenas em relação a mim e a meus semelhantes dotados de uma consciência capaz de interpretá-la.

Tanto quanto a subjetividade³⁷, a intencionalidade é um conteúdo mental qualitativo que caracteriza a consciência. Como parte integrante dos *qualia*, ela apresenta-se como uma objeção real à crença funcionalista, como a sustentada por Dennett, de que não há nenhuma consciência para além das propriedades computacionais. Talvez, uma boa justificativa para o reducionismo extremado de Dennett possa ser encontrada em sua convicção de que o desafio de uma teoria da mente é o de construir uma teoria dos eventos mentais compatível com o método científico e, portanto, possível exclusivamente sob ponto de vista da terceira pessoa (2008, p.131). Mas, diante disso, a principal objeção questiona até que ponto o verificacionismo científico é capaz de dar conta integralmente do funcionamento do mundo. Se, de fato, o mundo parece, pelo menos quase em sua integralidade, sempre passível de uma descrição objetiva, sob o ponto de vista da terceira pessoa, no fim das contas, isto não autoriza a ciência a excluir certos fenômenos

³⁷Embora Searle construa sua crítica ao funcionalismo computacional basicamente partindo do conteúdo intencional da consciência, o aspecto qualitativo da *subjetividade* também proporciona o suporte para várias objeções. Uma delas é inclusive apontada no texto de Fodor como os "qualia invertido", Mas podemos ainda citar, dentro da literatura, os argumentos "O que Mary não sabia", "O que Fred sabia e Mary não viu", o argumento da mente chinesa etc.

concretos da experiência humana com os quais ela é incapaz de lidar. Nem mesmo pode ser dito destes fenômenos, partindo de uma ótica científica reducionista, que tais não passam de ilusão, porque no que diz respeito à consciência, em seus aspectos qualitativos, a existência da aparência precisa ser concebida, em absoluto, como a própria realidade (2008, p. 130). Como Descartes antes já havia sugerido, parece ser verdade que a própria subjetividade deve ser assumida como uma experiência primária genuína. Portanto, torna-se possível concluir, a partir do argumento de Searle, que se estamos tratando com uma definição de ciência que nos proíbe de investigar uma parte importante do mundo, mesmo que misteriosa, parece que é esta definição que deve ser mudada, não o próprio mundo (2008, p.133).

Referências Bibliográficas

CHURCHLAND, P. M. *Matter and consciousness: a contemporary introduction to the philosophy of mind*. Cambridge, Mass. M.I.T. Press, 1984.

DENNETT, D. C. *Consciousness Explained*. Boston: Little, Brown, 1991.

FODOR, J. A. *O problema Mente e Corpo*. trad: Saulo de Freitas Araújo. Filosofia das Ciências Naturais (disciplina), São Paulo: FFLCH-USP, 2011.

NAGEL, T. Como é ser um Morcego? Trad.: Paulo Abrantes e Juliana Orione. *Cadernos de História e Filosofia da Ciência*. v.15, n.1. Campinas: Unicamp. 2005. p.245-262).

MASLIN, K. T. *Introdução à Filosofia da Mente*. São Paulo: Artmed, 2009.

SEARLE, J. R. *Mente, Cérebro e Ciência*. Trad.: Artur Morão, Lisboa, Portugal: Edições 70, 1984.

SEARLE, J. R. *A Redescoberta da Mente*. Trad.: Eduardo Pereira e Ferreira, São Paulo: Martins Fontes, 1997.

SEARLE, J. R. *O Mistério da Consciência e discussões com Daniel C. Dennett e J. Chalmers*. Trad.: André Yuji Pinheiro Uema e Vladimir Safatle. São Paulo: Paz e Terra, 1998.

Em que sentido há um problema mente-corpo?

In which sense is there a mind-body problem?

Cíntia Roso Oliveira³⁸

RESUMO: Afirmações do tipo “Essa alergia é psicológica” implicam em aceitar que estados mentais interferem nos estados físicos. O problema mente-corpo surge na medida em que também aceitamos dois pressupostos: 1. Que mente e corpo são distintos entre si; e 2. Que a “causação” implica uma produção ontológica na qual causa e efeito são semelhantes. Pois como poderia um estado mental, diferente de um estado físico, produzir mudanças físicas? O objetivo deste artigo é esclarecer em que sentido há um real problema mente-corpo. Para isso, é necessário distinguir algumas abordagens em relação à mente e à causação. Existem duas concepções referentes à causação: a ontológica e a epistêmica. Quanto ao conceito de mente, apontamos três abordagens: a eliminativista, a reducionista e a não-reducionista. Inicialmente, iremos apresentar a abordagem eliminativista de Churchland; em seguida, vamos apresentar a abordagem não-reducionista de Chalmers; e, por fim, vamos expor a abordagem reducionista de Dennett. Se compreendermos que causação é um princípio que permite compreender os fenômenos do mundo, mas que não se refere a nenhuma interferência real entre estados mentais e físicos, o problema mente-corpo desaparece. Assim como desaparece para as perspectivas eliminativistas e reducionistas, nas quais a mente ou é eliminada ou é reduzida aos processos físicos do corpo. Apenas se compreendermos a mente como tendo alguma propriedade irredutível aos processos físicos, e que tal propriedade tem o poder de produzir mudanças físicas, é que haverá um real problema mente-corpo a ser investigado.

PALAVRAS-CHAVE: Eliminativismo. Reduccionismo. Não-reduccionismo. Causação.

ABSTRACT: Sentences like “This allergy is psychological” imply to accept that mental states interfere in physical states. The mind-body problem arises insofar as we also accept two presuppositions: 1. That mind and body are different things; and 2. That “causation” implies an ontological production in which cause and effect are similar. Because how could a state of mind, other than a physical state, produce physical change? The purpose of this article is to clarify in what sense there is a real mind-body problem. For this, it is necessary to distinguish some approaches to mind and causation. There are two conceptions concerning causation: the ontological and the epistemic. In relation to the concept of mind, we point out three approaches: the eliminativist, the reductionist, and the non-reductionist. Initially, we will present the eliminativist approach of Churchland; next, we will present Chalmers’ non-reductionist approach; finally, we will expose Dennett’s reductionist approach. If we understand that causation is a principle that allows us to understand the phenomena of the world, but does not refer to any real interference between mental and physical states, the mind-body problem disappears. It occurs similarly into the eliminativist and reductionist perspectives, in which either the mind is eliminated, or it is reduced to the physical processes of the body. Only if we understand the mind as having some property irreducible to physical processes, and that such property has the power to produce physical change, there will be a real mind-body problem to be investigated.

KEYWORDS: Eliminativism. Reductionism. Non-reductionism. Causation.

³⁸ Doutoranda em Filosofia na UNISINOS.

Considerações iniciais

Quando dizemos que alguém está com enxaqueca devido a um estresse emocional, estamos pressupondo que um estado mental causou um estado físico. Mas, é possível um estado mental causar um estado físico? Muitas perspectivas de senso comum, seguindo perspectivas religiosas ou mesmo filosófico-platônicas de que a alma é que comanda o corpo, consideram que isso é inquestionável. Mas também cientistas e filósofos, como, por exemplo, David Bohm (2011, p. 208), afirmam que, assim como mudanças induzidas na química cerebral podem modificar os estados mentais, assim também “Pensamento, sentimento e intenção podem mudar radicalmente a distribuição do fluxo sanguíneo e variar os processos químicos do cérebro”.

No entanto, um grupo de filósofos chamados epifenomenistas defendem que estados mentais não têm poder causal sobre o físico, nem mesmo sobre outros estados mentais; para a perspectiva epifenomenista, os eventos mentais são causados por eventos físicos no cérebro, mas não têm nenhum poder causal sobre eventos físicos ou mentais (Cf. ROBINSON, 2015). Esse é um problema que permanece em aberto, sendo que muitos filósofos ainda investigam se há de fato uma relação mente-corpo e como ela se caracterizaria.

Diante disso e de tantas concepções diferentes sobre os conceitos envolvidos na relação mente-corpo, pretendemos esclarecer os pressupostos da investigação desse problema. Considerar essa relação um problema depende do que estamos entendendo por “mente” e de como caracterizamos a relação de “causação”. Em geral, quem considera a relação mente-corpo problemática aceita dois pressupostos: 1. Que o conceito de causa implica uma produção ontológica, na qual causa e efeito são semelhantes; e 2. Que mente e corpo são distintos entre si. Existem duas concepções referentes à causação: a ontológica e a epistêmica. A primeira compreende, de modo geral, a causação como um princípio de produção de coisas novas a partir de outras preexistentes. Para a segunda, causação seria um princípio que permite compreender os fenômenos do mundo, mas não se refere a nenhuma relação de produção entre os eventos. Quanto ao conceito de mente, apontamos três abordagens: a eliminativista, a reducionista e a não-reducionista. Para as duas primeiras não haveria um real problema mente-corpo, uma vez que a mente ou é eliminada (eliminativismo) ou é reduzida aos processos físicos do corpo

(reducionismo). Portanto, apenas se compreendermos a mente como tendo alguma propriedade irredutível aos processos físicos, é que haverá um real problema mente-corpo.

Inicialmente, iremos abordar o primeiro pressuposto, criticando a perspectiva epistêmica humeana em relação à causação com base em Bunge (1978), que defende uma perspectiva ontológica da causação; porém não vamos nos aprofundar na discussão de se a novidade gerada pela produção é apenas quantitativa ou pode ser qualitativa.³⁹ Em seguida, abordaremos o segundo pressuposto, investigando o conceito de mente; nessa parte, dividida em três, começamos apresentando a abordagem eliminativista de Churchland (2004); depois, pretendemos esclarecer a abordagem não-reducionista de Chalmers (1996; 2003); e, por último, iremos expor a abordagem reducionista de Dennett (1991;1992).

a. O conceito de “causa” implica uma produção ontológica entre causa e efeito?

O problema em relação à causação mente-corpo é explicar como a mente pode causar modificações no corpo. A esse tipo de interferência de eventos mentais sobre eventos físicos, chamamos causação mental. Isso torna-se um problema se compreendermos mente como algo diferente do corpo e se entendermos causação como uma produção ontológica de coisas semelhantes entre si. Quando afirmamos que algum comportamento ou alguma alteração física foi causada por determinados aspectos mentais, estamos buscando compreender os fenômenos do mundo estabelecendo as causas de determinados efeitos. Fazer isso é importante seja para evitar consequências indesejadas e tentar reproduzir efeitos benéficos, seja para simplesmente conhecer a natureza das coisas. Mas o que caracteriza uma relação causal?

Quando dizemos, por exemplo, que o copo quebrou *porque* o gato bateu nele e ele caiu no chão, queremos dizer que o fato de o copo ter quebrado foi “produzido ontologicamente” pelo fato de o gato ter batido no copo e feito ele cair? Ou a nossa afirmação de que a batida do gato no copo de vidro causou a sua queda e a sua ruptura existe apenas em virtude do nosso “hábito” de perceber que constantemente copos de vidro se quebram depois de caírem no chão? Enfim, quando dizemos que x causou y , x é a força geradora de y , no sentido de ser uma condição para a sua existência, ou apenas queremos dizer que nos habituamos a perceber y logo depois de x e os

³⁹ Apesar de isso ser importante esclarecer e investigar, deixaremos para outro momento.

associamos no pensamento para explicar logicamente os fenômenos e x e y não têm necessariamente relação alguma?

Para Aristóteles (384-322a.C.), a relação causal pertence a dimensão ontológica no sentido de que “A é a razão para B existir”, e essa razão, além de permitir que B possa ser deduzido racionalmente a partir de A, também fornece uma condição ontológica para B. A relação causal para o estagirita representa, portanto, uma relação necessária e absoluta, no sentido de que sempre e em todas as situações B é causado por A, em virtude de sua essência. Isso origina a ideia de que a relação causal é unívoca e que existe apenas uma única explicação para um fenômeno, o que permite prever “racionalmente” eventos futuros diante de determinadas condições.

Mas David Hume (1711-1776) nega que possamos conhecer, apenas pelo pensamento racional, que o movimento de uma bola de bilhar, por exemplo, pode movimentar outra. Para ele, o conhecimento da relação que se dá entre uma e outra bola do jeito que acontece, só é possível pela experiência e observação, pois muitos outros resultados seriam igualmente coerentes (*Tratado*, 1.3.1, 2). Para ele, a inferência feita entre causa e efeito não pode ser deduzida a partir do exame dos objetos particulares, pois não há nada num objeto que implique a existência de outro (*Tratado*, 1.3.6). Não há nada “no gato batendo no copo” que implique a existência “do copo caindo” e nada “no copo caindo” juntamente com “o copo ser de vidro” que implique a existência “do copo quebrado”.

Portanto, o efeito não seria ontologicamente conectado à causa e deduzido racionalmente dela, mas associado por hábito; a conexão entre o efeito e a causa seria meramente arbitrária. De acordo com a filosofia de Hume, a conexão causal seria dada a partir de um critério epistêmico fornecido pela repetição observável dos eventos, que formaria o hábito de ver que quando um acontece, o outro logo sucede.

Três seriam os princípios que permitiriam perceber a repetição dos eventos e formar o hábito que constituiria a ideia de causação: 1. A relação de *contiguidade* que se refere à proximidade da causa e seu efeito no espaço e tempo (*Tratado*, 1.3.2); 2. A *sucessão temporal* que prevê a prioridade temporal da causa em relação ao seu efeito; e 3. Para percebermos uma

relação de causação entre dois eventos, seria necessário percebermos uma *conjunção constante* entre eles.

Por exemplo, quando pensamos no fogo, logo lembramos as chamas. Isso acontece, segundo Hume, por causa da nossa *experiência passada* que associou a ideia das chamas à ideia do fogo. Mas pode acontecer de haver fogo e não haver chamas, como quando o fogo está em brasas. Isso significa que as experiências passadas só nos permitem ter uma “previsão provável”, mas não uma “previsão infalível” de que no futuro se tivermos a impressão do fogo, teremos também a das chamas (Cf. *Tratado*, 1.3.6).

Compreendida assim, a causação não poderia ser sinônimo de “produção”, pois dizer que um objeto ou ação produz outro implica em dizer que existe uma “conexão necessária” entre os dois objetos, sendo que um teria “poder” sobre o outro. Mesmo que um objeto da experiência passada tivesse o poder de causar outro, para Hume, nada nos daria a certeza de que objetos semelhantes no futuro teriam o mesmo poder (*Tratado*, 1.3.6). Esse é o ponto forte da crítica humeana à noção ontológica de causação. No entanto, para assumir a impossibilidade de termos certeza com relação às nossas descrições dos eventos não é necessário assumir que não haja relação real de produção entre as coisas.

Hume pretendia criticar que as coisas teriam uma ‘essência’ a ser conhecida racionalmente. Para ele, as possíveis conexões existentes entre os fenômenos se dariam apenas por princípios que associam umas ideias às outras na imaginação. A única garantia que teríamos para afirmar uma conexão entre a causa e o efeito seria a multiplicidade dos casos semelhantes que aconteceram no passado (*Tratado*, 1.3.14). E a semelhança que há entre os múltiplos casos é o que indica que a conexão é necessária, nada mais (*Tratado*, 1.3.14). Tal ideia de semelhança é resultado de uma impressão de reflexão gerada a partir de impressões sensíveis. Por isso, a causalidade, para Hume, é um princípio psicológico ou epistemológico e não ontológico.

Mário Bunge (1919 -) critica a perspectiva humeana por deixar de lado o caráter dinâmico da causação que tem a ver com a produtividade da causa eficiente, única das quatro causas aristotélicas a permanecer na modernidade (1978, p. 58). Ele apresenta contraexemplos para criticar os três princípios humeanos defendidos como a base da formação da ideia de causalidade. Para Bunge, pode-se perceber que o princípio da *conjunção constante* não é

adequado para caracterizar a causação, visto que os braços de uma alavanca estão sempre unidos e um não é a causa do outro⁴⁰. Quanto ao princípio da *contiguidade*, Bunge lembra que a ação da gravidade é uma ação à distância e, ainda assim, “produz” consequências; por exemplo, dizemos que os objetos caem, quando os jogamos para cima, por “causa” da gravidade, ainda que não exista contato direto entre causa e efeito. E quanto ao princípio da *sucessão temporal*, Bunge afirma que a existência da causa já basta para que o efeito se efetive, mas isso não implica pré-existência temporal. Segundo ele, a causação é compatível tanto com acontecimentos simultâneos quanto com acontecimentos distantes no tempo (BUNGE, 1978, p. 100). Aristóteles já havia defendido a possibilidade de a causa e o efeito serem simultâneos; seu exemplo: enquanto o pintor pinta, o pintor é a causa e a pintura é o efeito e estes são simultâneos, enquanto que o pintor como causa e o quadro como efeito não o são (*Fís.*, II 3, 195a – 195b).

Para Bunge, portanto, o critério imprescindível para que haja uma *relação causal* é que haja uma *relação de produção de coisas novas a partir de outras já existentes*. Diante disso, restaria investigar em que sentido essas coisas geradas são novas. A novidade seria quantitativa ou qualitativa? Para a relação mente-corpo, se a novidade for apenas quantitativa, mente e corpo devem ser semelhantes; porém se a novidade for qualitativa, mente e corpo podem ser diferentes entre si e mesmo assim estabelecerem relações. De qualquer forma, uma vez assumida a existência de uma conexão ontológica estabelecida pela relação de produção entre mente e corpo, pelo menos a investigação de como se dá tal relação não será em vão. E, quem sabe, teremos fundamentações teóricas mais apropriadas para sustentar escolhas de tratamentos de saúde e de atitudes para intensificar o próprio bem-estar.

b. Mente e corpo são distintos entre si?

A relação mente-corpo configura-se um problema desde o momento em que se considera mente e corpo como coisas totalmente distintas, mas que no ser humano interagem como se fossem uma só. Descartes, especialmente em *Meditações da filosofia primeira* (1641) e *As paixões da alma* (1649), ao sustentar um dualismo de substâncias e, ao mesmo tempo, um

⁴⁰ Segundo Osvaldo Pessoa Júnior (em diálogo pessoal durante o VIII Colóquio Internacional de Filosofia da Mente entre os dias 11 a 13 de abril de 2016 na UFMG), embora os contraexemplos de Bunge aos outros princípios sejam adequados, esse não o é. Mas não vamos nos aprofundar nesse aspecto aqui.

interacionismo entre a *res extensa* e a *res cogitans*, marca o início da discussão sobre o “problema” mente e corpo; ora, como duas coisas completamente distintas, podem se misturar e interagir entre si?

Nesta seção vamos apresentar algumas concepções sobre o que caracteriza a mente com o objetivo de esclarecer quais ainda concebem que existe um real problema mente-corpo e quais fazem o problema desaparecer por considerá-lo sem sentido. Inicialmente, vamos apresentar a abordagem eliminativista de Churchland (2004); em seguida, a abordagem não-reducionista de Chalmers (1996; 2003); e, por último, iremos expor a abordagem reducionista de Dennett (1991;1992).

2.1 A abordagem eliminativista de Paul Churchland

Descartes (AT VI, 55-60), na quinta parte de o *Discurso do método*, considera que o homem, além da máquina perfeita do corpo, também tem alma, pois isso explica a sua capacidade para compreender uma linguagem e a sua criatividade para agir de formas diversas em diferentes situações; enquanto que os outros animais teriam apenas um corpo que age somente pela disposição de seus órgãos. Porém, filósofos como Julien Offray de La Mettrie (1709-1751) e Barão D’Holbach (1723-1789) consideram o homem como mais um ser da natureza, assim como os outros animais, sendo a racionalidade um produto da evolução do sistema nervoso, e, portanto, algo puramente físico⁴¹. Eles assumiram uma posição monista materialista, segundo a qual tudo o que existe é físico ou material.

Nessa linha de pensamento se insere o casal de filósofos Paul Churchland (1942-) e Patrícia Smith Churchland (1943-). Eles compreendem que, por serem parte da natureza, os seres humanos não têm nada além do que pode ser cientificamente explicado; portanto para explicar seu comportamento é mais fácil e preciso simplesmente eliminar o vocabulário mentalista e usar

⁴¹ O que é muito sensato, mas não quer dizer que a racionalidade e o que consideramos processos e estados mentais não estivessem presentes em alguns graus nos outros animais. E, de fato, muitos filósofos, biólogos e psicólogos, dentre eles, Peter Singer (2002), Maturana e Varela (1995), Hanna e Maiese (2009) e Damásio (2000), consideram que animais não-humanos como chimpanzés tenham algum grau de racionalidade e autoconsciência. O evolucionismo não elimina o mental, apenas mostra que ele já poderia estar presente em níveis menos complexos nos nossos ancestrais.

apenas termos que se referem à neurofarmacologia, anatomia e atividades neuronais. Os Churchland argumentam que termos mentais como dor, prazer etc. perderam seu valor para a produção de conhecimento e devem ser eliminados; da mesma forma que termos como gnomo, bruxa, éter, fluido calorífico e flogisto foram úteis antigamente, mas atualmente não são mais usados.

A linguagem mentalista não é eficiente para explicar e prever comportamentos humanos; principalmente, porque ela não consegue explicar de maneira satisfatória distúrbios comportamentais e manipulá-los a fim de minimizar os problemas das pessoas que os apresentam, o que a neurociência faz bem. Mais ainda, para Churchland (2004, p. 79), “a estrutura psicológica de nosso senso comum é uma concepção falsa e racialmente enganosa das causas do comportamento humano e da natureza da atividade cognitiva”.

A vantagem dessa teoria é tentar simplificar a nossa ontologia; e, além disso, buscar uma linguagem precisa e clara para falar dos processos internos. Porém, essa visão apresenta muitas dificuldades, pois conforme esclarecem David Armstrong e Norman Malcolm (1984, p. 99), ela não eliminaria o discurso sobre sensação, na medida em que, para alguém dizer que tem dor, deveria dizer “minhas fibras C estão disparando” para dizer isso o sujeito estaria querendo dizer que tem alguma sensação específica. Então, no fim das contas, seria apenas mudar a terminologia. E apesar de pretender simplificar a nossa ontologia, essa teoria não iria simplificar a nossa linguagem ordinária, pois muito mais termos deveriam ser usados para falar das sensações. Para a linguagem técnica utilizada na área da saúde, talvez essa substituição fosse benéfica, na medida em que esclareceria e especificaria exatamente do que se está falando, mas isso não quer dizer que se devam eliminar os termos mentais como “dor”, “crença” etc. da linguagem do senso comum⁴². Até porque falar da “minha dor” não parece ser o mesmo que falar das “minhas propriedades físicas das fibras C disparando”.

Churchland (2004), ao defender que não existe nada no ser humano além dos processos físicos e ao sugerir a completa eliminação do vocabulário mentalista, pretende eliminar o problema mente-corpo na medida em que haveria apenas uma causação física entre as diferentes

⁴² Pensemos na consequência disso para a literatura e a arte em geral: a falta de ambiguidade na linguagem faria com que se perdesse o sentido e a graça de piadas, poemas etc.

estruturas anatômicas e fisiológicas no organismo humano. Mas será que apenas uma descrição nesses termos daria conta de explicar o nosso comportamento?

2.2 A abordagem não-reducionista de David Chalmers

Com os inúmeros avanços nas neurociências, nas ciências cognitivas etc., tem-se desvendado muitos dos funcionamentos do cérebro e do comportamento humano. Porém, a questão que fica é se o que as neurociências descrevem é tudo o que há para saber sobre o ser humano. Será que quando nos referimos a estados mentais não nos referimos a nada mais que estados do funcionamento cerebral descritos objetivamente?

David Chalmers, em seu livro *The conscious mind: in search of a theory of conscious experience*, distingue dois tipos de investigações sobre a mente: uma de fundo filosófico-metafísico e outra filosófico-científica. Para ele, investigar (1) como surgem as sensações e (2) qual o seu papel funcional no comportamento, são investigações sob o ponto de vista filosófico-científico. E, em grande parte, as Ciências Cognitivas, a Psicologia, as Neurociências etc. conseguiram grandes progressos nessas pesquisas. Porém, investigar (3) a razão de toda estimulação associada com a dor ser acompanhada por uma experiência de dor, é investigar a mente de um ponto de vista filosófico-metafísico. E é aí que está o ‘problema difícil da consciência’, ou o real problema a investigar sobre a relação mente-corpo (CHALMERS, 1996, p. 23).

Thomas Nagel (1974, p. 435-436) já havia mostrado que para refletirmos sobre as questões realmente interessantes sobre a mente não podemos deixar de considerar a *consciência*, no sentido de investigar ‘como é *ser* como algo’. Isso se refere ao caráter subjetivo da experiência consciente, ponto de vista acessado apenas em primeira pessoa. Com base nisso, Chalmers (1996, p. 4-5) afirma que o maior problema em relação à mente é explicar por que temos *qualia*, ou seja, por que nossas emoções, crenças, sensações etc., por exemplo, são acompanhadas por *uma experiência qualitativa subjetiva ‘consciente’*⁴³.

⁴³Segundo Tye (2016), o uso mais amplo do termo *qualia* deriva da noção de *quale* apresentada por Peirce, em 1866, para se referir a um caráter fenomênico que pressupõe focar a atenção, estar consciente (*aware*) de certas qualidades experienciadas. Mas existem ainda outros três usos mais restritos do termo: 1. *Qualia* como propriedades

Essa abordagem da mente refere-se ao *aspecto fenomênico da mente*, o qual caracteriza a mente pelo que ela sente ou por como as coisas se parecem para nós. Para Chalmers, a Psicologia, as Ciências Cognitivas, entre outras, já se dedicaram a investigar a mente, mas até então apenas se debruçaram a investigar o *aspecto psicológico* sobre o processo causal do comportamento, caracterizando a mente pelo que ela faz. Segundo Chalmers (1996, p. 11; p. 22), os aspectos fenomenais e psicológicos juntos é que ajudam a explicar o mental.

Chalmers (1996, p. 33-48) defende a existência dos *qualia* como algo diferente e que existe para além da constituição físico-biológica do ser humano. Para ele, há uma relação de superveniência local entre os *qualia* e os fatos físicos, o que significa que as experiências qualitativas subjetivas e conscientes dos seres humanos dependem dos processos físicos que ocorrem no corpo, mas isso não implica que ao saber tudo sobre os fatos físicos possamos deduzir o conhecimento dos fatos fenomênicos (superveniência lógica).

Para Chalmers (1996, p. 94-99), por ser teoricamente possível pensar na existência de um mundo físico sem a existência dos *qualia* (a possibilidade de pensar em um mundo zumbi⁴⁴), é possível concluir que os fatos fenomênicos não sobrevivem logicamente dos fatos físicos. Isso significa que há uma certa independência entre a consciência fenomênica e os processos físicos no corpo, no sentido de ela não poder ser reduzida a esses processos.

dos dados sensórios, influenciado pelo uso que Lewis faz do termo, em 1929, para se referir a propriedades dos dados dos sentidos que são intrínsecas, acessadas conscientemente, características não-representacionais dos dados dos sentidos e de objetos fenomênicos não-físicos; 2. *Qualia* como propriedades intrínsecas não-representacionais, semelhante à 1, mas que não pressupõem o endosso de uma Teoria dos Dados dos Sentidos—Nagel (1974) e Block (1990) são autores que seguem esse uso; e 3. *Qualia* como propriedades intrínsecas, não-físicas e inefáveis, que seria o conceito criticado por Dennett (1988; 1991) ao defender que existem experiências conscientes, mas não *qualia* neste sentido. Relacionado a esses usos podemos considerar, portanto, que existem, em geral, três perspectivas a respeito dos *qualia*: as reducionistas, as eliminativistas e as não-reducionistas. Diversos autores, tal como Place (1956), Smart (1959), Ryle (2009), Dennett (1988; 1991), tentam reduzir estados mentais aos estados físicos, incluindo aí os *qualia*. Já o casal Churchland defende a eliminação completa de todo e qualquer conceito mental, enquanto que diversos outros autores defendem que os conceitos mentais relacionados aos *qualia* sejam irredutíveis — é o caso de Block (1980), Shoemaker (1982), Jackson (1986) e Chalmers (1996). Kim (2000), apesar de propor uma perspectiva não-reducionista do mental, ainda é um eliminativista quanto aos *qualia*. Nesse sentido, da perspectiva de Chalmers, Kim não investiga o real problema mente-corpo.

⁴⁴ Chalmers diferencia zumbis psicológicos de zumbis fenomênicos. Os primeiros seriam os dos filmes de Hollywood, que são idênticos fisicamente aos seres humanos, mas são diferentes psicologicamente, pois agem diferentemente de nós ao não demonstrarem dor. Já os zumbis fenomênicos dos quais Chalmers fala, são idênticos física e psicologicamente aos seres humanos, ou seja, agem exatamente como os humanos, porém não têm experiências qualitativas (CHALMERS, 1996, p. 95).

Essa, dentre outras experiências de pensamento⁴⁵, é usada por Chalmers (1996) como argumento contra a superveniência lógica. Outra, já apresentada por Shoemaker (1982) e antes ainda por Locke (1999), é a do espectro invertido. Segundo essa experiência de pensamento, no nosso mundo, eu vejo o céu azul e a maçã vermelha, mas meu gêmeo no mundo gêmeo ao nosso, pode ver o céu vermelho e a maçã azul. Essa possibilidade também sugere uma falha da superveniência lógica da experiência consciente em relação ao físico, pois se ela existisse, como os mundos são idênticos fisicamente, os seres que vivem neles deveriam ter experiências conscientes idênticas.

Chalmers considera o experimento de pensamento elaborado por Frank Jackson (1986), conhecido por “Mary e o quarto preto e branco”, como o argumento mais claro que pode ser usado contra a superveniência lógica da consciência sobre o físico. Pois mostra que, mesmo que Mary conheça todos os fatos físicos sobre as cores, ela adquire conhecimento novo ao sair do quarto e ter contato direto com as cores do mundo. Logo, a experiência consciente não sobrevém logicamente do físico, caso contrário, ela não teria adquirido novo conhecimento ao sair do quarto (CHALMERS, 1996, p. 103-104).

É certo que Mary adquire um conhecimento novo ao sair do quarto, mas a ambiguidade existente com relação ao tipo de conhecimento adquirido (se um conhecimento proposicional, de habilidade ou por familiaridade) faz com que seja questionável se ele realmente se refere a algo diferente dos fatos físicos ou não. Entretanto, Chalmers conclui que a consciência fenomênica não sobrevém logicamente dos fatos físicos, mas que os fatos fenomênicos acompanham os fatos físicos como fatos adicionais e devem ser explicados por si mesmos (CHALMERS, 1996, p. 107). No entanto, será que realmente existe esse aspecto da mente que estamos chamando de *consciência fenomênica* relacionada aos *qualia*? E será que esse é o aspecto irreduzível que faz com que a mente seja diferente dos processos físicos

⁴⁵Os cinco argumentos usados por Chalmers contra a ideia de superveniência lógica da consciência fenomênica sobre o físico são: 1. O argumento dos zumbis; 2. O argumento do espectro invertido; 3. O argumento da assimetria epistêmica; 4. O argumento do conhecimento; e 5. O argumento da ausência de análise.

2.3 A abordagem reducionista de Daniel Dennett

Para alguns filósofos fisicalistas como Daniel Dennett (1988; 1991) as experiências conscientes não teriam *qualia*, enquanto propriedades qualitativas subjetivas, que constituiriam uma outra ontologia para além da existência dos estados físicos. Ele nega a existência ou a importância dos *qualia*, que seriam tidos como reais ou significativos, é essa a tese defendida em “Quining Qualia” (DENNETT, 1988, p. 381). Para esse filósofo, as experiências conscientes não têm nada de diferente dos processos físicos que ocorrem no corpo. Nesse artigo, Dennett usa 15 intuições para criticar as propriedades atribuídas pela tradição dualista aos *qualia*, e, com isso, pretende criticar toda e qualquer noção de *qualia*. Para ele, as experiências conscientes teriam, sim, propriedades, mas, ao invés de serem intrínsecas, privadas, inefáveis e apreendidas diretamente, são extrínsecas, públicas, relatáveis e mediadas; e isso não implica em aceitar a existência de algo para além dos processos físicos do corpo.

Dennett (1988, p. 384-385) explica que, segundo a tradição, dizer que os *qualia* são ‘inefáveis’ significa dizer que não se pode dizer exatamente o que alguém está vendo, sentindo pelo tato ou olfato etc., ou seja, que eles são não analisáveis. Dizer que os *qualia* são ‘intrínsecos’ significa dizer que eles são homogêneos, simples, e, poderíamos complementar, que pertencem à natureza ou essência de algo (Cf. ABBAGNANO, 2000, p. 423). Dizer que os *qualia* são ‘privados’ significa dizer que nenhuma comparação interpessoal deles é possível. Dizer que os *qualia* são diretamente apreendidos significa dizer que o sujeito está intimamente familiarizado com seus *qualia* e, associado a isso, a ideia de que o sujeito não poderia se enganar em relação a eles (DENNETT, 1988, p. 384-385).

Quanto à propriedade dos *qualia* de serem inefáveis, Dennett (1988, p. 409, tradução nossa) a critica dizendo que “seria um erro transformar o fato de que inevitavelmente há um limite na nossa capacidade de descrever as coisas como nós as experienciamos na suposição de que não existem propriedades absolutamente indescritíveis em nossa experiência”. Por exemplo, Dennett (1988, p. 387-388) pede para supor a situação de alguém que certo dia acorda e se depara com a grama vermelha, o céu amarelo etc. A pessoa questiona os outros sobre a alteração da cor da grama, do céu etc., mas ninguém nota nenhuma anormalidade. Acontece que, sem a pessoa saber, seus neurônios foram adulterados por neurocientistas. Ela consegue relatar suas experiências conscientes aos outros e parece que elas podem ser empiricamente verificadas. Isso

mostra que os *qualia* não existiriam enquanto algo inefável, porque a experiência consciente é relatável. Além disso, apesar de parecer impossível para nós compreendermos o que é o choro da água marinha, é possível descrevê-lo em livros, por exemplo. Portanto, os *qualia* como algo inefável não existem, porque a experiência consciente é relatável. Mas será que o fato de a experiência consciente ser relatável significa que é possível descrever exatamente como alguém experiência algo?

Quanto à propriedade dos *qualia* de serem intrínsecos, Dennett (1988, p. 397) a crítica dizendo que ninguém poderia gostar de cerveja na sua primeira degustação, pois gostar de cerveja é algo que acontece gradualmente. Com base nesse exemplo, Dennett argumenta que se aprendemos a gostar de algo, então não existe um *qualia* que seja intrínseco a esse algo. Da mesma forma, se podemos adaptar a nossa visão ao começar a usar óculos e isso não acontece de forma imediata ao colocar os óculos, reforça-se assim a tese de que não existem *qualia* intrínsecos às coisas (DENNETT, 1988, p. 399-400). Para Dennett (1988, p. 397-403), a experiência consciente é relacional e extrínseca, ela não depende apenas dos objetos que são experienciados, mas da nossa relação com eles. O exemplo da feniltiouréia, que é sentida como amarga por 70% da população mundial, enquanto o resto a sente como insípida, mostra que a forma como as coisas parecem a nós, muda conforme os nossos detectores sensitivos (DENNETT, 1988, p. 398). Mas além dos detectores sensitivos mudarem de uma pessoa para outra, fazendo com que elas percebam as coisas de forma diferente, a própria pessoa pode perceber algo como diferente em momentos diferentes da sua vida; por exemplo, um degustador de vinho pode modificar seu paladar à medida em que aprende a saborear diferentes tipos de vinho (DENNETT, 1988, p. 408-409). Isso mostra que não existe um *qualia* absoluto, mas que eles podem ser aprendidos. E que a experiência consciente é extrínseca e relacional. Mas, apesar de isso ser interessante, será que significa que não existe uma experiência intrínseca a um sujeito numa determinada situação? Será que a fragilidade da concepção de *qualia* da tradição, leva a conclusão de que qualquer concepção de *qualia* o seja?

Quanto à propriedade dos *qualia* de serem imediatamente apreendidos, Dennett (1988, p. 388) a crítica dizendo que o fato de uma alteração neurocirúrgica poder alterar os nossos *qualia* de duas formas diferentes (ou alterando os *qualia*; ou alterando nossa memória para reagir às disposições), e nós não sabermos o que teria acontecido conosco, mostra que nós não temos

acesso imediato aos *qualia*. Além disso, ele argumenta que o fato de degustadores de café saberem que seus gostos mudaram depois de 6 anos, porém não saberem se foram suas reações frente aos *qualia* que mudaram ou se foram pequenas alterações graduais que aconteceram em seus *qualia*, leva à conclusão de que as pessoas nem sempre estão completamente certas sobre seus próprios *qualia*, portanto, que eles não são infalíveis e incorrigíveis como se pensava. Para Dennett (1988, p. 396), a experiência consciente é mediada por processos físicos, pois dependendo de alterações nos neurônios, por exemplo, a experiência consciente pode mudar. Mas será que o fato de as experiências conscientes serem mediadas por processos físicos, significa que não exista uma qualidade subjetiva da experiência que seja apreendida imediatamente, sem que se saiba necessariamente algo sobre os processos físicos do nosso corpo?

Quanto à propriedade dos *qualia* de serem privados, Dennett (1988, p. 387) apresenta o caso do espectro invertido, já apresentado por Locke, segundo o qual haveria uma impossibilidade de compararmos as experiências conscientes e sabermos se são as mesmas. Porém, diferentemente de Locke, Dennett afirma que o que importa é se as nossas palavras são as mesmas diante do que estamos experienciando. Assim, ele nega que os *qualia* sejam privados, isso seria insignificante, pois o que importa de fato é o caráter público da experiência consciente. No entanto, dizer que o que importa é que as palavras combinem diante de algo, prova de fato que não existe algo de privado e subjetivo que pode ser diferente?

As quinze intuições apresentadas por Dennett (1988) parecem realmente minar a concepção tradicional de que os *qualia* são totalmente independentes do físico. Mas será que provam que não existem *qualia* como uma experiência fenomênica a respeito de como as coisas aparecem para nós?

Considerações finais

A relação mente-corpo configura-se como um problema se considerarmos que mente e corpo referem-se a coisas distintas que tem poder de se influenciar mutuamente. De acordo com Chalmers (1996, p. 4-5), o aspecto fenomênico da mente é irreduzível aos processos físicos do corpo. E, a partir disso, investigar o real problema mente-corpo implica em pesquisar como é

possível os *qualia* modificarem os (e serem modificados por) processos físicos do corpo e nossos comportamentos.

Vimos que para as perspectivas eliminativista e reducionista com relação à mente não existe um real problema mente-corpo, pois a primeira elimina o mental como um conceito obscuro e fica apenas com a causação física no corpo; ao passo que a segunda considera significativo manter o vocabulário mental, mas o reduz a descrições de processos físicos, ficando também apenas com a causação física. Nesse sentido, não haveria dificuldade alguma em explicar como processos mentais, entendidos no fim das contas como processos físicos, podem modificar outros processos físicos. Apenas se considerarmos que há algum tipo de causação entre mente e corpo como coisas diferentes ou contendo propriedades diferentes é que há um problema a ser solucionado e esclarecido.

Além disso, investigar o real problema mente-corpo implica em considerar que a mente, como algo diferente ou contendo propriedades irreduzíveis, interage com os processos físicos de modo a produzir alterações recíprocas. O que implica em pressupor que a causação é um princípio ontológico e não meramente psicológico ou epistemológico, como defende Hume. Isso significa que as afirmações de que “o estresse emocional gerou a enxaqueca” ou que “a alergia é psicológica” pretendem descrever exatamente como a realidade é.

Ainda que possamos ter explicações diferentes para esclarecer como se dá a interferência entre o mental e o físico, desde que compreendamos que o mental e o físico são irreduzíveis um ao outro, e que eles produzem modificações um no outro, estaremos diante de um mesmo problema a ser resolvido.

Referências

- ABBAGNANO, N. *Dicionário de filosofia*. Trad. Alfredo Bosi. São Paulo: Martins Fontes, 2000.
- ADAM, C.; TANNERY, P. (Orgs.). *Oevres de Descartes. Principia philosophae. Epistola Ad G. Voetium. Lettre apologetique. Notae in programma*. Paris: Librairie Philosophique J. Vrin, 1996. v. 8.
- ARISTÓTELES. *Física*. Trad. Guillermo R. de Echandía. Madrid: Gredos, 1995.

- _____. *Obras*. 2. ed. Trad. Francisco de P. Samaranch. Madrid: Aguillar, 1967.
- ARMSTRONG, D. M.; MALCOLM, N. *Consciousness and Causality*. Oxford: Blackwell, 1984.
- BLOCK, N. Inverted earth. *Philosophical Perspectives*, v. 4, p. 53-79, 1990. Disponível em: <<http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/1990.Inverted.Earth.pdf>>. Acesso em: 18 fev. 2017.
- _____. Troubles with functionalism. In: _____ (Ed.) *Readings in philosophy of psychology*. Cambridge, Massachusetts: Harvard University Press, 1980. p. 268-305. v. 1
- BOHM, D.; PEAT, D. F. *Science, order and creativity*. Londres; Nova York: Routledge Classics, 2011.
- BUNGE, M. *Causalidad: el principio de causalidad en la ciencia moderna*. Trad. Hernán Rodríguez. 4. ed. Buenos Aires: Universitaria de Buenos Aires, 1978.
- CHALMERS, D. J. Consciousness and its place in nature. In: STICH, S. P.; WARFIELD, T. A. (Eds.). *Blackwell guide to philosophy of mind*. [S. l.]: Blackwell, 2003. p. 1-46. Disponível em: <<http://consc.net/papers/nature.pdf>>. Acesso em: 01 out. 2016.
- _____. *The conscious mind: in search of a fundamental theory*. New York: Oxford University Press, 1996.
- CHURCHLAND, P. *Matéria e consciência: uma introdução contemporânea à filosofia da mente*. Trad. Maria Clara Cescato. São Paulo: Editora UNESP, 2004.
- DAMÁSIO, A. R. *O mistério da consciência: do corpo e das emoções ao conhecimento de si*. Trad. de Laura Teixeira Motta. São Paulo: Companhia das Letras, 2000.
- DENNETT, D. C. *La consciencia explicada: una teoría interdisciplinar*. Buenos Aires: Ediciones Paidós, 1991.
- _____. Quining qualia (1988). Reimpresso em MARCEL, A. J.; BISIACH, E. (Eds.) *Consciousness in contemporary science*. [S. l.]: Oxford University Press, 1992. As páginas citadas entre 381 e 414 são da versão disponível em: <<http://www.fflch.usp.br/df/opessoa/Dennett-Quining-Qualia.pdf>>. Acesso em: 01 nov. 2016.
- DESCARTES, R. *Discurso do Método. Meditações. Objeções e Respostas. As paixões da alma. Cartas*. Trad. J. Guinsburg e Bento Prado Júnior. São Paulo: Victor Civita, 1973.
- HANNA, R.; MAIESE, M. *Embodied minds in action*. New York: Oxford University Press, 2009.
- HOLBACH, B. de. *System of Nature*. [S. l.: s.n.], 1770.
- HUME, D. *Tratado da Natureza Humana*. Trad. Serafim da Silva Fontes. Coimbra: Fundação Calouste Gulbenkian, 2001.
- JACKSON, F. Postscript on qualia. In: LUDLOW, P.; NAGASAWA, Y.; STOLJAR, D. (Eds.). *There's something about Mary: essays on phenomenal consciousness and Frank Jackson's knowledge argument*. MIT Press. 2004. p. 417-420.
- _____. What Mary didn't know. *The Journal of Philosophy*, [S. l.], v.83, n. 5, p. 291-295, maio 1986.

- KIM, J. *Mind in a physical world: an essay on the mind-body problem and mental causation*. Cambridge: MIT Press, 2000.
- LOCKE, J. *Ensaio sobre o entendimento humano*. Trad. Eduardo Abranches de Soveral, Gualter Cunha e Ana Luísa Amaral. 2. ed. Lisboa: Fundação Calouste Gulbenkian, 1999.
- MATURANA, H.; VARELA, F. *A árvore do conhecimento: as bases biológicas do entendimento humano*. Trad. de Jonas Pereira dos Santos. Campinas: Editorial Psy II, 1995.
- NAGEL, T. What is it like to be a bat. *The Philosophical Review*, [S. l.], v. 83, n. 4, p. 435-450, out. 1974.
- NIDA-RÜMELIN, M. Qualia: the knowledge argument. In: ZALTA, E. N. (Ed.) *The Stanford Encyclopedia of Philosophy*. [online], 2015.
- PLACE, U. T. Is consciousness a brain process? *British Journal of Psychology*, [S. l.], v. 1, n. 47, p. 44-50, fev. 1956.
- RYLE, G. *The Concept of Mind*. London, New York: Routledge, 2009.
- ROBINSON, W. Epiphenomenalism. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*, 2015.
- SINGER, P. Tirar a vida: os animais. In: _____. *Ética prática*. Trad. de Jefferson Luiz Camargo. 3. ed. São Paulo: Martins Fontes, 2002. p. 119-143.
- SHOEMAKER, S. The inverted spectrum. *The Journal of Philosophy*, [S. l.], v. 79, n. 7, p. 357-381, jul. 1982.
- SMART, J. J. C. Sensations and brain processes. *Philosophical Review*, [S. l.], v. 68, n. 2, p. 141-156, abr. 1959.
- TYE, M. Qualia. *The Stanford Encyclopedia of Philosophy*, 2016. Disponível em: <<https://plato.stanford.edu/archives/win2016/entries/qualia/>>. Acesso em: 17 fev. 2017.

O desenvolvimento do conceito de intencionalidade:

Brentano, Husserl e Searle em diálogo

The development of the concept of intentionality:

Brentano, Husserl and Searle in dialogue

Danilo da Cunha Pontes⁴⁶

RESUMO: Para Husserl, a intencionalidade não é um elemento real e objetivo para a distinção entre o psíquico e o físico, portanto ele não atribuiu um significado intencional a todos os fenômenos psíquicos. Por conseguinte, não são todas as vivências intencionais. Husserl tece uma crítica à tese de Brentano a qual defendia que a intencionalidade continha imanentemente um objeto como representação psíquica do objeto real. Ao contrário do que Brentano sustentava, o objeto é transcendente à intencionalidade e se manifesta na própria intencionalidade somente como componente não real, isto é, como significado, sentido. Não-real significa ideal; reais são os atos subjetivos, componentes da intencionalidade que se correlacionam à intuição desta idealidade. Na doutrina da intencionalidade de Brentano, os objetos percebidos, fantasiados, julgados, desejados, etc. "entram na consciência", "são recebidos na consciência" ou, de maneira inversa, "a consciência" entra em relação com eles. Na tradição analítica, o termo reaparece com Searle na obra *Intencionalidade*, onde o autor dando continuidade às suas pesquisas na área de filosofia da linguagem – teoria dos atos de fala – recupera o termo intencionalidade acrescentando-lhe as ideias de *Network* e *Background*. Para o autor assim como as enunciações, as representações mentais possuem direção de ajuste e condição de satisfação.

PALAVRAS-CHAVE: Intencionalidade; consciência; Background.

ABSTRACT: For Husserl, intentionality it is not a real element and objective to the distinction between psychic and physical, thus, he did not assign an intentional meaning to all psychic phenomena. Thereafter, it is not all intentional experiences. Husserl build a critic to Brentano's thesis which defended that intentionality has immanently an object as psychic representation of the real object. On the contrary to what Brentano sustained, the object is transcendent to intentionality and manifests itself in its own intentionality only as a not real component, this is, as meaning, sense. Not real means ideal; real are the subjective acts, components of the intentionality that are correlated with intuition of this ideality. In Brentano's doctrine of intentionality, the perceived, fantasized, judged, desired objects "enter into the consciousness", "are received into the consciousness" or, on the reversed way, "the consciousness" enter in relation with them. In the analytic tradition, the term reappears with Searle in the work *Intentionality*, where the author, continuing with your researches on the area of philosophy of language - theory of speech acts – recovers the term intentionality adding the ideas of *Network* and *Background*. To the author, as well as the enunciations, the mental representations have direction of adjust and condition of satisfaction.

KEYWORDS: intentionality; consciousness; Background.

⁴⁶ Mestrando em Filosofia pela Universidade Federal de Goiás (UFG).Email: danilo_pontox@yahoo.com.br.

1.

Falar em intencionalidade é evocar um tema que grosso modo atravessa toda história da filosofia. No entanto é preciso atentar ao fato de que é uma tarefa quase impossível querer realizar uma abordagem sistemática, dado que as formulações e teorias propostas que encontramos ao longo da história tenha sido tão variadas.

O conceito clássico de intencionalidade que conhecemos é originário da filosofia escolástica, mas como é sabido, foi Brentano quem formulou a versão hoje clássica de intencionalidade em sua *Psicologia do ponto de vista empírico*. Husserl reapropria de tal termo inserindo-o na quinta de suas *Investigações Lógicas*, dando assim curso a tradição fenomenológica. Posteriormente R. Chisholm em um de seus artigos introduz a temática da intencionalidade no mundo da filosofia analítica, dando um enfoque linguístico – dessa forma a ideia de intencionalidade é trabalhada por diversos autores da tradição analítica e da filosofia da mente, como é o caso do filósofo norte americano John Searle que abordaremos mais adiante. Dessa maneira o problema da intencionalidade se converte em um dos pontos onde é mais fácil a comunicação entre duas tradições filosóficas – a fenomenológica e a analítica.

2.

Quando Brentano retoma o conceito de intencionalidade dos escolásticos, o autor busca utiliza-lo como aspecto distintivo entre os fenômenos psíquicos, contrapostos àqueles físicos. Dessa forma a ideia de intencionalidade para Brentano surge como “relação ao conteúdo ou direção ao objeto”. Chega-se dessa maneira a compreensão de que todo fenômeno psíquico se distingue por ser uma "consciência de alguma coisa".

Todo fenómeno psíquico está caracterizado por lo que los escolásticos de la Edad Media han llamado la inexistencia intencional (o mental) de un objeto, y que nosotros llamaríamos, si bien con expresiones no enteramente inequívocas, la referencia a un contenido, la dirección hacia un objeto (por el cual no hay que entender aquí una realidad), o la objetividad imanente⁴⁷. (BRENTANO, 1935, p.81)

⁴⁷Todo e qualquer fenómeno psíquico é caracterizado pelo que os escolásticos da Idade Média denominavam como inexistência intencional (ou também mental) de um objeto e que nós, se bem que com expressões não completamente inequívocas, poderíamos denominar como a referência a um conteúdo, a direção para um objeto (pelo qual não se deve entender uma realidade) ou a objetividade imanente. (Tradução nossa)

Husserl ao se apropriar do termo “brentaniano”, se distancia da concepção que seu mestre apresentava. Diante disso o desenvolvimento da intencionalidade para o fenomenólogo se diverge de Brentano em alguns pontos. O primeiro ponto consiste em que, para Husserl, a intencionalidade não é um elemento real e objetivo para a distinção entre o psíquico e o físico, portanto ele não atribuiu um significado intencional a todos os fenômenos psíquicos. Por conseguinte, não são todas as vivências intencionais. Um outro ponto se pauta no fato de que o objeto não é imanente à intencionalidade. Husserl desenvolve uma crítica à tese de Brentano que sustentava a ideia de que a intencionalidade continha de forma imanente um objeto como representação psíquica do objeto real. Ao contrário do que sustentava Brentano, o objeto é sempre transcendente à intencionalidade e se manifesta na própria intencionalidade somente como componente não real, isto é, se manifesta como significado, sentido. Não-real deve ser entendido aqui como ideal; reais são os atos subjetivos, que emergem como componentes da intencionalidade que se correlacionam à intuição desta idealidade. Nas palavras de Husserl:

É um grave erro estabelecer uma distinção real entre os objetos “meramente imanentes” ou “intencionais” e os objetos “transcendentes” ou “reais” que lhes corresponderiam eventualmente, ora se interprete essa distinção como uma distinção entre um signo ou uma imagem existente realmente na consciência e a coisa designada ou representada em imagem, ora se interprete o objeto imanente de outro modo qualquer, como um dado real da consciência [...]. O objeto intencional da representação é o mesmo que seu objeto real e – conforme o caso – que seu objeto exterior, e é um contrassenso distinguir entre ambos. O objeto transcendente não seria o objeto dessa representação, se não fosse seu objeto intencional. Em suma, se compreende que esta é uma proposição meramente analítica. (HUSSERL, 1929)

Na doutrina da intencionalidade proposta por Brentano, os objetos percebidos, fantasiados, julgados, desejados, etc. “entram na consciência”, “são recebidos na consciência” ou, de maneira inversa, “a consciência” entra em relação com eles. Husserl diante de tal doutrina sinaliza para duas possíveis falhas ou más interpretações: (i) que se trataria de uma relação real pertencente ao conteúdo real da vivência; (ii) que se trataria de uma relação entre duas coisas encontradas na consciência, de acordo com a expressão escolástica “in-existência intencional”.

No direcionamento intencional da consciência, isto é, na intencionalidade não há uma referência ou um processo real que tenha lugar entre a consciência de um lado e a coisa

consciente no outro. Não há também uma relação entre duas coisas que se encontram realmente da mesma maneira na consciência, isto é, um estado psíquico em que se encontram um objeto intencional e um objeto real, dois conteúdos da consciência conectados um ao outro. Apenas uma coisa é presente: a vivência intencional, cujo carácter descritivo essencial é precisamente a intenção.

A in-existência intencional apresentada por Brentano, compreendida em seu sentido original de "existência *in mentis*", necessita ser compreendida no sentido de existente em intenção, no entanto, essa existência não se dá de forma natural na mente, em ipseidade (carne e osso). Husserl deixa claro que as vivências intencionais referem-se de modos diferentes a objetos representados, no entanto, o "objeto é nelas 'visado', é 'tido em vista'". Devemos nos atentar ao fato de que não há duas coisas presentes na vivência: o objeto e a vivência intencional. O que temos é a vivência intencional de um dado objeto, e visá-lo é uma vivência, até mesmo pelo fato de que o objeto pode simplesmente não existir *extra mentis*. Para compreender melhor isso que estamos buscando esclarecer utilizaremos um exemplo similar ao que Husserl apresenta em suas *Investigações Lógicas*. Pensemos na figura de Odim. Se me represento Odim, o objeto representado tem in-existência intencional. Isso vale que tenho uma determinada vivência de representação daquilo que, em minha consciência, se consuma em um representar Odim, sem, no entanto, encontrar aí "naturalmente" algo como o próprio Odim. Mesmo que ele não seja algo *extra mentem*. O fundamental aqui é que, do ponto de vista da experiência fenomenológica, a existência ou não do objeto (no caso Odim) em nada muda a situação: "Para a consciência, o dado é essencialmente idêntico, quer o objeto representado exista, quer seja ficcionado, quer seja mesmo um contrassenso" (HUSSERL, [1900-1901] 1929).

Em suma, o fato de Husserl adotar a expressão "objeto intencional" está pautado na questão de que ela aponta para uma forma de representação que é real, na vivência, sem que tal implique necessariamente uma existência natural *extra mentis* dos objetos intencionados. Os objetos intencionais, enquanto visados, sempre existem na consciência. Dessa maneira, pode a consciência, portanto, exprimir-se sobre este ser, do modo como ele se denota, elucidando a forma pelo qual ela o visa. Para isto, não tem necessidade de sair de si, mas somente de proceder ao exame destes modos de intenção.

3.

Entrando agora no terreno da filosofia analítica buscaremos fazer uma rápida abordagem do tema da intencionalidade em Searle, onde tal teoria, ocupa um dos papéis centrais. Ela possui de algum modo, uma evolução paralela a tudo que Searle produziu em filosofia da mente. Mas não podemos deixar de lado sua relação estrita com todo o propósito investigativo desenvolvido pelo autor. Podemos dizer que a teoria da intencionalidade em Searle está situada em continuidade com sua teoria dos atos de fala. Compreender a ideia de intencionalidade do referido autor não é um trabalho fácil, dado que esse pressupõe um naturalismo biológico, dificultando na compreensão do que ele entende por consciência e o que ele compreende por intencionalidade⁴⁸. Segundo o *Dicionário de Obras Filosóficas* (2002) a teoria da intencionalidade apresentada por Searle, deve ser compreendida como “a capacidade biológica fundamental do espírito de pôr o organismo em relação com o mundo” (p. 304). Isto é, a intencionalidade emerge como aquilo que nos possibilita relacionarmos com o mundo, dado que se dirige a, ou se refere a algo. Tal direcionalidade representa a característica fundamental de todo e qualquer estado intencional. Afim de obter um estado intencional (seja ela qual for), faz-se necessário determinadas condições de satisfação, ou seja: veracidade (qualquer objeto intencional necessita existir), direção de ajuste (Existem dois modos, mente-mundo e mundo-mente), conteúdo (aquilo que da condição para que o estado intencional seja acerca de algo) e modo psicológico (crença, desejo, intenção, esperança etc.). Dessa maneira, pode haver estados intencionais que possuam o mesmo conteúdo e a mesma direção de ajuste, no entanto seus modos psicológicos são diferentes. Podemos exemplificar da seguinte maneira: o ato de levantar da cama e sair do quarto tem como seu conteúdo a ação (levantar da cama e sair do quarto) e como sua direção de ajuste mundo-mente, contudo o modo psicológico pode alterar, ainda que seja conservando esses dois aspectos. Levantamos da cama e saímos do quarto porque *cremos* que deitar no sofá na sala estará melhor, ou ainda, temos a *intenção* de ir ao banheiro, ou *esperamos* encontrar alguém e assim por diante.

⁴⁸Há algumas ambiguidades no pensamento de Searle nesse ponto, dado que há necessidade de pressupostos da neurobiologia, dos quais não é possível ser assegurado pelas ciências. Contudo surge uma confiança de certo modo ingênua de que a ciência neurobiológica venha em algum momento autenticar as teses de seu naturalismo biológico.

4.

Outro aspecto que já havíamos adiantado anteriormente e que faz jus ser novamente salientado aqui é a articulação que Searle faz entre intencionalidade e Naturalismo Biológico. O filósofo afirma que a intencionalidade é produto da seleção natural e, por conta disso, evoluiu juntamente com todo o sistema biológico humano, estando inserido dessa maneira no âmbito do mundo físico. Temos, portanto, que, se a consciência é compreendida como um fenômeno natural, necessariamente a intencionalidade deverá ser compreendida da mesma forma – pelo fato de ser um aspecto da própria consciência. Afim de enfatizar o aspecto biológico da intencionalidade, Searle nos apresenta, uma família de noções.

Em um primeiro momento, temos dentro dessas famílias de noções, as formas biologicamente basilares de intencionalidade, representadas pela: percepção e ação. Nas percepções, o que surge como o mais característico, — enfatizando dessa maneira que elas são características básicas da intencionalidade — são as condições de satisfação, decorrentes de qualquer ato perceptivo. Há sempre no ato perceptivo uma direcionalidade (toda percepção é sempre percepção de algo) bem como uma direção de ajuste (mente-mundo), constituindo dessa maneira, características essenciais para todo e qualquer ato intencional. Para além disso, em um âmbito valorativo, pode ser dito que a percepção estabelece uma íntima relação com o *background*⁴⁹, crenças e desejos. Percebemos objetos, e os percebemos da maneira como eles nos apresentam, contudo estabelecemos uma relação subjetiva com os mesmos. De outra maneira, mesmo sendo a percepção uma das formas fundamentais de intencionalidade, ela estabelece relações com outras formas mais “complexas”, ou superiores, de intencionalidade, como por exemplo as crenças e os desejos. Diante disso, como infere o autor, podemos aceitar que na percepção não temos apenas o ato perceptivo em si, mas, junto deste, há também um modo psicológico de apresentação de qualquer percepção (SEARLE, 1995).

⁴⁹“O *background* é um conjunto de capacidades mentais não-representacionais que permite a ocorrência de toda representação. Os estados intencionais apenas tem as condições de satisfação que têm e, portanto, apenas são os estados que são sobre um *Background* de capacidades que, em si mesmas não são estados intencionais. Para que eu possa ter agora estados os intencionais que tenho, preciso ter determinados tipos de saber práticos (*know-how*): preciso saber como as coisas são e preciso saber como fazer as coisas, mas esses tipos de “saber como” (*know-how*) em questão, não são, nesses caso, formas de “saber que” (*know-what*)” (SEARLE, 1995)

5.

No que tange às ações intencionais, Searle (1995) defende que da mesma forma como uma crença é satisfeita se, e somente se, o estado de coisas representado pelo conteúdo do desejo venha a ocorrer, bem como, uma intenção só é satisfeita se, e somente se, a ação representada por seu conteúdo vier de fato a ser realizada. Isso significa que uma ação intencional, consiste meramente às condições de satisfação de uma intenção. Tomemos o seguinte exemplo, derramar uma bebida não dispõe ser a condição de satisfação de uma intenção, dado que as pessoas geralmente não derramam bebidas intencionalmente, contudo, ainda assim, algo dessa natureza poderia ser uma ação intencional, uma vez que de alguma forma poderia ser a condição de satisfação de uma outra ação. Cabe nesse ponto uma questão demasiado importante. Como relacionar intenção e ação? Ou seja, toda ação seria intencional? Para Searle (1995) há dois tipos de ações intencionais, (i) ações em que a intenção é precedente à ação, e (ii) ações que transcorrem sem que haja uma intenção pregressa. No entanto qualquer ação intencional possui intenções em suas ações, sejam elas precedentes ou não. Isso fica mais nítido se pegarmos o exemplo que o próprio Searle nos apresenta: Bill pretende matar seu tio utilizando uma arma de fogo. Bill encontra-se muito nervoso com a referida situação e, no momento do disparo ele acaba errando seu alvo. Todavia, com o barulho emitido pelo disparo, uma vara de porcos se assusta e esses acabam, pisoteando acidentalmente o tio de Bill, resultando em seu óbito. Agora o objetivo de Bill está concretizado, contudo, não foi concretizado da maneira como ele havia planejado. A intenção de Bill não era de assustar os porcos, mas sim matar seu tio com a arma de fogo. Ao planejar a morte do tio, Bill tinha uma intenção em ação (atirar com a arma de fogo e, por consequência, matar seu tio), no entanto mesmo com sua ação malograda, o objetivo final fora satisfeito.

Essa questão nos leva automaticamente a um outro problema: no que se refere ações não intencionais, seriam essas destituídas de intencionalidade?

Não pode haver nenhuma ação, nem mesmo não-intencional, sem Intenção em ação. Portanto, as ações necessariamente contêm intenção em ação, mas não necessariamente são causadas por intenções prévias. (SEARLE, 1995).

E o exemplo oferecido por Searle da morte do tio do Bill nos auxilia na compreensão da questão acima colocada. Bill não tinha intenção de assustar a vara de porcos com seu disparo, todavia isso possibilitou com que seu objetivo final fosse alcançado. Diante disso, podemos

compreender que, mesmo ações completamente malogradas – que não ocasionem em um desfecho acidental, como no caso de Bill) têm intenções em ação. Para finalizar podemos tomar outro exemplo: derramar bebida. Normalmente esse tipo de fato não costuma ser uma ação intencional, no entanto podemos perceber a intenção em ação de se agir de uma outra maneira, mas que resultou em um acidente.

Referências

- BRENTANO, F. *Psicologia desde um ponto empírico*. Madri: Universidade Complutense, 1935.
- HUISMAN, D. *Dicionário de Obras Filosóficas*. São Paulo: Martins Fontes, 2ª ed., 2002.
- HUSSERL, Edmund. *A Ideia da fenomenologia*. Tradução A. Morão. Lisboa: Edições 70, [1907] 2008.
- _____. *Ideias para uma fenomenologia pura e para uma filosofia fenomenológica*. Tradução Márcio Suzuki. São Paulo: Ideias e Letras, [1913] 2006.
- _____. *Investigaciones Lógicas*, 1. Madri: Alianza Editorial, [1900-1901] 1929a.
- _____. *Investigaciones Lógicas*, 2. Madri: Alianza Editorial, [1900-1901] 1929b.
- SEARLE, J. *Intencionalidade*. São Paulo: Martins Fontes, 1995.
- _____. *Mente, Cérebro e Ciência*. Edições 70, Lisboa, 1997.
- _____. *O Mistério da Consciência*. São Paulo: Paz e Terra, 1998.
- _____. *A Redescoberta da Mente*. São Paulo: Martins Fontes, 2006.
- SILVA, P. Intencionalidade: Mecanismo e Interação. *Principia*, Santa Catarina, v.14, n.2, p. 255-278, ago.2010.

**Paul Ricoeur e a investigação sobre o sujeito:
considerações sobre o *cogito* ferido e a mediação da psicanálise**

*Paul Ricoeur and the research on the subject:
remarks about the wounded cogito and the mediation of psychoanalysis*

João Bosco Batista

RESUMO: O presente trabalho é resultado de uma pesquisa que busca investigar a temática do sujeito na “filosofia reflexiva” de Paul Ricoeur. Como procedimento metodológico, optamos pelo recorte da fase da obra do autor (1960-1969), na qual retoma o problema da subjetividade pelo viés psicanalítico, notadamente em *Da interpretação: ensaios sobre Freud*. Nossa argumentação consiste em demonstrar a relevância desta problemática tal como é apresentada por Ricoeur. A nosso ver, o autor, num rasgo de originalidade, consegue por meio da hermenêutica simbólica, por ele adotada, apresentar uma reflexão filosófica sobre a psicanálise e apropriar-se de um discurso que se beneficia dela, enquanto movimento de desconstrução da consciência imediata e, com destreza, aponta uma saída para a aporia em que se encontrava o sujeito. A denegação do sujeito, potencializada pela psicanálise, é habilmente submetida à reflexão crítica e instrumentalizada estrategicamente para servir ao discurso de valorização de uma nova concepção do sujeito. Diríamos que das cinzas do sujeito (ou cogito) cartesiano, deixadas pela desconstrução freudiana, Ricoeur faz renascer uma nova concepção da existência humana ou da subjetividade, depurada de sua falsidade ou imediatez. Que sujeito é esse que é resgatado pela hermenêutica ricoeuriana? Qual a importância dos símbolos culturais e religiosos para a empreitada reflexiva de construção de um novo conceito de sujeito ou de consciência em Ricoeur? Como a psicanálise serviu aos propósitos do filósofo? Tais problemas norteiam nossa investigação da filosofia “prática” que propõe o autor, na qual a consciência é uma tarefa inacabada, um esforço ou desejo de ser.

PALAVRAS-CHAVE: Ricoeur; Psicanálise; Sujeito; Hermenêutica.

ABSTRACT: This paper is a result of a research which sought to investigate the subject theme in the Paul Ricoeur’s “reflexive philosophy.” As the methodological procedure, we chose the period between the 1960’s and 1970’s of the author’s work, in which it takes over the subjectivity issue through the psychoanalytic bias, notably inside “From the interpretation: an essay about Freud. Our argumentation consists unto demonstrate the relevance of this problematic as it is presented by Ricoeur. In our view, the author, in an insight of creativity, achieves through the symbolic hermeneutics, by him adopted, to present a philosophy reflection about the psychoanalysis and to appropriate of a speech which benefits of that, as a movement of deconstruction of the immediate consciousness, and, with cunning, points to a way out of the aporia in which was the subject. The subject’s denial, potentialized by the psychoanalysis, is ably submitted into critical reflection, and strategically instrumentalized in order to serve a speech of valorization coming from a new subject’s conception. We could say, that, from the cartesian subject ashes (or cogito), left by Freud’s deconstruction, Ricoeur makes a new conception of the human being existence or subjectivity reborn, purified from his falseness or immediacy. What subject is this who is rescued by the Ricoeur’s hermeneutics? What is the importance of the religion and cultural symbols to the reflective endeavor of a new subject or consciousness’ concept by Ricoeur? How did the psychoanalysis serve to the philosopher proposes? Such problems guide our investigation of the philosophy’s “practice” proposed by the author, in which the consciousness is a work in progress, an effort or a wish of being.

KEYWORDS: Ricoeur; Psychoanalysis; Subject; Hermeneutics.

Introdução

Paul Ricoeur (1913-2005), um dos maiores representantes da “hermenêutica francesa”, teve uma participação singular na filosofia contemporânea quanto ao tema do sujeito, empreendendo uma hermenêutica do homem na história em suas multifárias manifestações, desde o problema da linguagem e da narrativa, até problemas de cunho mais existencial, como os temas da morte, da angústia, da culpabilidade, entre outros.

A sua hermenêutica procura superar o sujeito cartesiano dicotomizado da realidade e substancializado em *cogito*. Sua atenção volta-se prioritariamente para as questões do homem enquanto indivíduo situado e contextualizado por uma malha de signos que precisam ser interpretados, visando a uma maior compreensão da existência humana e de seus problemas fundamentais. Ricoeur é considerado um dos mais autorizados intérpretes da nossa estrutura cultural e do entendimento que o homem contemporâneo possui da própria vida. Ele aborda com profundidade problemas filosóficos e teológicos, colocando no cerne de suas indagações o problema da significação dos feitos e fatos humanos – o homem é o problema fundamental que permeia sua chamada “filosofia reflexiva”.

O objeto de nossa pesquisa situa-se na primeira etapa do pensamento de Ricoeur (1960 a 1969), onde aparecem dois temas bem delimitados: a hermenêutica dos símbolos e a psicanálise. Três são os livros que compreendem este período: *A simbólica do mal*, *Da interpretação: ensaio sobre Freud* e *O conflito das interpretações: ensaios de hermenêutica*. O livro *Da interpretação: ensaio sobre Freud*, será o objeto específico de nossa investigação, por ser o estudo no qual o autor apresenta o encontro da hermenêutica dos símbolos e da psicanálise. Nesta obra, Ricoeur apresenta uma leitura precisa da obra de Freud. Uma leitura que se caracteriza por enxergar na fala psicanalítica uma dualidade de discurso: Freud se referia à experiência humana em termos de “sentido” e “força”. A característica da psicanálise para Ricoeur se encontra na recusa em optar por uma análise do homem somente em termos de força, ou somente em termos de sentido. Para o filósofo, a psicanálise é tanto uma hermenêutica que compreende os sentidos quanto uma “energética” que explica as forças em jogo na psique humana.

Para um filósofo formado na escola da fenomenologia, das filosofias existenciais e do retorno aos estudos hegelianos, o encontro com a psicanálise constitui inegavelmente um abalo

considerável. Tal é o caso de Paul Ricoeur, cuja obra publicada em 1965, *De Interpretação: ensaio sobre Freud*, expressa, de forma significativa, o projeto da reflexão filosófica posto em questão em seu conjunto pela demolição psicanalítica: “o que é posto em questão de maneira geral e radical é aquilo que nos aparece, a nós bons fenomenólogos, como o campo, como o fundamento, como a própria origem de toda significação, quero dizer, a consciência” (RICOEUR, 1969, p. 100). Assim como Marx e Nietzsche, Freud levanta-se diante dos filósofos como um *mestre da suspeita* (expressão comum tanto a Michel Foucault quanto a Paul Ricoeur), fazendo despontar o problema da consciência como ilusão. Importa, na presente reflexão, avaliar a importância da leitura de Freud para o desenvolvimento do projeto hermenêutico em filosofia, tal como concebido por Ricoeur.

1. O *Cogito* partido: a consciência desapropriada de si mesma

Desde o início de sua trajetória em *O voluntário e o involuntário*, Ricoeur se situa num combate às versões mais exacerbadas do idealismo, em particular à pretensão de autossuficiência da consciência de si, para ressaltar os limites dessa tentativa. Isso se evidencia por meio da aceitação dos limites apresentados pela crítica kantiana, onde ficam expostas as demarcações intransponíveis da racionalidade e da linguagem humanas, sob pena de cair nas aporias ou, pior, na *hybris* (desmedida) de um pensamento que se auto institui em absoluto. Nessa empreitada de oposição aos exageros da tradição idealista, Ricoeur aponta como *leitmotiv* de sua filosofia, uma pesquisa simultaneamente sóbria e apaixonada das relações da consciência e do sujeito com o mundo que os circunscreve e os constitui por inúmeros laços. Diríamos, em termos heideggerianos, que Ricoeur se debruça sobre a pesquisa das inúmeras maneiras do sujeito humano habitar o mundo e torná-lo mais habitável. À “exaltação do Cogito” ele opõe um *Cogito* “desancorado” ou “partido” como escreve no prefácio a *O si-mesmo como um outro* (RICOEUR, 1991, p. 22), obra em que ele busca desenvolver uma “hermenêutica do si”.

A abordagem de Ricoeur da problemática do sujeito é eminentemente hermenêutica, interessa-lhe elaborar o próprio conceito de interpretação. Ele observa que seus primeiros escritos repousavam sobre uma noção “amplificante” de interpretação: uma interpretação atenta ao acréscimo de sentido incluído no símbolo e que a reflexão tinha por tarefa liberar. O

confronto com os “três grandes mestres da suspeita”, Freud, Marx e Nietzsche – através do diálogo com a psicanálise, com o estruturalismo, entre outros marxistas, e com o pós-heideggerianismo – leva-o a admitir outra possibilidade de interpretação: uma interpretação redutora, isto é, de denúncia das ilusões, genealógica no sentido de Nietzsche, de crítica ideológica em Marx, de descoberta do recalque e da repressão em Freud. Dessas leituras, portanto, Ricoeur tira ainda mais munição para sua empresa de desmistificação das pretensões teóricas totalizantes. Por essa mediação dos mestres da suspeita, seu pensamento se depara com instrumentos privilegiados de análise da relação temporal e histórica que subjaz à prática hermenêutica. Como as manifestações culturais, individuais ou coletivas não se constituem a partir de uma produção linear e tranquila de sentidos acumulados, mas surgem também de conflitos, de deslocamentos, de disfarces e de transferências, assim também a relação entre o presente do intérprete e o passado mais ou menos longínquo da obra interpretada não se resume à mera relação de aceitação e de transmissão. A própria transmissão da tradição obedece a motivos e interesses diversos, explícitos ou implícitos, tematizados ou inconscientes, que interferem no processo hermenêutico enquanto tal. No processo interpretativo confrontam-se sempre dois mundos, o da obra e o do intérprete. Ambos devem ser refletidos. A dinâmica da compreensão comporta, porém, certo apagamento do intérprete em favor da obra; uma “desapropriação de si” para deixar o texto, por exemplo, nos interpelar na sua estranheza e não só nos tranquilizar naquilo que nele projetamos, mas também produzir, graças ao confronto entre o universo do intérprete e o universo interpretado, uma transformação de ambos. O processo hermenêutico, poderíamos dizer, desapropria duplamente o sujeito da interpretação: obriga-o a uma ascese primeira diante da alteridade da obra; e, num segundo momento, desaloja-o de sua identidade primeira para abri-lo a novas possibilidades de habitar o mundo (GAGNEBIN, 1997).

É nessa perspectiva que Ricoeur coloca-se diante da psicanálise: como hermeneuta e filósofo da reflexão. O seu encontro com a obra de Freud é permeado pelo anseio interpretativo com vistas à ampliação do dimensionamento do homem no mundo da cultura, da linguagem e dos símbolos: ou seja, do sujeito concreto em um mundo sempre aberto a novas possibilidades e diferentes leituras. Horizontalidade e verticalidade se dialetizam na concreção da existência: o horizonte da história e sua inerência na verticalidade do sagrado. Ricoeur explora a influência que a psicanálise exerce sobre a hermenêutica e vice-versa. Ou seja, além de uma interpretação

da psicanálise, Ricoeur pensa ser possível falar de uma psicanálise da hermenêutica, isto é, uma elucidação de uma possível base pulsional nas disciplinas interpretativas. Há, portanto, entre psicanálise e hermenêutica uma relação de reciprocidade.

O freudismo apresenta uma difícil questão para o filósofo: o questionamento do sujeito pensante. O sujeito jamais é aquilo que se pensa crer que ele seja. Há um desapossamento radical do *Cogito* cartesiano. Freud apresenta a nudez da consciência narcísica imediata, tal como aparece em Descartes e perpassa o pensamento moderno. Em outras palavras, a hermenêutica da suspeita do sujeito desmorona a apoditicidade do “eu penso” que envolve, por sua vez, a dúvida quanto à natureza do “eu sou”. O resultado da hermenêutica psicanalítica desvela a impossibilidade de uma apreensão direta da consciência de si e por si.

Ricoeur reconhece a justeza da crítica psicanalítica. Não pensa que seja possível continuar a fazer filosofia como antes. Mas não pensa que seja necessário encerrar a reflexão filosófica. Para ele, depois do advento da psicanálise, a reflexão filosófica precisa incorporar os recursos do próprio freudismo como instrumento da reflexão e da crítica da consciência imediata. Ele percebe que através da reflexão pós-analítica, a filosofia precisa abandonar a ingenuidade e a segurança da consciência imediata. A consciência passaria a ser vista como uma entre outras localidades psíquicas. A reflexão deve agora se apropriar do instrumental interpretativo oferecido pela metapsicologia – com isto, a interpretação do desejo passaria a ser um dos instrumentos da reflexão.

Nessa etapa do descentramento da consciência operada pela psicanálise, além de se ter que reconhecer o inconsciente como localidade psíquica, torna-se também necessário reconhecer a presença do que se denominou de *narcisismo*. Com a introdução do narcisismo, o ego deixaria de ser o sujeito do *Cogito* e passaria a ser considerado como um objeto de desejo. O narcisismo surge como uma grandeza metafísica, um gênio maligno que resiste profundamente à pretensa verdade sobre o indivíduo. Uma das verdades que o narcisismo nos impede de ver é a de não conhecermos quase nada acerca de nós mesmos e que não somos senhores de nossa própria psique. Descobre-se, assim, o paradoxo de que só mediante um processo de humilhação, é possível ao ego aprender alguma coisa sobre si mesmo.

Com a desconstrução da consciência imediata, realizada pela psicanálise, fica patente que a consciência humana não pode pensar-se como puro princípio ou ponto de partida. Pelo contrário, interpreta Ricoeur, ela é apenas uma antecipação que deve como tal realizar-se como “desejo de ser”, sabendo de antemão que tal esforço é uma tarefa inacabada. Devido à sua condição encarnada, o eu surge disseminado e perdido; isto significa justamente que não possui desde logo aquilo que é. Por isso é obrigado a investir-se em signos contingentes e opacos, que o mediatizam em seu ato de ser.

Vimos que a metapsicologia de Freud sugere à filosofia o desapossamento da consciência como caminho e o tornar-se consciente como *tarefa*. Caso este projeto seja aceito pela filosofia, então o único *Cogito* possível é um *Cogito* que não se possui, que aceita não compreender a sua própria verdade originária, um *Cogito* que lute contra a ilusão e a mentira da consciência atual. Ricoeur reconhece a inadequação do conceito de consciência imediata, mas não se contenta com tal constatação, pensa que é preciso continuar com a reflexão. Depois da “humilhação” da consciência, não é mais possível vislumbrá-la como algo em si e existente por si mesmo. O modelo pulsional da psicanálise sugere que ela deve ser entendida pelo desejo que a sustenta. Dessa forma, a consciência experimenta seu primeiro “trabalho de luto”, pois é humilhada e abandonada como centro da estrutura da existência humana. Diante da inexistência da consciência imediata, evidencia-se a necessidade da interpretação para a verdadeira autocompreensão. Fica claro para Ricoeur que o ato de existir afirma-se na diferença e na relação com outros atos, exprime-se por meio de obras e sinais.

2. A mediação da psicanálise

Em 1965 Paul Ricoeur publicou o livro *Da interpretação*, que destaca a questão da consciência. Segundo o autor, é com os *mestres da suspeita* que a consciência será pensada por um prisma radicalmente distinto daquele oferecido pela filosofia moderna. Nesta nova abordagem, o sujeito deve aprender a duvidar da consciência, assim como Descartes havia ensinado a duvidar de tudo. Ricoeur designa esta mudança de *antifenomenologia*, a qual realiza uma *epoché ao contrário*. Enquanto a *epoché* husserliana era uma redução de tudo o que é transcendente, uma redução à consciência, Freud realiza uma redução da consciência enquanto

tal, colocando em questão a posição central da consciência imediata como origem da significação. Esta reviravolta inflige uma “ferida” ao *Cogito* – isso representa a humilhação do sujeito da reflexão. A redução fenomenológica ainda atesta a mestria do ego, enquanto este suspende ou coloca entre parêntesis tudo o que está fora de si. Por sua vez, a suspensão freudiana é mais radical, pois coloca entre parêntesis a própria consciência, chamada de consciência imediata ou autoevidente. Nos “mestres da suspeita” as análises a respeito da realidade e de sua constituição, são conduzidas por caminhos divergentes, mas em todos eles o resultado é semelhante: a consciência é questionada em sua própria constituição ontológica – o que comumente compreendemos como consciência não passa de uma falsa consciência.

A busca de Ricoeur por uma interpretação filosófica da psicanálise deve-se ao reconhecimento da necessidade deste teste rigoroso para as ilusões da consciência, que permite então à reflexão sair da abstração e prosseguir o caminho da autocompreensão concreta. Ricoeur utiliza, portanto, da hermenêutica de Freud como um guia num duplo processo para chegar à sua meta.

A primeira fase é a do despojamento, um caminho regressivo que vai desde a pretensão da autoevidência da consciência até à origem do sentido no desejo, anterior à divisão entre sujeito e objeto. Este caminho requer a renúncia da consciência como detentora do sentido; a “humilhação” do *Cogito* é necessária se quisermos ser verdadeiramente despojados das nossas ilusões acerca da imediatez da autocompreensão. Apercebermo-nos, portanto, que não existem atalhos para o conceito e a consequente compreensão do sujeito; há apenas o *longo caminho da consciencialização*. Este longo caminho constitui a segunda fase, a da reapropriação.

Nesta etapa, Ricoeur avança em seu trabalho analítico com o intuito de redescobrir de alguma forma o sujeito ou a subjetividade que fundamenta o ser do homem. Como veremos adiante, para Ricoeur não há uma subjetividade em si, pura. Ela se dá e se apresenta sempre na mediação com o outro – o mundo. O filósofo é aquele que busca um sentido para a existência do sujeito ou do homem no mundo. O sentido terá que ser arrancado das ruínas da auto ilusão, através da atividade da interpretação, ou seja, de uma hermenêutica da redescoberta do sujeito pela mediação do mundo. Vejamos que o novo “sentido” a que almeja Ricoeur não pode mais ser descrito em termos de autoevidência ou apoditicidade. O sentido não é imediato ou *a priori*, mas o resultado de todas as mediações pelas quais deve passar o trabalho da interpretação. Em outras

palavras, o sentido não é um frente a frente de um ego isolado ou encapsulado, mas, ao contrário, o produto de um processo intersubjetivo de reconhecimento da alteridade como constitutiva de sua “identidade”.

É por meio da leitura de Freud que Ricoeur aborda o conflito das interpretações. Este é o ponto crucial da sua obra *O conflito das interpretações: ensaios de hermenêutica*, publicada em 1969. Neste livro, Ricoeur nos mostra que os símbolos possuem uma relação importante com a interpretação, além de proporcionarem um conhecimento de nós mesmos. Os símbolos encontram-se sempre referidos a alguma coisa e revelam uma verdade que através da linguagem, torna-se instrumento de mediação. Para que se possa compreender um símbolo é necessário interpretá-lo, e é pela interpretação que a diversidade dos sentidos se torna manifesta e com ela, a multiplicidade dos sentidos.

A reflexão de Ricoeur suscita a importância de se levantar o problema da subjetividade nos dias atuais. Ousamos afirmar que o filósofo francês apresenta um contributo relevante à tematização do sujeito ou da subjetividade em meio a um contexto que apresenta a crise do humanismo e, conseqüentemente, do sujeito. Levantamos uma indagação, a nosso ver pertinente, para se avaliar a presença do filósofo na atualidade: é ainda possível falarmos de sujeito e de subjetividade no pensamento atual que superou a noção de *Cogito*, tal como fundamentou o pensamento moderno, oriundo da reflexão cartesiana? Certamente, esta é uma questão que não passa despercebida no pensamento ricoeuriano. Ele teve plena consciência de que estava pisando em um terreno minado ao abordar o tema do sujeito pelo viés da hermenêutica. Não pretendeu “ressuscitar” o sujeito cartesiano e suas características modernas: obviamente o autor levou em consideração a crítica de Heidegger à metafísica e ao humanismo ilustrado como constituinte da estrutura da *weltanschauung* moderna, na qual o método científico assumiu a orientação e o papel modelar do pensamento, da verdade e da visão histórica e cultural até o século XIX.

Para o filósofo francês, o homem é acima de tudo aquele que fala, aquele que *é* fala. O problema da linguagem tem um lugar fundamental no percurso hermenêutico que constitui o arcabouço de seu pensamento. A linguagem para Ricoeur é de importância crucial para a compreensão das manifestações humanas no mundo. Pode-se dizer que ela antecede a cada um de nós, pois ela sempre já está aí presente num mundo falante – fenomenologicamente

afirmamos que o mundo nos fala. A linguagem recolhe o olhar mudo e articula seu sentido. Tal expressividade de sentido é o que, afinal, determina a linguagem em sua essência.

2.1 Ricoeur analisa Freud.

É importante frisar que Ricoeur lê Freud como quem lê um pensador; ele faz filosofia, não psicologia. Ricoeur não dispõe da experiência analítica nem como terapeuta, nem como paciente. Ele investiga a condição de discurso comunicável tal como está disposta na escrita da psicanálise. Fundamentalmente, trabalha a consistência do discurso freudiano. Lê e interpreta os escritos de Freud, como quem lê e interpreta um pensador clássico. Para ele, ainda que Freud não tenha a formação de filósofo, muitos de seus conceitos podem ser comparados com a tradição reflexiva ocidental. Portanto, para entender a leitura que Ricoeur faz de Freud, é preciso compreender também que esta leitura está circundada por um projeto filosófico. O filósofo francês estuda Freud com o intento de elaborar hipóteses e enriquecer sua condição de intérprete dos símbolos e filósofo da reflexão. A psicanálise, enquanto obra escrita, o fascina, pois ela tem uma maneira própria de ler e interpretar os símbolos e a fala humana, além de propor difíceis questões à filosofia a partir de sua descoberta do inconsciente.

Não é do interesse de Ricoeur tratar do Freud pré-psicanalítico ou do período de formação da psicanálise. O Freud que trata exclusivamente de problemas neurológicos, no período da anatomia e da fisiologia, não interessa ao estudo de Ricoeur.

Para Ricoeur a “primeira tópica” é solipsista, pois não leva em conta a condição social humana. O aparelho psíquico só estaria em condição social na “segunda tópica”. Apesar desta consideração, Ricoeur não nega que a psicanálise leva em conta o outro e a cultura desde seu início, ainda que de modo latente. Por exemplo, a situação transferencial da terapia e a trama edípica, marcos fortes da psicanálise, não podem ser pensadas desde um enfoque solipsista e nem tampouco nasceram com a segunda tópica. Não há como negar que o surgimento da psicanálise coloca em xeque a noção de “interior” e de “exterior” da psique. Aliás, o ganho teórico da psicanálise é exatamente este: falar de uma psique sempre em condição social. A cultura é decisiva no “funcionamento” da mente humana. Talvez o que Ricoeur quisesse dizer é que tal estatuto psicanalítico só se fez representar, de fato, plenamente na teoria, a partir da segunda

tópica. De qualquer forma, temos para nós que a psicanálise é interpretação da cultura, ainda que esta seja vista como um elemento no balanço energético do aparelho psíquico do indivíduo.

Ricoeur deixa evidente sua posição crítica sobre a validade e o limite da validade da interpretação psicanalítica. Evidentemente, a psicanálise não pode simplesmente ser aplicada a qualquer fenômeno humano, pessoal ou social; sua interpretação, ainda que penetrante, é reducionista na exata medida de sua teoria. Ela só considera seus objetos pelo modelo tópico-econômico. Os fenômenos são sempre vistos dentro da ótica da satisfação do desejo, como no paradigma do sonho. De toda forma, a psicanálise tem seu lugar garantido entre outras formas ou correntes de interpretação da cultura; ela não é a única, mas uma a mais dentre as interpretações da cultura. Ricoeur desmascara qualquer pretensão totalizante do discurso psicanalítico; o mesmo se aplica aos outros discursos.

O estudo de Ricoeur evidencia o fato de que no fim da obra de Freud, a psicanálise passa a sustentar uma cosmovisão. Para Freud, a vida é encarada como uma disputa entre *Eros* e *Tanatos*. A experiência de viver é marcada pela resignação em face do trágico, da necessidade e da própria morte.

Dissemos que Ricoeur possui uma leitura da obra de Freud que se caracteriza por enxergar na fala psicanalítica uma dualidade de discurso: Freud se referia ao funcionamento do aparelho psíquico em termos de *sentido e força*. A psicanálise é tanto uma hermenêutica que compreende os sentidos, quanto uma energética que explica as forças em jogo na psique humana. Esta é a tese de Ricoeur.

Tal tese, na verdade, representa uma solução para uma disputa entre duas leituras concorrentes da obra de Freud. Há quem veja Freud como um naturalista. Segundo essa leitura, sua metodologia de trabalho deveria ser comparada à da biologia ou da física. O homem freudiano, segundo essa corrente, é em essência um complexo pulsional, um produto passivo dos impulsos que vêm do corpo. Nessa visão, de viés positivista, o psiquismo seria “infiltrado” e determinado pelo corpo. Para os adeptos desta leitura, a linguagem mecanicista de Freud é privilegiada. O homem deve ser compreendido a partir de um jogo de forças cegas.

Os críticos dessa visão naturalista de homem hão de dizer que tal concepção deixa de fora exatamente o que há de mais essencial e humano no homem; reducionista, esta visão não

trabalha o sentido da existência humana. Não se trata de dizer apenas que esteja faltando algo na visão naturalista. De fato, outros tantos leitores de Freud vão dizer que sua obra é uma pesquisa constante do sentido, uma busca de significações ocultas. A própria relação terapêutica na psicanálise só pode ser pensada como uma produção de sentidos. O sentido dos sintomas está oculto aos pacientes, as significações que a sessão produz tratam exatamente de devolver o sentido ao paciente. Essa segunda abordagem quer privilegiar a decifração de sentido na obra de Freud. Desta forma, vemos instaurado o conflito: será Freud um naturalista ou será um hermeneuta? A psicanálise deve ser compreendida como um jogo de forças ou como uma decifração de sentidos?

Diante destes dois padrões de discurso, as correntes acima mencionadas têm tentado separar um Freud de outro. Quem quisesse fazer uma leitura seletiva de Freud, não teria dificuldade para sustentar uma ou outra dessas visões. Na verdade, tal separação de “dois Freuds” representa uma enorme violência ao texto freudiano. As duas linguagens estão ali o tempo todo. Nem se pode dizer que haja uma direção ou um movimento aonde a linguagem vá superando a outra. Há, de fato, momentos mais energéticos e momentos mais hermenêuticos. Mas os momentos energéticos são seguidos por momentos hermenêuticos, e vice-versa. O texto freudiano não recomenda a separação nem a exclusão de uma das linguagens. Pelo contrário, Freud parece se sentir bem à vontade com seus dois estilos de falar. Parece que para ele não há contradição entre decifração e explicação energética. Ricoeur embrenha-se por este caminho perceptível nos escritos de Freud.

Para Ricoeur, Freud não era ingênuo do ponto de vista epistemológico. Ele acredita que Freud, deliberadamente, reuniu duas ordens de discurso diversas em seu arcabouço teórico: a linguagem da força e a linguagem do sentido. Dito isso, o problema agora não se localiza em ter que escolher uma linguagem, em detrimento da outra. O problema real é que Freud não possuía as condições necessárias para demonstrar a articulação destas duas linguagens, deixando a questão em aberto.

Não é a pretensão de Ricoeur preencher as duas lacunas deixadas por Freud. Como filósofo e leitor crítico de Freud, ele busca a dialética entre as linguagens da força e do sentido. Quando ele estuda o *Projeto* (1890) de Freud, percebe que este fala de um estado energético de sua teoria, um estado em que não comparece a hermenêutica. A antítese propriamente entre as

duas linguagens emerge em *A interpretação dos sonhos* (1900), onde a hermenêutica floresce. Segundo Ricoeur, o equilíbrio entre força e sentido, seria encontrado nos escritos metapsicológicos. O filósofo pretende mostrar a irredutibilidade do discurso freudiano à linguagem quer do sentido, quer da força.

Para mostrar como estas duas ordens se articulam, Ricoeur se refere ao processo de interpretação dos sonhos, criado pela psicanálise. Desta forma, Ricoeur pretende demonstrar que não há conflito entre as duas ordens de linguagem. Aponta o conceito de *pulsão* em Freud como o lugar privilegiado desta harmonia. Nesse conceito freudiano, ele vê reunido tanto o sentido quanto a força.

No entanto, para compreender a envergadura da tese de Ricoeur é preciso não se ater somente à psicanálise. De fato, como afirmamos acima, o estudo que o filósofo francês empreende da psicanálise é um passo a mais em seu grande projeto filosófico. Na verdade, a psicanálise se insere como um modelo de interpretação que mediatiza o seu plano maior que consiste em ampliar sua filosofia reflexiva. A psicanálise é importante e entremeia sua hermenêutica, principalmente quando essa lhe apresenta os conceitos de causalidade e motivação. Isso lhe propicia verificar a diferença e ao mesmo tempo a relação que há entre explicação e compreensão. Ricoeur pode, ao lado de Freud, reapresentar o problema da ação humana em termos de causa e motivo, ou seja, por meio da explicação e da compreensão. Não há dúvida que a explicação causal tem seu lugar garantido, quando se pretende explicitar o comportamento humano. Por exemplo, sempre que houver constrangimento e força, o comportamento humano deve ser explicado causalmente, mesmo quando o constrangimento é interno. Os motivos inconscientes freudianos, que são interpretados pelo modelo energético, estão muito próximos da causalidade constrangedora. Especialmente em casos patológicos, a compulsão pode ser de tal ordem, na qual o sujeito se sente passivo, uma vítima de seus impulsos internos. Por outro lado, é possível encontrar comportamentos humanos que não são regidos deste modo. Toda vez que surge um legítimo projeto humano, não há uma relação de causa e efeito, mas motivação. Todavia, percebe Ricoeur, junto a Freud, que no dia a dia, o fenômeno humano se situa em um espaço intermediário, entre a causalidade e a motivação, entre a explicação de causas e a compreensão de motivos. Em outras palavras, o ser humano se move por desejo e por necessidade. Ele é membro da natureza e da cultura. Tem um corpo que o une

ao natural e ao mesmo tempo (sem dicotomizar), é uma experiência existencial que gera pensamento e amor. A psicanálise é um instrumento de interpretação que auxiliará Ricoeur na busca de se desvendar o fenômeno humano, o mistério da existência. Sua abordagem da psicanálise está baseada numa hipótese antropológica muito mais abrangente. De forma contundente, o pensador francês pretende evidenciar que os achados freudianos se relacionam e contribuem para o debate filosófico que procura desvelar na filosofia e na ciência os mistérios da vida humana.

Ricoeur explora a influência que a psicanálise exerce sobre a hermenêutica e vice-versa. Ou seja, além de uma interpretação da psicanálise, Ricoeur pensa ser possível falar de uma psicanálise da hermenêutica, isto é, da elucidação de uma possível base pulsional nas disciplinas interpretativas. Há, portanto, entre psicanálise e hermenêutica uma relação de reciprocidade.

No livro III da obra *Da interpretação*, intitulado “A dialética: uma interpretação filosófica de Freud”, Ricoeur procura investigar e solucionar os problemas levantados após uma passagem por Freud. Ricoeur constata o conflito existente entre as hermenêuticas que disputam uma interpretação conclusiva e excludente da temática da cultura. Ele aposta no poder de arbitragem da filosofia nesta disputa das “hermenêuticas concorrentes”. Para ele há uma hermenêutica da suspeita, assim como há uma hermenêutica da restauração do sentido. Elas se encontram em permanente conflito. A psicanálise é o paradigma da hermenêutica da suspeita, embora se mencione os pensamentos de Marx, Nietzsche e Feuerbach, onde cada um à sua maneira se alinha à suspeita como ponto de partida metodológico. Por outro lado, a religião é o paradigma da hermenêutica da restauração do sentido. Ricoeur espera de algum modo, superar tal antítese.

Ambas as interpretações têm um sentido e uma razão de ser. Por exemplo, como abandonar a relevância da crítica da institucionalização, da burocratização, da ilusão e da mentira elaborada pela hermenêutica da suspeita? Por outro lado, como abrir mão da perspectiva do sentido e da esperança, anunciada pela hermenêutica da restauração? Por que ficar só com uma parte e perder a outra, pressionado pelas hipóteses deste ou daquele modelo hermenêutico? Ricoeur espera uma superação. Espera que a filosofia seja capaz de integrar todo e qualquer processo de interpretação dentro dos horizontes da filosofia reflexiva. Para ele a interpretação precisa alcançar um *status* filosófico, sem, porém, perder a concretude; ela não pode ser uma

mera reflexão abstrata, é preciso estar “encarnada” na realidade concreta. Ricoeur procura no “símbolo” o *locus* desta concretude. Para ele, o símbolo além de revelar as experiências reais do ser humano, possui uma textura que remete a uma base pulsional. Por meio do símbolo, a reflexão permanece concreta. No símbolo há uma remissão para frente (futuro) e para trás (passado), abrindo espaço para a interpretação entre a hermenêutica crítica e a hermenêutica afirmadora do sentido da vida.

Quanto à hermenêutica da suspeita, representada pela psicanálise, Ricoeur busca estudá-la por meio de uma filosofia deliberadamente reflexiva (vincula a hermenêutica à reflexão). Trata-se, portanto, de uma reflexão que procura elevar ao nível propriamente filosófico os resultados das descobertas freudianas. Para isso, Ricoeur procura uma ideia guia para pensar o discurso analítico e propõe para isto o conceito de “arqueologia do sujeito”. Tal ideia não deve ser entendida como uma imposição à leitura de Freud, mas como um conceito que ajude a trabalhar o legado freudiano e, ao mesmo tempo, propicie a realização do projeto filosófico do pensador francês.

Conclusão

Com a desconstrução do conceito de consciência imediata, realizada pela psicanálise, patenteia-se que a consciência humana não pode mais pensar-se como puro princípio ou ponto de partida. Pelo contrário, interpreta Ricoeur, ela é apenas uma antecipação que deve como tal realizar-se, isto é, o conceito de consciência tem como base originária o *desejo de ser* (o esforço para ser) e como tal, é sempre uma tarefa inacabada. Devido a sua condição de inconcluso, o eu surge disseminado e perdido – isto significa justamente que não possui desde logo aquilo que é. Por isso é obrigado a investir-se em signos contingentes e opacos (os elementos constituintes da cotidianidade), que mediatizam, realizando-o em seu ato de ser. O eu ou o sujeito não se compreende como uma cápsula abstrata, mas se vê desde sempre enraizado num corpo e na historicidade, na qual este corpo é fundamental e constituinte da tarefa de compreensão de si mesmo.

Referências bibliográficas

- ANDRADE Abrahão Andrade. *Ricoeur e a formação do sujeito*. Porto alegre: EDIPUCRS, 2000.
- CESAR, Constança Marcondes (org.). *A hermenêutica francesa: Paul Ricoeur*. Porto alegre: EDIPUCRS, 2002.
- _____. *Paul Ricoeur: Ensaio*. São Paulo: Paulus, 1998.
- FRANCO, Sérgio de Gouvêa. *Hermenêutica e psicanálise na obra de Paul Ricoeur*. São Paulo: Loyola, 1995.
- GUIMARÃES, André Eduardo. *A paternidade no confronto entre psicanálise da religião e fé*. Porto alegre: EDIPUCRS, 1999.
- RICOEUR, Paul. *Da interpretação: ensaios sobre Freud*. Rio de Janeiro: Imago, 1977.
- _____. *Finitud Y culpabilidad*. Madrid: Editorial Trotta, 2004.
- _____. *O conflito das interpretações: ensaios de hermenêutica*. Rio de Janeiro: Imago, 1978.
- SILVA, Maria Luíza Portocarrero da. *A hermenêutica do conflito em Paul Ricoeur*. Coimbra: Livraria Minerva, 1992.

**Sobre estados mentais relacionais e *qualia*:
o totalmente subjetivo não pode ser conhecido**

*On relational mental states and qualia:
what is completely subjective cannot be known*

Juliana de Orione Arraes Fagundes⁵⁰

RESUMO: Conforme um elenco elaborado por Dennett das características dos *qualia*, eles são estados mentais simultaneamente inefáveis, intrínsecos, privados e diretamente apreensíveis pelo sujeito. Contudo, esse conjunto de características os isola dos outros estados mentais e do comportamento, mantendo-os inescrutáveis a qualquer abordagem objetiva. Disso resultam dois problemas epistemológicos que possuem a mesma origem: o problema das outras mentes e o problema da autoridade do sujeito acerca de seus próprios estados mentais. A origem de ambos os problemas está na tradicional ruptura filosófica entre os mundos objetivo e subjetivo, considerando-os como realidades isoladas. Porém, tal problema pode se desfazer se os *qualia* forem vistos como estados mentais relacionais. Em defesa disso, este artigo busca estender o argumento da triangulação de Davidson para todos os estados mentais. Além disso, apresenta o método heterofenomenológico de Dennett.

PALAVRAS-CHAVE: *qualia*; problemas epistemológicos; triangulação; heterofenomenologia.

ABSTRACT: According to a list of qualia characteristics developed by Dennett, they are mental states that are ineffable, intrinsic, private and directly apprehensible by the subject at the same time. However, this set of characteristics isolates them of the other mental states and from behavior, keeping them inscrutable from any objective point of view. This leads to two epistemological problems which have the same origin: the problem of other minds and the problem of first person authority about her own mental states. The origin of both problems is in the traditional philosophical rupture between objective and subjective worlds, considering them as isolated realities. Yet, such a problem can be minimized if qualia are viewed as relational mental states. In defense of this point, this work tries to extend Davidson's triangulation argument to all mental states. In addition, Dennett's hetero-phenomenological method will be presented.

KEYWORDS: qualia; epistemological problems; triangulation; hetero-phenomenology.

⁵⁰ Professora do DFCH / UESB; Doutoranda do PPGF / UFBA. julianadeorione@hotmail.com.

1. Considerações iniciais: *qualia* e outras mentes

Quando falamos de nossos próprios pensamentos, temos um grau bem alto de convicção, mas quando atribuímos pensamentos aos outros, a dúvida e a possibilidade do erro se fazem presentes. O filósofo estadunidense Donald Davidson (1987/2003) reconhece a existência de uma assimetria entre o conhecimento de primeira pessoa acerca dos próprios estados mentais e o conhecimento que temos acerca dos estados mentais alheios. Nesse artigo, ele pretende explicar como é possível conciliar essa autoridade de primeira pessoa com seu externalismo⁵¹. Seus argumentos se restringem apenas aos estados mentais chamados de atitudes proposicionais⁵². Mesmo assim, é possível que suas ideias possam também, de alguma forma, iluminar a questão dos supostos estados mentais não proposicionais, os chamados *qualia*. Este trabalho pretende olhar para tal possibilidade.

A conexão entre o problema da autoridade de primeira pessoa e o problema tradicional das outras mentes é evidente, mas como levanto o primeiro problema, existem duas diferenças importantes. A autoridade de primeira pessoa é o problema mais restrito, uma vez que o tratarei apenas na medida em que se aplica a atitudes proposicionais como a crença, o desejo e a intenção; estar satisfeito, admirado, assustado ou orgulhoso de que algo seja o caso. Mas não analisarei o que frequentemente se considera que é central no problema das outras mentes: as dores e outras sensações e o conhecimento, a memória, a atenção e a percepção tanto dirigidas a objetos como a pessoas, ruas, cidades, cometas e outras entidades não proposicionais. *Parece-me que o que se mantém para as atitudes proposicionais há de ser relevante para as sensações e o restante, mas não vou explorar aqui essa conexão* (DAVIDSON, 1984, p. 25-6, grifo nosso)⁵³.

Na citação acima, observa-se que, embora Davidson não amplie seus argumentos para os estados mentais não proposicionais, ele entende que tal ampliação pode ser feita. É o que se pretende aqui discutir. Será importante, então, compreender qual é o problema apresentado por

⁵¹Na segunda seção deste trabalho, explicaremos brevemente o termo “externalismo” aplicado aos estados mentais. Falando de maneira abrangente e rápida, externalismo é a tese de que nossos conteúdos mentais ou o significado de nossos discursos são, ao menos parcialmente, determinados por fatores externos a nós.

⁵²Atitudes proposicionais são estados mentais cujo conteúdo pode ser expresso por uma proposição particular, como crença, desejo e dúvida (Huang, 2012, p. 251). Por exemplo, na sentença “Ana acredita que o atual presidente do Brasil alcançou esse cargo por meio de um golpe”, observa-se uma crença de Ana dirigida a uma proposição, ou seja, que o atual presidente do Brasil alcançou esse cargo por meio de um golpe. Para muitos, as atitudes proposicionais se contrastam com os *qualia*, aspectos individuais de nossas experiências subjetivas, sem qualquer manifestação linguística.

⁵³As citações presentes neste trabalho são traduções livres.

Davidson acerca da autoridade de primeira pessoa e a sua proposta para lidar com isso⁵⁴. Em seguida, mostraremos que um problema análogo se aplica a certa concepção do que sejam *qualia*. Por último, tentaremos estender a resposta de Davidson, baseada no aspecto relacional dos estados mentais proposicionais, aos *qualia*. Também será abordada uma proposta de Dennett para lidar com a questão da subjetividade: o método heterofenômico. Tanto na resposta de Davidson quanto na de Dennett, o ponto de vista do intérprete ganha um papel relevante na compreensão do mental.

O que são *qualia*? Há várias concepções do termo, mas buscaremos aqui uma que reúna características bastante difundidas entre os filósofos. Dennett (1988) apresenta algumas propriedades que são em geral atribuídas conjuntamente aos estados mentais denominados *qualia*: são estados inefáveis, intrínsecos, privados e diretamente apreensíveis pelo sujeito. Muitas vezes associados às experiências sensoriais, mas também associados às emoções, ao prazer e à dor - *qualia* seriam o que constitui as nossas experiências subjetivas, algo que só pode ser vivido pelo sujeito, que não se compartilha, só meu. A ideia dos *qualia* como constituinte básico da vida mental de um sujeito tem sido defendida ao longo da história recente da filosofia da mente a partir de uma série de experimentos de pensamento filosóficos. Neste artigo, dois desses experimentos serão brevemente abordados: o argumento do zumbi e o do espectro invertido. Ora, que coisa é essa que só eu, sujeito, sei? Como eu poderia, por meio de uma linguagem compartilhada e publicamente constituída, estar falando dessa coisa? Que parâmetro eu tenho para saber que estou falando da coisa certa?

Locke (1690/1999) denomina “ideia” aquilo que está na mente, e não nos objetos, embora os objetos tenham a potência de provocar ideias nas mentes. Para ele, há dois tipos de qualidades nos objetos que percebemos: as qualidades primárias, que estariam nos objetos independentemente dos nossos sentidos (extensão, figura, número e movimento) e as qualidades secundárias, isto é, as que se apresentam aos nossos sentidos e estão nos objetos apenas enquanto

⁵⁴Conforme Leclerc (2005), o artigo de Davidson (1987) tem dois objetivos: o primeiro, conciliar o externalismo com o fisicalismo de “token”, fracassa. Porém o segundo, conciliar o externalismo com a autoridade de primeira pessoa acerca dos estados mentais, obtém sucesso. Aqui, aborda-se apenas essa segunda meta do artigo. Entende-se “token” como um caso particular de algo, em oposição ao “type”, que seria essa coisa multiplamente instanciada (Huang, 2012, p. 304-305). O fisicalismo de “token” é a posição segundo a qual cada estado mental particular se identifica com um estado físico particular. Opõe-se ao fisicalismo de “type”, para o qual há tipos de estados mentais que, quando instanciados, se identificam com certos tipos de estados físicos, por exemplo, “dor” (em geral) é “ativação da fibra C”. Talvez possamos traduzir “token” por “exemplar” ou “caso particular” e “type”, podemos traduzir por “tipo” ou “modelo geral”. Não entraremos nos detalhes da argumentação do professor Leclerc. Para isso, basta ler o artigo dele aqui referenciado.

potências capazes de nos provocar tais experiências (cores, gostos, cheiros etc.). As qualidades primárias estariam restritas à mente do sujeito de tal forma que, na concepção de Locke, não é possível ao sujeito saber com certeza o que se passa na mente de outro sujeito.

Nem traria qualquer imputação de falsidade para nossas ideias simples se, através da diferente estrutura de nossos órgãos, estivesse ordenado que o mesmo objeto devesse produzir ideias diferentes nas mentes de diversos homens em simultâneo, por exemplo, se a ideia que uma violeta produziu na mente de um homem, através de seus olhos, fosse a mesma que uma calêndula, [amarela] produziu num outro homem e vice-versa⁵⁵.

No trecho acima, encontramos o argumento segundo o qual se uma pessoa tiver experiências subjetivas de cores radicalmente diferentes das minhas (ou das de outra pessoa qualquer), não teremos como detectar essa diferença, pois a experiência é absolutamente restrita ao sujeito, ao passo que a linguagem é compartilhada. Esse argumento é denominado “argumento do espectro invertido”⁵⁶.

As qualidades secundárias defendidas por Locke nos trazem uma boa noção do que comumente se considera que sejam os *qualia*: algo que se restringe à mente subjetiva e que não pode ser comunicado; algo que o sujeito sabe sempre sobre si mesmo, mas nunca sobre o outro. Vale notar que, se as outras pessoas podem possuir experiências subjetivas diferentes das minhas sem que eu tenha condições de perceber, talvez também possam simplesmente não possuir quaisquer experiências. Como eu posso ter certeza de que minha amiga possui experiências? Seguindo essa linha de raciocínio, ao que parece, eu só poderia saber uma única coisa: que *eu* possuo mente. Quanto às mentes das outras pessoas, eu nada poderia saber. Então, caminhando lado a lado com esse tipo de concepção do mental como algo isolado e de acesso restrito ao sujeito, está o ceticismo acerca das outras mentes. Mas há alguma possibilidade de evitarmos esse tipo de ceticismo? Afinal, todo o meu comportamento social depende de que eu atribua aos outros uma vida mental e isso já parece um bom motivo para se buscar uma superação do ceticismo acerca das outras mentes.

⁵⁵ Livro II, cap. XXXII. Documento “on line” sem paginação.

⁵⁶ Atualmente, há pesquisas em psicologia mostrando que temos uma série de reações comportamentais diante das experiências de cores. Ainda assim, tais argumentos de base científica acabam não sendo acolhidos pelos filósofos que tendem a separar a experiência subjetiva de qualquer expressão comportamental.

2. O problema das outras mentes e o problema da autoridade de primeira pessoa

É possível que uma nova abordagem acerca do mental seja capaz de diluir um pouco o problema das outras mentes, ou ao menos desfocá-lo. Externalismo acerca do mental é a posição segundo a qual o significado ou o conteúdo de certos estados mentais é ao menos parcialmente determinado pelo ambiente externo. De acordo com Davidson (1987/2003), alguns autores (como Putnam, Burge e ele próprio) defendem que os estados mentais proposicionais, por possuírem conteúdo linguístico, só podem ser formados por meio de relações com o ambiente externo, inclusive o ambiente social. Ora, se aceitamos que tais estados mentais dependem de algum tipo de relação externa com o ambiente socialmente compartilhado com outros sujeitos, a perspectiva cética acerca das outras mentes resta enfraquecida. Contudo - afirma Davidson - Putnam e Burge fazem uma diferenciação entre estados mentais intrínsecos e relacionais. Certos estados mentais, para eles, ainda pertenceriam exclusivamente ao sujeito e a autoridade de primeira pessoa se conservaria apenas em relação a tais estados mentais. Haveria, então, na concepção desses autores, estados mentais completamente isolados do mundo externo, resistentes a qualquer tipo de manifestação comportamental.

Esse tipo de externalismo dá origem ao seguinte problema epistemológico: se o conteúdo de meus estados mentais proposicionais depende, ainda que parcialmente, de fatores externos ao sujeito, então como pode o sujeito saber com certeza o que ele próprio está pensando? Um exemplo de Parent (2013): se para Descartes saber que está pensando sobre carvalhos ele precisa saber o que são carvalhos no mundo exterior, então ele não tem como determinar o conteúdo de seus próprios pensamentos apenas sentado em sua poltrona. Esse é o problema de conciliar o externalismo e o autoconhecimento acerca dos estados mentais.

A tradição cartesiana é internalista. A atribuição a um sujeito de uma propriedade psicológica não depende de nada que seja externo a esse sujeito. Mas isso dá origem ao ceticismo acerca das outras mentes. Ora, se duvidamos de que os outros possuam mentes, toda nossa comunicação e vida social também ficam em cheque e, no fim das contas, talvez eu possa ser simplesmente um cérebro numa cuba dentro do laboratório de um cientista maligno que me controla por meio de eletrodos e me faz crer que sou uma professora diante de um computador escrevendo justamente sobre isso (PUTNAM, 1992). Em outras palavras, talvez nada mais exista além da minha própria mente.

Em uma perspectiva externalista, por outro lado, o problema das outras mentes pode se obscurecer, pois os estados mentais se constituem parcialmente a partir de suas relações causais com o mundo externo. Contudo, um novo embaraço filosófico pode surgir: o problema de se a pessoa sabe em que estado mental ela própria está. Em outras palavras, enquanto se apaga o problema das outras mentes, acende-se o problema da minha própria mente. Nesse sentido ocorre um deslocamento do ceticismo para um lado ainda mais contra intuitivo. No fundo, esses dois problemas têm a mesma origem: a tradição de ruptura entre os mundos objetivo e subjetivo.

Davidson, embora desenvolva seu argumento em torno dos estados mentais proposicionais, não se compromete com a existência de outros estados mentais que sejam isolados e restritos ao sujeito. Ao contrário, como podemos notar na última frase da primeira citação deste trabalho, ele pensa que suas conclusões podem ser ampliadas para tais estados comumente considerados intrínsecos, mas não se ocupará com o desenvolvimento desse ponto. O que ele pretende mostrar é que a sua versão do externalismo acerca dos estados mentais proposicionais não compromete a defesa da ideia de que o sujeito possui uma autoridade especial acerca daquilo que ele mesmo crê ou pensa.

Para Davidson (1988/2003, p. 79), a melhor maneira de compreendermos a maneira como ocorre essa conciliação é a partir dos fatos por meio dos quais aprendemos a linguagem e interpretamos as palavras. Essa aprendizagem provém dos objetos e circunstâncias que cercam o aprendiz. Há uma relação de triangulação entre aprendiz, instrutor e mundo.

De acordo com Leclerc (2005, p. 159-160), há três apresentações da triangulação na obra de Davidson. A primeira apresenta duas criaturas carentes de linguagem e de pensamento conceitual reagindo ao mesmo evento e tentando cada uma associar as reações da outra a esse evento. A segunda envolve o aprendiz de uma linguagem em um ambiente social que contém professores. Na terceira, dois falantes de línguas diferentes e que não possuem uma língua em comum interagem. As duas últimas são bastante claras para compreendermos a noção de triangulação, pois a maneira como o aprendiz ou os falantes de línguas diferentes conseguem ajustar suas reações linguísticas à realidade recorre às definições ostensivas, o que evidencia a necessidade de conexões entre linguagem, indivíduo e mundo externo, sem as quais não é possível a linguagem⁵⁷.

⁵⁷ Seria interessante mostrar que os estados mentais dos animais não humanos também possuem características relacionais. O argumento da triangulação é muito interessante e Davidson (2003/1997) o aplica a animais não

Eu preciso, por exemplo, me comunicar com uma pessoa estrangeira que chegou ao Brasil. Nesse caso, ela e eu começamos a estabelecer conexões entre os sons emitidos por cada uma e o mundo objetivo ao nosso redor. O mundo objetivo é um parâmetro fundamental sem o qual não podemos interpretar as crenças uma da outra. Além disso, precisamos ter como pressuposto que cada uma está fazendo essas conexões entre suas palavras e o mundo, pois a comunicação não será possível se eu acreditar que minha interlocutora é louca e que seu discurso está totalmente desconectado do mundo.

A individualização das crenças e pensamentos, mas também de significados, conceitos e outros estados mentais, só pode ser concebida a partir de conexões causais sistemáticas na triangulação entre o indivíduo, o outro falante com quem ele interage e objetos ou eventos no mundo (SILVA FILHO, 2007, p. 122).

Assim, Davidson (1990/2003) argumenta que não é possível sequer ter crenças e nenhum tipo de estado mental proposicional se não há a relação dinâmica que envolve o sujeito, seu interlocutor e o mundo objetivo. Quando aprendemos a falar, os nossos diversos professores estabelecem essas relações com o mundo. Portanto, o fato de a linguagem estar aí para pensarmos sobre o ceticismo traz consigo fortes indícios de que o mundo existe e de que os interlocutores possuem estados mentais apropriados para falar dele. Afinal, é apenas no contexto do debate que podemos pensar sobre o ceticismo (SILVA FILHO, 2002, p. 161).

Outro ponto que merece ser ressaltado é que, após o surgimento da linguagem convencional, o pleno funcionamento do processo de triangulação depende de que os envolvidos no diálogo atribuam crenças majoritariamente verdadeiras uns aos outros. Cada um precisará partir do pressuposto que os sinais linguísticos emitidos pelo outro possuem uma conexão apropriada com o mundo. Apenas assim é que a comunicação linguística se torna possível. Essa postura se estende a qualquer situação de comunicação. Além disso, se conseguimos algum sucesso ao nos comunicarmos, é porque de fato o que se passa na mente de nosso interlocutor coincide em grande medida com o que pensamos que se passa lá.

A perspectiva do intérprete que atribui crenças ao outro para poder compreender o que ele diz, assim, é bastante confiável. Ainda que o intérprete atribua muitas crenças falsas ao seu interlocutor, só poderá atribuí-las sobre um pano de fundo de crenças majoritariamente

humanos, pois a triangulação é anterior à linguagem. Assim, o aspecto relacional dos estados mentais não depende de linguagem convencional e, portanto, pode se aplicar a qualquer tipo de estado mental, independentemente de ser um estado de uma mente humana.

verdadeiras. O resultado final será uma atribuição de crenças quase sempre correspondente ao que o outro de fato crê. Em outras palavras, dentro do conjunto total de crenças do usuário de uma língua, a grande maioria delas deverá ser verdadeira. Caso contrário, a comunicação não seria possível. De modo bem geral, é a isso que se denomina “princípio de caridade”, um pressuposto para a interpretação bastante desenvolvido no pensamento de Davidson.

Diante dessa perspectiva externalista de Davidson, sua estratégia para salvaguardar a autoridade de primeira pessoa acerca dos estados mentais consiste em apresentar o mental como sendo relacional. Deve-se abandonar a ideia de que os conteúdos mentais sejam entidades perante a mente ou que os estados mentais proposicionais requeiram a existência de objetos mentais correspondentes. “A solução, no caso dos estados mentais (...) se trata de desfazer-se da metáfora dos objetos ante à mente” (1987/2003, p. 68).

Uma vez que tenhamos abandonado o mito do subjetivo – a ideia de que os pensamentos requerem objetos mentais – é natural pensar que a autoridade de primeira pessoa, o caráter social da linguagem e os determinantes externos do pensamento e o significado se deem conjuntamente. (DAVIDSON, 1987/2003, p. 43)

A proposta é que se abandonem as concepções de atitudes proposicionais como coisas que se apresentam perante o teatro mental⁵⁸. Se houvesse objetos na mente que fossem isolados do mundo externo, não poderíamos atribuir estados mentais a outras pessoas, não teríamos uma base para fazer isso. Ainda que haja estados mentais que não estejam diretamente conectados com o mundo exterior, eles estão conectados indiretamente, via outros estados mentais. Desse modo, torna-se possível o sujeito enganar-se acerca do mundo sem que esteja ameaçada a certeza que tem acerca de seus próprios estados mentais. Não é que Davidson defenda uma completa impossibilidade de erro do sujeito perante seus estados mentais. É possível haver erro, sim. Porém, a possibilidade de erro é pequena quando se trata do que o sujeito crê que pensa.

⁵⁸Dennett (1991) chamaria essa perspectiva de “Teatro Cartesiano”. Um dos problemas desse tipo de abordagem é causar uma regressão ao infinito. Torna-se necessário uma audiência para assistir às representações que se passam diante desse teatro, mas então, o sujeito é essa audiência? Isso não responderia ao problema e levaria à necessidade de uma audiência interna ao sujeito para receber essas representações, e assim sucessivamente. Esse problema é chamado também de “problema do homúnculo”, isto é, a ideia de que existiria dentro do cérebro uma área responsável por integrar as experiências subjetivas, uma espécie de pequeno sujeito dentro do cérebro, mas então, como seriam integradas as experiências desse homúnculo?

3. Uma concepção de *qualia* a ser evitada

Se aceitarmos os estados mentais proposicionais como relacionais e parcialmente determinados por aspectos externos ao sujeito, sobra algo privado dentro da mente? Ocorre que o aprisionamento dos *qualia* como objetos perante a mente dá origem a um problema epistemológico análogo ao problema do ceticismo acerca da própria mente. Assim, analogamente, tal problema também pode se diluir por meio de uma perspectiva relacional dos *qualia*. O problema do ceticismo acerca da própria mente tem origem na ruptura entre os mundos objetivo e o subjetivo, pois isso isola os estados mentais em um dos lados do abismo.

Há muitas concepções do termo *qualia*. A que pretendemos evitar é aquela segundo a qual *há algo para o sujeito que é como estar em um determinado estado mental*. Essa concepção, tornada célebre no artigo “Como é ser um morcego” de Nagel (1974/2005), parece carregar um comprometimento ontológico com algo que se manifesta perante o sujeito. De acordo com essa visão, os *qualia* são o aspecto fundamental da vida consciente e estão sempre disponíveis para o sujeito, mas só para ele, criando um abismo entre dois aspectos da realidade: o aspecto objetivo e o aspecto subjetivo. É justamente essa ideia de *entidade* isolada dos demais estados mentais e cerebrais que pretendemos evitar. Dennett (2013, p. 296-297) considera um erro a ideia de que os *qualia* sejam estados mentais isolados. As nossas experiências sensoriais se combinam com outras experiências, com desejos crenças e comportamentos diversos, verbais ou não (nesse sentido, o argumento também se estende aos animais não-humanos, ainda que desprovidos de linguagem). Ao compreendê-los de forma relacional, não sobra nada encarcerado na mente do sujeito.

Como mencionado no início deste artigo, em um artigo de 1988, Dennett propõe algumas condições para tentar unificar o conceito a que se referem os defensores dos *qualia*. De acordo com ele, os *qualia* seriam simultaneamente (a) inefáveis, (b) intrínsecos, (c) privados e (d) diretamente apreensíveis para o sujeito. Em outras palavras, (a) eles só aparecem por meio da introspecção, (b) são indescritíveis; (c) não são relacionais “a cor vermelha pode ser provocadora de ansiedade para algumas pessoas, mas essa disposição subjetiva não é um *quale* de vermelho” (2013, p. 298) e (d) aparecem apenas para o sujeito, não podem ser compartilhados; estão presentes na consciência e são o que há de mais íntimo.

Essa reunião de características é o que provoca determinadas aporias, evitáveis a partir de uma compreensão relacional dos *qualia*. Por isso, tentemos olhar para as experiências sensoriais de outra forma. Os diversos estados mentais podem ser vistos como conectados entre si, não havendo nenhum tipo de estado mental isolado, nem mesmo os *qualia*. Nesse modo de encarar, as atitudes proposicionais fazem parte de uma rede de relações com nossas sensações e comportamentos, além de estarem conectadas também ao mundo externo. Isso não significa negar a autoridade de primeira pessoa acerca dos próprios estados mentais, mas é uma ponte que conecta o sujeito ao mundo e às outras mentes.

4. O problema epistemológico dos estados mentais isolados

O argumento do zumbi é comumente utilizado para defender a existência dos *qualia* como algo resistente a qualquer abordagem de um ponto de vista objetivo e para tentar aguçar a intuição com vistas a esclarecer o que significa a noção de *qualia*. De acordo com Chalmers (1996, p. 94), por exemplo, é logicamente possível a existência de “alguém ou algo fisicamente idêntico a mim (ou a qualquer outro ser consciente), mas completamente destituído de experiências conscientes”. Vale a pena, aqui, explicar que o zumbi não é simplesmente uma duplicata molecular de um ser consciente da pele para dentro. Ele mantém com o mundo físico as mesmas relações que um ser consciente e apresenta os mesmos comportamentos, inclusive os verbais.

Ocorre que, para poder dizer as mesmas coisas e exibir os mesmos comportamentos em geral, presume-se que o zumbi possua as mesmas atitudes proposicionais da sua contraparte consciente. Afinal, para que um ser possa ser considerado um zumbi, ele precisa ser capaz de falar sinceramente acerca de suas crenças, tal qual um ser consciente. Suponha que eu tenha um zumbi gêmea. Ela é minha substituta em um mundo no qual eu não existo. Lá, ela faz tudo o que eu faço aqui. Então, se eu digo que tenho plena confiança nas instituições jurídicas brasileiras, ela diz a mesma coisa. Se formos ambas submetidas a um teste no detector de mentiras, seremos igualmente desmascaradas. Em tudo, nossos corpos se manifestam exatamente da mesma maneira.

O que nos provoca tais reações perante o detector de mentiras não é dizermos algo em que não acreditamos? Ora, mentir é justamente afirmar algo em que não se crê. O

comportamento de mentir está sempre ligado a uma crença. Então, a minha contraparte zumbi pensa como eu, possui as mesmas crenças, as mesmas dúvidas, os mesmos desejos e gostos que eu... Mas nós temos uma enorme diferença em relação uma à outra: eu sou consciente e ela não é, no seguinte sentido: eu possuo *qualia* e ela não possui. Ao entrar na sala do detector de mentiras, eu fico ansiosa, sinto as mãos frias e minha testa fica úmida. A testa dela também fica úmida, as mãos gelam, mas ela não tem a sensação do frio, embora ela creia que tem a sensação e diga sinceramente que está sentindo as mãos frias. Ao ver uma barata andando pela sala, o estômago dela revira, os pelos se arrepiam e, assim como eu, ela afirma sinceramente que está sentindo muito nojo, mas com a diferença que ela não tem a experiência subjetiva do nojo.

Um aspecto curioso dessa diferença tão radical entre mim e ela é que ninguém, absolutamente ninguém, é capaz de nota-la. Qualquer pessoa que observe nossos comportamentos tenderá a atribuir a ambas a mesma vida subjetiva. Se um dia eu e ela pudéssemos marcar um encontro, eu não teria meios para saber que ela carece de experiências e ela não poderia saber que eu tenho uma vida interna rica em experiências que lhe carecem. Isso é um aspecto importante do argumento. Sua força reside no fato de que nós nunca podemos afirmar com certeza o que se passa na vida subjetiva alheia.

Contudo, também é justamente isso o que conduz o argumento a uma aporia. Assim como ela não tem como saber que eu tenho uma vida subjetiva abundante, também não tem como saber que sua vida subjetiva é vazia. Se o saber passa pela crença e minha gêmea zumbi crê que possui experiências tão variadas e abundantes como as minhas, então ela não sabe de sua ausência de experiências.

Aqui, surge novamente o problema da autoridade acerca dos próprios estados mentais e, aparentemente, ele surge com ainda mais força. Mas o motivo porque ele surge é o mesmo: considerar certos estados mentais (nesse caso, os *qualia*) como objetos existentes perante a mente e isolados dos demais estados mentais e do mundo objetivo. O argumento do zumbi tem como pressuposto a existência desses aspectos do mental que seriam as experiências subjetivas como algo isolado dos demais estados mentais e do comportamento. É isso que dá origem a tal conclusão paradoxal.

Ora, mas que coisa é essa que chamamos de *qualia*, que cremos que possuímos, que uma contraparte zumbi nossa também crê que possui, mas nós estamos certos e ela está errada em sua crença? Esse argumento não me conduz a uma insegurança acerca de se eu mesma possuo

estados mentais? Diante de minhas limitações epistêmicas, nada me poderia garantir que eu mesma não seja um zumbi. Essa conclusão não parece ser o que os defensores do argumento do zumbi desejariam defender. Eles querem defender a existência de estados subjetivos diretamente apreensíveis pelo sujeito consciente e absolutamente internos, sem expressão comportamental. Porém, o que eles conseguem é nos levar à conclusão de que não sabemos se somos ou não somos zumbis. Nesse sentido, o argumento do zumbi, comumente usado para defender os *qualia*, talvez fosse melhor orientado se utilizado para rejeitar a noção de *qualia*.

Vale notar que esse problema não emerge apenas do argumento do zumbi. Ele é um exemplo de um problema maior. O ceticismo oscila entre as outras mentes e a própria mente como resultado de uma concepção que isola os estados mentais do mundo objetivo. O problema da própria mente e o das outras mentes são dois lados da mesma moeda. A mente não se constitui de um teatro frente ao qual uma audiência subjetiva possui uma série de experiências. É possível que uma concepção diferente do mental, ainda que possa parecer contra intuitiva à primeira vista⁵⁹, traga resultados mais frutíferos.

5. Considerações finais: *Qualia* relacionais

Dennett (1991) defende que os estados mentais de alguém podem ser estudados a partir dos relatos que a pessoa faz de suas experiências, denomina esse método de “heterofenomenologia”. Há a possibilidade de uma pessoa se enganar acerca das suas próprias experiências subjetivas, isso não precisa ser descartado⁶⁰. Por isso, ao tentar interpretar o relato de um sujeito acerca de suas experiências, o intérprete deve lhe atribuir racionalidade e coerência. Além disso, é necessário que o intérprete atribua ao sujeito crenças, desejos e demais estados mentais para que os sons emitidos façam algum sentido e sejam compreendidos como algo que ele *queria* dizer. Para Dennett, ao interpretarmos os relatos do sujeito, permitimos que eles constituam o universo a ser pesquisado, que é constituído a partir da interpretação que fazemos daquele discurso. É um universo ao mesmo tempo subjetivo e intersubjetivo, pois se

⁵⁹É provável que haja apenas uma aparência de algo contra intuitivo quando tentamos pensar todos os estados mentais como inter-relacionados. Essa aparência deriva de hábitos de pensamento adquiridos do pensamento filosófico moderno.

⁶⁰ Por exemplo, a prosopagnosia é uma desordem neuronal na qual o sujeito deixa de reconhecer os rostos das pessoas, inclusive de pessoas familiares, embora tenha a visão normal.

constitui no discurso. Então, embora o sujeito seja uma autoridade para saber e falar de seus estados mentais, ele não é a única autoridade.

Na concepção de Dennett, é pouco provável que haja referência concreta na mente para aquilo de que o sujeito fala. A mente é feita de abstrações e nela se encontram os sonhos, as alucinações e as fantasias, não há objetos internos aos quais a mente se direciona e sobre os quais constrói os seus discursos. Outro aspecto interessante desse método é que o discurso do sujeito acerca de suas experiências internas é apenas uma forma de descrevê-las. É uma forma bastante eficaz, mas não é a única. Algumas formas de descrever o que se passa na vida subjetiva de alguém podem não estar disponíveis para o próprio sujeito: por exemplo, a descrição do que acontece no seu cérebro a nível neuronal quando ele entra em um ambiente completamente branco⁶¹. Isso não significa que Dennett (2005) desconsidere a autoridade de primeira pessoa acerca de seus estados mentais. Na realidade, a descrição de primeira pessoa acerca do fenômeno mental é um material crucial para se compreender a subjetividade daquela pessoa. Ocorre, aqui, que a descrição feita pelo sujeito acerca de seu mundo íntimo é apenas uma forma de descrição, sendo que outras abordagens descritivas seriam possíveis. Tais abordagens, reunidas, poderiam trazer uma compreensão mais ampla e completa do que seja a mente.

Considerando a proposta de Davidson de que as crenças atribuídas a alguém por seu intérprete são majoritariamente verdadeiras e apropriadamente conectadas ao mundo objetivo, podemos nos perguntar sobre a possibilidade de os *qualia* serem vinculados a essa cadeia causal. Nesse caso, certamente, teríamos que adotar uma concepção de *qualia* bastante diferente daquela apresentada acima e comumente adotada pelos defensores dos *qualia*. Porém, dadas as aporias epistêmicas a que fomos conduzidos pela concepção de estados qualitativos como inefáveis, privados, não relacionais e diretamente apreensíveis pelo sujeito, talvez uma reformulação do conceito seja bem-vinda.

Negar que existam *qualia* perante a mente reunindo essas características não significa dizer que tenhamos uma vida subjetiva desbotada e insossa. Temos uma série de experiências e essas experiências trazem informações sobre o ambiente externo. Por isso, são capazes de ajustar nossos comportamentos, contribuir na constituição de nossas crenças e desejos e de se fazerem

⁶¹ Vale notar que há experiências das quais não podemos falar, mas elas nos provocam outras reações físicas ou comportamentais. Ao pensar isso, olhei pela janela e reparei que o céu está colorido de rosa, amarelo e azul. É lindo, está diante dos meus olhos, mas eu não teria podido falar sobre isso a menos que tivesse parado um pouco para prestar atenção ao que está acontecendo. Ao mesmo tempo, enquanto observo o céu, outras experiências passam por mim e posso mesmo dizer que meu corpo reage a elas, mas não posso falar sobre elas, pois não as percebo.

presentes em nossas vidas sociais. Ao falarmos acerca de nossas experiências, nossa fala é constituída a partir da compreensão que temos dos significados das palavras que utilizamos e essa compreensão é constituída a partir de nossas relações com o mundo externo. Se digo sinceramente a alguém que estou sentindo calor, a fala se constitui a partir do que aprendi acerca do significado da palavra “calor”. Aprendi isso por meio de relações com as pessoas que me ensinaram a linguagem e com o ambiente externo. Então, a sensação que meu interlocutor me atribui também deriva de suas relações com o mundo, a linguagem e suas próprias experiências internas sendo, portanto, se não totalmente correta, suficientemente próxima daquilo que eu realmente sinto. Caso contrário, não seria possível me compreender nem se fazer compreendido por mim. As interações entre o sujeito, o mundo e os outros sujeitos tendem a nos conduzir a crenças majoritariamente verdadeiras sobre o mundo, sobre minha mente e sobre as outras mentes⁶².

A subjetividade se revela a partir do fato de que os diversos estados mentais estão relacionados entre si, influenciam o comportamento e, no caso humano, constituem a linguagem e são constituídos por ela. Então, se é possível compreendermos as sensações a partir de suas relações com as atitudes proposicionais, outros estados mentais e comportamentos em geral, o ideal é que a compreensão do sujeito envolva o discurso do próprio sujeito acerca de seus estados mentais subjetivos. Mas esse discurso é constituído a partir das relações que o sujeito possui com os outros e com o mundo, precisa ser interpretado pelo interlocutor para que possa fazer algum sentido. Se encararmos pela perspectiva de Davidson, a interpretação dada pelo interlocutor tem grandes chances de ser correta. Nesse sentido, uma abordagem relacional é capaz de se desviar do ceticismo acerca das outras mentes sem descambar para o ceticismo acerca da própria mente.

Referências

CHALMERS, D. *The conscious Mind*. Oxford: Oxford University Press, 1996.

DAVIDSON, D. La autoridad de primera persona (1984). In: _____. *Subjetivo, intersubjetivo, objetivo*. Madrid: Ediciones Cátedra, 2003. p. 25-40. (Traducción de Olga Fernández Prat.)

_____. Conocer nuestra propia mente (1987). In: _____. *Subjetivo, intersubjetivo, objetivo*. Madrid: Ediciones Cátedra, 2003. p. 41-71. (Traducción de Olga Fernández Prat.)

_____. El mito de lo subjetivo (1988). In: _____. *Subjetivo, intersubjetivo, objetivo*. Madrid: Ediciones Cátedra, 2003. p. 72-90. (Traducción de Olga Fernández Prat.)

⁶² Dennett, contudo, não está comprometido com a veracidade das crenças. Ele tem uma abordagem instrumental da subjetividade.

_____. La externalización de La epistemología (1990). In: _____. *Subjetivo, intersubjetivo, objetivo*. Madrid: Ediciones Cátedra, 2003. p. 265-279. (Traducción de Olga Fernández Prat.)

_____. La aparición del pensamiento. _____. *Subjetivo, Intersubjetivo, Objetivo*. Madrid: Ediciones Cátedra, 2003. p. 176-192. (Traducción de Olga Fernández Prat.)

DENNETT, D. Quining qualia. In: MARCEL, A. BISIACH, E. (eds.) *Consciousness in Modern Science*. Oxford: Oxford University Press, 1988. p. 42-77.

_____. *Consciousness explained*. London: Penguin Books, 1991.

_____. *Intuition Pumps and Other tools for Thinking*. New York; London: W. W. Norton & Company, 2013.

LECLERC, A. Davidson's externalism and swanpman's troublesome biography. *Principia*, Florianópolis, v. 1-2, n. 9, p. 159-175, jun./dez. 2005.

LOCKE, J. *Ensaio sobre o entendimento humano* (1690). Lisboa: Fundação Calouste Gulbenkian, 1999. (Excerto selecionado por O. Pessoa Jr. Disponível em: <http://www.fflch.usp.br/df/opessoa/Locke-Qualia-2.pdf>. Acesso em 08 out. 2016. Publicado em 1999.)

NAGEL, T. Como é ser um morcego? (1974). *Cadernos de história e Filosofia da Ciência*. Campinas, v. 15, n. 1, p. 245-262, jan./jun. 2005.

PARENT, T. Externalism and Self-knowledge. ZALTA, E. (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2013 Edition). Disponível em: <http://plato.stanford.edu/archives/sum2013/entries/self-knowledge-externalism/>. Acesso em 07 out. 2016.

PUTNAM, H. *Razão, verdade e história*. Lisboa: Don Quixote, 1992.

SILVA FILHO, Waldomiro J. Externalismo sem dogmas. *O que nos faz pensar*. Rio de Janeiro, v. 16, n. 21, p. 113-131, jun. 2007.

_____. Davidson, Sócrates e os instrumentos da Filosofia. In: DAVIDSON, Donald. *Ensaaios sobre a verdade*. São Paulo: Unimarco Editora, 2002. (Organizado por Paulo Ghiraldelli Jr., Pedro F. Bendassolli e Waldomiro José da Silva Filho. Tradução de Pedro Fernando Bendassolli.)

Há filosofia da mente no pensamento de Guilherme de Ockham?

Is there philosophy of mind in William of Ockham?

Laiza Rodrigues de Souza⁶³

RESUMO: Neste trabalho, propomos investigar a possibilidade de interpretar alguns aspectos da filosofia de Guilherme de Ockham à luz do externalismo e da Filosofia da Mente contemporânea. Há algo como uma filosofia da mente no pensamento de Ockham? A teoria do conhecimento ockhamista é construída sobre uma base linguística e, portanto, reflete a necessidade de que as proposições que expressam nossas crenças correspondam à realidade do mundo que nos circunda. Assim, a condição de possibilidade de uma ciência está na correspondência entre seu corpo teórico e a realidade. De forma semelhante, a intelecção que antecede e compõe a crença privada de um sujeito diz respeito a um complexo mecanismo de atos intelectivos cujo início é desencadeado sempre a partir de um fator externo. O conhecimento é sempre o conhecimento de algo. Assim, o conteúdo de nossos pensamentos, situado no que Ockham postula como “linguagem mental” tanto em relação ao conhecimento, quanto em relação à linguagem, deve fazer referência ao mundo extramental. Neste sentido, buscamos viabilizar a possibilidade de aproximar teorias externalistas contemporâneas e a filosofia de Ockham.

PALAVRAS-CHAVE: Guilherme de Ockham; Linguagem Mental; Filosofia da Mente.

ABSTRACT: In this work, we propose to investigate the possibility of interpreting some aspects of Guillermo de Ockham's philosophy in the light of externalism and the Contemporary Philosophy of Mind. Is there something like a philosophy of mind in Ockham's thought? The theory of ockhamist knowledge is built on a linguistic basis and therefore reflects the need for the propositions expressing our beliefs to correspond to the reality of the world around us. Thus, the condition of possibility of a science lies in the correspondence between its theoretical body and reality. Similarly, the intellection that precedes and composes the private belief of a subject concerns a complex mechanism of intellectual acts whose onset is always triggered by an external factor. Knowledge is always the knowledge of something. Thus, the content of our thoughts, situated in what Ockham posits as "mental language" both in relation to knowledge and in relation to language, must make reference to the extramental world. In this sense, we seek to make feasible the possibility of approaching contemporary externalist theories and Ockham's philosophy.

KEYWORDS: William of Ockham; Mental Thought; Philosophy of Mind.

⁶³ Doutoranda em Filosofia na UFRGS. Email: la-izarodrigues@hotmail.com.

Introdução

A relação mente cérebro tem intrigado a humanidade desde o início das reflexões que o homem fez acerca do mundo. A existência da mente é um ponto que podemos chamar de consensual. Entretanto, na discussão filosófica pergunta-se que tipo de existência tem a mente e que tipo de relação ela tem com o cérebro é o que falta ser explicado. Se, por um lado, tivemos avanços no que diz respeito às descobertas do funcionamento do cérebro, de outro temos a mente como uma estrutura impossível de ser conhecida do mesmo modo que uma estrutura material como o cérebro. A assimetria entre mente e cérebro foi um ponto discutido por Descartes e deu origem à concepção moderna de dualismo cartesiano, que separa mente e cérebro de forma quase irreconciliável. Mas esta concepção está longe de ser livre de problemas. O dualismo cartesiano apenas divide as estruturas, mas não explica sua relação ou funcionamento. A divisão entre mente e corpo norteia uma série de questões filosóficas acerca de como essas duas estruturas se relacionam e a resposta para a maioria delas permanece obscurecida. As questões levantadas a partir da divisão de Descartes são o ponto de partida da Filosofia da Mente que investiga principalmente a relação entre mente e cérebro. Segundo Teixeira (2008, p. 16) foi Descartes que provou que o pensamento ocorre no cérebro. Pois até o século XVII não se tinha certeza sabia disso ainda.

1. Filosofia da Mente Contemporânea

O surgimento efetivo da filosofia da mente como a conhecemos surgiu na segunda metade do século XIX quando foi feita uma retomada entre as teorias metafísicas acerca dos fenômenos mentais e levou-se também em conta as novas descobertas da neurociência. Gilbert Ryle em 1949 lançou uma interpretação inovadora acerca do problema da mente e cérebro. Ryle encarou a dificuldade da relação entre mente e corpo como um pseudo-problema que se origina de uma confusão teórica acerca do emprego linguístico. Dispomos de um vocabulário para tratar o físico e outro para tratar o mental e esperamos que seus respectivos referentes sejam coisas distintas e incompatíveis. Ele defende que uma análise linguística é o bastante para que o pseudo-problema seja desfeito. Embora o campo das discussões da filosofia da mente seja

extremamente rico, é precisamente a análise linguística sobre o problema mente-corpo o nosso foco neste trabalho.

1.1 Linguagem do pensamento

O problema central da filosofia da mente pode ser reduzido à dificuldade de estabelecer a relação entre mente e corpo. Entretanto, do ponto de vista semântico, o problema questiona como se dá a ligação entre os conceitos, que expressam estados mentais, com os objetos do mundo extramental. Qual é a relação entre termos e objetos? Como a descrição dos estados mentais pode ser significativa através da linguagem comum se os estados mentais têm um conteúdo privado? Qual é precisamente a característica que torna estados mentais subjetivos comunicáveis acerca do mundo exterior? Existirá algo como uma linguagem do pensamento que permita que o conteúdo dos meus estados mentais acerca do mundo exterior seja exatamente acessível aos meus interlocutores? A este respeito Dennett (2006, p. 142) fez a seguinte formulação acerca da existência de uma linguagem do pensamento:

Nós e outras criaturas exibimos comportamento inteligente e, uma vez que a produção regular de tal comportamento requer pensamento, e o pensamento requer representação, e nada pode representar exceto um sistema, devemos ser dotados de um sistema de representação interna e utilizá-lo, um sistema que possua sua própria “gramática” e “vocabulário”, e que poderíamos chamar de linguagem do pensamento.

A primeira tese acerca da linguagem mental ou linguagem do pensamento foi escrita por Jerry Fodor intitulada *The Language of Thought*⁶⁴. A noção de linguagem do pensamento foi desenvolvida através de uma comparação com programação computacional⁶⁵. Fodor defendeu a existência dessa linguagem baseado na ideia de que o processo computacional consiste numa espécie de manipulação de símbolos que ocorre de forma ordenada e infere, a partir disso, que os processos cognitivos humanos também se constituem como uma manipulação ordenada de símbolos. A manipulação ordenada de símbolos que seria a linguagem do pensamento tem um

⁶⁴ “A linguagem do pensamento” em livre tradução.

⁶⁵ A linguagem que o usuário usa com a máquina é uma linguagem input/output que será codificada e traduzida numa linguagem que será acessível a outros usuários, a linguagem máquina, linguagem do sistema operacional do computador. A linguagem convencional, seria na visão de Fodor, uma linguagem input/output que permitiria que os outros seres comunicassem a linguagem universal inata em todos os seres humanos, seria análoga a linguagem máquina.

caráter inatista, de modo que a aquisição de qualquer linguagem convencional ou idioma pressupõe a capacidade humana de manipulação ordenada de símbolos, como o que D’Arc (2010, p. 163-164) chama de “qualidade intrínseca do processo cognitivo”. Entretanto, a linguagem mental não se restringe a uma estrutura inata que fundamenta a possibilidade de aquisição e manipulação de símbolos. Para além disso, qualquer um que pretenda defender a teoria da linguagem mental precisa explicar como se dá a interação dos símbolos da linguagem com os objetos do mundo exterior. A relação entre objetos e símbolos da linguagem pode ser melhor esclarecida através das noções de intencionalidade e extensionalidade.

1.2 Intencionalidade

Dizemos que a intencionalidade é uma característica dos estados mentais (crenças, desejos, eventos linguísticos e elocuições) serem dirigidos ou serem acerca de determinado objeto. Segundo D’Arc (2010, p. 164), ela pode ser entendida como uma espécie de relação a ser estabelecida entre dois objetos diferentes quando um é acerca do outro. Mas, como a linguagem pode significar e fazer referência a coisas que são universalmente compreendidas, se ela expressa estados mentais subjetivos? A reflexão acerca da intencionalidade não dá conta sozinha do problema da referência dos termos linguísticos. Para abranger este problema é preciso então falar em termos de referencialidade e extensionalidade. A extensão de uma proposição diz respeito aos objetos do mundo ligados a uma determinada expressão linguística, isto é, o objeto que tal expressão representa. O conceito expresso pela proposição se chama “intenção linguística” ou significado. Intenção e extensão são determinadas uma pela outra, de modo que não pode haver expressões com a mesma extensão e intenções diferentes. A partir da noção de intencionalidade, compreendemos como um estado mental pode ter um conteúdo que se refere a uma realidade extramental e podemos expressar esse conteúdo fazendo referência a essa realidade. Isto seria um modo de dizer que os nossos conteúdos mentais representam o mundo à nossa volta. Mas como é possível que diferentes experiências subjetivas tenham uma mesma representação ou mesmo sejam unificadas através de uma linguagem?

Dennett sugere ser necessária uma teoria completamente geral da representação. Fodor afirma, por sua vez, que tal teoria só pode se desenvolver fundamentada num sistema de

representação interna, uma linguagem do pensamento. Segundo Putnam (Cf. 1981, p. 2) aquilo que as palavras significam depende de algo mais do que aquilo que está no cérebro. Segundo a crítica de Putnam a afirmação de uma estrutura inata que garante a formação de linguagens dotadas de significado e é expressa mediante pensamentos tem de ser identificada em parte com acontecimentos ou objetos exteriores à mente dessas pessoas. De um lado, temos a possibilidade de existência de uma Linguagem do Pensamento que supostamente consegue explicar como formamos nossos signos e conceitos, de outro, a possibilidade da existência dessa estrutura não demonstra de que forma esse pensamento remete a objetos extramentais ou mesmo se depende deles, de alguma forma, para a formação desses signos e conceitos. Afinal, existe ou não uma conexão entre a linguagem do pensamento e os objetos extramentais para a realização de nossas estruturas inatas de formação de signos e conceitos? Assim formulado este é o debate do internalismo e externalismo.

Que tipo de acesso temos aos eventos que ocorrem em nossa mente e que tipo de acesso temos a eventos que ocorrem fora da nossa mente? Que tipo de acesso temos aos nossos estados mentais e que tipo de acesso temos aos objetos e/ou eventos do mundo físico extramental. Para o internalista o significado ou conteúdo mental dependem principalmente do indivíduo. Já o externalismo defende que o significado ou conteúdo mental depende principalmente do mundo, do meio no qual o indivíduo se encontra. Além de investigar se o significado dos conteúdos mentais - e o próprio conteúdo mental - depende do mundo externo ou se restringe à estados internos do indivíduo, o debate também se estende à justificação epistêmica das crenças: somente os estados internos de um indivíduo são suficientes para justificar suas crenças ou é necessário que suas crenças sejam suportadas por dados objetivos do mundo exterior?

Para os externalistas, as crenças são epistemicamente justificadas através do contato com a estrutura causal do mundo (mesmo que o agente cognoscitivo não tenha acesso imediato a essa estrutura). Já os internalistas acreditam que o sujeito tem que possuir acesso imediato a tudo que permite justificar suas crenças. Fatos e eventos do mundo físico são externos à mente e não configuram ou permitem um acesso imediato de modo que os internalistas consideram que não devem fazer parte da justificação epistêmica. Para os internalistas, todas as condições que determinam a justificação de uma crença de uma pessoa envolvem de forma necessária os estados mentais dessa pessoa, aos quais o sujeito teria acesso imediato. Os externalistas não

acreditam que os estados mentais sejam necessariamente a única condição que possibilite a justificação epistêmica. Essas teorias pressupõem que eventos internos e externos à mente são coisas distintas e que o conteúdo dos estados mentais é uma propriedade extrínseca aos próprios estados mentais.

O que interessa ao nosso propósito é precisamente a visão do externalismo que não se trata de uma corrente estritamente consensual. Existe o consenso geral de que os conteúdos de nossos estados mentais intencionais (como o são as crenças e os pensamentos) são, pelo menos em parte, determinados por nossa relação com o mundo externo físico. De outro lado, não há um consenso acerca de se é possível compatibilizar esse sistema com a crença que temos um conhecimento direto dos nossos próprios estados mentais.

Para a filosofia da mente, nossos estados mentais, como crenças e vontade, são estados mentais intencionais com conteúdo ou significado proposicional específico. Os estados mentais são chamados de atitudes proposicionais e expressam uma atitude específica de um sujeito em relação a uma proposição específica. Deste modo podemos dizer que temos “uma crença de que” ou um “pensamento de que”. Ou seja, o sujeito tem uma relação com o conteúdo ou atitude expressa por seus estados mentais. Os estados mentais têm conteúdo representacional e são constituídos por características intencionais uma vez que cada uma das palavras usadas para expressar esse conteúdo representacional remete a algo além de seu próprio significado.

O debate externalista do século XX, apoiando-se na suposição de que os estados mentais tenham sempre conteúdo externo, questionava: como uma frase, pensamento ou uma crença teria conteúdo ou um significado determinado e externo ao próprio estado mental? A este respeito, existem duas perspectivas de resposta, a saber, a perspectiva descritivista e a perspectiva não descritivista. A perspectiva descritivista afirma que quando um indivíduo pensa em um objeto mediante uma descrição, seu pensamento depende necessariamente do objeto particular referido pela descrição: “as descrições da mente podem ser referidas de tal modo que não tenham uma referência específica de como as coisas são no mundo” (D’Arc, 2010, p.173). Tal percepção foca-se somente nas percepções, sentimentos, significados e conceitos interiores de um indivíduo. A perspectiva não descritivista afirma que um sujeito pensaria em um objeto não pela via da descrição, mas pelo fato de o sujeito estar de alguma forma em relação real ou causal com

o objeto. Os pensamentos conectam-se com seus objetos mediante a relação na qual o sujeito estabelece para com aqueles objetos.

Num sentido amplo, do ponto de vista externalista, o significado de nossas frases e palavras e nossos estados mentais intencionais não está “na cabeça”, mas relacionados com o mundo. Tudo o que pensamos, os objetos de nossa experiência e os objetos de nossa fala dependem, ao menos de certa forma, do mundo exterior à mente do sujeito cognoscente, ou melhor, pode-se em algumas variantes, afirmar que eles seriam causados pelo mundo exterior. E se os estados mentais de um sujeito são causados pelo mundo exterior, a condição de existência de estados mentais é fundamentada na existência do sujeito num mundo. Os pensamentos e atitudes proposicionais do sujeito cognoscente são individuados por estados subjetivos e por objetos, acontecimentos, ou eventos do mundo exterior à mente subjetiva. Segundo D’Arc (2010, p. 174) “o externalista não nega os dados intrínsecos, ele simplesmente afirma que os pensamentos do sujeito cognoscente são individualizados parcialmente pelo entorno que o envolve”. O que o sujeito pensa depende de seu entorno. A partir da noção de que a relação com o mundo é a causa dos nossos pensamentos nos aproximamos do pensamento de Ockham e consequentemente de nosso objetivo principal, o de estabelecer se existe uma filosofia da mente em seu pensamento.

2. O pensamento de Ockham

Ockham foi um pensador do século XIV lembrado por sua navalha. Ele desenvolveu seu discurso interior por ocasião da discussão dos universais na Idade Média. Suas primeiras reflexões acerca da linguagem mental resultaram dos esforços de resolver o problema do conhecimento dos universais. Naquele momento, ele se deparou com a dificuldade de conciliar a noção aristotélica de que não há ciência senão do universal com seu posicionamento nominalista segundo o qual o universal não é uma realidade fora do espírito. Em sua teoria do conhecimento, o objeto do saber científico não são as coisas exteriores ao espírito ou à linguagem, mas as proposições, sejam orais, escritas ou mentais. Esta concepção parece resolver seu dilema entre a concepção de ciência aristotélica e sua própria visão dos universais pois nesse sistema as proposições são compostas de termos gerais, isto é, os objetos das proposições são universais.

Surge então a questão de se a ciência é sobre os universais, como ela pode ser acerca do real, uma vez que o real é composto somente por indivíduos segundo o nominalismo de Ockham? Pois é através do que é conhecido pelo vocabulário técnico da semântica terminista por suposição que os termos gerais, sejam eles falados, escritos ou pensados podem representar coisas exteriores. Supor seria tomar ou ocupar o lugar das coisas numa proposição. O objeto do saber ou seu conteúdo, isto é, aquilo mesmo que é sabido é da ordem da proposição, então saber é “saber que ...”. Nesta medida sua concepção de conhecimento é intencional.

Essa abordagem proposicionalista é refutada pelo relativismo linguístico segundo o qual o saber difere de conteúdo segundo a língua em que é formulado (Cf. PANACCIO, 1998, p. 256). Mas, sem dúvida a crítica mais forte a esta aproximação é o ceticismo que emerge do fato de o conhecimento não tratar das coisas mesmas, mas somente de suas representações mentais agrupadas em proposições. Ockham acaba por evitar o relativismo linguístico através de sua concepção de linguagem mental na qual as proposições mentais não são de nenhuma língua. E, por outro lado, seu sistema não tem uma consequência cética devido a atribuição da função referencial da suposição aos termos constitutivos das proposições.

Para Ockham o conhecimento intuitivo do individual ocorre sem recorrer à abstração ou qualquer entidade oculta, forma ou conceito. Isso ilustra o seu princípio de economia do pensamento conhecido como “navalha de Ockham”. Para sustentar essa teoria de um conhecimento intuitivo sem abstração, Ockham tem de formular sua própria teoria do conhecimento. A base da teoria do conhecimento de Ockham é o conhecimento intuitivo do singular, chamada de *notitia intellectualis*. Nesta teoria, todo conhecimento abstrativo (*notitia abstrativa*) que se acrescenta ao conhecimento intuitivo não supõe nenhuma nova operação do entendimento para a formação do conceito. Por que então chama-se de conhecimento abstrativo? Porque esse conhecimento abstrai, prescinde da existência do indivíduo. O objeto é representado por um signo, mas não como uma abstração do objeto.

Na lógica de Ockham, os termos fazem parte de um sistema de signos linguísticos. Signo escrito, falado e mental (OCKHAM, SL *Summa Logicae*, I. 1)⁶⁶. O mental é o conceito (signo mental) que é universal porque representa uma pluralidade de indivíduos. Porém, ele se refere a

⁶⁶ A partir daqui utilizaremos a sigla SL para se referir à *Summa Logicae* de Ockham.

coisas individuais existentes. Os termos escritos e falados são convencionais e não são universais naturalmente. A referência dos termos falados ou escritos em relação aos objetos é o seu significado. Eles significam os objetos. O significado é explicado através da propriedade da suposição, a capacidade que um signo tem de ocupar o lugar de um objeto ou de uma coleção de objetos numa proposição.

Suposição divide-se em pessoal, simples e material (OCKHAM, SL, I. 63). Pessoal: um termo toma o lugar das coisas singulares às quais se aplica, como em “O cavalo é um mamífero”. Simples: o termo toma o lugar dele mesmo como conceito. Como “cavalo” em “‘cavalo’ é um conceito da espécie natural”. Material: o termo toma o lugar da palavra oral ou escrita ao qual corresponde. Como “cavalo” em “‘cavalo’ é uma palavra de seis letras”. Dito isso, passemos à análise dos conceitos.

2.1 O conceito como ato do entendimento

Um ponto fundamental é o status ontológico do conceito. Ockham percebeu que, para ser sujeito ou predicado numa proposição, é preciso apenas ser signo, ter uma significação. E ter uma significação garante a atribuição das funções referenciais da teoria da suposição. O conceito enquanto signo pode ser identificado como um ato do intelecto e ainda assim preservar todas essas funções. Assim, o ato de intelecção passa a ser visto como um signo que desempenha todos os papéis semânticos desejáveis, principalmente os de receber as funções referenciais da teoria da suposição.

A semântica do conceito como ato começa pela gramática. Ockham entende o trabalho feito por outros gramáticos como uma gramaticalização do pensamento e estende essa gramaticalização que era feita pelos greco-romanos aos nomes e verbos a outras partes do discurso como pronomes, participios, advérbios, conjunções, preposições (SL, I. 3). Além disso, também fazem parte da linguagem mental a distinção entre singular e plural, nominativo, genitivo e demais casos do latim como modo, tempo, etc. Assim, o escopo da linguagem mental é quase igual ao das línguas convencionais.

Mas o paralelo entre a linguagem mental e linguagem convencional não é total. Ficam de fora distinções entre masculino e feminino e terminações nominais e conjugações verbais. O critério de Ockham para a divisão é de ordem semântica. A intenção é que a linguagem mental possua uma capacidade expressiva pelo menos tão grande quanto a linguagem convencional – falada ou escrita. Na linguagem mental devem haver tantas distinções gramaticais quantas necessidades de significação. Neste sentido os sinônimos são considerados supérfluos (SL, I, 3, p. 10) porque para fins de significação, tudo o que pode ser significado por um sinônimo pode ser significado por outro. A distinção entre masculino, neutro e feminino também é puramente ornamental.

Ockham pega emprestado da lógica terminista não somente a gramática, mas também divisões lógicas no que diz respeito à significação dos termos. Tanto os termos que significam por natureza quanto signos que significam convencionalmente. A distinção maior é entre termos categoremáticos e sincategoremáticos. Os primeiros possuem significação definida como os substantivos. Já os segundos não possuem uma significação definida como “todo”, “nenhum”, “algum” e só adquirem significação junto de um termo categoremático.

Os termos categoremáticos se dividem entre termos absolutos e conotativos. Os termos absolutos são os conceitos que significam naturalmente e sobre os quais iremos nos deter mais.⁶⁷ Os termos absolutos são os considerados conceitos de espécie natural pois remetem a todos os seus significados da mesma maneira sem estabelecer nenhuma hierarquia entre eles. O conceito de ‘cavalo’, por exemplo, significa todos os cavalos e nada além disso. De forma que pode supor por qualquer cavalo num contexto proposicional. Panaccio aponta que segundo alguns comentadores⁶⁸ os termos categoremáticos e absolutos deveriam ser os únicos a figurar na linguagem mental. Entretanto, ele acredita que o “mentales” como ele chama a linguagem mental de Ockham não constitui uma linguagem ideal como a proposta por Frege “das quais as fontes

⁶⁷Um termo conotativo é dito como um termo que possui uma significação primária e uma significação secundária. Eles compartilham com os termos absolutos a característica de poderem supor e serem verdadeiramente preditados daquilo que significam primariamente. Quando o termo é verdadeiramente preditado e supõe pelo que significa, aquilo que é significado por ele constitui sua significação primária. Entretanto, não podem supor nem são verdadeiramente preditados daquilo que significam secundariamente. Assim, o que caracteriza o termo conotativo é um tipo de significação não predicativa. Ex: o termo ‘branco’ supõe por todos as coisas brancas primariamente e, secundariamente, supõe pela qualidade da brancura. Assim também o termo ‘pai’ supõe por todos os pais primariamente e, secundariamente, supõe pelos filhos que tem um pai.

⁶⁸ (SPADE, 1975, 1980 e Normore, 1990 In: PANACCIO, 1998, p. 268)

semânticas seriam reduzidas a *strict minimun*” (1998, p. 268), mas apenas uma linguagem que elimina os erros mais grosseiros que levam ao equívoco habitual das linguagens ordinárias.

O elemento mais relevante retido por Ockham da lógica terminista para a formulação de sua teoria da linguagem mental é primeiramente a função da suposição que pode ser aplicada tanto aos conceitos absolutos quanto aos termos conotativos da linguagem mental. A significação é uma propriedade invariável do conceito categoremático enquanto a suposição é uma função que o termo adquire quando tomado como sujeito ou predicado de uma proposição. A suposição varia conforme o contexto. O conceito “cavalo” conserva o mesmo significado, mas não a mesma suposição nas seguintes frases: “todo cavalo é um mamífero” que está em suposição simples e na qual “cavalo” é um conceito de espécie. Enquanto em “um cavalo galopa”, “cavalo” está em suposição pessoal. Assim enquanto o significado do conceito absoluto é fixo, a suposição depende do contexto proposicional em que o termo é inserido. Neste sentido, a diferença entre significação absoluta e significação primeira e segunda explica porque um mesmo nome ora figura como conceito, ora figura como um nome simples.

A teoria da suposição também desempenha um papel vital para o estabelecimento dos valores de verdade das proposições. Nas proposições mentais elementares da forma “sujeito + cópula + predicado” as condições de verdade são fundadas na relação entre a suposição do sujeito e a do predicado, isto é, a verdade da proposição depende de ambos suporem pela mesma coisa, terem o mesmo *suppositum*. A condição necessária e suficiente para a verdade de uma afirmativa singular como “bucéfalo é um cavalo”, por exemplo, é que o predicado suponha pela mesma coisa que o sujeito. Uma proposição universal negativa como “nenhum humano é um cavalo” é verdadeira se, e somente se o predicado não supõe por nada que o sujeito supõe. Embora a teoria da suposição tenha uma função essencial no sistema do *Venerabilis Inceptor*, podemos dizer que a significação é a base sobre a qual se constrói toda a sua estrutura. A significação demarca a distinção entre categoremáticos e sincategoremáticos, distingue os termos absolutos e os termos conotativos e é também a base da qual deriva a suposição.

Na linguagem mental, os conceitos são signos, o que significa que os signos linguísticos significam não os conceitos, mas as coisas mesmas. Segundo Panaccio (1998, p. 272) esta noção só é transposta mais facilmente sem incorrer em equivocidade para a ordem mental. A distinção entre significação primeira e segunda é fundamental porque é através desse recurso que sabemos

que as palavras significam conceitos apenas secundariamente, caso contrário, se os conceitos fossem os significados primários das palavras, não poderíamos tomar o conceito como signo e extinguiríamos a função de signo atribuída aos conceitos de supor por seus significados. Na visão de Ockham as palavras estão subordinadas aos conceitos:

Os sons vocais são signos subordinados aos conceitos ou intenções da alma (...)
As palavras são criadas por imposição para significar as coisas mesmas que são
significadas pelos conceitos do espírito. (SL, I. 1)

Aqui compreendemos melhor a significação natural e também a distinção entre significação primária e secundária. O conceito significa qualquer coisa naturalmente enquanto o som vocal significa esta mesma coisa de maneira segunda. Significar naturalmente quer dizer que o conceito é um termo absoluto enquanto a palavra falada (que é diretamente subordinada ao conceito) é um termo conotativo cuja principal característica é significar primariamente o conceito a que é subordinada e secundariamente a coisa significada naturalmente pelo conceito.

A forma como o conceito-signo e a forma significada se relacionam naturalmente é um tema intrigante. Ockham não nos fornece grandes explicações acerca da naturalidade da significação, mas no âmbito da linguagem mental a significação dos termos singulares é explicada em termos de causalidade. Já em relação aos termos gerais, a significação é explicada através da similitude. Os termos singulares são apreendidos convencionalmente por imposição e por ostensão e causados pelo contato empírico com a coisa singular. Já os termos gerais, isto é, os conceitos, significam naturalmente porque são similares às coisas que significam. Ockham não explica exatamente a natureza específica da similaridade entre os conceitos e as coisas que significam. Na teoria madura os termos singulares são atos da intuição intelectual pelos quais o intelecto apreende de forma direta a existência das entidades singulares fora da mente. O intelecto apreende intuitivamente uma coisa singular e forma o conhecimento intuitivo desta coisa e este conhecimento – esta intuição intelectual do singular – é apreendida e formada como um signo capaz de supor por esta coisa singular. Em relação aos termos gerais Ockham explica a similitude entre o signo mental e as coisas exteriores apenas como um modo de assimilação. Panaccio atribui a este caso um isomorfismo que “supõe que o ato mental é dotado por natureza de uma certa estrutura interna capaz de reproduzir, de uma maneira ou de outra, aquela da coisa

que ele representa”⁶⁹ (1998, p. 275). Os atos da intuição são os únicos termos simples da linguagem mental e somente sua significação é determinada causalmente – pelos objetos dos quais são signos. O conceito abstrato é uma similitude que representa de maneira igual todos os indivíduos semelhantes que significa.

A relevância da linguagem mental desenvolvida por Ockham é o desenvolvimento da estrutura de uma linguagem do pensamento que pode ser sistematicamente analisada através das categorias gramaticais e semânticas disponíveis em sua época. O refinamento do seu sistema está no fato dele ter identificado as unidades que eram identificadas como predicado e sujeito nas proposições mentais com atos intelectuais que têm o potencial para atuar como signos capazes de significar, conotar e receber a função referencial da suposição dentro de proposições pelas coisas singulares existentes no mundo. Ele desenvolveu uma linguagem do pensamento que se estrutura e se articula com um instrumento de análise lógica e semântica que estabelece e regula os valores de verdade através da suposição e se mantém fiel à sua motivação principal de evitar o recurso a entidades metafísicas indesejáveis como gêneros e espécies, seja no discurso científico-filosófico, seja no discurso mental.

3. Há filosofia da mente em Guilherme de Ockham?

De tudo quanto dissemos brevemente a respeito dos problemas fundamentais e da teoria da linguagem mental de Ockham, é preciso explicitar alguns pontos antes de responder a esta pergunta. Partimos do princípio de que na filosofia de Ockham existe a noção de mente representada principalmente através da sua teoria da linguagem mental e da sua teoria do conhecimento como um todo que lida com operações do intelecto que ocorrem de maneira privada no que podemos traduzir em termos contemporâneos como mente. Ockham não desenvolveu propriamente reflexões acerca da mente como fez Descartes, mas o desenvolvimento de sua linguagem mental traz implicitamente uma certa noção de mente que seria melhor descrita em termos de linguagem interior. Graças a sua influência aristotélica, aquilo que chamaríamos hodiernamente de conteúdo mental seria descrito em termos de

⁶⁹ “(...) *ce qui suppose l’act mental dote par nature d’une certaine structure interne capable de reproduire, d’une façon ou d’une autre, celle de la chose qu’il représente*”. Tradução nossa.

intenções da alma. Entretanto, não estamos afirmando que exista em sua filosofia nenhuma preocupação ou comprometimento específico em desenvolver uma teoria que responda aos problemas e questionamentos postos pela filosofia da mente contemporânea. Fazer uma leitura da filosofia da mente de Ockham sob este paradigma consistiria num anacronismo e é preciso compreender suas ideias à luz do contexto de problemas filosóficos específicos com que ele lidava em sua época, a maioria dos quais bastante distantes dos temas discutidos na filosofia da mente contemporânea. Dito isto, embora não se possa dizer que a teoria da linguagem mental de Ockham vise cumprir a “agenda” da Filosofia da Mente contemporânea – aquela de defender por exemplo, que exista uma mente e que ela tem uma linguagem específica pela qual a conhecemos – é inegável que sua teoria deva ser levada em conta no debate contemporâneo principalmente no momento em que se começa a discutir a possibilidade de haver uma linguagem própria do pensamento que seja universalmente a todos os seres humanos. Embora a linguagem mental de Ockham não vise responder a este tópico de discussão específico, ela serve para o debate contemporâneo para provar a possibilidade de uma linguagem própria do pensamento.

O ponto de confluência que fizemos entre o sistema de Ockham e a filosofia da mente diz respeito principalmente ao externalismo. Principalmente com três vertentes específicas o externalismo linguístico, o externalismo do conteúdo mental e o externalismo epistêmico. Essas correntes debatem principalmente a relação e influência do mundo exterior nas nossas crenças e estados mentais. O externalismo linguístico é aquele no qual o falante, no momento de sua enunciação, depende de alguns elementos externos a ele, elementos que iriam além de suas estruturas internas e o conduziram a elementos do seu entorno. Para o externalismo de conteúdo mental o verdadeiro conteúdo que compõe as crenças de um agente não depende unicamente de estados internos desse agente cognoscente, mas, também, de alguns fatores externos à sua mente. O externalismo epistêmico, postula que o que um agente acredita saber não depende exclusivamente de determinados estados internos do agente, mas, também, de alguns dados exteriores. Existe no sistema ockhamista uma relação com o externalismo linguístico e também com o externalismo de conteúdo mental.

O primeiro indicativo que nos leva a acreditar que existe uma filosofia da mente em Ockham é a existência de uma linguagem mental. O segundo ponto é que a concepção de ciência de Ockham é proposicional – saber é “saber que ...” Isso expressa que a noção de conhecimento

ockhamiana é intencional. Ao mesmo tempo todo o aparato da linguagem mental possui função referencial, isto é, se refere diretamente a objetos exteriores. Neste sistema o conteúdo mental se refere ao mundo exterior e existe uma preocupação de que as proposições mentais façam referência à realidade fora da alma de maneira inequívoca. De modo que poderíamos aproximar muito a concepção ockhamiana de uma visão externalista.

Do ponto de vista epistemológico a aquisição do conhecimento em Ockham é explicada através da teoria dos atos mentais, uma teoria essencialmente intencional segundo a qual as proposições mentais estão relacionadas ao mundo, porém a aquisição do conhecimento se dá principalmente através de operações do intelecto. A teoria de Ockham forneceria uma solução para o debate externalista de como a linguagem pode comunicar conteúdos do mundo exterior através de estados mentais que são subjetivos. Longe de cair em qualquer tipo de subjetivismo forte, a noção de Ockham preserva a universalidade da linguagem porque postula uma linguagem mental que é inata e anterior às línguas. A estrutura que compõe a linguagem mental é extremamente semelhante à linguagem das línguas convencionais, porém bastante reduzida, o que garante a universalidade da comunicação dos estados subjetivos através da linguagem convencional. Tanto a comunicação quanto a aquisição de qualquer língua são possíveis graças a estrutura interna inata da linguagem mental do nosso autor.

Conclusão

A filosofia da mente presente no pensamento de Ockham diz respeito a considerações gerais sobre a mente e seu funcionamento que não visa responder aos interesses específicos da filosofia da mente contemporânea. Contudo, inserida no debate sua teoria pode fornecer uma perspectiva crítica considerável para tratar sobretudo do problema existência da linguagem do pensamento e algumas dificuldades do externalismo contemporâneo.

Referências

D'ARC, A. Ferreira. *A linguagem mental de Guilherme de Ockham*: Contribuição frente ao externalismo contemporâneo. 2010, 220 p.

FODOR, J. A. *The language of thought*. Mass.: Cambridge, 1981.

OCKHAM, William. *Ockham's theory of Terms: part I - of the Summa Logicae*. Trad. and Introd. by Michael J. Loux. Indiana: St. Augustine's Press, 1998.

PANACCIO, Claude. *Le discours intérieur: de Platon à Guillaume de Ockham*. Paris: Éditions du Seul, 1998.

PUTNAM, Hilary. *Reason, truth and history*. Cambridge University Press, Cambridge, 1981, pp. 1-21.

RYLE, Gilbert. *The concept of mind*. New York: Routledge, 2009

TEIXEIRA, João de Fernandes. *Como ler a filosofia da mente*. São Paulo Paulus, 2008. 72 P.

O conceito de mente na filosofia indiana: breves tópicos introdutórios

The concept of mind in the Indian philosophy: short introductory topics

Lincoln T. Zacconi⁷⁰

RESUMO: Nas palavras de Dasgupta; a antiga civilização indiana gerou uma unidade concreta, formada por grandes desenvolvimentos em todas as áreas do conhecimento humano, mas a mais importante realização do pensamento indiano foi a filosofia. O desenvolvimento da filosofia na Índia ocorreu na transição do final do século VII para o século VI AEC, com a formação de diversas correntes de pensamento. As primeiras concepções sobre as possíveis funções da mente na Índia vêm através de suas obras clássicas; os *Vedas* (literalmente o conhecimento - esta obra é um conjunto de hinos, cantos, poemas, relatos que constituem uma amostragem do que era a Índia em seus primórdios, datando sua primeira versão escrita, por volta de 2000 AEC) e os *Upaniṣads*. Uma das referências mais interessantes a condição da mente se encontra em um dos textos do *Upaniṣads* que diz que a mente é a fonte da escravidão ou da libertação do indivíduo. Desta forma este trabalho mostrará uma breve descrição sobre o conceito de mente, desviando totalmente do eixo ocidental, e lançando algumas luzes sobre um sistema filosófico, que apesar de sua antiguidade, ainda apresenta novidades em seu estudo. A organização da apresentação se dará em função das ideias apresentadas pelos principais *darśanas* a saber *Buddhismo*, *Nyāya-Vaiśeṣika*, *Sāṃkhya*, *Mīmāṃsā*, *Vedānta*, *Cārvāka* e Jainismo. Reforçando que apenas é uma visão muito introdutória de um assunto extremamente vasto.

PALAVRAS-CHAVE: Mente, Filosofia Indiana, Filosofia da Mente, Mente na Filosofia Indiana

ABSTRACT: In the words of Dasgupta; The ancient Indian civilization generated a concrete unity, formed by great developments in all areas of human knowledge, but the most important achievement of Indian thought was philosophy. The development of philosophy in India occurred in the transition from the late seventh to the sixth century BCE, with the formation of diverse currents of thought. Early conceptions about the possible functions of mind in India come through their classical works; The *Vedas* (literally knowledge - this work is a collection of hymns, chants, poems, accounts that constitute a sampling of what India was in its early days, dating to its first written version, around 2000 BCE) and the *Upaniṣads*. One of the most interesting references to the condition of the mind is found in the *Upaniṣads* texts which says that the mind is the source of slavery or liberation of the individual. In this way, this work will show a brief description of the concept of mind, totally diverging from the western axis, and throwing some light on a philosophical system that, despite its antiquity, still presents novelties in its study. The organization of the presentation will take place according to the ideas presented by the main *darśanas* namely Buddhism, *Nyāya-Vaiśeṣika*, *Sāṃkhya*, *Mīmāṃsā*, *Vedānta*, *Cārvāka* and Jainism. He emphasizes that it is only a very introductory view of an extremely broad subject.

KEYWORDS: Mind, Indian Philosophy, Philosophy of Mind, Mind in Indian Philosophy

⁷⁰ Mestrando em Filosofia na FFLCH / USP. Email: zacconi@gmail.com / zacconi@usp.br

Introdução

Uma característica bem marcante da filosofia indiana é a de abranger vários campos de estudo de forma metódica, para que dessa forma se crie uma estrutura coesa, sistêmica e logicamente estruturada de todos esse conjunto de saberes.

Por abranger vários conhecimentos, pode se afirmar, que não existiu um único sistema filosófico indiano que não tenha tratado sobre o conceito de mente, alguns de maneira muito semelhante entre si, outros de maneira diferenciada, alguns de modo epistemológico, outros de modo psicológico. Apesar dessas diferenças sempre foi feito um esforço contínuo para relacionar estas duas últimas visões e conseguir chegar em uma explicação satisfatória sobre a natureza da mente.

O que ocorre no estudo da filosofia indiana é que alguns problemas se apresentam ao se adentrar nesse universo. Aqui cabe apresentar uma pequena lista desses elementos, a começar com a linguagem, uma vez que esses trabalhos estão em sânscrito e escritos em um estilo muito particular⁷¹, a outra é o uso de termos que são específicos e técnicos em cada corrente de pensamento e outro (mas não o último) é que muitos dos textos foram perdidos. Outro ponto que tem que ser destacado é o problema quanto a datação dos textos em sânscrito, desta forma em vez de usar um possível ano (que pode mudar devido a um continuado estudo sobre essas datações, o que pode fazer recuar a data de vários textos) utiliza-se como referência o século onde essa data pode estar contida.

Existem várias visões, conforme cada corrente de pensamento filosófico da antiga Índia, sobre o conceito de mente, desta forma, dentro do escopo desse artigo, se torna um trabalho insustentável tentar explicar cada um desses conjuntos de visões de forma detalhada⁷², desta forma o que irá ser apresentado serão tópicos sucintos, sobre o tema, a partir da perspectiva de cada um dos *darśanas*⁷³.

71 Este estilo é denominado sūtra, que se trata de uma forma extremamente concisa de tratar os temas, é tal qual um índice sumarizado dos tópicos a serem tratados e analisados em profundidade em outros textos paralelos.

72 Para uma compreensão sobre cada uma das interpretações existe a necessidade de se recorrer aos textos canônicos de cada uma das correntes de pensamento, esse corpus canônico é formado pelos *sūtras* (vide nota 1) e pelo conjunto de textos dos vários comentadores que vão se sucedendo ao longo da história de cada uma das correntes, o que forma uma vastíssima coleção de textos.

73 Termo, em sânscrito, que denomina cada uma das correntes de pensamento tanto filosófica, quanto científica, na Índia, e pode ser livremente traduzido como ponto de vista estruturado ou ponto de vista obtido pela razão.

Antes da apresentação das sínteses, alguns pontos têm de ser esclarecidos para a compreensão do desenvolvimento dos sistemas filosóficos indianos. Todas essas correntes tiveram como base duas obras que são lapidares dentro da cultura indiana, os *Vedas* e os *Upaniṣads*.

Os *Vedas* não são textos filosóficos, na verdade são um conjunto de hinos, narrativas, tradições, comportamentos que passaram da tradição oral para a escrita. Mas mesmo não sendo um texto técnico existem algumas considerações que podem ser vistas como o preâmbulo de futuro pensamento filosófico. Em sequência a este conjunto de textos, surge uma outra obra denominada *Upaniṣads* que apresentam uma nova visão sobre muitos temas apresentado nos textos védicos e já possuindo preocupações filosóficas bem estruturadas. Por ser uma obra posterior a obra védica, ela também é denominada como *Vedānta* ou, em uma tradução literal, o fim dos vedas (a parte final).

Enquanto os *Vedas* procuram descrever os fenômenos do mundo, os *Upaniṣads* passam a se concentrar nas interpretações destes fenômenos e as correntes filosóficas procuram explicações racionais (CHENNAKESAVAN 1960 p. vii, DASGUPTA 2010 pp. 258 - 273).

1. *Vedas*

Dentro dos *Vedas* não existe uma definição exata sobre a mente e está geralmente é usada como sinônimo de consciência.

O que ocorre nestes textos é o uso de vários termos para designar a mente, tais como o estado de consciência (*saṃjñāna*), compreensão (*saṃjñā*), entendimento (*avagamaṇa*), discernimento (*drṣṭi*), resolução (*dhṛti*), reflexão (*manas*), impulso (*jūt*), vontade (*saṃkalpa*), ignorância (*ajñānam*) e sabedoria (*vijñānam*).

Os textos védicos podem ser compreendidos como a primeira forma de organizar um conjunto de conhecimentos oriundos da tradição oral, fazendo partes deste conjunto de textos hinos religiosos, comentários sobre práticas e costumes, pontos históricos, rituais e tradições. O que ocorre é que não existe uma preocupação em entender o conceito de mente, e sim de relacionar as possíveis atividades mentais (CHENNAKESAVAN, 1960 p. 47, SIDDHANTALANKAR, 1969 pp. 82-100).

2. *Upaniṣads*

Os textos dos *Upaniṣads* passam a concentrar suas interpretações sobre o conceito de mente, onde os instrumentos do conhecimento passam a ser os objetos de uma primeira forma de pesquisa investigativa. Assim, ao longo do texto os seguintes conceitos vão sendo usados e associados com a mente:

Manas: literalmente mente

Prāṇa: instrução, aquilo que se sabe

Samkalpa: vontade; intenção; convicção; resolução

Citta: conhecimento, mente, o intelecto, a razão, sabedoria

Os mais antigos textos vedantinos não fazem uma categorização em separado das funções mentais, sendo que em uma primeira análise poderia ser considerado que a mente possuía duas formas: a primeira seria a forma mais pura da mente, que não é consciente, e a segunda uma forma sutil de matéria⁷⁴. Dentro de um dos vários textos que formam os *Upaniṣads*, naquele denominado *Pingala Upaniṣads* existe uma certa hierarquia das funções mentais que é assim expressa: os órgãos dos sentidos, os órgãos motores (os órgãos da fala, membros superiores, membros inferiores, sistema excretor e sistema reprodutivo), a mente perceptiva (que coordena os órgãos da percepção), o intelecto, e a mente subconsciente que formaria um depósito de impressões passadas.

A mente era caracterizada por apresentar duas características, ela poderia ser denominada pura ou impura. A mente desprovida de desejo é denominada pura, enquanto aquela denominada impura é aquela dirigida pelo desejo e a vontade. A mente sozinha é ao mesmo tempo o elemento que amarra ou que liberta o ser humano.

O sistema nervoso do corpo prove os canais através dos quais a mente viaja; a direção na qual ela se move é determinada por seu desejo. Quando a mente se torna pura e sem desejos ela toma o sentido ascendente, e na hora da morte ela deixa o corpo. Enquanto a mente se mantém repleta de desejos seu curso é descendente em direção aos campos onde estes possam ser satisfeitos.

Ainda segundo os conceitos vedantinos, o coração possui 101 nervos sendo que um deles penetraria o centro da cabeça, ao se tomar o sentido ascendente deste, o indivíduo alcançaria a

74 Segundo a tradição constante dos textos vedantinos, a comida ingerida é quebrada em três instâncias; a mais baixa é excretada, a intermediária nos nutre e a sutil forma a mente.

imortalidade. Os outros 100 nervos levariam a outros diferentes mundos (SARMA, 1975 pp. 1-29).

3. *Vaiśeṣika*

Dentro da corrente de pensamento *Vaiśeṣika*, o ser humano é constituído de corpo, órgãos sensitivos de conhecimento e ação, *manas* e *ātman/self*⁷⁵. Tanto a mente como *ātman/self* são infinitas em número e eternas em existência.

Existe uma relação unívoca entre *ātman/self* e mente, isto quer dizer que cada *ātman* é associado a uma única mente. A conexão entre esses dois elementos é estabelecida em duas etapas, os objetos de conhecimento são percebidos pelos órgãos de sentidos e conhecimento, estas percepções chegam à mente que transmite para *ātman/self* todas essas informações.

Ātman/Self é uma substância com dois tipos de qualidades, as denominadas genéricas (que podem ser compartilhadas com outras substâncias) que são: número, magnitude, conjunção e disjunção e distinção e as específicas, que são: conhecimento, prazer, dor, desejo, aversão, vontade, mérito, demérito e impressões. A consciência não é uma qualidade essencial pertencente a ela.

A mente (*manas*) é tal qual um aparelho para obter o conhecimento do que se passa no *ātman/self*. Assim ela passa a ser como o sentido, através do qual a introspecção dos estados mais internos é possível e por esse motivo é denominado de sentido interno. A mente também é o meio através do qual as impressões dos sentidos externos alcançam a *ātman/self*. Ela é atômica⁷⁶ em extensão, mas não no mesmo sentido que os átomos físicos, formando um tipo de substância. As características da mente é que ela é imaterial, atômica, inconsciente e capaz de ações ou movimentos. Percepção de objetos externos tem lugar com a ajuda da mente. Cognição, prazer, dor, desejo, aversão e vontade são percebidas através da mente. A mente é um órgão interno pelo meio do qual *ātman/self* recorda, tem dúvidas e sonha (RAO 2002 p. 319, CHENNAKESAVAN, 1960 pp. 10-14).

75 O conceito de *ātman* está muito mais próximo ao conceito contemporâneo de Self (principalmente a visão junguiana), do que a um conceito de “alma” tal qual é compreendido no ocidente. Mesmo assim, em cada corrente filosófica sua interpretação é única, por este motivo sempre será usado o binômio *ātman/self* ao longo do texto para indicar sua múltipla interpretação, fato esse que escapa ao escopo do artigo.

76 O conceito de átomo dentro do *darśana vaiśeṣika* é que são elementos cuja dimensão é não observável, tem uma variedade de elementos finita e a combinação desses elementos é que formam toda a matéria conhecida, assim ao se falar em dimensão atômica implica em algo não mensurável.

4. Nyāya

Dentro da visão Nyāya⁷⁷ a constituição do ser humano é a mente (*manas*), alma/self (*ātman*), os órgãos dos sentidos (*indriya*) e o corpo (*śarīra*).

O corpo é constituído atomicamente em uma estrutura que permite seu crescimento, e modificação, é apto a se desintegrar e ser destruído. É a base para os órgãos dos sentidos e um meio para *ātman*/self ter contato com o meio externo. Os órgãos dos sentidos são elementos destinados para um tipo específico de conhecimento e atividade. E são divididos em dois grupos os *jñānendriya* e *karmendriya*.

Os *jñānendriya* são os responsáveis pela forma de como o corpo recebe conhecimento do mundo, ou seja, a audição, o tato, a visão, o paladar o olfato e os sentimentos internos, assim a mente também está associada aos órgãos sensoriais. Os *karmendriya* são os órgãos de ação, são aqueles que desempenham algum tipo de atividade que são: movimentos em geral, o ato de pegar, o falar, a satisfação, e a excreção.

A mente também é responsável em realizar o conhecimento servindo também como um meio para levar as percepções desde os órgãos dos sentidos até o *ātman*/self. A mente também tem a capacidade de apreender os estados internos de prazer, dor, etc. É um elemento de tamanho atômico e é único para cada corpo.

Manas também é considerada como um órgão de atenção, com a restrição de que somente se ocupa de um elemento por vez, mas em rápidas sucessões o que gera uma corrente de pensamentos ou a sensação de um foco contínuo ou uma corrente de consciência. *Ātman*/Self é considerado como o real conhecedor, o sensor e ator por trás do mecanismo da mente, sentidos e corpo. Esses elementos são meros instrumentos de *ātman*/self.

Enquanto a mente é atômica em tamanho, o *ātman*/self é ilimitado em espaço e tempo (*vibhu* - eterno e *nitya* - invariável). Apesar de *ātman*/self ser capaz de saber, sentir, e agir ele nada pode fazer sem a intermediação, ou a instrumentação da mente, órgãos dos sentidos e corpo (CHAKRABARTI, 1999 pp. 103-113, CHENNAKESAVAN, 1960 pp. 69-75).

77 Muitos elementos que constituem a estrutura de conhecimento do *darśana Nyāya* é proveniente do *darśana vaiśeṣika* (teoria atômica e metafísica principalmente) uma vez que as duas correntes se fundiram (DASGUPTA 2010 pp. 274-363, MATILAL 1977).

5. *Sāṃkhya*

O sistema filosófico *Sāṃkhya* se utiliza da descrição de todo um processo evolutivo para determinar o surgimento de várias instâncias de consciência, até a determinação da mente. Todo esse processo evolutivo tanto de seres animados como inanimados, tem início com um estado denominado *prakṛiti*, que é considerado como a matéria primordial (física).

Em um primeiro momento *prakṛiti* está em um estado de dormência, em um equilíbrio, indefinido e não manifesto. Após sofrer uma singularidade este estado se altera e passa a ser denominado *mahat* (vasto, grande) que leva a criação da próxima categoria, *ahamkara* (autoconsciência) a primeira manifestação de elementos individuais com identidade distinta. Deste ponto surge o conhecimento e os órgãos para adquirir o conhecimento e realizar ações. *Manas* vem a partir desta autoconsciência e se mantém em um estado de equilíbrio. Assim, os órgãos de ação e de conhecimento, a autoconsciência e *manas* formam a componente psíquica do ser humano.

De acordo com esta corrente filosófica existe uma sucessão de princípios evolutivos que, em uma ordem cosmogônica de ações, geram os seres vivos e suas capacidades. São eles:

Prakṛiti - É a raiz causal de tudo que pertence ao mundo objetivo (fenômenos em constante mudança). Seu princípio é caracterizado por estar em um estado de dormência, em equilíbrio, indefinido e não manifesto com a estabilidade dos três *guṇas* (qualidades)⁷⁸. Qualquer forma de distúrbio nesse estado (singularidade) resulta uma criação no sentido de um processo evolutivo de uma outra ordem fora desta.

Mahat - Inteligência cósmica, a primeira manifestação de *pradhana* (natureza primordial, o ator principal).

Ahamkara - O segundo manifesto que ascende a partir de *mahat*. Esse processo gera a autoconsciência, o surgimento do ego.

Tanmāta - Substância, os cinco elementos básicos (fogo, água, terra, ar e *ākāśa*) que se manifestam, surgindo a partir de *Ahamkara*.

⁷⁸ São elas: *sattva* - a pura essência, primeiro *guṇa* da natureza, a pureza da verdade; *rajas* - a paixão, segunda *guṇa* da natureza, essência ativa de poder e desejo; *tamas* - as trevas, terceira *guṇa* da natureza, essência passiva da ignorância e inércia.

Vikṛti - Mudança, transformação, degradação a partir do mundo fenomênico.

Destas estâncias surgem as classes de elementos (que também tem a denominação de *guṇas*) que não dão origem à evolução de quaisquer novas ordens de ser. Esses elementos acima descritos (os instrumentos da ação e conhecimento) evoluem a partir de *Ahaṁkāra*. São eles:

Maha-bhutas - Os cinco elementos (fogo, água, terra, ar e *ākāśa*) decorrentes de seus princípios básicos (princípios que caracterizam cada elemento).

Puruṣa - Não é nem causa nem efeito de nada. Princípio estático dentro da ordem cósmica que é ativado pelo princípio dinâmico da natureza. Sistema pelo qual não existe a capacidade de um corpo existir sem um outro corpo. Esse princípio pode assim ser descrito; existem duas possíveis formas, concomitantes, da existência de *puruṣa*, liberto e não liberto. A primeira forma é uma pura consciência, enquanto a segunda forma é a consciência ligada ao corpo.

O sistema psicológico do *sāṁkhya* se divide em duas frentes: instrumentos de conhecimento e funções dos instrumentos de conhecimento. Os instrumentos são: o intelecto (*buddhi*), o ego (*ahaṁkāra*), o sentido interno (*manas*) e os órgãos dos sentidos (órgãos de conhecimento).

Os instrumentos de conhecimento, denominados como *antaḥkaraṇa*, são constituídos pelo intelecto, ego, o sentido interno, e os sentidos externos (audição, paladar, tato, olfato e visão). A função do intelecto (*buddhi*) é a de verificação ou determinação da real natureza de tudo. O ego (*ahaṁkāra*) tem a finalidade de perceber o indivíduo como algo único dentro da natureza.

A função de *manas* é confirmar a natureza das informações recebida pelos órgãos sensoriais e dos órgãos de ações envolvidos (quais as reações do corpo físico como resposta as percepções dos órgãos dos sentidos). Além de perceber o mundo, *manas* é responsável por duvidar dessas informações. Outra função da mente é a formação do desejo por algo, sendo esse ato executado pelos órgãos de ação (CHENNAKESAVAN, 1960 pp. 56-62, RAO, 2012, p. 318).

6. *Yoga*

O *Yoga* pode ser considerado como uma filosofia psicológica, sendo a ciência da disciplina mental. A mente recebe o nome de *citta* que pode ser traduzido por conhecimento, pensamento, ideia. Dentro deste sistema *citta* denomina todo o mecanismo de conhecimento e por sua vez incorpora três outros aspectos desse processo:

Buddhi - A capacidade para a percepção, determinação e convicção. Fonte da virtude e sabedoria, sendo que manifesta a si mesmo por meio da determinação e resolução em pensamento e ações, formação e retenção de conceitos e generalizações. Enquanto *manas* está registrando os objetos de pensamento, é em *buddhi* que ocorre a discriminação, determinação e reconhecimento.

Ahaṁkāra - Representa o princípio do ego. Ele arroga para si mesmo a experiência realizada por *manas* e a passa para o *buddhi* para ser determinado. Ele é o sujeito consciente de toda a experiência psicológica.

Manas - É a força diretiva por trás de todas as ações, sendo que ele possui, dentre outras, a capacidade de atenção, seleção e assim por diante, além de sintetizar as várias sensações. Ele pode perceber, mas não pode transmitir, como faz o *buddhi*.

Estes três elementos constituem *citta* como um todo, formando o que se pode chamar de mente. Praticamente não existem muitas diferenças entre o sistema *Yoga* e o *Sāṁkhya* quanto a mente (suas diferenças ficam entre como suas estruturas ligadas a lógica, epistemologia, ontologia, ética e psicologia são fundadas) (CHENNAKESAVAN, 1960 pp. 119-125, RAO, 2012 p.319).

7. *Mīmāṃsā*

Dentro do pensamento *Mīmāṃsā* o ser humano é constituído do corpo físico, órgãos dos sentidos, mente (*manas*), consciência e *ātman/self*. Todos esses elementos possuem funções distintas que são realizadas por meio da influência de *apūrva* (literalmente – aquilo que não existia antes), que se apresenta como um tipo de força gerada pelos atos virtuosos e pelos vícios.

Desta forma tem-se que o indivíduo é a interação da mente (*manas*) com o self (*ātman*) conforme as definições:

Manas - É um órgão interno pelo qual é feita a apreensão do que é recebido pelos órgãos dos sentidos assim como a dor, prazer, cognição e outras qualidades internas do self. A mente é sutil e imóvel, sendo uma substância intangível que não é nem um efeito nem uma causa de qualquer outra coisa, apesar de ser permeável é limitada pelo corpo. A mente serve como um órgão de percepção interno que funciona junto ao self.

Ātman - É não material, eterno e totalmente permeável, sendo distinta do corpo, dos órgãos de percepção e processos cognitivos. É o conhecedor, o agente ativo, e o desfrutador. A cognição, dor, prazer, desejo, aversão, impressões, as volições são modificações de *ātman*, que não são eternas e sem ações durante o sono.

A percepção se dá por meio da seguinte corrente de ações: mundo real → órgãos dos sentidos → *manas* → *ātman*. Supervisionado por *manas* os órgãos dos sentidos produzem as sensações de textura, odor, paladar, gerando assim a cognição de prazer, dor, aversão, etc. (VERPOORTEN, 1987 pp. 6-21, CHENNAKESAVAN, 1960 p.107).

8. *Advaita Vedanta*

Segundo a corrente filosófica *Advaita Vedanta* a mente é uma substância sutil, não tem dimensão atômica nem dimensão infinita, mas é dita ser neutra e possuidora de possibilidades latentes e capaz de permear todo o corpo do indivíduo ao qual ela pertence, ou seja, cada indivíduo possui uma mente particular e diferente das demais. A mente é denominada como um órgão interno, produzida a partir da essência mais pura dos cinco elementos e pode ser denominado por quatro nomes diferentes de acordo com sua função, são eles:

<i>Manas</i> -	Cognição
<i>Buddhi</i> -	Intelecto
<i>Citta</i> -	Armazenar experiências na memória

Dentro do sistema *Advaita Vedanta*, tem-se o conceito de *antaḥkaraṇa* (literalmente capacidade de pensamento) que pode ser compreendido como a somatória da cognição, intelecto, o ego e o armazenamento das experiências. A união de todos esses elementos é, como tudo a nossa volta, um produto de *māyā* (o mundo das aparências).

Um outro modo de interpretar essa situação seria a percepção de que a mente é apenas a totalidade de estados e processos conscientes. O self sozinho é consciente, e a consciência não pode florescer como um fator contingente da relação sujeito-objeto. O self que não é nem mente nem matéria, é a base tanto dos estados mentais quanto dos estados físicos da existência, é a base de toda a experiência seja ela física ou mental (CHENNAKESAVAN, 1960, pp. 66-71).

9. *Cārvāka*

É um erro pensar que a Índia é a terra da espiritualidade, várias correntes filosóficas, quando de seu estudo mais aprofundado, mostram isso, mas principalmente o darśana *Cārvāka* que é uma corrente totalmente materialista.

Para essa linha de pensamento filosófico a mente é um subproduto do corpo e a consciência é o resultado dela. Assim na morte, tanto corpo como mente e consciência são desintegrados. Os cinco órgãos dos sentidos pertencem ao corpo e todo conhecimento provem somente destes órgãos. A mente não é um órgão de conhecimento, ela apenas regula as sensações, que, por sua vez, levam a percepção. A continuidade dessas percepções levanta a falsa impressão de uma substância permanente denominada mente.

Dentro do pensamento *Cārvāka* existe a importância da percepção direta como fonte de conhecimento, não havendo a necessidade de outros elementos tais como inferência e testemunho.

Para essa corrente o universo é uma combinação fortuita de partículas elementares de matéria, assim sendo a vida tem origem pela associação de átomos sem nenhum propósito e não existe nenhum desenho por trás de sua criação. O mundo é uma dança sem significado de átomos, desta forma a morte é apenas a dissociação dos átomos (RAO, 2012, p. 322, KAR, 2013 pp. 75-86).

10. *Buddhismo*

Para o buddhismo nada permanente existe no campo físico ou mental. A personalidade humana passa a ser uma coleção de vários agregados. Tudo é um fluxo perpétuo com nada se mantendo o mesmo através de sucessivos movimentos. Existe uma continuidade em tudo, em constante mudança, mas sem identidade. A personalidade humana é um agregado dos seguintes elementos:

Rupa - O corpo físico e todos os órgãos.

Vedana - Compreende três tipos de sentimentos de prazer, dor e sentimentos naturais.

Samjñā - Incluem percepções dos objetos, que possuem nomes e todo o conhecimento articulado dos objetos.

Saṅkhāra - Inclui todos estados mentais.

Vijñāna - Toda forma de consciência dos objetos e autoconsciência.

Dentro do Buddhismo existem, pelo menos, três termos para indicar mente:

Manas - poder mental ou faculdade mental.

Vijñāna - discernimento ou faculdade de consciência.

Citta - mente ou pensamento.

Essas diferentes nomenclaturas são dependentes de suas funções, mas não importa por qual nome seja referido sempre será uma referência ao órgão interno de conhecimento. A existência deste órgão em razão de que se os órgãos de percepção, *ātman* e os objetos fossem suficientes para a percepção, essa seria eterna, uma vez que *ātman* é eterno e os órgãos de sentidos e os objetos estão em contínua interação, sem discernimento do que ocorre uma vez que existe um contínuo fluxo de entrada de informações. Por esse motivo existe a necessidade de um órgão interno que controle esse fluxo e que estabeleça uma espécie de dualidade de informações tais como a atenção e a não atenção, a percepção e a não percepção, por exemplo, elementos estes que não existiriam em um fluxo constante de informações.

O buddhismo possui uma complexa e contra intuitiva teoria da mente e fenômeno mental. Existe a visão de que o domínio do mental engloba um conjunto de faculdades e processos inter-relacionados, mas não associam fenômenos mentais à atividade de um eu substancial, independente e duradouro ou agente. Em vez disso, as teorias budistas da mente se centram na doutrina do não-eu que postula que os seres humanos são redutíveis aos constituintes físicos e psicológicos e aos processos que os compõem (DASGUPTA, 2010 pp. 78-166, CHENNAKESAVAN, 1960 pp. 65, 82, 90, 97,98).

11. *Jaina*

Os jainistas atribuem que todas as atividades mentais tais como a cognição, afeição e conação pertencem ao *self*. A mente teria uma função limitada e o *ātman* é o princípio permanente da consciência. A mente não é material e é diferente da constituição e independente do corpo e do cérebro, sendo a agente e fonte de todas atividades mentais e é diretamente experimentado pelo eu, também é a fonte da consciência.

Como um instrumento de conhecimento o corpo é formado por cinco sentidos e manas. Cada órgão de sentindo apreende objetos específicos enquanto manas reconhece todos os objetos de todos os sentidos, assim sendo ele não pode ser considerado como os outros sentidos. Ele é um instrumento interno que auxilia o *ātman* a fazer a cognição dos objetos percebidos pelos sentidos e os estados internos.

Todo senso de percepção é resultado do contato entre os órgãos de sentido e seus objetos. A filosofia jainista não sustenta que a mente se move para fora para o objeto na forma de modificações para tomar forma de objetos para conhecê-los, mas por outro lado sustenta que os objetos em si entram em contato com os sentidos. A mente não entra em contato direto com os objetos. Conhece esses objetos, que já são percebidos pelos sentidos. Além disso, manas conhecem as atividades internas e modificações como prazer, simples, amor, aversão, reconhecimento, etc., que são realizadas pela própria mente (DASGUPTA 2010 pp. 169-207, KACHHARA, 2011 pp. 1-14).

Conclusão

Determinar um conceito único sobre mente e filosofia da mente dentro da filosofia indiana é um trabalho enriquecedor, uma vez que todas as correntes de pensamento filosófico (e aqui apresentadas as principais) contaram em seus sistemas com esse estudo. Impossível poder mergulhar sobre cada aspecto de maneira concisa, uma vez que são textos muito extensos portadores de uma linguagem altamente técnica e particular de cada sistema filosófico. O intuito deste artigo foi mostrar o básico de algumas características para que se possa gerar estudos e discussões posteriores.

Referências

- CHAKRABARTI, K. K. Classical Indian philosophy of mind. Delhi. Motilal Banarsidass Publishers, 1999
- CHENNAKESAVAN S. The concept of mind in Indian philosophy. London: Asia Publishing House, 1960.
- DASGUPTA, S. A history of Indian philosophy v. 1. Delhi: Motilal Banarsidass Publishers, 2010.
- KACHHARA, Narayan Lal. Philosophy of Mind: A Jain Perspective. US-China Education Review. vol. 8(3), pp. 1-14, 2011
- KAR, B. The philosophy of Lokāyata. Delhi. Motilal Banarsidass Publishers, 2013.
- MALVANIA, D., SONI, J. Jain philosophy. Delhi: Motilal Banarsidass Publishers, 2007
- MATILAL, B. K. A History of Indian Literature - Nyāya-Vaiśeṣika, Wiesbaden: Otto Harrassowitz Verlag, 1977
- _____. Perception: An essay on classical Indian theories of knowledge. Oxford: Clarendon Press, 1992
- POTTER, K. H. Indian metaphysics and epistemology. Delhi: Motilal Banarsidass Publishers, 2011
- PRADHAN, R. C. “*Language and Mind in the Upanisads*”, *Language and Mind: The Classical Indian Perspective*, ed. K. S. Prasad, Hyderabad Studies in Philosophy no. 5, New Delhi: Decent Books, 2008, pp.51-68
- RAO, A.V. “Mind” in Indian philosophy. Indian journal of psychiatry, vol. 44(4) pp. 315 – 325. 2002.
- SARMA, D. S. The Upaniṣads – An anthology. Bombay: Bharatiya Vidya Bhavan, 1975

SIDDHANTALANKAR, S. Heritage of Vedic Culture. Bombay: D. B. Taraporewata, 1969.

SINHA, J. Indian psychology (3 volumes). Delhi: Motilal Banarsidass Publishers, 2008.

VERPOORTEN, Jean-Marie. A history of Indian literature - Mīmāṃsā literature. Wiesbaden: Harrassowitz Verlag, 1987

Filosofia da mente no pensamento pré-platônico: o caso de Parmênides

Philosophy of mind in early Greek thought: the case of Parmenides

Nicola Stefano Galgano⁷⁹

RESUMO: Se compulsarmos um manual introdutório à Filosofia da Mente em busca de Parmênides, encontraremos muito pouco ou até mesmo nada. E acredito que o problema seja a associação de nossa noção de mente principalmente com a palavra grega *psykhe*, uma palavra que não consta em Parmênides. Todavia, Parmênides foi um observador sutil da mente. No proêmio de seu poema ele apresenta o assunto de sua pesquisa: conhecer tanto as afirmações verdadeiras quanto aquelas opinativas nas quais não há convicção verdadeira. Isso implica saber distinguir umas das outras e, de fato, nos versos sucessivos Parmênides explica um método – baseado num preceito que constitui o primeiro enunciado do princípio de não contradição – que permite essa distinção. A suportar essa concepção há uma teoria da mente. O artigo argumenta que é possível encontrar uma teoria da mente em Parmênides, se atentarmos à sua noção peculiar de *não ser*. De fato, essa noção é uma autêntica criação da mente, já que – pela definição do próprio Parmênides – ela não se refere a nada real. Dada essa noção, a mente que investiga estaria num caminho errado se usasse o *não ser* nos argumentos, pois o *não ser* é impensável e indizível. A discussão sofisticada que Parmênides faz das funções da mente alcança o ponto alto quando indica os hábitos culturais como a causa da maneira de pensar comum dos povos.

PALAVRAS-CHAVE: Parmênides, filosofia da mente, história da psicologia, teoria da mente, eleatismo

ABSTRACT: If we read an introductory handbook of Philosophy of Mind in search for Parmenides we find very little or nothing at all. The problem, I believe, is the association of our notion of mind mostly with the Greek word *psykhe*, a word that does not appear in Parmenides. Notwithstanding, Parmenides was a subtle observer of the behavior of mind. In the proemium of his poem, he presents the topic of his research: to know both true statements and opinions in which there is no true conviction. This implies to know how to distinguish one another, and, in fact, in the next verses, Parmenides explains a method – based on a precept that is the very first enunciation of the Principle of non-contradiction – that makes this distinction. On the background of such a conception there is a theory of mind. The paper argues that it is possible to find out a theory of mind in Parmenides if we analyze his peculiar notion of non-being. In fact, this notion is a true creation of mind, since – by definition of Parmenides himself – it does not refer to anything real at all. Given this notion, the inquiring mind would be on a wrong path if uses non-being in arguments, for non-being is unthinkable and unknowable. Parmenides' sophisticated discussion of functions of mind achieves a high point when points out cultural habits as the causes of the common mind-set of peoples.

KEYWORDS: Parmenides, philosophy of mind, history of psychology, theory of mind, eleatism.

⁷⁹ Pesquisador de Pós-Doutorado na USP. Email: nicolagalgano@usp.br.

1. Introdução. Desde Hegel, Parmênides de Eleia, que viveu do final do sec. VI até a metade do V a.C., é amplamente conhecido como o ‘filósofo do ser’. Porém, essa visão hegeliana, que se tornou até mesmo um clichê nos manuais de história da filosofia, está recebendo complementos (às vezes até mesmo críticas) pelos estudos parmenidianos mais recentes, os quais estão revelando muitos aspectos para além daquele lugar comum de ‘filósofo do ser’. Em termos muito gerais, são dois os principais rumos percorridos pela pesquisa recente; o primeiro deles é a requalificação de Parmênides como estudioso da natureza, portanto, como cientista, em oposição à imagem tradicional de um pensador que nega a experiência dos sentidos e a multiplicidade dos entes; e o outro, dedicado mais à parte filosófica, é a reavaliação das interpretações tradicionais procurando eliminar a pátina de platonismo depositada ao longo de cerca de 24 séculos. Com o termo platonismo quero me referir à influência dos platonistas, ou seja, os seguidores de Platão, e não ao próprio Platão. Com efeito, Platão discutiu profundamente e sutilmente a filosofia parmenidiana e extraiu suas próprias conclusões. Estas, porém, não são definitivas e exaustivas, pois muitas outras sugestões filosóficas podem ser encontradas no poema parmenidiano. Todavia, os platonistas e comentadores de Platão nos legaram a interpretação platônica de Parmênides como única possível e acabaram configurando uma imagem que, do ponto de vista histórico não é completa, para dizer o mínimo. Platão, no diálogo *O Sofista*, discute o não ser de Parmênides e, em parte, sua concepção de ser; a seguir, discute a possibilidade e configuração de uma dimensão ideal no conhecimento humano, aquilo que é conhecido, um tanto impropriamente, como o “mundo platônico das ideias”. Todavia, o poema de Parmênides, a parte que chegou até nós, não se restringe a uma discussão sobre ser e não ser, mas se estende a vários outros assuntos. Um deles – que a herança platonista deixou na sombra – é exatamente uma discussão sobre o papel da mente no conhecimento. No poema, os elementos dessa discussão são tantos e tão consistentes que se pode falar sem receio de uma autêntica teoria da mente. Neste breve ensaio, procurarei mostrar esses elementos e também a ideia de mente que Parmênides construiu, uma ideia avançada para seus tempos e, a bem da verdade, ainda moderna e instigante. Antes de entrar propriamente na teoria da mente, esboçarei o contexto em que Parmênides expôs sua concepção.

2. A necessidade de um critério de verdade. Diz Diógenes Laércio (D. L., *Vitae*, 9, 21, 7-8.) que Parmênides teve, entre seus mestres, Xenófanés de Colofão. Este, que foi um poeta que viajara em toda a Hélade, notara que: “*Os mortais acreditam que os deuses são gerados, que como eles se vestem e têm voz e corpo.*” E também: “*Os egípcios dizem que os deuses têm nariz chato e são negros, os trácios, que eles têm olhos verdes e cabelos ruivos.*” (DK 21 B 14 e 16) Afora a questão especificamente teológica, sobre os deuses, e antropológica, sobre os comportamentos dos povos, resta ainda a questão cognitiva e de forma mais geral, a questão filosófica: os mesmos deuses são representados de forma diferente. O argumento implícito de Xenófanés é que os mesmos deuses não podem se apresentar com traços fisionômicos diferentes; ou seja, ele contesta a veracidade das crenças populares não por ter ele mesmo o conhecimento verdadeiro, mas contestando a lógica interna das várias afirmações; ele utiliza implicitamente o princípio de não contradição. Afinal, se pode dizer que, ao menos em relação ao que escapa à sua experiência imediata o homem comete erros e acredita ser verdadeiro o que verdadeiro não é. A lição ficou para Parmênides que, indo além de seu mestre, escreveu um poema para mostrar que é necessário distinguir entre afirmações verdadeiras e afirmações não verdadeiras.

3.O poema de Parmênides: o fragmento 1. Parmênides pertence àquele grupo de autores antigos que escreveram uma só obra. Ele escreveu em versos e, portanto, se trata de um poema, conhecido como *periphyseos*, “*Da natureza*”. O poema foi reconstruído a partir de citações de doxógrafos e comentadores, que foram dispostas numa ordem provável, com cerca de 160 versos formando fragmentos de um todo que não devia conter menos do que 500. Pelas palavras do próprio Parmênides, e também pelos testemunhos dos comentadores que conheciam o poema inteiro, a obra estava dividida em duas partes, precedidas por um proêmio. Essas duas partes são conhecidas como a parte da verdade e a parte da opinião. A segunda parte se refere a teorias científicas da época e não nos interessa aqui. Aqui falarei somente do proêmio e da primeira parte. No proêmio, numerado segundo a edição mais famosa, a de Diels-Kranz (1951), como fragmento 1, Parmênides conta a viagem de um jovem discípulo, o *kouros*, para um lugar onde uma deusa anônima o recebe com estas palavras⁸⁰:

*O jovem, companheiro de aurigas imortais,
tu que assim conduzido chegas à nossa morada,
salve! Pois não foi mal destino que te mandou perlustrar*

⁸⁰ Todas as traduções do poema de Parmênides reportadas aqui são de Cavalcante de Souza (1978).

*esta via (pois ela está fora da senda dos homens),
mas lei divina e justiça; é preciso que de tudo te instruas,
do âmago inabalável da verdade bem redonda,
e de opiniões de mortais, em que não há fé verdadeira. (DK 1.24-30)*

As palavras que nos interessam são as finais. Diz a deusa (versos 1.28-30): “*é preciso que de tudo te instruas, tanto do âmago da verdade bem redonda, quanto das opiniões dos mortais, nas quais não há fé verdadeira*”. Ou seja, a deusa diz que o discípulo tem que aprender tudo (*panta*, todas as coisas), tanto algo relacionado com a verdade, quanto algo relacionado com a opinião. O algo relacionado com a opinião é fácil de traduzir e entender: trata-se das opiniões dos mortais, isto é, da pessoa comum; essas opiniões não são confiáveis, são crenças. Mas a primeira parte, a relacionada com a verdade apresenta palavras misteriosas: “*âmago inabalável da verdade bem redonda*” (v. 1.29, *alēthēies eukykleos atremes ētor*). É uma imagem bem bonita que atçou e atiza a fantasia dos estudiosos (Cfr. Galgano, 2012). A interpretação mais comum é de que há uma verdade, que é bem redonda, ou seja, metaforicamente perfeita, e que ademais possui um âmago que é inabalável, firme, sólido e indestrutível. Forma-se assim uma ideia de verdade com Vê maiúsculo, ou seja, a Verdade ideal e abstrata. A palavra que o tradutor traduziu com ‘âmago’ é *ētor* que quer dizer ‘coração’, ‘peito’, e que o tradutor, acompanhando a versão de todos os demais tradutores, entendeu como metáfora de ‘centro’, portanto, ‘âmago’, ‘núcleo’, a parte mais sólida da verdade bem redonda. Esse é um dos exemplos daquilo que eu chamava camada de platonismo (referido aos seguidores de Platão e não a ele mesmo). Todavia, se traduzirmos literalmente, *atremes ētor* significaria ‘coração que não treme’, o que não é uma boa metáfora, pois um coração que não treme, que não oscila, é um coração morto. Por outro lado, na época de Parmênides não existia a noção de verdade como a entendemos hoje, abstrata e absoluta. Essa noção começará com o idealismo platônico e ficará consolidada somente, provavelmente com Agostinho. Na época de Parmênides ‘verdade’, *alētheia*, era uma expressão ligada ao relato presenciado por alguma testemunha ocular, então significa ‘contar tudo’, ‘não esquecer de contar nada’ (de *a-lēthē*, não esquecimento) (Cfr. Levet 1976). Nas inscrições de algumas lâminas áureas funerárias contendo orações órficas, aproximadamente da época e região de Parmênides, no sul da Itália, *alētheia* é a palavra que não deve ser esquecida pelo morto, porque vai lhe permitir passar por certos lugares do além (Bernabé, 2013). É o segredo dos órficos, que não pode ser esquecido; hoje nós diríamos: é uma *password* que não deve ser esquecida, sob pena de perder o acesso desejado. Então, a expressão o ‘coração intrépido da

verdade bem redonda’ é uma expressão que não faz sentido nem como metáfora, pois se trataria de um coração morto, nem como semântica, pois a noção de verdade abstrata e perfeita não existia naquela época. As coisas começam a se esclarecer se atentarmos para os conhecimentos fisiológicos da época. Na época de Parmênides (e antes), a fisiologia conhecida considerava que o órgão do pensamento era o peito. Assim, se a expressão ‘coração que não treme’ não faz sentido, a expressão ‘órgão do pensamento (mente) que não treme’ faz todo sentido: trata-se da sensação mental de decisão, de ausência de dúvida, de firmeza mental. Então podemos traduzir o verso 1.29 dessa forma, ‘a mente que não oscila (que não tem dúvida) originada por um relato (ou conto ou palavra ou discurso) bem redondo’, onde o ‘bem redondo’ (*eukyklos*) é um adjetivo que se aplicava à circularidade protetiva⁸¹. Em outras palavras ‘bem redondo’ significa bem amarrado, bem fechado, bem protegido porque bem conexo. Agora sim, as palavras da deusa fazem sentido:

*é preciso que de tudo te instruas,
da mente inabalável originada de uma fala verdadeira bem conexa,
e de opiniões de mortais, em que não há convicção verdadeira.*

Parmênides está dizendo que o discípulo aprenderá tudo, tanto as verdades quanto as opiniões. Por isso, é necessário um método que possa realizar essa distinção.

4. O fragmento 2. De fato, na sequência, naquilo que é chamado fragmento 2, a deusa diz:

*Pois bem, eu te direi, e tu recebe a palavra que ouviste,
os únicos caminhos de inquérito que são a pensar:
o primeiro, que é e portanto que não é não ser,
de Persuasão é o caminho (pois à verdade acompanha);
o outro, que não é e portanto que é preciso não ser,
este então, eu te digo, é atalho de todo incrível;
pois nem conhecerias o que não é (pois não é exequível),
nem o dirias...*

Estas palavras são importantíssimas na história do pensamento ocidental. A deusa, que é uma deusa anônima, diz ao discípulo: quando se pensa o pensar investigativo, o de inquérito – portanto, não o pensar comum, mas o pensar culto de pesquisa – os pensamentos podem seguir dois caminhos; o primeiro caminho é de persuasão que segue junto com a verdade, já o segundo

⁸¹ De fato, encontramos para *kyklos* os seguintes significados: um círculo de caçadores que circundava sem escapatória um leão (*Od.*4.792), construção militar (*Hdt.*1.98.;*Th.* 2.13.7.4-5), uma paliçada de defesa em volta do acampamento militar (*X. An.* 6.4.20.2), os aros metálicos que fechavam e prendiam todas as peças do escudo (*Il.* 11.33).

caminho é um caminho inviável. A primeira coisa a notar é que Parmênides está falando do pensamento e de como o pensamento se comporta, ou seja, do pensar. Não se tem notícia de uma preocupação com o pensar e nem com nada similar antes dele. Depois, ele fala do pensar epistêmico, aquele que busca a verdade das coisas, que redonda no discurso bem conexo. Depois, e aqui estamos já numa dimensão de observação muito profunda, Parmênides fala de ‘persuasão que acompanha a verdade’, ou seja, a persuasão verdadeira em oposição às opiniões dos mortais onde não há convicção verdadeira (1.30). Começamos de esse último ponto. Como vimos, Xenófanés já apontara a incongruência dos saberes tradicionais. Parmênides, instruído nesses conhecimentos, põe então o problema: o homem pode se convencer tanto de teorias verdadeiras quanto de falsas⁸². Por exemplo, no caso citado por Xenófanés, relatos de povos diferentes dizem que os deuses têm cabelo preto ou têm cabelo ruivo, por conseguinte, desses relatos, ou um é falso ou ambos os são. No entanto, é um fato que os povos estão persuadidos de suas concepções. Isto significa que o homem pode se persuadir do fato verdadeiro mas também do fato não verdadeiro, logo, há uma persuasão verdadeira e outra não verdadeira. Ou seja, entre a verdade dos fatos e o pensamento que o homem tem dos fatos há um intermediário. Esse intermediário é uma instância onde se dá a persuasão; a essa instância nós chamamos ‘mente’ e Parmênides ‘peito’: *ētor*, no fr. 1, ou o *stēthos*, como veremos, no fr. 6.

A mente pode ser convencida de convicção não verdadeira, como no exemplo trazido por Xenófanés, ou de persuasão verdadeira, como depois Parmênides irá exemplificar no fr. 8. Mas antes de apresentar as conclusões, Parmênides expõe o procedimento que deve ser seguido pelo pensar investigativo. Esse pensar, pode seguir por dois caminhos na mente: o primeiro é o caminho do “*pensar que é e que não é não ser*”. O que caracteriza esse caminho de persuasão verdadeira? Parmênides não dá argumentos, só o descreve. Já para o segundo caminho, “*pensar que não é e portanto que é preciso não ser*” ele diz coisas importantes: é um caminho que não pode ser seguido até o fim, porque não é possível conhecer *o que não é*. Em relação ao conhecimento do *não ser*, excluo aqui as leituras banais e as leituras predicamentais, que alguns fazem, do tipo que se algo não está aqui presente não posso conhecê-lo. Não, Parmênides não está falando disso. Parmênides está falando do *não ser* absoluto e, pelas palavras referidas aos

⁸² Os deuses representavam também as explicações para os fenômenos naturais, explicações que hoje chamamos teorias.

caminhos do pensamento, podemos entender que ele tentou pensar o *não ser* e chegou à conclusão de que isto é impossível⁸³.

Sendo assim, considerando que o *não ser* absoluto não pode ser pensado, o pensamento pode somente pensar o *ser*, ou como ele diz “*pensar que é*” e além disso, esse ‘*é*’ é algo que não pode ser negado absolutamente, portanto esse ‘*é*’ “*não pode não ser*”. Façamos algumas considerações. A primeira é que Parmênides é o primeiro a falar do *não ser* no ocidente e, em sentido filosófico, ou seja, de forma racional, é o primeiro no mundo, pelo que se sabe até agora. De fato a noção de *não ser* se encontra nos *Vedas* mais antigos (*Rig Veda*, 10.129), mas é tratada de forma religiosa e sem nenhuma argumentação. A segunda consideração é que Parmênides descobre que ‘*não ser*’ é um estranho objeto da mente, pois embora a expressão ‘*não ser*’ pareça se referir a algo, ou seja à negação absoluta, essa negação não se realiza, é impossível; por conseguinte, a expressão não cumpre o que promete. Nós, hoje, podemos acrescentar que a noção de ‘*não ser*’ é a prova de que a mente existe e tem função ativa e criativa, pois, ‘*não ser*’ não é reflexo de nenhum objeto real, sensível ou inteligível, pois, por definição ‘*não ser*’ não corresponde a nenhum objeto real, seja ele sensível ou inteligível. Então, diz Parmênides, quando usamos a noção de ‘*não ser*’ no discurso, nós achamos que estamos dizendo algo, mas na verdade não estamos dizendo nada que tenha sentido real. Isto significa que esta noção, por ser sem sentido, acaba tornando vão o discurso que quer ter sentido de persuasão verdadeira, isto é o discurso epistêmico. Por conseguinte, o *não ser* não deveria ser usado nesse tipo de discurso. Em resumo, o *não ser* não pode ser pensado e também não pode ser expresso.

Atente-se para as palavras finais do fr. 2 “*nem o dirias*” (*oute phrasais*). O *não ser* não deve ser dito, porque afinal não tem sentido e se não tem sentido é algo que quando é dito, não diz; ora, o dito que não diz é contra o dizer, é contraditório. Então, nessas palavras está escondida sob o véu poético, a descoberta da contradição. Para Parmênides, a contradição é um fato da mente, pois os homens confundem ser e não ser e, por isso não conseguem expor discursos confiáveis a respeito do mundo.

⁸³ Tentar pensar o não-ser absoluto é um exercício praticado por algumas escolas de filosofia. O primeiro a realizar esse exercício, de quem se tem notícia, certamente foi Parmênides, pois nesse fr. 2 ele fala do caminho do pensar ‘*não é*’ (*oukesti*), o *não ser* (*mēeinai*) e ‘*o que não é*’ (*tomē eon*). É conhecido também como a meditação do não ser. O leitor pode fazer esse exercício por ele mesmo: tente pensar o *não ser* absoluto, ou seja, a não existência de absolutamente todo ente; a seguir, descreva suas reflexões.

5. O fragmento 6. Vejamos este último ponto no fr. 6:

*Necessário é o dizer e pensar que (o) ente é; pois é ser,
e nada não é; isto eu te mando considerar.
Pois primeiro desta via de inquérito eu te <afasto>⁸⁴,
mas depois daquela outra, em que mortais que nada sabem
erram, duplas cabeças, pois o imediato em seus
peitos dirige errante pensamento; e são levados
como surdos e cegos, perplexas, indecisas massas,
para os quais ser e não ser é reputado o mesmo
e não o mesmo, e de tudo é reversível o caminho.*

Nos dois primeiros versos temos enunciado o princípio de não contradição, em seu enunciado lógico. Notem que aqui também Parmênides diz: “*Necessário é dizer e pensar que o ser é e o nada não é*”. Ou seja, esse princípio pertence à esfera do pensar. Mas deixemos o princípio de não contradição de lado. Vejamos os versos seguintes. A deusa diz para o *kouros*, o discípulo, que ele deve se manter <afastado> da via seguida pelos mortais (cujas opiniões não são de convicção verdadeira, fr. 1.30). Como eu disse antes, os mortais são as pessoas comuns. Eles “*nada sabem*” (isto é, não são instruídos) e o “*imediato*” em seus peitos dirige errante pensamento. Aqui tenho que fazer uma correção na tradução, pois a palavra que Cavalcante de Souza traduziu como ‘imediato’, em grego é *amēkhaniē* (Cfr. Galgano, 2016a). Tal palavra é uma negação, com o alfa privativo, de *mēkhanē*, que significa ‘recurso’, ‘ardil’, ‘mecanismo que resolve um problema’, aliás a nossa palavra portuguesa ‘mecanismo’ deriva do grego *mēkhanē*. Então *amēkhaniē* significa ‘falta de recursos’, falta de algum mecanismo. Para Parmênides, os mortais têm a falta de um certo mecanismo, e esta falta conduz em seus peitos (*stēthesin*, que agora sabemos que significa ‘em suas mentes’) um pensamento errante, ou seja, que vaga sem direção e sem rumo certo. Eles, continua Parmênides, são levados como ‘surdos’ e ‘cegos’, ‘atordoados’ e ‘indecisos’. Aqui é preciso entender esses significados; ‘como surdos e cegos’ (*kōphoiomōstyphloi*) significa que eles se comportam como se o fossem, ou seja como se os dados que vêm dos órgãos dos sentidos não estivessem sendo computados. Os mortais não são surdos e cegos, mas é como se o fossem, porque suas mentes não conseguem decifrar os dados que vem da sensibilidade, pois o pensamento em suas mentes, vagando sem direção, não os

⁸⁴ Nesse ponto, todos os manuscritos têm uma lacuna, portanto, qualquer palavra que a preencha é uma conjectura. A conjectura reportada aqui é de Diels, o famoso editor dos pré-socráticos. Todavia, a escolha de “afasto” como conjectura revelou-se infeliz porque acabou gerando uma longuíssima discussão filológica sobre a existência ou não de um terceiro caminho além daqueles que Parmênides descreve no fr. 2. Considero essa conjectura imprópria, mas, por razões didáticas, a mantenho, pois não interfere diretamente na discussão do assunto do artigo. Para maiores informações sobre a conjectura de Diels e suas consequências, veja-se Cordero, 2005.

coordena, porque lhe falta um certo recurso, um certo mecanismo. O resultado é que os mortais são confusos e indecisos e, para eles, ser e não ser são o mesmo e não o mesmo. Ou seja, não distinguem ser de não ser, acham que ‘o *ser* é’ e também que ‘o *não ser* é’. Nós diríamos com nossa linguagem que eles são contraditórios; Parmênides diz, na linguagem dele, que eles consideram todo caminho reversível (*palintropos*), que significa exatamente – é uma expressão homérica (*Il.* 9.56) – contradizer o discurso de alguém.

6. O fragmento 7. No fragmento 7, ainda há referências ao comportamento mental dos mortais:

*Não, impossível que isto prevaleça, ser (o) não ente.
Tu porém desta via de inquérito afasta o pensamento;
nem o hábito multiexperiente por esta via te force,
exercer sem visão um olho, e ressoante um ouvido, e a língua,
a julgar pelo discurso controversa tese
por mim exposta.*

De novo encontramos o imperativo da deusa (Cfr. Galgano, 2015b): o discípulo tem que se afastar da via onde se pensa que os não entes são. Aqui, no primeiro verso (7.1), temos a exposição cosmológica do princípio de não contradição e no segundo (7.2) a deusa pede que o discípulo se afaste da via onde se considera que os não entes sejam. Mais uma vez, vamos deixar o princípio de não contradição de lado e nos dedicar aos versos seguintes. Em 7.3, a deusa diz: “*o hábito das repetidas experiências não te force nesta via*”: que via? Aquela onde se exerce uma visão que não vê, e onde se exerce um ouvido e língua confusos. Veja-se que, nessa via, o que força a usar indevidamente os sentidos é o hábito que resulta de experiências repetidas: Parmênides percebeu que o modo de pensar depende de um hábito cultural adquirido. Ao se repetir muitas vezes um contrassenso intelectual se acaba acreditando nele e se configura uma maneira de pensar que vê, mas não enxerga, algo que nós poderíamos chamar hoje de dogmatismo. Ora, o discípulo tem que se manter afastado deste hábito, pois ele força nessa via. Vê-se claramente a consciência antropológica de Parmênides, que afinal percebe que a maneira de pensar é uma herança cultural (cujos indícios já se encontravam em Xenófanés), resultado de um certo tipo de educação; portanto, percebe que é necessário levar o discípulo por outro caminho, por outra educação.

7. O fragmento 8. No fr. 8 Parmênides apresenta a sua concepção de ser (*eon*), articulada numa complexa sequência de demonstrações⁸⁵. O ser, para ele, não é evidente, mas há sinais (*semata*) que permitem identificá-lo; esses sinais são como que características e não podem ser de outro jeito, pois a demonstração rigorosa (que Parmênides chama *correntes de Dike*, a deusa da justiça) amarra as partes do argumento, assim como as leis do cosmo amarram as partes do mundo num contínuo homogêneo. Essas amarras são do mundo e deveriam ser também da mente que reflete sobre o mundo. Por isso, Parmênides diz que o pensamento epistêmico a respeito do mundo deve estar submetido às mesmas leis do cosmo. Vejamos, por exemplo, a demonstração do primeiro e mais importante dos *semata* demonstrados no fr. 8, a impossibilidade da geração e da corrupção. Eis os versos:

*pois que geração procurarias dele?
Por onde, donde crescido? Nem de não ente permitirei
que digas e pense; pois não dizível nem pensável
é que não é;*

Parmênides está falando do *eon*, o ser, e afirmara anteriormente que ele é ingênito e também imperecível (8.3), onde ingênito e imperecível são dois dos *semata* do ser. A seguir, então argumenta e mostra (mas, se aceitarmos a lógica antiga, poderia se dizer que demonstra) que o *eon* é ingênito. A forma argumentativa é bastante diferente da nossa, mas é possível identificar o seguinte; 1) a tese: o *eon* é ingênito; 2) o problema: o *eon* é gerado? 3) hipótese: se fosse gerado seria gerado a partir do *não ser*; 4) o método: o *não ser* é indizível e impensável, logo não deve ser usado na argumentação; 5) conclusão: confirma-se a tese, o ser é ingênito, *quod erat demonstrandum* (*QED*). O que nos interessa aqui⁸⁶ é o preceito da deusa, que diz ao discípulo: “*nem de não ente permitirei que digas e pense*”. Lembrando que é uma deusa falando, a forma imperativa da expressão tem o sentido da expressão de uma lei universal e não particular, ou seja, não é só o discípulo que não tem a permissão, mas é universalmente não permitido dizer e pensar o não ser. Todavia, esse impedimento é fornecido externamente (uma lei dada por uma deusa), porque espontaneamente, a mente comum acredita que o *não ser* seja pensável e, portanto, também dizível. Por isso, já que a mente espontaneamente tende a acreditar que o *não*

⁸⁵ A estrutura argumentativa do fr. 8, assim como a do restante do poema, pode ser encontrada em artigo de Rossetti (2010), onde o autor expõe também as expressões, já usadas por Parmênides, correspondentes à clássica expressão conclusiva dos geômetras: *quod erat demonstrandum*.

⁸⁶ Para uma discussão dessa passagem cfr. Galgano, 2015a e Galgano, 2017, p. 145-161.

seré pensável, é necessário um recurso externo para alertar que ‘*não ser*’ é uma virtualidade da mente, ou seja, uma sua criação que de fato, e por definição, na realidade não existe.

Vejamos agora uma outra passagem do fr. 8 e precisamente aquela constituída dos versos 34 a 38:

*O mesmo é pensar e em vista de que é pensamento.
Pois não sem o que é, no qual é revelado em palavra,
acharás o pensar; pois nem era ou é ou será
outro fora do que é, pois Moira o encadeou
a ser inteiro e imóvel;*

Estes versos apresentam muitos problemas, tanto filológicos quanto filosóficos. Tentaremos escapar da necessidade de atravessar este difícil território concentrando nosso foco sobre o primeiro verso. O sentido geral da inteira passagem, embora com algumas discordâncias entre os estudiosos, pode ser considerado este: pensar (ou conhecer) e o objeto do pensar (ou do conhecer) são o mesmo, porque sem o ser é impossível pensar ou conhecer algo, isto é, é impossível pensar ou conhecer o que não é, porque, sem o ser não há nada.

Creio que o acento sobre o aspecto psicológico do pensamento parmenidiano pode ajudar a esclarecer estas palavras. Antes de tudo, a expressão ‘o mesmo’ (*tauton*) não pode ser entendida como ‘idêntico’. Trata-se do ‘mesmo’ na semântica arcaica que significa pertencer ao mesmo ramo, genealogia, povo, tribo. Por exemplo, todas as pessoas de Roma são o mesmo, isto é, são Romanos. No caso de Parmênides, ele está tentando descrever dois fatos que parecem ser diferentes mas, ele diz, eles pertencem ao mesmo fato, ao mesmo acontecimento. Numa linguagem muito moderna, e levando em conta que Parmênides está falando do mundo físico (*eon*), podemos traduzir com: eles pertencem ao mesmo evento. O ‘evento’ naturalmente é o *eon*, para o qual o fragmento 8 é dedicado. Como em outras partes da filosofia parmenidiana, a descrição do mundo físico inclui a lei que governa aquele fato físico. Assim, ‘o mesmo’ significa que os fatos que ele está descrevendo são submetidos à mesma lei de comportamento físico.

Pensar(que aqui significa *operar* com a parte cognitiva da mente) e aquilo pelo qual se dá o pensar (a *operação cognitiva*) pertencem ao mesmo evento. De fato, se algo é afirmado (“*revelado em palavra*”) do mundo, este algo possui realidade, porque a afirmação aqui não é uma afirmação qualquer. Para entendermos isso, temos que ter presente que, quando fala sobre afirmação de algo nesse contexto cosmológico, Parmênides quer se referir à afirmação

verdadeira, ao enunciado verdadeiro e ao discurso verdadeiro; de fato, a fala verdadeira segue junto com a persuasão verdadeira. Assim, no discurso verdadeiro, quando ele fala sobre algo, faz afirmações verdadeiras porque 1) o pensamento verdadeiro sobre algo e 2) a *operação cognitiva* operada com o ‘recurso’ (os vários preceitos oferecidos pela deusa) pertencem ao mesmo ‘evento’, isto é à mesma realidade verdadeira, a realidade do *eon*. O ‘recurso’ aqui é mostrado pela imagem de uma outra deusa, a Moira (o Fado), acorrentando todas as coisas num inteiro, de uma forma que, sem passado e sem futuro (que introduziriam o *não ser*, pois as coisas que devêm viriam do *não ser* e iriam para o *não ser*) não há ser que venha e vá e não há outra coisa exceto o ser que já existe. Aqui Parmênides cumpre um grande passo no desenvolvimento do conhecimento humano, porque ele afirma a identidade de função entre o pensar humano verdadeiro e o pensamento verdadeiro a respeito de algo. Ele diz que o sujeito e o objeto pertencem à mesma ordem da *physis*, e é esta pertinência que permite a possibilidade do conhecimento⁸⁷. Esta mesma ordem não é aquela adquirida pelo hábito multiexperiente, mas é a ordem do percurso cognitivo do ‘é’. Atender a esta ordem permite o conhecimento verdadeiro, porque atender aos preceitos da deusa permite a relação entre o pensar (isto é, aquele pensador que pensa com as *operações cognitivas* verdadeiras) e o pensamento do objeto expresso em declarações verdadeiras; o sujeito pensador (o homem sábio que segue os preceitos da Deusa) pode conhecer a verdade do objeto porque o mundo (que inclui sujeito e objeto) é ordenado pelo mesmo preceito que Dike e Moira põem em ação acorrentando o *eon* num todo.

8. O fragmento DK 4. Ainda no fr. 8, convém colocar o aquele trecho que por Diels foi enumerado como fr. 4. Esta mudança na ordem dos fragmentos não é unânime entre os estudiosos, mas muitos acreditam que ele tenha que pertencer ao fr. 8 porque apresenta uma discussão similar àquela ali realizada e com linguagem também similar. Eu aceito essa colocação e passo a falar dos versos sem adentrar a exegese filológica, impossível de ser reportada aqui. Vamos aos versos:

*Mas olha embora ausentes à mente presentes firmemente;
pois não deceparás o que é de aderir ao que é,*

⁸⁷ Vale lembrar que anteriormente a Parmênides, o conhecimento verdadeiro era reservado só aos deuses (Xenófanés, DK 21 B 18 e 24; Alcmeon, DK 24 B 1). Ao estabelecer uma mesma lei entre o pensar e o ente que é pensado, Parmênides estabelece um princípio pelo qual o mundo (que naquela época incluía os deuses, diferentemente do pensamento cristão que põe o deus fora do mundo) e o pensar o mundo obedecem à mesma lei, autorizando epistemologicamente, desta forma, o pensamento humano a alcançar e incluir o mundo inteiro no seu objeto de conhecimento verdadeiro.

*nem dispersado em tudo totalmente pelo
cosmo, nem concentrado...*

A grande problemática da tradução é bem conhecida pelos estudiosos e se reflete na insatisfação geral com os resultados, que são muito diferentes para cada interprete⁸⁸. Nesse caso, nem o sentido geral recebe um acordo dos estudiosos; todavia, eu penso que se o trecho for lido desde o ponto de vista do *não ser* e pelo viés de um comportamento mental, ele não só resulta claro mas até mesmo luminoso, pois Parmênides, assim me parece, está descrevendo, pela primeira vez na história da filosofia ocidental, o fenômeno mental da abstração. Façamos a paráfrase: “Observe com a mente (*noōi*, pelo pensar, pelas *operações cognitivas da mente*) as coisas ausentes como (*omōs*) sendo presentes firmemente, pois [subentendido: vóos, a mente ou o pensar ou, mais especificamente, *a operação cognitiva*] não dividirá (*ou garapotmēxi*) o ente de aderir (*ekhesthai*) ao ente (*toueontos*), nem pelo espalhamento em todas as direções em todo lugar, nem pela união.” Do ponto de uma teoria da mente, a interpretação deste fragmento resulta simples e clara se fizermos algumas considerações. Antes de tudo, nós temos os instrumentos hermenêuticos apresentados acima, os preceitos da deusa anônima e as correntes das deusas Dike e Moira; depois, sabemos que estas correntes precisam estar tanto no pensar (as *operações cognitivas* verdadeiras) quanto no mundo. Mas, se o mundo não nos parece acorrentado é porque nós estamos olhando para ele sem o recurso (*amēkhaniē*), ou seja, como o olham os mortais. A primeira palavra que encontramos é *leusse*, que é a forma imperativa do verbo *leussō*, observar, fitar. A deusa diz ao discípulo: olhe atentamente! Não é um convite, é uma ordem imperativa, como muitas outras ao longo do poema, para que o discípulo obedeça e faça aquilo. O que ele tem que fitar? A deusa diz que ele tem que olhar atentamente as coisas ausentes (ou as coisas afastadas) como firmemente presentes (ou próximas). E a deusa acrescenta que o *kouros* precisa fazer isto com a mente (*noōi*, dativo instrumental). Ora, nós já sabemos que este olhar atento é aquele com o recurso (aquele que atende aos preceitos da deusa), diferente do olhar dos mortais que são forçados a exercer um olho sem visão (fr. 7) e a andar como cegos (fr. 6); o próximo passo consiste em entender o que poderia significar o fato de que coisas ausentes devem ser vistas como presentes.

Ao longo do poema, Parmênides apresenta uma sequência argumentativa específica que segue

⁸⁸ Este fragmento raramente é estudado em si, mas quase sempre em relação à doutrina geral do ser no poema. Uma exposição de algumas interpretações se encontra em Marques, 2007.

um mesmo padrão: primeiro apresenta o resultado do argumento e depois fornece as razões. Portanto, as razões para o imperativo da deusa se encontram nos versos seguintes. No segundo verso não há um sujeito claro para a oração⁸⁹, todavia, por razões filológicas não explicitadas aqui, qualquer um dos dois sujeitos possíveis redundaria no mesmo significado: tomarei então como sujeito implícito o *noos* do verso anterior, para permitir que o tom imperativo da deusa seja mais universal: o *noos* (a mente, aqui em seu operar cognitivo) não separará o ente de aderir ao ente. Isto significa que o pensar epistêmico (as operações cognitivas da mente) verdadeiro não cortará as correntes que amarram todas as coisas num todo. As coisas não podem ser separadas, e o olhar (com recurso) das operações cognitivas verdadeiras não separará o ente do ente, ou também, o ente do todo ao qual pertence. Pela mesma razão, as coisas não podem ser vistas como espalhadas ou concentradas. Até agora tudo é muito claro, mas resta o problema de entender ‘coisas ausentes’ e ‘coisas presentes’ (*apeontae pareonta*). Para poder entender como as coisas ausentes devem estar presentes para o olho da mente com recurso, precisamos nos perguntar: em quais circunstâncias as coisas presentes parecem estar ausentes para a mente comum? Esta me parece uma observação psicológica surpreendente de Parmênides. No nosso comportamento psicológico comum nós observamos as coisas como se fossem isoladas umas das outras. Se eu observo um livro, eu considero este livro como uma coisa isolada do resto do mundo, e por isso eu digo: o livro está sobre a mesa. Considero o objeto isolado ‘livro’ sobre o objeto isolado ‘mesa’. Este isolamento dos objetos é o resultado de um processo normal de nossa cognição que nós chamamos atenção. A atenção é uma maneira dos nossos processos perceptivos e cognitivos pela qual uma coisa, que é o objeto de nossas atenções, é “separada” do “resto do mundo”; este permanece em segundo plano em relação ao objeto de nossa atenção. Assim, pelo comportamento psicológico, eu seleciono ‘objetos’ para a minha atenção, um livro, uma mesa, uma xícara e os considero um de cada vez para as minhas finalidades. Quando esta atenção é organizada numa consideração consciente de um único objeto, eu digo: “um livro”. Assim, a minha ideia de livro é extraída da realidade e o livro real revive em minha mente não como aquele livro concreto, relacionado naquele momento com aquela mesa, naquele ambiente etc.,

⁸⁹O verbo *apotmēxei* pode ser da 2ª pessoa em voz média ou da 3ª pessoa em voz ativa do futuro de *apotmēgō*: se nós aceitarmos como 2ª pessoa, o sujeito é a mesma pessoa para a qual o imperativo é direto, isto é, o *kouros*; se nós aceitarmos como 3ª pessoa, o sujeito é o mesmo da primeira oração, isto é, o *noos*, isto é, o pensar, as operações cognitivas da mente. Levando em conta que as operações cognitivas são referidas ao *kouros*, do nosso ponto de vista, esta ambiguidade gramatical não é importante e o sentido geral permanece o mesmo: o sujeito é um operador cognitivo, isto é, as operações cognitivas do *kouros*.

mas como uma ideia. A ideia de livro é algo extraído da realidade e nós chamamos esta extração de ‘abstração’ (do latim *abtraho*, levo embora). A realidade é um conjunto de coisas sempre juntas, todas elas. Mas a mente comum “vê” as coisas como que separadas umas das outras. Quando penso ‘livro’ eu faço uma ideia como se o livro fosse algo rodeado pelo nada: isto é o que a mente comum faz quando pensa de forma abstrata. Mas o livro concreto está sempre mergulhado no todo da realidade e eu não posso nunca encontrar um livro isolado, rodeado por uma espécie de coisa real que poderia ser ‘o nada’. De fato, diz Parmênides, não há dispersão nem concentração no mundo ordenado (*katakosmon*) e quando nós vemos um mundo com dispersão e concentração nós estamos olhando como surdos e cegos, como as pessoas sem recurso. Mas as operações cognitivas verdadeiras da mente não cortarão as correntes que ligam todas as coisas no *ser*, no todo, no *eon*. Portanto, as operações cognitivas da mente devem olhar, junto com o livro, que a mente comum “vê” como presente, também a mesa e a xícara e as outras coisas, as quais, embora pareçam estar ausentes, de fato, na realidade estão presentes. Esta fatal inteira presença deve ser vista (é necessário que seja vista, ‘*leusse!*’, imperativo) pelo homem com recurso. Se esta interpretação estiver correta, estamos diante da primeira descrição do processo de abstração da cultura ocidental. Isto tornaria Parmênides um observador realmente extraordinário do comportamento da mente, desde o comportamento psicológico, até as implicações cognitivas, epistemológicas e filosóficas.

9. Conclusão. Vamos tomar os elementos que encontramos até aqui e vamos colocá-los numa outra ordem. Parmênides, a partir do ensinamento de seus mestres, entende que é possível ao homem se convencer de noções verdadeiras e não verdadeiras. Mas, se tal se dá, significa que a percepção imediata não é suficiente, pois de fato, pode se ter olhos e não enxergar, pois o enxergar e o ouvir acontecem não no olho e no ouvido, mas no peito (*ētor*, *stēthos*), isto é, na mente. Então, há dois tipos de homens, o homem que sabe e os mortais que nada sabem, estes últimos tem uma *amēkhaniē*, uma falta de recursos em seus peitos (nas suas mentes) pela qual os pensamentos, confundindo ser e não ser, acabam sendo contraditórios. Já o homem que sabe tem essa *mēkhanē* e por essa *mēkhanē* o discurso se mantém bem conexo, de uma conexão tal que mantém o pensamento firme e sólido na mente, há uma adesão do pensamento à mente e a mente não vacila. A deusa conhece esse mecanismo e se dispõe a ensiná-lo; assim o discípulo aprenderá tanto a mente firme originada do argumento bem conexo quanto as opiniões dos mortais, originadas por um hábito cultural, repetido muitas vezes, que torna embaçada e visão e ressoante

o ouvido. Então a deusa ensina esse mecanismo, pois ele não vem naturalmente, é aprendido. Ela diz: há duas maneiras de pensar o pensar de inquérito, por uma via, *pensar que é e que não é não ser*, este é o caminho da persuasão verdadeira que precisa ser seguido. Depois há um outro caminho, de que é necessário que se saiba de sua existência, pois, apesar de não levar a lugar nenhum, se não se identifica como caminho sem saída, se acaba por entrar nele e se acaba confundindo *ser* e *não ser*. Portanto, é necessário que se saiba que o *não ser* é o *não ser*, uma expressão fruto do engano da mente, uma virtualidade a ser evitada. O *não ser* não pode ser pensado e não pode ser dito (Cfr. Galgano, 2016b), ele é contraditório. Por fim, quando os preceitos da deusa são aplicados, se consegue “ver” o *eon*, a realidade das coisas, na qual elas não tem geração e nem corrupção, pois se houvesse o devir, os entes supostamente viriam do *não ser*; mas não há tal coisa, o *não ser*, pois o *não ser* é apenas de uma virtualidade da mente e portanto algo que não tem sentido na realidade. O pensar epistêmico verdadeiro apresenta uma correspondência com a realidade, porque o pensar verdadeiro e a realidade são o mesmo, ou seja, pertencem ao mesmo evento, o *eon*. O que é pensado sem o recurso é um algo que pertence à esfera mental, mas não à realidade, são invenções dos mortais que forjam explicações incorretas (como os etíopes e os trácios dos fragmentos de Xenófanés). O mesmo acontece com o fenômeno comum das ideias produzidas pela mente; as ideias se apresentam na mente como que avulsas e isoladas de sua realidade concreta. O processo mental da abstração nos faz acreditar que as coisas sejam isoladas umas das outras, mas, na realidade, tudo está amarrado e não há carência de ser; na realidade, nenhum ente é separado de outro e o olhar com recurso deve prestar atenção e não separar na mente o que não é separado na realidade. Se tal separação aparece é porque a mente foi além do limite da realidade. Então, a mente epistêmica tem um limite (tem que ter), assim como o *eon* também tem; ambos são limitados e para lá desse limite se acredita pensar o impensável, e se diz o indizível, o que não deveria ser dito, o contraditório.

Eu penso que estamos diante de uma teoria completa da mente, com a clara descrição de seu papel de intermediário entre a realidade do mundo e a realidade do pensamento, de sua função, de seu comportamento, de sua formação pelo hábito cultural, de suas operações cognitivas impróprias e da maneira de superá-las através do princípio de não contradição, um mecanismo extramental para compensar a falta natural de recurso. É o primeiro exemplo de teoria da mente na história do pensamento ocidental, e ademais, de altíssimo valor filosófico, valor que aqui consegui apenas apontar mas não descrever, nem sequer parcialmente. Acrescento

ainda que estou firmemente convencido do valor histórico da proposta parmenidiana no campo da filosofia da mente e, para tanto, gostaria de sugerir que Parmênides seja declarado patrono desse campo de estudo, apenas uma maneira de homenageá-lo e de manter viva a memória dos grandes pensadores que nos antecederam e forjaram os pressupostos de nossa cultura.

Bibliografia

- Bernabé, A. (2013) Filosofia e mistérios: leitura do proêmio de Parmênides. In *Archai*, n. 10, jan-jul, p. 37-58. Brasília.
- Cavalcante de Souza, J. (dir.) (1978) *Os pré-socráticos*. Ed. Abril Cultural, São Paulo.
- Cordero, N.-L. (2005), *Siendo, se es*. Ed. Editorial Biblos, Buenos Aires.
- Diels, H. e Kranz, W. (1951) *Die Fragmente der Vorsokratiker*. Reimpressão da 6a edição, 1989. Ed. Weidmann, Zurich.
- Galgano, N. S. (2012) DK 28 1.29 – A verdade tem um coração intrépido? In AAVV *Una mirada actual a la filosofía griega*. Ponências del II Congreso Internacional de Filosofía Griega, Ed. Ediciones de la SIFG, Madrid-Mallorca.
- Galgano, N. S. (2015a) Depois ou antes: Parmênides e o tempo em DK 8. 9-10. In *Anais de Filosofia Clássica*, vol. 9 nº 17, UFRJ.
- Galgano, N. S. (2015b) Estudo do imperativo no Poema de Parmênides. In: *Filosofia grega e helenística*. 1 ed. São Paulo: ANPOF, 2015, p. 21-30.
- Galgano, N. S. (2016a) Amēkhaniē em Parmênides DK B 6.5. *Journal of Ancient Philosophy*, v.10, p.1-12. (USP).
- Galgano, N. S. (2016b) Os limites da palavra: Parmênides e o indizível. *Ética e Filosofia Política*, v. II, p.4 – 24, UFJF.
- Galgano, N. S. (2017) *I precettidelladea. Non essere e contraddizione in Parmenide di Elea*. Ed. DiogeneMultimedia, Bologna.
- Levet, J. (1976) *Le vrai et Le faux dans La pensée greque archaïque – Étude de vocabulaire. Tome 1*. Ed. LesBellesLettres, Paris.
- Marques, M. P. (2007) O frag. 4 de Parmênides. In *Anais de filosofia clássica*, vol. 1, n. 2, Rio de Janeiro.

A mente e a evolução da ficção: estado atual do problema

The mind and the evolution of fiction: current state of the problem

Pedro Dolabela Chagas⁹⁰

RESUMO: Discussão das funções que fomentaram a evolução da ficção como prática cultural, observando como o seu processamento cognitivo permitiu que ela se estabilizasse como hábito individual. Argumenta-se que esses dois pontos se entrelaçam: a sinergia entre funções de grande apelo coletivo e um processamento cognitivo acessível teria favorecido a popularidade e o sucesso histórico da ficção. Quanto ao primeiro ponto, sugere-se que o seu surgimento e estabilização como prática social foi motivado pela sua disposição como *affordance* cultural de função homeostática, nos níveis individual e coletivo. Quanto ao segundo, discute-se a sugestão da psicologia experimental de que a leitura do texto ficcional não demanda quaisquer capacidades cognitivas excepcionais: a ficção se adapta às nossas predisposições mentais, e não o contrário; a sua leitura não pressupõe capacidades mentais desenvolvidas especialmente para ela. Defende-se que esses dois argumentos convergem na explicação das origens e do sucesso evolutivo e histórico da ficção, aqui definida não apenas como um atributo de certas produções humanas ou um conjunto de produções dotadas de características comuns, mas como uma prática cultural tradicionalizada.

PALAVRAS-CHAVE: teoria da ficção; cognição; evolução cultural; teoria das *affordances*; homeostase.

ABSTRACT: Discussion of the functions that fostered the evolution of fiction as a cultural practice, observing how its cognitive processing enabled its stabilization as an individual habit. It is argued that these two points are intertwined: the synergy between functions of great collective appeal and an accessible cognitive processing could have favored fiction's popularity and historical success. About the first point, it is suggested that the emergence and stabilization of fiction as a social practice was motivated by its disposition as a cultural affordance of homeostatic function, at the individual and collective levels. As to the second, we discuss experimental psychology's suggestion that the reading of fictional texts does not demand any special cognitive capacities: fiction adapts to our mental predispositions, not the other way around; its reading does not presuppose any mental capacities specially developed for it. It is argued that these two arguments converge in the explanation of the origins and of the evolutionary and historical success of fiction, defined not only as an attribute of certain human productions or as a set of objects endowed with common characteristics, but as a kind of traditionalized cultural practice.

KEYWORDS: theory of fiction; cognition; cultural evolution; affordance theory; homeostasis.

⁹⁰ Professor Adjunto de Teoria Literária / UFPR. Pesquisador do CNPq. Email: dolabelachagas@gmail.com.

Em livro recente, Felix Martin propõe que o dinheiro não é “uma coisa, e sim uma tecnologia social: um conjunto de ideias e práticas que põe ordem naquilo que produzimos e consumimos e na forma como vivemos em sociedade” (MARTIN, 2016, p. 46). Em relação ao dinheiro em si, e não aos objetos e símbolos que o representam, não haveria “nada físico para observar”, noção que evoca a proposição de Searle pela qual “um tipo de fenômeno só é dinheiro se pensarmos nele como dinheiro[, e] o que se aplica ao dinheiro aplica-se à realidade social e institucional em geral[:] linguagem, propriedade, casamento, governo, universidades, [...] são todos, em parte – mas não inteiramente –, assim descritos porque os consideramos dessa maneira” (SEARLE, 2000, p. 63). “Conjunto de ideias e práticas”, a fruição da ficção pressupõe a própria noção de ficção no “pano de fundo” cognitivo do leitor. Se ela “põe ordem na forma como vivemos em sociedade”, é porque os seus mundos imaginários conferem ordenação à instabilidade e ao caos informacional da vida social. E se a sua prática subsiste como hábito e crença, tal como no caso do dinheiro isso pressupõe instituições (teatros, cinemas, editoras, museus...) que lhe deem subsistência: sua “função de status” é o que sustenta a sua presença social naturalizada.

Uma obra só é ficcional, pois, quando assim a compreendemos. Mas tal como não é preciso “compreender” o dinheiro para utilizá-lo, a circulação da ficção parece igualmente naturalizada. Assediada pela crítica moral e epistemológica, ela demoraria a se afirmar como prática, mas a familiaridade da nossa relação com objetos ficcionais demonstra a sua integração à cultura – entendida como o conjunto de valores, crenças, normas, ideias, hábitos e práticas transmitidas entre as gerações, sob linhas de continuidade e diferenciação. A identificação dessa confluência de hábito e crença na ontologia da ficção desloca a sua teorização da tentativa de identificar a inscrição da ficcionalidade na materialidade dos objetos (ela não seria redutível às propriedades materiais dos objetos, em sua heterogeneidade radical), para a investigação das condições que permitem que eles sejam processados cognitivamente e circulem socialmente de maneira naturalizada: veremos que isso pressupõe a convergência de uma visada evolutiva, que discuta porque a ficção surgiu e se rotinizou como prática, e uma visada psicológica, que explique como ela pode ser processada com tanta facilidade.

Com isso em mente, se o nosso título fala do “estado atual do problema”, pelo menos na interface entre a teoria literária e a filosofia da mente não é difícil defini-lo. Tomando-se a literatura como paradigma-objeto, trata-se de substituir a noção pela qual a ficcionalidade seria uma propriedade que, inscrita na materialidade da obra, a distinguiria ontologicamente da não-

ficção, pelo tratamento da ficção como hábito individual e prática cultural. Muitas dificuldades relativas à descrição e conceitualização do fenômeno se devem à propensão a confundir o processo e o produto, tomando-se como paradigma-objeto produções que emergiram em ambientes onde a ficção já estava rotinizada – para então experimentar o pasmo diante de algo que não tem função aparente mas é muito desejado, cuja linguagem soa peculiar mas é processada sem empecilhos, que não fala daquilo que é mas é compreendido como se o fizesse... É a dificuldade de explicar o fenômeno já amadurecido que motiva a declaração da sua radical singularidade, noção que será desacreditada neste artigo – em particular, pela descrição do seu processamento cognitivo e pela condição relacional (e *fuzzy*) da sua percepção material e funcional.

Aqui a ficção não será pensada primordialmente como “propriedade” ou “atributo” (associada ao adjetivo “ficcional”), nem como conjunto delimitado de objetos (associada ao substantivo “ficção”). Essa escolha é uma questão de ênfase: não há como pensar a ficção como prática cultural sem remeter àquelas duas acepções; essa terceira via é, em última análise, um procedimento heurístico para levar a teoria a um nível superior de generalidade e abrangência. Trata-se de destacar que, mesmo que a sua prática varie muito – a depender do lugar, da época, da mídia, do grau de institucionalização... –, a ficção não é apenas o livro, mas o deitar-se no sofá para a leitura; é o filme, mas é também sair de casa, pagar o ingresso e entrar no cinema – ela implica, em suma, a abertura na rotina de espaço e tempo para a sua fruição. Práticas tradicionalizadas, o livro, o filme, a revista em quadrinho detêm hoje um lugar que prescinde de justificção; se alguém vai ao cinema, ninguém pergunta as razões dessa decisão. Como chegamos a esse ponto?

Eis, afinal, o tema deste artigo: discutir as funções sociais que motivaram a evolução da ficção como prática cultural, e o tipo de apelo às nossas capacidades mentais que permitiu que ela se estabilizasse como hábito individual. Os dois pontos se entrelaçam na minha proposição, pela qual a sinergia entre funções de apelo grande coletivo e um processamento cognitivo acessível teria favorecido o seu sucesso histórico. A psicologia experimental tem sugerido que as suas produções se adaptam às nossas predisposições mentais, sem demandar capacidades desenvolvidas especialmente para elas. Mesmo que as convenções ou tradições comunicacionais dos “gêneros” e “estilos” co-determinem o processamento da informação – ao predispor os leitores a certos padrões de relação com o texto –, e se, nessa medida, o aprendizado cultural

fomenta capacidades específicas de leitura, mesmo esses códigos apelam a capacidades mentais inatas e/ou desenvolvidas espontaneamente em nossa experiência do mundo.

Quanto às suas funções evolutivas, é certo que o romance grego do início da era cristã, registro material mais antigo da ficção no Ocidente, revela descontinuidades importantes em relação a códigos e funções comunicacionais previamente institucionalizados. Eram narrativas sobre personagens que o leitor só passava a conhecer no contato com o texto, desprovidas de funções autoevidentes, sem reputação no campo letrado, e desejadas pelo prazer da leitura – condições ainda hoje associadas à ficção, e às quais eu remeterei ao utilizar o termo. Mas a psicologia evolutiva sugere que essa autonomização em relação ao *mythos* e à verdade histórica, com o desenvolvimento correlato de códigos próprios, não eliminava a exploração de funções ancestrais de ficções que não deixaram registro material; neste artigo, tais funções serão aglutinadas na noção de “homeostase” (individual e coletiva). Se ficções não trouxessem benefícios de algum tipo, a sua prática não teria se estabilizado; elas só existem porque há pessoas que as produzem e que as desejam e, se a prática se tradicionalizou, é porque o tempo e o trabalho investidos nessa produção e recepção foram – em sociedades suficientemente complexas – reiteradamente recompensados, o que dificilmente teria ocorrido se a sua função se limitasse ao prazer da fruição.

Esse processo de popularização trouxe uma consequência: para o leitor atual, praticar a ficção é procurar textos que, de antemão, ele sabe serem ficcionais. Teorias que se preocupam com as condições da identificação da ficcionalidade nas características imanescentes dos textos costumam insistir num falso problema, pois em geral a sua condição ficcional é informada ao leitor de antemão. Decerto as suas convenções comunicacionais também a indicam, mas há séculos a circulação social das obras pressupõe o anúncio da sua própria ficcionalidade – como toda prática cultural, a ficção é dotada de espaço próprio em meio a produções circunvizinhas, a ponto de toda “subversão” dessa fronteira pressupor a fronteira no próprio ato da subversão (como ocorre, hoje, com a chamada “autoficção”). É sabido que a institucionalização desse espaço se consolidou apenas no século XVIII, mas esse tema não concerne à gênese e à evolução inicial da ficção, e sim à sua história. Se ela custou a se autonomizar, para tanto contou a longa série acusações morais e epistemológicas que ela recebeu ao longo dos séculos, e que custariam a ser (ao menos parcialmente) equacionadas.

Bem antes disso, porém, ela já se estabilizara como prática cultural. Como isso aconteceu? Quais foram os passos que levaram, no passado remoto, das cripto-ficções do pensamento hipotético ao mito e à epopeia, e daí à produção de histórias prazerosas dirigidas ao público contemporâneo? Não sabemos exatamente, mas este artigo aposta que a prática não teria se rotinizado se ela não trouxesse benefícios individuais e coletivos, nem se o seu processamento cognitivo fosse difícil a ponto de restringi-la a públicos seletos. Tratarei os benefícios como um tipo peculiar de homeostase; quanto ao processamento cognitivo, a proposição é que a ficção continuamente se adapta (e tem que se adaptar) à mente humana: mesmo as metaficções de Luciano, Cervantes, Sterne e Pirandello, na aparência apartadas das prováveis motivações biológico-culturais que deram origem à prática, pressupõem uma capacidade de leitura (pela narratologia cognitiva de uma Monika Fludernik, pela neurociência da leitura de um Stanislas Dehaene...) biologicamente condicionada.

A concepção da ficção como prática cultural se apoiando na análise do seu processamento cognitivo: essa convergência tanto complexifica quanto deflaciona a sua teorização. Ela a complexifica, pois muitos elementos devem ser inscritos na explicação; ela a deflaciona, pois a ficção perde em especificidade para tornar-se uma exaptação de capacidades e atividades que não evoluíram para ela. Vejamos como isso funciona.

Como a ficção veio a existir, como ela assegurou a sua subsistência? Entendo que a sua origem pressupõe a *collective intentionality* de Michael Tomasello (2014), que teria evoluído, com a complexificação das sociedades humanas, para possibilitar a comunicação de pessoas estranhas entre si, mas que compartilhavam realidades comuns. À diferença da *shared intentionality*, que fora pano de fundo das comunicações pragmáticas entre pessoas da mesma tribo e que teria fundamentado a origem do pensamento humano (ao permitir, entre outras coisas, a enunciação de categorias), a *collective intentionality* regularia o convívio entre desconhecidos que se encontravam regularmente em sociedades maiores e internamente diferenciadas. Teria emergido uma relação de *feedback*: crenças embrionárias seriam associadas a símbolos materiais e signos linguísticos que, com o tempo, passariam a ser processados como manifestações diretas das crenças, colaborando para fortalecer a sua disseminação e, com isso, a própria complexificação social, com sua necessária regulação de larga escala apoiada na difusão e

internalização de símbolos, valores, crenças e categorias coletivamente aceitas – sem as quais a ficção não existe...

Já a elaboração de “mundos possíveis” de autoria coletivizada, como os mitos e epopeias, pressupunha descrições tipificadas (de lugares, personagens, situações), categorias e valores compartilhados, padrões comuns de expectativa, ou seja: o “pano de fundo” de que fala Searle, “conjunto de capacidades, habilidades, tendências, hábitos, disposições, pressuposições e ‘know-how’” que confere estabilidade à nossa relação com o mundo, determinando as condições de satisfação de “todos os nossos estados intencionais, todas as nossas crenças, esperanças e medos pessoais” (SEARLE, 2000, p.61). Pré-intencional, infra-consciente e em parte culturalizado, sem “pano de fundo” os mitos não existiriam – e menos ainda os mundos ficcionais criados por autores individualizados, que explorariam o “pano de fundo” para antecipar as respostas do público a estímulos criados para a sua fruição.

Esses mundos imaginários viveriam em meio a outros usos da linguagem. O atributo da ficcionalidade tem um claro componente relacional; a ficção se define em contraste com a não-ficção. Esse dado relacional dificulta a indução, das propriedades das obras, das características que seriam suficientes para a atribuição de ficcionalidade a todo e qualquer objeto; ao longo do tempo, pode até ocorrer que uma mudança no “pano de fundo” faça com que textos originalmente não-ficcionais passem a sê-lo – como a *Odisséia* e a *Ilíada*, ou mesmo (para os ateus) a Bíblia. Não surpreende, pois, que também as funções da ficção tenham um viés relacional: é em meio às funções de outros discursos e práticas que as suas especificidades vão transparecer. Daí que, se o desenvolvimento de “panos de fundo” culturalizados foi uma condição para a evolução e o processamento cognitivo da ficção, ele não explica o interesse por ela; em si, o apoio a crenças, símbolos e valores compartilhados não a difere, por exemplo, do mito e do pensamento moral. Sua especificidade envolveria o direcionamento ao presente: enquanto o mito e a moral miram a atemporalidade, a ficção fala do aqui e agora.

Em última análise, o interesse por ela parece ser movido por um uso emotivo da linguagem que permite a simulação mental, durante a leitura, de ações e situações que o leitor pode se imaginar vivendo, provocando efeitos como a empatia pelas personagens, o juízo moral sobre suas ações, o aprendizado pela vivência simulada (e portanto segura) de vivências alheias, o puro prazer da gargalhada... A ficção seria um simulador mental, emocionalmente carregado, de vivências alheias aproximadas por empatia, implicadas em enredos atraentes e, não raro,

perpassados por problemas socialmente relevantes, mas de equacionamento incerto. Esse último ponto nos põe diante da função que, no meu entender, a notabiliza: ser uma *affordance* cultural que oferece orientação infra-consciente para a autocolocação de indivíduos e coletividades no mundo, colaborando para a sua homeostase. Nos termos de António Damásio, a arte, em geral,

pode ter começado como um expediente homeostático para o artista e os que desfrutassem de sua arte, e também como um meio de comunicação. Por fim, tanto para o artista como para sua plateia, os usos diversificaram-se. A arte passou a ser um meio privilegiado de trocar informações a respeito de fatos e emoções considerados importantes para os indivíduos e para a sociedade. [...] Também se transformou em um modo de induzir emoções e sentimentos alentadores, [e] tornou-se um modo de explorar a própria mente e a mente dos outros, uma maneira de ensaiar aspectos específicos da vida e um modo de exercitar juízos morais e ações morais. [...]

Em suma, as artes prevaleceram na evolução por terem valor para a sobrevivência e contribuírem para o desenvolvimento da noção de bem-estar. Elas ajudaram a dar coesão aos grupos sociais e a promover a organização social, auxiliaram na comunicação, compensaram desequilíbrios emocionais decorrentes do medo, raiva, desejo e tristeza. Provavelmente, também inauguraram o longo processo de criar registros externos da vida cultural. [...]

As artes foram uma compensação imperfeita para o sofrimento humano, para a felicidade não alcançada, para a inocência perdida, mas ainda assim alguma compensação elas trouxeram e ainda trazem, como um consolo diante das calamidades provocadas pela natureza e do mal causado pelos homens. Elas são uma das maravilhosas dádivas da consciência ao ser humano. (DAMÁSIO, 2012, p. 359-360)

Se a arte é assim importante, é porque a vida humana está em constante desequilíbrio. Para além do sofrimento individual (e existencial) de que fala Damásio, na vida social não é fácil circular entre expectativas, normas, jogos de atribuição de poder e status, crenças e valores ambíguos, idiossincrasias individuais... O ambiente social não assegura sequer a estabilidade das suas características (culturais, políticas, demográficas...) aparentemente mais cristalizadas, o que sugeriu a Ellen Spolsky (2015) a função homeostática que especificamente a ficção teria assumido. Spolsky propõe que, periodicamente, o aumento da instabilidade das normas e crenças reguladoras, com o conseqüente tensionamento das relações públicas e privadas, faz com que os “contratos sociais” hegemônicos devam ser renegociados. Contratos sociais são tão mais importantes quanto menos visíveis (ou enunciados) eles forem – eles se tornam efetivos ao fundamentarem expectativas silenciosamente compartilhadas. Daí que, quando a instabilidade do mundo social sob a sua regência impõe a sua renegociação, a dificuldade será enunciá-la, com alguma consistência, os próprios termos do contrato e os seus problemas relativos. Especialmente os problemas recentes costumam ser mal discernidos, emergindo como

“insatisfação”, “incômodo”, “desconforto”, “angústia”, sentimentos que, não por acaso, a ficção adora explorar: ela sabe produzir coerência da ambiguidade e da incerteza.

Para que a sua função homeostática tenha efeito, é imperativo que as obras atraiam a atenção do público num universo informacional competitivo. Ficções competem por atenção entre si e com a não-ficção (nesse caso com a desvantagem, jamais completamente superada, da acusação de “desperdício de tempo” e “extravio moral”). Nesse desafio, prevalecem duas estratégias complementares: abrigar temas de interesse contextual em conflitos de interesse universal; recorrer a códigos comunicacionais que seduzam o público visado. Sobre o primeiro ponto, a psicologia evolutiva sugere um elenco de temas de interesse trans-histórico: conflitos de status, cooperação e competição; problemas e tabus relativos ao sexo, à morte e à doença; as posições de inimigos e grupos desfavorecidos; amizade e rivalidade; a ordem material e familiar e o lugar da coletividade no mundo... Steven Pinker (1998) não exagera ao dizer que a história da literatura universal passeia por temas como esses – que interessam diretamente, afinal, à nossa homeostase social e psíquica. Em cada época e lugar, eles são mobilizados para abordar questões contextualmente inscritas: em enredos que apresentam conflitos de interesse universal, problemas locais são co-extensivamente evocados – e vice-versa.

Quanto ao segundo ponto, tudo parece mais simples: para cativar a atenção do público, a estória deve ser bem contada. Mas o ato de narrar é carregado de implicações: ao invés de abordar um tema de maneira abstrata, o enredo o radica numa estória pessoal; se o tema é a escravidão, ele será mostrado na vida de um escravo. Na competição com a não-ficção, essa individuação favorece a ficção ao aumentar o enlace afetivo do público: se Kahneman sugere que generalidades são pouco efetivas para a absorção de uma nova informação, sermos “surpreendidos por casos individuais” favorece o aprendizado por inferência:

nem mesmo estatísticas causais atraentes vão mudar crenças longamente acalentadas [ou] enraizadas na experiência pessoal. Por outro lado, casos individuais surpreendentes têm um impacto poderoso e constituem ferramenta mais eficaz para ensinar [...] porque a incongruência deve ser solucionada [e] embutida numa história causal. (KAHNEMAN, 2011, p. 220-1)

Nesse aprendizado (consciente ou infra-consciente) e nas conversas que a ficção motiva processa-se a sua função homeostática, que só se efetivará, porém, se a obra permitir um processamento cognitivo que energize o envolvimento do público. As técnicas mais favoráveis são aquelas que reciclam convenções comunicacionais já difundidas – os chamados “gêneros”,

sejam eles ficcionais ou não. Mesmo que a sua delimitação seja *fuzzy*, um gênero é passível de reconhecimento, contando com um público habituado a ele e dotado de um repertório de exemplos internalizados como memória (individual e compartilhada). Essa memória será acionada em cada nova leitura: como em qualquer situação inédita, a mente do leitor tentará identificar padrões de fundo (do enredo, no caso) que expliquem os comportamentos dos personagens e as situações em que eles se encontram. Para atrair atenção para esses personagens e situações, a narrativa recorrerá à familiaridade do leitor com personagens e situações semelhantes, no mundo real ou na tradição ficcional (pois seres ficcionais são tão inspirados em outros seres ficcionais quanto em figuras reais). Nesse apelo da obra à tradição ficcional internalizada como memória, duas situações predominam. Se o leitor se depara com uma obra que segue convenções de gênero, ele terá a expectativa, tal como num jogo esportivo cujas regras constitutivas ele conhece bem, de que aquele enredo – aquela partida em particular – se mostre diferente de enredos anteriores, i.e. de que ele seja único à sua maneira: então, os elementos que lhe conferirem singularidade prenderão a sua atenção e permanecerão na sua memória de curto (e talvez de longo) prazo. Por outro lado, se a obra impuser o aprendizado das regras constitutivas do mundo ficcional (como *O processo* ou *Admirável mundo novo*), tal aprendizado por si conferirá singularidade ao jogo, prendendo a atenção e permanecendo na memória.

Seja qual for o caso, a tradição ficcional orientará o processamento do texto. Ficcionistas exploram convenções ficcionais e não-ficcionais: a reciclagem é a regra, na imitação e na inovação. Esse apelo a convenções de gênero pode ser entendido como uma estratégia de redução do caos informacional e direcionamento da interpretação. Gêneros são instrumentos de desaceleração da história; tendo evoluído (rápida ou lentamente) por tentativa-e-erro, nas relações diretas ou mediadas entre autores e leitores, eles se consolidaram pela imitação reiterada dos códigos de obras inovadoras de grande sucesso, o que levaria à formação de convenções disseminadas e internalizadas por autores e leitores durante certa duração de tempo. Essas convenções servirão de moldura para o processamento do texto: assim que o leitor discerne o gênero em questão, ele afina as suas expectativas e o seu comportamento de acordo – e aqui retornamos a Ellen Spolsky e ao tema da homeostase.

Pois o caso é que aquelas convenções podem ser mobilizadas para colocar em perspectiva os limites dos contratos sociais hegemônicos. Se Spolsky trata os gêneros como *affordances* culturais, é porque, ao orientarem a postura do público diante do tema abordado, eles também

pode ser utilizados para subverter posturas convencionais. Assim como o recurso reiterativo confirma os contratos ao reiterar um modo convencional de enquadramento dos temas, o recurso diferenciador, ao lançar um jogo que tensiona as convenções de gênero às quais ele mesmo remete, pode perturbar aquele enquadramento usual. Ao usar o humor e a sátira para narrar a história gaúcha em *Incidente em Antares*, Érico Veríssimo manipulava a expectativa pela qual o objeto de riso não é digno de importância: em regra, rimos daquilo que é baixo o suficiente para merecer a nossa risada. Mas ali o tema era a ditadura militar brasileira, sobre o pano de fundo da história de mandonismo, patrimonialismo e violência da política local – nada muito engraçado, pois. Em 1971, porém, misturar as *affordances* do humor e da narrativa histórica, tratando a história como piada, era uma estratégia para denunciar a participação alegre da sociedade civil nas operações de um Estado organicamente corrupto: velhas convenções, novas *affordances*.

A proposição geral é que ficções são *affordances* culturais cuja função homeostática se manifesta na abordagem de temas que incomodam, mas que ainda não sabemos bem como articular, e menos ainda como enfrentar. Quando já sabemos lidar com ele, a ficção muda de assunto: em 1937, Octavio de Faria (em *Mundos mortos*) criticava a opressão da moral católica sobre a formação do jovem brasileiro, e hoje o tema soa ocioso a ponto de ter sumido do nosso romance; em 2016, por sua vez, Michel Laub (em *O Tribunal da quinta-feira*) abordou as implicações morais da destruição de uma reputação nas redes sociais, tema que ainda não estamos coletivamente preparados para enfrentar – não há consenso ético, repertório de experiências amplamente compartilhadas, aprendizado histórico consistente, jurisprudência consolidada... Na abordagem de angústias urgentes, mas de apreensão e equacionamento ainda incerto, a ficção é estratégica: mesmo que apenas durante a leitura, ela fomenta certa estabilização da vivência emocional do tema, sob a orientação subliminar dos juízos que a narração lhe atribui. Se o leitor vivencia imaginativamente o tema de maneira organizada, sob impactos emocionais direcionados, isso pode (na aceitação ou na rejeição dos valores sugestionados), suscitar-lhe algum aprendizado, consciente ou infra-consciente.

Não é que o leitor termine a leitura sabendo exatamente o que pensar sobre o tema. Mas se ele se envolveu de maneira emocionalmente intensa com a sua ficcionalização, é possível que essa experiência altere, nalguma medida, o seu padrão de relação com ele. Tal seria um possível efeito permanente da leitura: um aprendizado não racionalizado ou, pelo vocabulário de Kahneman, uma mudança no “Sistema 1” do leitor que altere o seu padrão de atribuição de valor

ao tema pela alteração dos vieses heurísticos que ele ordinariamente lhe aplica. É legítimo supor que todo ficcionista quer manipular as associações mentais do leitor entre os conteúdos evocados pelo texto e as suas noções pessoais sobre o real, para tanto sugestionando valores como implicaturas: se, por exemplo, o autor manipula efeitos de *priming* para associar valores morais às representações textuais, Kahneman (v. 2011, p.163) sugeriria que a condição infra-consciente da apreensão desses juízos torna-os mais propensos a serem absorvidos sem a interposição racionalizada do “Sistema 2”, e assim alterarem os vieses heurísticos que orientam as atribuições de valor do leitor em sua vida cotidiana.

Essa ocorrência de efeitos duradouros, questão recorrente na teoria da ficção, está implicada na proposição da sua função homeostática. Para aprofundá-la, darei um giro pela segunda proposição deste artigo: que a ficção não teria se consolidado como prática cultural se ela exigisse capacidades mentais especificamente desenvolvidas para ela, ao invés de explorar aquelas que o leitor já possui – como poderia ter sido diferente? A hipótese é que a ficção foi uma exaptação de práticas culturais preexistentes, mediante a exploração de capacidades mentais preexistentes. Aqui as teorias sobre a sua origem evolutiva e o seu processamento cognitivo dão-se as mãos: entendo que apenas o argumento de que a ficção se adapta à mente humana, e não o contrário, explica o seu sucesso evolutivo; nosso interlocutor, a partir de agora, será o Richard Gerrig de *Experiencing narrative worlds* (1993).

Gerrig acredita que os teóricos da ficção erraram ao apostar numa “taxonomia” que a separa de maneira radical da não-ficção, o que os levou a postular que capacidades mentais específicas são necessárias para a sua leitura. Ao invés disso, ele entende que a narração ficcional se apropria de um padrão comunicacional trivial ao fazer do leitor um *side-participant* de situações em que ele não é solicitado a participar, sem demandar capacidades diferentes daquelas que são acionadas quando testemunhamos, num lugar público, uma cena na qual não participaremos em nenhuma medida. Como falantes, temos vasta experiência em informar os outros através de enunciados que não são dirigidos a eles e em sermos informados por enunciados não dirigidos a nós. Ora o falante exclui o ouvinte deliberadamente, ora ela apenas não tem a intenção de dirigir-se a ele: esses ouvintes “excluídos” serão *overhearers* do diálogo em curso, tal como os ouvintes ou leitores dos diálogos de Shakespeare – na condição de *overhearers* das falas de *A Tempestade*, o fato de sabermos que aquelas personagens não são

reais não impede que as compreendamos como se estivéssemos diante de estranhos num elevador.

A ideia geral é que a prática conversacional que adquirimos na vida, somada à intuição de que o autor quer que saibamos das informações que o narrador nos comunica e que os personagens comunicam uns aos outros, e à intuição de que essas informações são verdadeiras e confiáveis no mundo ficcional, bastam para processarmos os enunciados ficcionais. Assim apoiado na sugestão de que o processamento da ficção apela a predisposições formadas no nosso aprendizado cognitivo global, Gerrig critica o pressuposto da contraposição consciente, durante a leitura, entre realidade e ficção: tal contraposição não seria *a priori* relevante, uma vez que, durante a imersão na leitura, a realidade fica como pano de fundo, sem emergir necessariamente à consciência – o que faz com que as remissões do texto ao real e os seus aspectos propriamente ficcionais se misturem continuamente na mente do leitor. Gerrig entende que o princípio de Coleridge da “suspensão voluntária da descrença” (ainda o “modelo padrão” da teoria da ficção), ao postular que a consciência da diferença ontológica entre realidade e ficção determina a relação do leitor com o texto, prevê que aquela consciência subjaz continuamente à leitura, determinando os efeitos que ela produz – o que, no entanto, não seria corroborado pelas evidências.

Gerrig critica especialmente as “teorias da alternância”, pelas quais o leitor processa a ficção como um mundo à parte, “indo e voltando” do real à ficção no decorrer da leitura – ele vai ao real quando o processamento da obra solicita informações externas ao texto, e volta à ficção para fruir os conteúdos ficcionais. Isso é supor que a consciência da ficcionalidade é constantemente ativada, levando o leitor a restringir, também conscientemente, a validade de certos conteúdos ao mundo ficcional. Mas se a validade das informações ficcionais é restrita ao mundo ficcional, isso teria por corolário que elas não podem influenciar a relação do leitor com o real, ao qual ele sabe que elas não se aplicam. Assim, ao postularem que a consciência da distinção ontológica entre fato e ficção é co-extensiva à leitura, as “teorias da alternância” redundam na negação à ficção de efeitos práticos de relevo – uma vez que, à diferença de D. Quixote e Madame Bovary, o leitor não confunde a ficção com a realidade...

Mas a descrição momento-a-momento do processamento mental do texto revela outra coisa. Experimentos sugerem que a leitura não é alternante, pois isso implicaria que o leitor sabe exatamente quando alterar sua relação com a informação, passando da consciência da ficcionalidade durante a avaliação da informação, para o esquecimento dessa condição durante a

fruição do texto. Gerrig aponta que isso implicaria “ligar e desligar” conjuntos inteiros de convenções e processos cognitivos, e em seu lugar ele sugere que o processamento do texto envolve o “escrutínio ativo” apenas das informações nele contidas, sem qualquer remissão necessária, contígua ou contínua, ao mundo real e a informações factuais.

Em questão está o modo como a informação apresentada na obra é representada na memória. As informações já presentes na memória do leitor fazem com que a informação ficcional nunca seja integralmente nova: no mínimo, ela evocará categorias genéricas ao descrever situações, lugares, pessoas... Mas como as informações factuais e ficcionais são registradas e armazenadas no cérebro? Uma hipótese é que haveria compartimentalização entre elas, com o respectivo armazenamento em lugares diferentes: nesse caso, a informação ficcional não seria associada a informações já armazenadas na memória – Gerrig entende que a teoria da alternância pressupõe essa compartimentalização, ao postular que os dois tipos de informação não se misturam. Há indícios, porém, de que ao fim da leitura o leitor tende a se esquecer de vários detalhes contextuais do enredo, mas retém informações cuja validade não se restringia ao mundo ficcional: ele se esquece de informações estritamente ficcionais (com as quais ele se envolvia durante a leitura), e retém informações transmitidas pela ficção, mas não circunscritas ao mundo ficcional – é como se, terminada a leitura de *Vidas secas*, o leitor se esquecesse dos detalhes da vida de Fabiano, mas se recordasse da representação geral da condição do retirante. Disso Gerrig deduz que a informação ficcional é incorporada a estruturas preexistentes na memória de longo prazo: há mistura entre elas, pois de outro modo não se explicaria como, da leitura de textos ficcionais, permanecem na memória tantos conteúdos não-ficcionais – o que ocorre, no seu entender, porque informações não restritas ao mundo ficcional têm valor potencial fora dele, em contraste com informações que, importantes apenas para o enredo em processamento, podem ser logo esquecidas.

Durante a leitura haveria, sim, alternância representacional, mas entre informações textuais e informações preexistentes na memória do leitor. A alternância não se dá entre “informação ficcional” e “fato real”, mas entre a informação ficcional e as representações mentais do leitor sobre o conteúdo em questão (a partir dos seus conteúdos de memória). Em geral, conteúdos de memória e informações ficcionais são associados de maneira coerente, com a memória balizando as representações mentais motivadas pelos conteúdos do texto. Texto e memória se complementam e, enquanto o leitor estiver imerso na leitura, as informações presentes no texto

bastarão para o seu processamento, sem que seja necessário o recurso a informações externas (“reais” ou “factuais”). Na leitura de uma estória de Sherlock Holmes, Londres poderá se resumir à Londres ficcional (a cidade como ela aparece na obra), sem que o leitor seja remetido à cidade real. Pode acontecer que ele jamais contraponha conscientemente uma Londres à outra, e nesses termos ele poderá, sim, armazenar informações sobre a Londres ficcional à parte das informações que ele já possuía sobre a cidade real, pelo menos num sentido específico: a descrição ficcional pode nunca chegar a influenciar as suas noções regulares sobre a Londres real. No entanto, se as informações ficcionais estabelecerem uma conexão forte não com a Londres real, mas com o conhecimento que dela tem o leitor, esse conhecimento poderá ser acessado conscientemente para balizar o seu juízo sobre a Londres ficcional: pode chamar-lhe a atenção, por exemplo, que um bairro fosse descrito, mais de cem anos atrás, com uma aparência distinta da que ele hoje apresenta. E o processo pode ocorrer no sentido inverso: a Londres ficcional pode balizar o seu conhecimento sobre a Londres real, revelando-lhe aspectos ainda não apreciados da cidade – nesse caso, a ficção estará alterando a sua percepção do real, causando um efeito potencialmente duradouro.

Há momentos, então, em que a informação ficcional incide sobre a memória provocando a remissão consciente ao real. O que parece determinar o grau e a frequência em que a informação ficcional motivará representações mentais conscientes do real e, em sentido inverso, o grau e a frequência em que o conhecimento do real afetará o processamento da informação ficcional, não é a força da associação entre informações factuais (passíveis de corroboração intersubjetiva) e ficcionais (disponíveis no texto), mas entre conceitos evocados pela ficção e conceitos preexistentes na memória. Não se trata, portanto, da associação entre a ficção e o que existe fora dela, mas entre a ficção e o que o leitor sabe ou acredita saber sobre o que existe fora dela: o real não interessa “em si”, mas enquanto internalizado como memória. É nessa associação que a informação ficcional poderá, eventualmente, influenciar as crenças do leitor sobre o mundo – potencial cuja realização dependerá da postura que ele assume diante da informação ficcional.

Segundo Gerrig, a postura padrão é deixar-se convencer pelos conteúdos de verdade apresentados no texto: apenas passagens que suscitarem o escrutínio da informação comunicada serão tratadas de maneira distanciada. Apenas então o leitor aplicará o rótulo “ficcional” à informação que, ademais, tenderá a ser aceita tão logo seja compreendida. A não aceitação seria um evento ocasional e subsequente à compreensão da informação: em regra, o resultado é a

crença na informação lida, as evidências sugerindo que o mero fato de que o leitor entenda uma frase (escrita ou falada) aumenta a chance de que ele a tome como correta. Não haveria, pois, evidências de que, nas representações mentais decorrentes momento-a-momento da leitura, seja compulsória a rotulação da informação como ficcional: o leitor tende a validar toda informação lida. Isso só muda quando o seu envolvimento com o tema faz com que ele se distancie da representação: aí ele tenderá a avaliar de maneira atenta (e talvez crítica) as informações que lhe parecerem relevantes para o seu mundo pessoal, aumentando a probabilidade de que ele estabeleça uma associação consciente entre elas e o seu saber sobre o real. Mas se, em regra, o leitor tende a acreditar na informação ficcional, a ficcionalidade não constitui, no decorrer da leitura, uma categoria dotada *a priori* de coerência psicológica: a informação só será rotulada como ficcional apenas quando ela estiver sob o escrutínio consciente do leitor, e mesmo nesse caso é improvável que ela venha a ser representada na memória de maneira diferente de outros tipos de informação que ele tiver trabalhado para desacreditar (como as que ele considera “falsas” ou “erradas”).

Desse giro retornamos à discussão sobre os efeitos duradouros da leitura da ficção. Se eles acontecem, é porque informações ficcionais resgatam da memória eventos reais e possibilidades causais: mesmo que o leitor sabia que o texto é ficcional, a estrutura da crença acessada sob a mediação da informação ficcional não é alterada pela consciência da ficcionalidade dessa informação mediadora – que o horror da escravidão seja evocado por um enredo ficcional, isso não altera nem a vivacidade da evocação, nem a crença na facticidade do horror evocado. Se nem sempre permanece na memória a informação ficcional, mas as estruturas de crença que vieram à mente do leitor pelo intermédio da ficção, pode acontecer que essas estruturas sejam acessadas para “desmentir” as informações ficcionais, mas também o contrário – i.e. que a ficção “desvele verdades” sobre o mundo real. Quando o enredo apresenta quadros suficientemente parecidos com o real – ou seja: confirmados pelas estruturas de crença do leitor –, esse leitor poderá aceitá-los como proposições consistentes, e talvez inéditas, sobre o real evocado. Gerrig sugere que, em situações desse tipo, pode parecer que o autor está oferecendo “evidências” sobre o real que levam o leitor a se indagar, ou mesmo a acatar, que aquelas representações podem ser tomadas como verdade num sentido estrito: é como se o autor tivesse feito uma descoberta sobre o real, e então oferecesse ao leitor um meio para que ele fizesse a mesma descoberta por si mesmo.

O estado atual da teoria da ficção: da perspectiva deste artigo, ele envolve a conexão entre a evolução da ficção (suas funções e motivações), as condições da sua permanência no tempo (e eventual institucionalização), seu processamento cognitivo, e seus efeitos práticos (duradouros ou não). Dessa interface emerge não uma teorização direcionada às especificidades que distinguiriam (de maneira absoluta) a ficção e a não-ficção, mas uma compreensão relacional de um fenômeno cujas características, efeitos e condições de efetivação são compartilhadas por produções circunvizinhas. A ficcionalidade não comporta uma distinção ontológica precisa baseada na materialidade dos objetos, e dissociada da pragmática da recepção; que, de ordinário, não tenhamos dificuldades em distinguir as ficções em meio a outros gêneros e discursos, isso se daria porque a sua prática subsiste há tempo suficiente para ter desenvolvido convenções comunicacionais tradicionalizadas – que serão evocadas mesmo nas obras inovadoras. Que a ficção tenha se estabelecido como prática, a sugestão é que isso se deveu à sua exploração de códigos atraentes e de processamento acessível, mobilizados para que ela se dispusesse como *affordance* cultural de função homeostática, nos planos individual e coletivo. E que essa função priorize problemas atuais, contextuais e urgentes, isso pode explicar porque a ficção, na sucessão das épocas, mídias, lugares e culturas, sempre se transforma – mas não desaparece.

Referências bibliográficas

- DAMÁSIO, A. *E o cérebro criou o homem*. São Paulo: Companhia das Letras, 2012.
- DEHAENE, S. *Os neurônios da leitura*. Porto Alegre: Penso, 2012.
- FLUDERNIK, M. *Towards a 'Natural' Narratology*. Nova Iorque: Routledge, 1996.
- GERRIG, R. *Experiencing narrative worlds. On the psychological activities of reading*. New Haven: Yale University Press, 1993.
- KAHNEMAN, D. *Rápido e devagar – duas formas de pensar*. Rio de Janeiro: Objetiva, 2011.
- MARTIN, F. *Dinheiro – uma biografia não autorizada*. São Paulo: Portfolio/Penguin, 2016.
- PINKER, S. *Como a mente funciona*. São Paulo: Companhia das Letras, 1998.
- SEARLE, J. *Mente, linguagem e sociedade*. Rio de Janeiro: Rocco, 2000.
- SPOLSKY, E. *The contracts of fiction. Cognition, culture, community*. Nova Iorque: Oxford University Press, 2015.
- TOMASELLO, M. *A natural history of human thinking*. Cambridge: Harvard Press, 2014.

Se a consciência é indispensável para que computadores sejam considerados artisticamente criativos

Whether consciousness is indispensable for computers to be considered artistically creative

Ricardo de Oliveira Toledo⁹¹

RESUMO: Este trabalho discute se a consciência em computadores é um quesito fundamental para que um programa computacional seja artisticamente criativo. Intenta-se saber se apenas um comportamento criativo não bastaria para que a criatividade seja atribuída a uma máquina. Como ponto de partida, confronta-se a posição de Georg Wilhelm Friedrich Hegel (1770-1831), que afirma que a obra de arte é manifestação espiritual e, portanto, de seres conscientes para outros seres conscientes. Para tanto, o artigo aborda recentes discussões sobre criatividade computacional, especialmente, aquelas que se encontram no livro *Computational Creativity Research: Towards Creative Machines*, de 2015. Baseando-se nas reflexões de Hubert Lederer Dreyfus (1929-2017) e John Rogers Searle sobre Inteligência Artificial, verifica-se a suposta ausência de compreensão semântica por programas computacionais e o impacto disso na geração de objetos estéticos. Concomitantemente, pela perspectiva dos objetos gerados, debate-se se estes apresentam uma espontaneidade criativa ou se demonstram apenas obediência a padrões e comandos externos.

PALAVRAS-CHAVE: Inteligência Artificial, Consciência, Criatividade, Arte.

ABSTRACT: This paper discusses whether consciousness on computers is indispensable for a computer program to be artistically creative. The research tries to know if only a creative behavior would not be sufficient for creativity to be attributed to a machine. As a starting point, the position of Georg Wilhelm Friedrich Hegel (1770-1831) is confronted, which states that a work of art is a spiritual manifestation, i.e., from a conscious being to other conscious beings. For this, the paper addresses recent discussions about computational creativity, especially those found in the book *Computational Creativity Research: Towards Creative Machines*, 2015. Based on the reflections of Hubert Lederer Dreyfus (1929-2017) and John Rogers Searle on Artificial Intelligence, the supposed absence of semantic understanding by computer programs and the impact of this in the generation of aesthetic objects are verified. Concurrently, from the perspective of the generated objects, it is debated whether they present a creative spontaneity or if they demonstrate only obedience to external patterns and commands.

KEYWORDS: Artificial Intelligence, Consciousness, Creativity, Art.

⁹¹ Doutor em Filosofia pela UERJ. Pesquisador independente. Email: ricardotoledo1979@hotmail.com.

Problematizando a suposta criatividade de sistemas computacionais

Ao se deparar desavisadamente com o quadro *Decorative Panel*, um óleo sobre tela de 1992, um espectador pode, de antemão, considerar estar inequivocamente diante de uma obra de arte. Possivelmente, algumas pessoas julgarão as qualidades artísticas que se exprimem nos elementos da pintura, como os traçados que envolvem suas formas. Outras tentarão enquadrar a obra em um estilo ou escola das artes plásticas. Um terceiro grupo se esforçará para descobrir os motivos que inspiraram o artista, perguntando-se sobre a importância do que foi retratado. Todavia, ainda mais relevante seria a realização de um dos objetivos de Harold Cohen (1928-2016), a saber, que alguém dissesse que a pintura era fruto da atividade artística de um ser humano e não de um programa de computador. Cohen desejava transformar seu programa computacional, AARON, num gerador autônomo de imagens, diferentemente de outras ferramentas computacionais utilizadas por usuários para elaboração digital de desenhos. Para tanto, AARON deveria estar habilitado por seu programador para decidir sobre cada passo da tarefa a ser feita.⁹²De acordo com Cohen, “AARON é uma entidade, não uma pessoa; e seu estilo artístico inconfundível é produto da sua entidadiade (*entitality*), se posso cunhar este termo, não de sua personalidade” (Cohen, 1999, p. 9). Cohen parecia ter certo grau de confiança de que as obras de AARON derivam de sua criatividade, ou de um comportamento criativo que não pode ser negligenciado.

Deixe-me começar por refletir que AARON é capaz de fazer o que seres humanos inteligentes fazem, e faz isso num nível significativo de inteligência, sem o sistema visual no qual os seres humanos confiam e sem o total alcance de conhecimento experimental que eles suportam, neste caso para colorir. Vale ressaltar também que a resposta que o seu trabalho é capaz de evocar no espectador parece não ser muito prejudicada pela própria falta de uma vida emocional do programa. No mínimo, esses fatos devem levantar questões interessantes que claramente não estão sendo abordadas por aqueles que se concentram apenas em saber se a máquina pode ser dita ter inteligência, ou consciência, ou, em último sentido, uma alma. No entanto, meu objetivo neste artigo não era simplesmente demonstrar a proeza de AARON como um

⁹²Cohen, além de criador do projeto AARON, programa computacional escrito para gerar imagens autonomamente, era um artista britânico que, em 1968, tornou-se professor e pesquisador da Universidade da Califórnia, em San Diego. Após 1991, passou a trabalhar no departamento de artes visuais e no Centro de pesquisas em computação e artes da mesma instituição (Cohen, 1994).

criador de imagens, mas especificamente para procurar provas de um comportamento X dentro do programa.⁹³

O referido “comportamento X” se relaciona com o que vários pesquisadores e programadores do campo da Inteligência Artificial (IA) chamam de “criatividade artificial”. AARON é apenas uma das várias tentativas de se desenvolver competências criativas em sistemas computacionais. Tendo isso em vista, as questões que se levantam são: a) se máquinas poderiam ser dotadas de algo similar ou idêntico ao que se entende como criatividade (humana); b) sendo a resposta de “a” razoavelmente viável, cabe saber quais seriam os quesitos para a construção de um sistema artificial criativo; c) encontrados e aplicados os quesitos requeridos em “b”, resta saber se tal sistema poderia criar obras genuinamente de arte. Permeando as três questões anteriores, inclui-se o problema que concerne ao suposto papel da consciência para que uma entidade qualquer detenha algo como a criatividade artística.

As reflexões aqui presentes partem de um ponto de embate considerado emblemático nos estudos de estética e filosofia da arte: os *Cursos de estética de Hegel*. Segundo o plano estético hegeliano, para se atribuir ao produto de uma atividade criativa qualquer o status de obra de arte é imprescindível que esta seja espiritual ou, noutros termos, executada de maneira direcionada por um ser autoconsciente, o qual estabelece sentidos para suas experiências, sejam elas individuais ou sociais. Nesta perspectiva, o grau elementar de consciência indispensável para produzir arte é a autoconsciência. O segundo grau seria a consciência de algo, seja um conteúdo interno, isto é, mental, ou externo, designadamente, outros entes e, talvez, possíveis relações entre eles. O terceiro, e último a ser abordado, é a consciência de se poder gerar um objeto de fruição estética por meio de uma atividade que tenha como finalidade apresentar artisticamente os próprios conteúdos mentais. Isto porque, de acordo com Hegel, a arte não tem seu acabamento na experiência mental, mas proporciona uma experiência estética – que se comunica. Ela se dá como evento sensível, porém, pode revelar seu conteúdo espiritual e, consequentemente, consciente. Tanto na criação quanto na contemplação da arte, apresenta-se e significa-se a autoconsciência do ser humano, bem como suas experiências estéticas. O trecho abaixo dá indícios claros de como Georg W. F. Hegel pensa as atividades humanas como reconhecimento de si enquanto espírito autoconsciente, diferente dos demais entes.

Trata-se da diferença infinita que, por exemplo, separa o homem do animal. O homem é animal, mas mesmo em suas funções

⁹³ Cohen, Ibidem, p. 13.

animais não permanece preso a um em-si como o animal, pois toma consciência delas, as reconhece e as eleva à ciência autoconsciente, tal como faz, por exemplo, com o processo da digestão. Por meio disso, o homem soluciona o limite da sua imediatez de existente em si, de tal modo que, pelo fato de saber que é animal, deixa de sê-lo e se dá o saber de si mesmo como espírito. – Se o em-si do estágio anterior, a unidade da natureza humana e divina, é elevada de uma unidade imediata para uma unidade consciente, então o verdadeiro elemento para a realidade deste conteúdo não é mais a existência sensível e imediata do espírito, a forma humana corporal, mas a interioridade autoconsciente (Hegel, 2001, p. 94).

Com o avanço das pesquisas e inovações no campo da IA, forma-se um coro de defensores da criatividade artificial que rechaça a consciência como quesito indispensável para a construção de máquinas criativas. Isto não significa que a preocupação com a “consciência artificial” tenha sido abandonada. Como ressaltou Hubert L. Dreyfus (1992), pesquisas no campo da cibernética e da IA fizeram crescer a esperança de se reproduzir faculdades mentais em máquinas computacionais. Nesse sentido, a discussão sobre o que é a consciência, quais são seus possíveis graus e como seria reconhecida em entidades ficou cada vez mais intensa e a filosofia adentrou ao campo de investigação científica com base fisiológica de referência, cibernética e IA.⁹⁴ Entretanto, desconsiderando-se atuais indagações a respeito da “consciência artificial”, salienta-se que tem havido um empenho em IA para se aliar as pesquisas em epistemologia, psicologia, linguística e estética ao desenvolvimento de uma espécie de código da criatividade, que seria replicável em máquinas de modo a torná-las supostamente criativas.

Serão tratados neste trabalho os desafios que a IA encara para que computadores apresentem comportamentos criativos, particularmente, na construção de sistemas com maior autonomia nos campos da semântica e no estabelecimento de regras e padrões que não advenham meramente de sua programação.

⁹⁴ Este debate encontra ecos na “teoria funcionalista estado-máquina”, discutida por Hilary Putman (2008), que postula que a mente funcionaria como um sistema computacional, o que ficou conhecido como “teoria funcionalista estado-máquina”. De certa forma, os estados mentais poderiam ser reproduzidos em tipos físicos de sistemas que não fossem o cérebro humano. Um exame da viabilidade da construção de protótipos de consciência artificial não é objetivo do presente estudo.

Breves considerações sobre Inteligência Artificial

Um dos marcos iniciais da Inteligência Artificial se deu com a proposição do *Teste de Turing*, inscrito na revista filosófica *Mind*, em 1950, num artigo intitulado *Compute machinery and intelligence*, escrito por Alan Turing (1912-1954). O teste serviria para verificar se uma máquina poderia apresentar um comportamento que viesse a ser percebido/confundido como resultante de uma entidade inteligente e pensante, ou seja, um ser humano. Se provado que sim, o próximo passo seria saber se tal máquina poderia ser considerada inteligente ou, até mesmo, pensante. Ressalta-se que Turing não escolheu o termo “teste” em seu artigo, mas a expressão “jogo da imitação”. Concomitantemente, desvinculou a pergunta sobre se as máquinas podem pensar da questão quanto à possibilidade de elas possuírem consciência ou introspecção. Da publicação do artigo até a segunda década do século XXI, muito se evoluiu na interação entre seres humanos e as máquinas.

Peter Norvig e Stuart Russell definem o campo de pesquisas em inteligência artificial da seguinte maneira:

Denominamos nossa espécie *Homo sapiens* — homem sábio — porque nossa inteligência é tão importante para nós. Durante milhares de anos, procuramos entender *como pensamos*, isto é, como um mero punhado de matéria pode perceber, compreender, prever e manipular um mundo muito maior e mais complicado que ela própria. O campo da inteligência artificial, ou IA, vai ainda mais além: ele tenta não apenas compreender, mas também *construir* entidades inteligentes. [...] Atualmente, a IA abrange uma enorme variedade de subcampos, do geral (aprendizagem e percepção) até tarefas específicas, como jogos de xadrez, demonstração de teoremas matemáticos, criação de poesia, direção de um carro em estrada movimentada e diagnóstico de doenças. A IA é relevante para qualquer tarefa intelectual; é verdadeiramente um campo universal. (Norvig; Russel, 2013, p. 26).

Ambos evitam o problema da tradição cartesiana referente à consciência – “*The hard problem*” - ou da introspecção em sistemas de Inteligência Artificial e reconhecem que o *Teste de Turing* (ou jogo da imitação) lançou os pilares dos estudos sobre inteligência artificial, que podem se resumir em:

Processamento de linguagem natural para permitir que ele [o computador] se comunique com sucesso em um idioma natural. Representação de conhecimento para armazenar o que sabe ou ouve. Raciocínio automatizado para usar as informações armazenadas com a finalidade de responder a perguntas e tirar

novas conclusões. Aprendizado de máquina para se adaptar a novas circunstâncias e para detectar e extrapolar padrões. [...] Visão computacional para perceber objetos. Robótica para manipular objetos e movimentar-se.⁹⁵

Todavia, como afirmam, nenhum exemplar que abarcasse todas essas exigências foi construído, e os pesquisadores em IA acabam se concentrando em um ou outro dos quesitos do *Teste de Turing*.

De acordo com João de Fernandes Teixeira (2009), as pretensões quanto à inteligência artificial podem ser categorizadas em três tipos. No primeiro o ser humano imitaria as máquinas, como no episódio da suposta máquina de jogar xadrez do século XIX do barão de Kempelen, que no final das contas teria sido desmascarada quando o anão que se escondia em seu interior e fazia suas jogadas foi descoberto num ato de ira de Napoleão Bonaparte (1769-1821). Nesse caso a inteligência artificial seria meramente uma fraude.

No segundo tipo estariam as máquinas que pudessem imitar um ser humano. Como exemplo, um programa computacional dedicado a jogar xadrez que deveria simular os raciocínios e desenvolver estratégias de jogo assim como um jogador profissional.

Finalmente, no terceiro tipo se encontrariam as máquinas destinadas a ultrapassar a mente humana, como ocorreu com o Deep Blue, um computador construído pela IBM e que, em 1997, venceu o então campeão mundial de xadrez, Gary Kasparov. Explicações puramente técnicas sobre os processos de funcionamento de dispositivos de IA não são de interesse deste trabalho. Satisfaz-se apenas em informar que, de acordo com Teixeira, o êxito de uma máquina como o Deep Blue, que ultrapassa a mente humana numa tarefa, está em seu poder de operar a partir da “força bruta”, isto é, na capacidade de averiguar num tempo mínimo todas as possibilidades existentes para a conclusão de certa operação, ao passo que as máquinas que imitam a mente humana realizam sempre uma espécie de raciocínio seletivo, traçando estratégias e eliminando as possibilidades menos consistentes. Este procedimento é denominado “heurística”.

Todavia, entre as principais novidades da IA está a “*machine learning*”, com destaque para as “redes neurais”.⁹⁶ Diferentemente da força bruta, os novos recursos possibilitam uma

⁹⁵ Ibidem, p. 28.

⁹⁶ A proposta das Redes Neurais Artificiais não é tão recente. Aparece em publicações como McCulloch e Pitts (1943), Hebb (1949) e Rosembatt (1958). Posteriormente, é lida, por exemplo, nos trabalhos de Hopfield (1982). Devido ao rápido desenvolvimento deste campo da IA, não é simples aqui cobrir os desdobramentos de suas pesquisas, importando apenas trazer ao debate esta breve apresentação, que contribuirá para as reflexões sobre a criatividade artificial.

capacidade de processamento superior e um vasto banco de dados capaz de gerar *outputs* que não podem ser previstos pelos programadores originais. Noutras palavras, esses sistemas têm a capacidade de promover soluções nunca propostas por seres humanos. No caso das redes neuronais, propõe-se um modelo matemático aplicado a sistemas artificiais baseado na estrutura neural de organismos inteligentes. Como aponta Simon Haikin (2001), numa rede neuronal um dispositivo de processamento é chamado de neurônio. Cada neurônio tem dois ou mais receptores de *inputs* nos quais percebem sinais ou informações, bem como conexões de *outputs*, pelos quais envia a resposta de seu processamento a outros neurônios. A ligação de vários neurônios em rede, cada um cumprindo uma função, amplifica a capacidade de processamento e de resposta de um sistema. Logo, numa Rede Neuronal Artificial (RNA) os neurônios (ou unidades) processam interligados, todavia, em paralelo, desempenhando tarefas específicas. Tal modelo é reconhecidamente uma técnica estatística não linear que pode resolver uma série de problemas de alta complexidade. De acordo com John J. Hopfield (1982; 1985), uma das competências primárias do cérebro humano seria a memória associativa (*associative memory*), que possibilita o reconhecimento ou a reconstrução de um padrão ainda que este, ao ser percebido, esteja corrompido ou incompleto. Para Marley M. B. R. Vellasco, o cérebro humano possui as características desejáveis em sistemas computacionais, como o poder de lidar com informações inconsistentes e/ou probabilísticas, de se adaptar a situações aparentemente pouco definidas e tolerância a falhas. De forma bastante sucinta, as etapas de modelagem de uma RNA são: a) treinamento e aprendizado, obtidos pelo ambiente gerador dos dados; b) associação, que se define como o reconhecimento de padrões distintos e c) generalização, isto é, a capacidade da rede de “reconhecer com sucesso o ambiente que origina os dados e não propriamente os dados utilizados no treinamento” (Vellasco, 2007, p. 9). Para que uma RNA tenha condições de modelar problemas difíceis de especificar, demanda-se uma grande gama de dados, bem como observações suficientes e representativas para que seu aprendizado ocorra com sucesso. O conhecimento é passado para rede por um algoritmo de treinamento. O aprendizado é transformado e armazenado, sendo, portanto, resultado de muitas apresentações de um conjunto de exemplos de treinamento. O objetivo do treinamento é fazer com que se obtenha um conjunto de *outputs* desejados ou minimamente consistentes a partir da aplicação de um conjunto de *inputs*. Em suma, a maneira como uma RNA processa permite que ela não seja meramente um

sistema computacional que roda programas, mas uma entidade artificial que aprende e realiza operações com resultados bastante complexos.⁹⁷

Um olhar filosófico sobre a suposta criatividade artificial

Pode-se dizer que a “criatividade artificial” é um ramo da IA com consequências naquilo que se convencionou chamar de humanização de computadores, isto é, a elaboração de máquinas que apresentem comportamentos e/ou tenham processos que se assemelhem aos estados mentais humanos. Exemplar dos desdobramentos de seus esforços teóricos é o livro *Computational Creativity Research: Towards Creative Machines* (Besold et al., 2015), composto por vários artigos/capítulos escritos por pesquisadores da área. O primeiro deles apresenta os “*stake holder groups*” em criatividade computacional e suas práticas. Neles estariam os programadores, o público em geral, artistas e demais domínios criativos, psicólogos, filósofos, educadores, críticos, jornalistas e governantes. De forma pragmática, os autores expõem como estes grupos que se relacionam com a “criatividade artificial” encaram os possíveis sistemas artificiais criativos a partir do que pressupõem ganhar ou perder com eles. Por outro viés, analisam visões filosóficas sobre a criatividade humana como sendo possivelmente replicável em computadores. Nota-se no artigo uma tentativa de se definir criatividade para atender aos fins da - ou justificar a - “criatividade artificial”.

Defendemos que a criatividade é uma qualidade secundária e essencialmente questionável em uma pessoa, e que o uso linguístico relacionado à criatividade pode frequentemente constituir atos de fala ilocutivos declarativos. [...] Primeiramente, nós acreditamos que atribuições de criatividade são contextuais, não tendo valor de verdade que seja independente do contexto, percepção e interpretação. Desta forma, nós vemos as atribuições de criatividade em analogia à noção lockeana de uma qualidade secundária. Locke distingue qualidades primárias e secundárias, no que as primeiras são tomadas como intrínsecas ao objeto, por exemplo, sua massa, e as últimas são entendidas como dependentes da percepção, por exemplo, a cor. [...] A postura intencional de Dennett também é interessante aqui: nós podemos adotar uma “postura criativa” em direção a uma pessoa e interpretar os trabalhos dela como se fossem criativos, com o objetivo de entender melhor seu comportamento. Igualmente, nós podemos

⁹⁷Um exemplo de RNA é aquela impulsionada pela Google DeepMind e com a colaboração de vários pesquisadores em IA, destinada a aprender a jogar videogames de maneira similar aos seres humanos, dentre outros objetivos (Antonoglou et al., 2016).

crer que a “postura criativa” permite uma nova forma de entender o comportamento de uma peça de software que vai além dos detalhes físicos do programa.⁹⁸Galileu apresentou os conceitos essencialmente questionáveis como aqueles pelos quais o próprio uso inevitavelmente envolve infindáveis disputas [...] da parte de seus usuários [...] não podendo estar baseados no apelo por evidências empíricas, no uso da linguagem ou nos padrões da lógica simplesmente. [...] Claramente, certas noções como arte são conceitos essencialmente questionáveis, considerando-se a multidão de artigos escritos a cada ano [...] perguntando: mas o que é arte? (Colton et al., 2015, p. 5-6).

Assim, os autores revelam que não existe um acordo filosófico para o que seria a criatividade, e que reconhecer algo como sendo criativo é muito mais tarefa de um espectador do que uma qualidade intrínseca e indispensável da coisa. Logo, para alguém, tanto uma pessoa pode parecer não criativa quanto um software parecer criativo. Adiante criticam: a “criatividade computacional está numa posição difícil [...]. Em certas compreensões, a questão ‘as máquinas podem ser criativas?’ deve ser respondida negativamente, sem um debate posterior”.⁹⁹Isto se dá, sobretudo, por aqueles que advogam a consciência como elemento essencial para a criatividade. De maneira semelhante aos cientistas que enfrentaram forte oposição política, cultural e religiosa para provar suas posições ao longo da história da ciência, os pesquisadores em criatividade artificial encaram as indisposições quanto ao tema não apenas incrementando seus computadores, mas, também, no campo teórico e filosófico.

O segundo capítulo do livro é dedicado à distinção entre criatividade (artificial) forte e fraca. Para a empreitada, Mohammad Majid al-Rifaie e Mark Bishop fazem uso do “argumento do quarto chinês”, de John R. Searle, ao qual se recorre frequentemente para confrontar os defensores da IA forte. O conceito de “*Strong AI*” se refere à possibilidade de se instanciar genuíno entendimento numa máquina, de forma que esta poderia pensar, entender algo e ter outros estados cognitivos, dentre os quais um tipo de consciência, bem como intencionalidade. Em contraste, a “*Weak AI*” (IA fraca) usualmente não iria além da expectativa de simulação da inteligência humana. Como afirma Searle (1980), o apelo por uma IA forte supõe um computador apropriadamente programado para tal finalidade não como ferramenta para estudo da mente, mas como uma mente em si, no sentido em que inserida nele uma programação correta

⁹⁸A postura intencional é exemplificada por Dennett(1987) com o argumento da partida de xadrez entre um ser humano e um computador, na qual aquele observa que existe neste a intenção de vencer a partida, a despeito de creditar à máquina algum tipo de consciência.

⁹⁹Colton et al., Ibidem, 2015, p. 11.

seria capaz de literalmente entender e ter outros estados cognitivos. Uma das objeções de Searle (1989) é que, grosso modo, computadores são hábeis para operações de demandas sintáticas, não para as que envolvam a semântica. Os símbolos formais executados segundo regras sintáticas, em si mesmos, não bastariam para seu entendimento, a saber, sua explicação ou interpretação, exceto na medida em que algo fora do sistema lhes explicasse ou interpretasse. Um computador poderia processar um dado e gerar uma saída aparentemente adequada, mas isso não significa que houve um entendimento genuíno no sistema computacional daquilo que foi processado. Não se intenta, aqui, discutir os contra-argumentos em relação às conclusões de Searle, mas relatar a reflexão de Majid al-Rifaiee Bishop, que segue com o filósofo apenas até este ponto. Para ambos, há uma analogia entre a IA forte e a criatividade (artificial) forte, bem como entre a IA fraca e a criatividade (artificial) fraca:

Em concordância, na criatividade computacional forte, argumentamos que um computador não é meramente uma ferramenta para estudo da criatividade, mas, se apropriadamente programado ele é realmente criativo, no sentido em que dados os programas corretos ele pode literalmente entender sua criação e ter outros estados cognitivos (por exemplo, teleológicos e emotivos) associados com processos criativos humanos (Majid al-Rifaiee Bishop, 2015, p. 48).

Aqueles que defendem a criatividade (artificial) fraca afirmam que enquanto um ser humano criativo sabe o que está fazendo ao escrever uma poesia, pintar um quadro, esculpir uma estátua, projetar e construir um edifício, compor ou coreografar uma peça musical, dentre outras espécies de criação, um computador somente segue as regras de sua programação, não podendo saber que não está fazendo outra coisa além de manipular símbolos. Majid al-Rifaiee Bishop também mencionam Dennett, com objetivo diferente daquele usado por Colton et al., dizendo que, na perspectiva da criatividade computacional fraca, é o observador externo quem atribui ao computador uma suposta criatividade ao julgar seu mérito artístico com a expressão “como se fosse criativo”. Tal julgamento se baseia no contexto social e nas ambições estéticas do observador, não na certeza da intencionalidade criativa numa produção do sistema. Noutros termos, ainda que seja projetado para produzir obras de arte que sejam legitimadas pelo *Teste de Turing*, o computador continuará não sabendo que estava fazendo arte. Concomitantemente, pesa o fato de que certos estados cognitivos humanos, como os emocionais, estariam sempre ausentes na criatividade artificial. Como se nota, retorna-se à questão ontológica sobre “o que é a criatividade” e que parecia resolvida de forma relativista por Colton et al. (2015), isto é, que se

não há acordo tácito sobre sua definição, um computador poderia ser encarado como sendo criativo.

Dreyfus (1992), em *What computers still can't do*, em concordância com Searle, apesar de sua investigação ser anterior, a saber, da edição de 1972 da referida obra, assinala como relevante saber se um computador pode interpretar os aspectos físicos dos quais é informado, construindo uma rede de sentidos, mas com carga semântica coerente. A mente humana não se relaciona com o mundo reduzindo cada fato estanque a dados que sejam categorizáveis atomisticamente a instâncias lógicas com objetivos meramente inferenciais. Antes, elabora uma rede complexa de relações e sentidos entre os fatos mentais e do mundo. Isto é verificado na concepção de situação, que não pode ser confundida com o estado físico do universo. Por exemplo, a palavra lar não designa meramente um lugar ou um objeto dentre os demais no mundo, mas algo que só é compreensível em seu aspecto situacional. Seu sentido é cambiável, apontando, num momento, para uma edificação específica e, noutros, para algo mais amplo, como a terra natal de alguém. Para o trabalhador cansado é almejado como o refúgio de descanso, para um contribuinte é a casa pela qual pagará impostos ao governo, para o viajante o lugar onde estão seus familiares mais próximos. Logo, uma situação não é um conjunto de fatos desconectados entre si, pois um fato só é situado quando seu papel é interpretado numa rede de sentidos na qual todos os outros fatos também adquirem papéis significativos. Não parece simples encontrar um meio para que uma máquina seja capacitada para estabelecer redes de sentidos complexas como as de uma situação, usando termos ou dados com cargas semânticas apropriadas em todos os contextos que forem sugeridos para ela. Mais precisamente:

A característica situacional funciona em dois sentidos: em qualquer situação particular um número indefinido de fatos é possivelmente relevante e um grande número irrelevante. Já que um computador não está numa situação, logo, pode tratar todos os fatos como sempre relevantes (Dreyfus, 1996, p. 169-170).

Como contraponto, poder-se-ia argumentar que uma situação qualquer se tornaria passível de ser formalizada em equações se suas interseções e os estágios de sua evolução fossem minuciosamente reproduzidos algoritmicamente, permitindo que o computador inferisse e significasse ao se deparar com questões contendo alto nível de variáveis e probabilidade. Porém, como questiona Dreyfus, resta o dilema de qual seria o limite de procedimentos *ad hoc* necessários para que uma máquina fosse capaz de gerar autonomamente sentidos *ad infinitum* derivados de uma quantidade finita de informações.

Testes com geradores artificiais de poemas indicam a plausibilidade dos argumentos de Dreyfus e de Searle. Num deles, o Poetry Me, que gera poemas instantaneamente a partir de “palavras-semente” inseridas por seu usuário, esta pesquisa observou as seguintes limitações que deverão ser enfrentadas para seu melhoramento. Embora algumas sentenças/versos pareçam semanticamente coerentes para o leitor humano, considerando-se este o juiz da coerência do texto, outras apresentam termos aparentemente sem qualquer conexão semântica com os demais. Alguns termos não se adequam a parte alguma do restante do poema – ou contexto no qual este poderia estar inserido. Outra limitação é verificada quando novas tentativas solicitando o uso das mesmas palavras-semente têm como resultado repetições de sentenças de tentativas anteriores em pontos diferentes do poema. A recorrência costuma não ser pequena. Tal ocorrência sugere certa randomização do sistema e, por conseguinte, prejudica a simulação de uma intencionalidade poética no programa. Nesse sentido, o computador não estaria demonstrando uma criatividade accidental ao invés de uma criatividade espontânea? A identificação de padrões ou de suposta randomização não se restringe ao Poetry Me.¹⁰⁰ Uma terceira limitação diz respeito ao uso inadequado dos termos sugeridos pelo usuário no contexto poético. É o caso de quando foram inseridas as palavras-semente “computador”, “inteligência”, “criatividade” e “artificial”. Esta última apareceu sempre relacionada a fingimento – em várias tentativas. Deslocamentos ou rupturas semânticas não são incomuns em poemas escritos por humanos, mas isso seria espontâneo ou derivado irrevogavelmente de uma limitação? Não se quer, aqui, afirmar que o programa não possa ser ensinado para minimizar ou até eliminar os impactos causados pelas limitações reportadas acima, e sim que este é seu estado atual e que, portanto, ainda dá fôlego aos críticos da criatividade artificial. Nem que computadores com competências literárias

¹⁰⁰ De acordo com Joanna M. B. Indurkha (2014), elaborar um sistema que produz poemas esteticamente razoáveis baseados em modelos predefinidos não é uma tarefa difícil. Há vários programas geradores de poesia, como o RACTER, Cyberneticpoet, ASPERA e outros, que geram uma linguagem modelo a partir de um conjunto de poemas inseridos pelo usuário e compõem novos no mesmo estilo. Existem sistemas que geram poemas substituindo palavras de um poema terminado. Contudo, a tarefa realmente desafiadora é escrever um programa gerador de poemas com uma forma aparentemente intencional. Existem projetos que tentam incorporar afetos emocionais no processo de criação, como em poemas que falam sobre o humor considerando as notícias do dia. Contudo, o humor é definido como bom ou mal, sem qualquer tipo de refinamento. Outro projeto interessante seria MASTER. Neste sistema, uma sociedade de agentes com vários estados emocionais influenciam o humor uns dos outros com suas peças de poesia. O poema final é o resultado de uma aprendizagem social. No entanto, os poemas produzidos pelo sistema não teriam significação num sentido usual, mas consistiriam na repetição de palavras e sons que criariam poeticidade. Dentre os sistemas mencionados, distinguem-se duas diferentes abordagens em direção à modelagem da “personalidade dos sistemas”. Na primeira, o comportamento do sistema é determinado por parâmetros predefinidos, como no MASTER, em que os agentes definem o humor e as palavras iniciais. Na segunda, a alternativa é adaptar o estado emocional aos fatores do ambiente. Para tanto, as palavras de seu banco devem contar com valores (bom/mau) para que se relacionem aos dados externos com os quais terão que lidar.

se restrinjam aos poetas artificiais. Fala-se hoje do “*robot-journalism*”¹⁰¹ e de geradores artificiais de romances.¹⁰² Contudo, sabe-se que há diferença não só na estrutura textual de narrativas e poemas, como também, no material com o qual podem ser compostos. Ainda que seja simplista, isso é percebido nas seguintes construções: “a chuva de ontem destruiu vinte casas” e “Um dia de chuva é tão belo como um dia de sol. Ambos existem; cada um como é” (Alberto Caieiro/Fernando Pessoa). Ou seja, estados cognitivos, bem como relações semânticas complexas (com alto nível de abstração ou adjetivação) não são indispensáveis em narrativas. Talvez a pergunta correta não seja “o que este computador está escrevendo?”, e sim, “por que ele escreveu isso?”, ou “o que ele quis dizer ao escrever isso?”.¹⁰³

A criatividade computacional não está restrita ao universo literário. Há o caso de programas computacionais compositores de música com impressionantes resultados, como o Emily Howell, desenvolvido por David Cope. Todavia, deles se espera que sejam capazes de ultrapassar as regras de composição baseadas num estilo que lhe foi ensinado (ou com o qual foram treinados) para que criem um corpo próprio de regras e, por conseguinte, novos estilos. Esta é uma marca da criatividade ao longo da história da arte: o vanguardismo. Como escreve Dreyfus (1992), a mente humana sempre parece estar infringindo regras, ainda que comumente estas sejam procuradas em tudo que seja resultado de um raciocínio.

Considerações finais

Talvez, a definição mais sensata para criatividade artificial possa tomar como base a expressão de Cohen “comportamento X” para afirmar que, a despeito de uma consciência em computadores, a criatividade computacional pesquisa comportamentos em máquinas que possam ser considerados criativos se apresentados a seres humanos. O tom relativista que ecoa nas definições contemporâneas sobre criatividade e, consequentemente, de arte, e que lança para o público a tarefa de dizer o que é ou não criativo, o que é ou não arte, reforça o juízo de que computadores podem ser vistos como criativos.

¹⁰¹Ver Clerwall, 2014.

¹⁰²Links para romances gerados por programas. 1) <http://nickm.com/poems/megawatt.pdf>. 2) <http://aparrish.github.io/nanogenmo2014/final.pdf>. Acessos em 05/06/2017.

¹⁰³ Uma das pinturas de AARON é *LibertyandFriends* (1985), na qual estão oito representações da Estátua da Liberdade. Porém, de acordo com o programador, Cohen, apenas uma imagem seria a Estátua da Liberdade, sendo as demais suas amigas. Até que ponto os elementos retratados e relacionados a ideias como liberdade e amizade seriam acessíveis a AARON? Quem estabeleceu os sentidos da obra, o programador ou o programa?

O desenvolvimento de supostos processos computacionais criativos não garante que se esteja replicando ou decodificando os processos mentais relacionados à criatividade humana. No estágio atual das pesquisas em IA e em neurociência, é prematuro prever se um dia máquinas e a mente humana contarão com estados criativos internos idênticos ou semelhantes. O que se pode constatar é que a criatividade artificial, na perspectiva dos objetos gerados, ainda parece incipiente, embora a IA trabalhe para um futuro promissor. São os principais limites para a criatividade artificial abordados nesta pesquisa: a) a dificuldade que programas computacionais têm para gerar resultados com elevada coerência semântica que seja reconhecidamente não acidental ou orientada pelos programadores - esta é uma frente que tem sido amplamente enfrentada pela IA, como se nota em “*machinelearning*” – e b) a capacitação para que máquinas consigam romper com as regras aprendidas, criando, por conseguinte, novos padrões que possam constituir propostas vanguardistas. O vanguardismo é basicamente um processo de ruptura e abertura de novos caminhos para a arte. Não há como negar que as vanguardas e a ideia socialmente vigente de criatividade estejam imbricadas, pelo menos para uma parte considerável do público. A fronteira entre a utilização das regras aprendidas e algo que possa ser considerado vanguardista em obras geradas por computadores é um problema a ser debatido entre aqueles que se relacionam com a criatividade computacional, sejam programadores, filósofos, psicólogos, neurocientistas, críticos da arte e artistas.

Retornando-se ao debate com Hegel, pode-se concluir que máquinas têm demonstrado alguma competência para a geração de objetos estéticos para a fruição humana, e que perseverar demandando consciência de máquinas para que sejam criativas tem gerado fortes ataques de parte das comunidades que se relacionam com o universo da criatividade computacional. A pergunta sobre quem é o verdadeiro artista, o programador ou o computador, perde o propósito se a capacidade de aprendizagem e treinamento dos programas computacionais puder torná-los mais autônomos. Se os computadores sabem o que estão fazendo, ou se fruem de sua arte, não é algo que, do ponto de vista do desenvolvimento da criatividade computacional, seja determinante para considerá-los criativos. Para os defensores da criatividade artificial, o que conta é o fato de que seres humanos possam se beneficiar prática ou esteticamente de seus produtos. Enfim, fundamental não seria a qualidade das criações, e sim, o poder de criar, ainda que este se refira a uma criatividade artificial. Porém, com as crescentes inovações em IA, imagina-se que o debate sobre as questões enfrentadas aqui adquirirá novas perspectivas filosóficas, como aquela de uma

ontologia da arte e da criatividade, no intuito de combater relativismos exagerados sobre criatividade artificial.

Referências bibliográficas

ANTONOGLOU, Joannis; CHRIS, Maddison; DIELEMAN, Sander; DRIESSCHE, George; GRAEPEL, Thore; GREWE, Domink; GUEZ, Arthur; HASSABIS, Demis; KALCHBRENNER, Nal; KAVUKCUOGLU, Koray; LANCTOT, Marc; LEACH, Madeleine; LILLICRAP, Timothy; NHAM, John PANNEERSHELVAM, Veda; SCHRITTWIESER, Julian; SIFRE, Laurent; SILVER, David & SUTSKEVER, Ilya. Mastering the game of Go with deep neural networks and tree search. In. *Nature*. Macmillian Publishers Limited, Vol. 529, Jan. 2016.

CLERWALL, Christer. Enter the Robot Journalist: User's perception of automated content. In. *Journalism Practice Journal*. Taylor & Francis: Vol. 8, No. 5, 2014, p. 519-531.

COHEN, Harold. The further exploits of AARON, painter. In. *Stanford Humanities Review*. Stanford University, 1994, vol.3.

_____. *Colouring without seeing: a problem in machine creativity. An Diego: University of California: 1999.*

COLTON, Simon; COOK, Michael; CORNELI, Joseph; HEPWORTH, Rose; PEASE, Alison and VENTURA, Dan. Stakeholder groups in computational creativity research and practice. In. *Computational creativity research: Towards creative machines. Osnabrück: Atlantis Press, 2015, p. 3-26.*

DENNETT, Daniel Clement. *The intentional stance*. Cambridge, The MIT Press, 1987.

DREYFUS, Hubert. *What computers still can't do: A critique of artificial reason*. Cambridge, MIT Press, 1992.

HAIKIN, Simon. *Redes neurais: princípios e prática*. Trad. Paulo M. Engel. 2 ed. Porto Alegre: Bookman, 2001.

HEGEL, G. W. *Cursos de Estética, volume I*. Trad. Marcos A. Werle. São Paulo: EDUSP, 2001.

HOPFIELD, John Joseph. Neural networks and physical systems with emergent collective computational abilities. In. *Proceedings of the National Academy of Sciences*, Vol. 79, 1982, pp. 2554-2558.

_____. Neurons with graded response have collective computational properties like those of two-state neurons. In. *Proceedings of the National Academy of Sciences*, Vol. 81, 1984, pp. 3088- 3092.

MISZTALE, Joanna; INDURKHYA, Bipin. Poetry generation system with an emotional personality. In. *Proceedings of 5th International Conference on Computational Creativity, ICC*, 2014.

NORVIG, P; RUSSEL, S. *Inteligência Artificial*. Trad. Regina de Macedo. Rio de Janeiro: Elsevier& Campus, 2013.

PUTNAM, Hilary. *Corda Tripla - Mente, Corpo e Mundo*. Trad. Adail Sobral. Aparecida: Ideias e Letras, 2008.

SEARLE, John. Minds, brains and programs. In. *Behavioral and Brain Sciences*. Cambridge University Press, Vol.3, no. 3, p. 417-457, 1980.

_____. Artificial Intelligence and the Chinese Room: An Exchange. In. *New York Review Books*, Vol. 36, no. 2, 1989.

TEIXEIRA, João Fernandes. *Inteligência Artificial: Uma odisseia da mente*. São Paulo: Paulus, 2009.

VELLASCO, Marley Maria Bernardes Rebuzzi. *Redes Neurais Artificiais*. PUC-Rio/Laboratório de Inteligência Computacional Aplicada, 2007.

Bergson: o problema mente-cérebro sob uma perspectiva pragmática

Bergson: the mind-brain problem from a pragmatic perspective

Ricardo Pinho Souto¹⁰⁴

RESUMO: O presente trabalho apresenta o problema mente-cérebro a partir do pensamento do pensador francês Henri Bergson. Partindo do dualismo de substância de Descartes, Bergson elabora uma Teoria da Ação, através da qual a memória e a percepção ocupam lugar destacado. Neste percurso, o cérebro é tomado como órgão que seleciona dentre as lembranças/memórias aquelas úteis à ação. Neste momento o dualismo cartesiano assume o aspecto pragmático destacado no artigo. Ao final, pesquisas recentes acerca do cérebro são revisadas e confrontadas com as ideias anteriormente expostas.

PALAVRAS-CHAVE: Epistemologia, Bergson, mente-cérebro, pragmatismo.

ABSTRACT: The present work presents the mind-brain problem from the thought of the French thinker Henri Bergson. Starting from Descartes' dualism of substance, Bergson elaborates a Theory of Action, through which memory and perception occupy a prominent place. In this course, the brain is taken as the organ that selects from those memories / memories those useful to action. At this point the Cartesian dualism assumes the pragmatic aspect highlighted in the article. In the end, recent research on the brain is reviewed and confronted with the ideas discussed above.

KEYWORDS: Epistemology, Bergson, mind-brain, pragmatism.

¹⁰⁴ Doutor em Psicologia Cognitiva. Professor do Departamento de Filosofia da Universidade Católica de Pernambuco. Email: Ricardopinho11@unicap.br.

Considerações iniciais

O presente trabalho tem por objetivo apresentar a tese defendida por Bergson acerca da função do cérebro e sua relação com a memória. Bergson (1999, p. 276) defende a tese de que a memória não é uma função do cérebro, e que entre percepção e memória não existe apenas uma diferença de intensidade (quantidade), mas sim de qualidade. Bergson (1999, p. 14) concebe o cérebro não como produtor das imagens nem como o local no qual a memória estaria acumulada. O cérebro é o órgão que seleciona representações e memórias visando o agir. Assim, o pensamento bergsoniano aproxima-se do pragmatismo de James (1979, p. 18.). Como salienta Zunino (2010, p. 222) “[...] poderíamos sugerir que das influências recíprocas entre bergsonismo e pragmatismo não se desprende apenas um pensamento que nos incita à ação, senão também uma concepção de filosofia como prática”.

Bergson (1999, p.14) concebe o corpo como um “como um centro de ação” e trabalha com a ideia de que a percepção mescla-se com a memória e nos coloca em contato com a matéria. Bergson (1999, p. 212) volta-se contra o que ele denomina de “dualismo vulgar”, no qual a ideia central encontra-se na tentativa de compreender como o espírito age sobre o corpo e o corpo sobre o espírito. Tal concepção derivada diretamente do dualismo de substância cartesiano deve dar lugar a uma nova forma de dualismo, um dualismo baseado na percepção, no qual a memória (realidade espiritual) mescla-se com a em função de possibilitar a ação do sujeito no mundo material.

Bergson (1999, p. 131) rompe com a concepção reinante a sua época através da qual o cérebro era visto como depositário da memória. No sentido inverso, o cérebro é tomado enquanto órgão que atua como espécie de “filtro da memória”. Em suma, o cérebro tem o caráter de órgão selecionador da memória, alimentando nossas percepções e possibilitando nossas ações com suas implicações no mundo físico. McNab e Klingberg (2007, p 104) respaldam as ideias de Bergson (1999, p. 131) ao realizarem pesquisas nas quais se trabalha com a hipótese de haver áreas específicas no cérebro, *gânglios basais* e *córtex pré-frontal*, responsáveis por filtrar memórias e informações irrelevantes incrementando processos cognitivos elementares e conformando o sujeito para realização de suas ações no mundo.

1 Crítica ao paralelismo psicofísico

Ao realizar a crítica à tese do paralelismo psicofísico, Bergson (2009, p. 202) elabora a ideia do cérebro enquanto imagem, Tal tese tem origem no dualismo de substância cartesiano.

Bergson(2009, p. 205) argumenta que o paralelismo psicofísico encontra-se apoiado em uma base inconsistente, pois tal tese só se sustenta se adotarmos, ao mesmo tempo, dois sistemas de notações filosóficos que se excluem mutuamente. No crivo da crítica bergsoniana, realismo e idealismo são sistemas metafísicos que se diferenciam pela forma de compreensão do real. De forma resumida, o realismo tende a versar acerca do mundo apoiando-se em coisas, enquanto o idealismo fala da realidade em termos de representações.

Do ponto de vista do idealismo a realidade confina-se ao que aparece à consciência. Assim, toda matéria com suas propriedades e as articulações do real são representação. Já para o realismo, para além das representações, há causas não acessíveis à sensibilidade (no sentido kantiano do termo). É suposto que além da percepção fenomênica há a realidade numênica. Assim, um objeto qualquer não se resume ao rol de suas propriedades aparentes. Trabalha-se, no realismo, com essências que escapam ao mostrar-se do objeto.

No entanto, os dois sistemas filosóficos valem-se de princípios adotados ao mesmo tempo por um e por outro. Em uma passagem elucidativa, Bergson (2009, p. 201) esclarece:

[...] idealismo tem por essência o fato de se deter no que está dado no espaço e nas divisões espaciais, enquanto o realismo tem estes dados por superficiais e estas divisões por artificiais: ele concebe, por trás das representações justapostas, um sistema de ações recíprocas, e conseqüentemente uma implicação das representações umas nas outras. Como, por outro lado, nosso conhecimento da matéria não pode derivar inteiramente do espaço, e como a implicação recíproca de que se trata, por mais profunda que seja, não se pode tornar extra-espacial sem tornar-se extra-científica, o realismo não pode ultrapassar o idealismo em suas explicações. Estamos sempre mais ou menos no idealismo (tal como o definimos) quando somos cientistas: caso contrário, não pensaríamos sequer em considerar partes isoladas da realidade para condicioná-las umas em relação às outras, o que constitui a própria ciência. A hipótese do realismo não é mais do que um ideal destinado a lembrar-nos que nunca aprofundaremos suficientemente a explicação da realidade, e que deveremos estabelecer relações cada vez mais íntimas entre as partes do real que se justapõem, a nossos olhos, no espaço. Mas o realista não se pode impedir de hipostasiar este ideal. Ele o hipostasia nas representações dadas que seriam para o idealista a própria

realidade. Estas representações tornam-se então, para ele, outras coisas, isto é, reservatórios contendo virtualidades ocultas: o que lhe permitirá considerar os movimentos intracerebrais (erigidos desta vez em coisas e não mais em simples representações) como encerrando potencialmente a totalidade da representação. Nisto consistirá sua afirmação do paralelismo psicofisiológico. Ele esquece que havia situado o reservatório fora da representação e não dentro dela, fora do espaço e não no espaço, e que, em todo caso, sua hipótese consistia em supor a realidade ou indivisível ou articulada de forma diferente da representação. Fazendo corresponder a cada parte da representação uma parte da realidade, ele articula o real da mesma forma que a representação, ele desdobra a realidade no espaço, e abandona seu realismo pelo idealismo, em que a relação do cérebro ao restante da representação é evidentemente a da parte ao todo.

Para evitar esta manifesta contradição, Bergson (1999, p. 2) adota como ponto de partida a visão do senso comum ao afirmar que se deve considerar a “a matéria antes da dissociação que o idealismo e o realismo operaram entre sua existência e sua aparência”. No bojo desta discussão, uma “teoria da imagem” é proposta visando a solver a contradição acima. Por imagem, Bergson (1999, p. 2) entende:

[...] uma certa existência que é mais do que aquilo que o idealista chama uma representação, porém menos do que aquilo que o realista chama uma coisa - uma existência situada a meio caminho entre a “coisa” e a “representação”. (BERGSON, 1999, p. 2).

Assim, da crítica ao paralelismo psicofísico e da não aceitação das contradições intrínsecas ao realismo e idealismo, tem-se uma concepção de imagem que supera os problemas mencionados, desde que a imagem é mais de que uma representação e menos do que uma coisa. O que há, então, são imagens. A partir daí abre-se caminho para pensar o problema espírito-corpo a partir de outra perspectiva.

1.1 O cérebro como “imagem” prevalente

Dentre as diversas imagens, uma se mostra prevalente: o nosso corpo. Isto porque, no conjunto de imagens que compõe o Universo, nada se produz de realmente novo e criativo que não seja pela mediação daquela imagem particular, cuja forma é fornecida pelo corpo. Ora, nesta cadeia de raciocínio, o cérebro é também uma imagem. No entanto, supor tal imagem – o cérebro, como criadora das demais imagens leva a outra contradição, pois:

É o cérebro que faz parte do mundo material, e não o mundo material que faz parte do cérebro. Suprima a imagem que leva o nome de mundo material, e você aniquilará de uma só vez o cérebro e o estímulo cerebral que fazem parte dele. Suponha, ao contrário, que essas duas imagens, o cérebro e o estímulo cerebral, desapareçam: por hipótese, somente elas irão se apagar, ou seja, muito pouca coisa, um detalhe insignificante num imenso quadro. O quadro em seu conjunto, isto é, o universo, subsiste integralmente. (BERGSON, 1999, p. 13-14).

Originalmente, Bergson (1999, p. 17) concebe o corpo como um centro de ação, que não tem o poder de fazer surgir representações. E, ainda, as imagens que cercam o corpo já trazem potencialmente sua ação. Seria absurdo colocar o cérebro como condição de produção da totalidade das imagens, desde que o cérebro é uma parte dessas imagens. Em poucas palavras, a parte não pode conter o todo. Neste ponto, o propósito primeiro da obra *Matéria e Memória* ocupa o centro das questões. O problema metafísico acerca das relações mente-corpo ganha nova dimensão e o conceito de memória passa a ocupar lugar destacado.

2 Memória e percepção

Para adentarmos no tema da memória, devemos primeiramente identificá-la com a ideia que ocupa a centralidade na obra de Bergson: a ideia de duração. Deleuze (1999, p. 39) afirma haver uma identidade entre duração, memória, consciência e liberdade. Pode-se conjecturar que a ideia de duração em Bergson (2011, p. 87) passa pela interpenetração de momentos heterogêneos. Então, duração é essencialmente mudança qualitativa. Duração e por consequência memória é a realidade temporal desobstruída das categorias próprias para apreensão do espaço. Aqui temos o dualismo que perpassa e é fio condutor do pensamento bergsoniano: o dualismo do espaço e do tempo irá gerar uma série de outros dualismos. Assim, o desdobramento do dualismo base faz surgir pares que apontam para direções diametralmente opostas. Temos a análise e a intuição; a quantidade e qualidade; o tempo da ciência e o tempo da consciência; a moral fechada e a moral aberta e o dualismo escrutinado no presente trabalho: o dualismo da matéria e memória.

Bergson (2011, p. 96) defende que a duração não consiste em uma série de instantes homogêneos, desde que o momento seguinte contém sempre, além daquele que o precedeu, a

lembrança deixada por este. Além disso, como um momento ainda não desapareceu quando o outro já desponta, eles se contraem e se condensam um no outro.

Bergson (2011a, p. 47) nos diz que a “duração é o progresso contínuo do passado que rói o porvir e incha à medida que avança”. O passado tanto cresce incessantemente quanto se conserva indefinidamente. O passado se conserva naturalmente, de maneira que nossos quereres, sentimentos e pensamentos se prolongam e reúnem-se compondo nossa memória, sempre dinâmica. A memória é o ponto de partida no qual nos colocamos em nossas atividades corriqueiras. Para que uma lembrança reapareça à consciência, é preciso que ela desça das alturas da memória pura. Em última análise, Bergson (1999, p. 278) pretende demonstrar que “a memória é algo diferente de uma função do cérebro” (BERGSON, 1999, p. 278). Para tal demonstração resta definir a ideia de percepção e tornar claro qual é a distinção entre memória e percepção e, ainda, mostrar o ponto de confluência no qual a memória e percepção se entrelaçam.

Para chegarmos à ideia de percepção, primeiro é importante afirmar que entre memória e percepção não há uma simples diferença de intensidade. Memória e percepção distinguem-se por natureza. Ou seja, as duas realidades diferenciam-se qualitativamente.

Mas essa ilusão recobre ainda uma outra, que se estende à teoria do conhecimento em geral. O que constitui o mundo material, dissemos, são objetos, ou, se preferirem, imagens, cujas partes agem e reagem todas através de movimentos umas sobre as outras. E o que constitui nossa percepção pura é, no seio mesmo dessas imagens, nossa ação nascente que se desenha. A *atualidade* de nossa percepção consiste portanto em sua *atividade*, nos movimentos que a prolongam, e não em sua maior intensidade. (BERGSON, 1999, p. 72).

Assim, o caráter verdadeiro da percepção consiste em ser um sistema de ações nascentes que se encontra enraizado no real. Novamente desponta a concepção do corpo como “centro de ação”.

2.1 Da relação entre memória e percepção

Poderíamos pensar nas relações entre memória e percepção a começar pela interação que essas duas realidades têm com o mundo material e a esfera espiritual. Desta forma, a percepção está para o mundo da matéria, assim como a memória está para as

coisas do espírito. No entanto, tal distinção não basta, pois não apresenta o ponto de confluência entre memória e percepção.

A originalidade do pensamento bergsoniano, nesta questão, encontra-se na reaproximação entre o inextenso e o extenso, entre a qualidade e a quantidade, entre matéria e memória. E é na ação que a aproximação entre matéria e espírito acontece de forma clara:

Para que uma lembrança reapareça à consciência, é preciso com efeito que ela desça das alturas da memória pura até o ponto preciso onde se realiza a *ação*. Em outras palavras, é do presente que parte o apelo ao qual a lembrança responde, e é dos elementos sensório-motores da ação presente que a lembrança retira o calor que lhe confere vida. (BERGSON, 1999, p. 179).

A memória insere-se em nossos hábitos motores, no instante em que a percepção se dá. Tudo se passa como se a consciência fosse buscar na memória as lembranças, para materializá-las progressivamente em contato com a percepção presente. Então, de nossa memória integral, o corpo em suas diversas atividades, através do cérebro, trabalha com recortes que visam à inserção na teia da matéria, no exato momento em que a percepção engendra a ação. Nesse percurso, o dualismo assumido entre matéria e memória atenua-se, uma vez que “a teoria da percepção pura de um lado, da memória pura de outro, prepararia então o caminho para uma reaproximação entre o inextenso e o extenso, entre a qualidade e a quantidade”. (BERGSON, 1999, p. 212).

Neste ponto, a dissolução do problema mente-corpo vem de uma abordagem centrada na categoria da ação. O dualismo clássico é aqui tratado como o dualismo da matéria e memória é abordado sob a égide do pragmatismo – filosofia examinada com certa simpatia por Bergson, sobretudo no que diz respeito à concepção de verdade enquanto criatividade e invenção. (BERGSON, 2006, p. 253).

Ao final, Bergson (1999, p. 263-291) estabelece proposições conclusivas: i) “o nosso corpo é um instrumento de ação e somente de ação. Em nenhum grau, em nenhum sentido, sob nenhum aspecto ele serve para preparar, e muito menos explicar, uma representação”; ii) “o que se explica pelo cérebro em nossas percepções são nossas ações começadas, ou preparadas, ou sugeridas, e não nossas percepções mesmas”; iii) “o corpo conserva hábitos motores capazes de desempenhar de novo o passado; pode retomar atitudes em que o passado irá se inserir”; iv) “em

nenhum caso o cérebro armazenará lembranças ou imagens”; v) “a memória de um ser vivo parece medir antes de tudo a capacidade de sua ação sobre as coisas”; vi) “A percepção então não é mais que uma seleção. Ela não cria nada; seu papel, ao contrário, é liminar do conjunto das imagens todas aquelas sobre as quais eu não teria nenhuma influência”; vii) “tudo se passará como se deixássemos filtrar a ação real das coisas exteriores para deter e reter delas a ação virtual”; viii) “essa ação virtual das coisas sobre nosso corpo e de nosso corpo sobre as coisas é propriamente a nossa percepção”; ix) “Certamente a seleção das percepções entre as imagens em geral é o efeito de um discernimento que anuncia já o espírito” e, x) “O espírito retira da matéria as percepções que serão seu alimento, e as devolve a ela na forma de movimento, em que imprimiu sua liberdade”.

No prefácio da 7ª edição de *Matéria e Memória*, Bergson indica o objetivo principal da obra: afirmar a realidade do espírito, a realidade da matéria, e determinar a relação entre as duas realidades, a partir de um exemplo concreto: o da memória. Pensamos que o atual trabalho tenha contribuído para tal fim. No entanto, ao final do atual trabalho, apresentaremos recentes pesquisas que corroboram a tese defendida por Bergson (1999, p. 26).

3 O cérebro como filtro da memória

Começamos a parte final do presente artigo com uma indagação: Se nossos centros de percepção são bombardeados por um conjunto incomensurável de informações visuais, auditivas, táteis, olfativas, gustativas, entre outras, e não podemos trabalhar ao mesmo tempo com todo este rol de estímulos, haveria algum dispositivo ou mesmo estruturas no cérebro que selecionam determinados estímulos em detrimento a outros?

Esta questão não está completamente resolvida entre neurocientistas. Sendo assim, enfatizaremos algumas pesquisas que estão alinhadas com a tese defendida por Bergson (1999, p. 263). A tese consiste em propor o cérebro como órgão que seleciona informações visando nossa ação no mundo. Eichenbaum (2012, p. 14) discute as dificuldades de relacionar áreas específicas no cérebro às determinadas funções, como o armazenamento das memórias, a despeito de asseverar a importância do papel do hipocampo frente aos processos da memória. Para ilustrar as dificuldades nesse particular, Eichenbaum (2012, p. 14) cita Karl Lashley e as diversas tentativas

falhas no intuito de localizar áreas cerebrais relacionadas às funções mentais superiores. Dado que corrobora com as ideias bergsonianas.

Basar e Duzgun (2016, p. 202) propõem que para o entendimento do “funcionamento neurofisiológico e o processamento cognitivo das populações neurais” é necessária uma nova descrição da memória. Esta nova descrição abrangeria todos os estados da memória – a hipermemória. Chamamos atenção para a proximidade do conceito de duração exposto em Bergson (2011, p. 69) e o conceito de hipermemória. Hipermemória, sendo o conjunto das memórias particulares, equivale à ideia de duração. Fato relevante no contexto final do presente artigo, desde que conceitos filosóficos ganham agora suporte empírico. Porém, dentre os estudos recentes no campo das neurociências que respaldam a tese bergsoniana de que o cérebro é eminentemente o órgão que impede com que a memória se mostrasse em sua totalidade, destacamos estudos que sinalizam nesta direção.

McNab e Klingberg (2007, p 106) demonstraram que a capacidade de armazenar e recuperar informações varia de acordo com o grau de relevância de nossas memórias. A questão posta em destaque diz respeito à forma como a seleção de itens relevantes acontece. Os autores concluíram que o cérebro não pode lidar com a totalidade de informações advindas das vivências. Logo é necessário haver uma espécie de filtragem das informações.

McNab e Klingberg (2007, p 105) utilizaram recurso de ressonância magnética em 25 sujeitos saudáveis. Os voluntários foram submetidos a um teste computadorizado, no qual eram instados a responder a certos tipos de imagens, apresentados ora com fatores irrelevantes para realização da tarefa, ora sem os mesmos. Na pesquisa, os cientistas observaram que pessoas que têm boa memória registraram maior atividade nos gânglios basais, mesmo quando sofrem interferência de elementos causadores de distrações,

Os cientistas descobriram que quando as imagens acompanhadas dos elementos tidos como causadores de distrações apareciam, havia um aumento na atividade nervosa nos gânglios basais e no córtex pré-frontal, o que sugere que o cérebro estava se preparando para "filtrar" estas informações visuais. No estudo referido, mostrou-se que a atividade no córtex pré-frontal e nos gânglios basais precedeu a filtragem de informações irrelevantes. Estes dados estão em conformidade com a tese apresentada por Bergson (1999, p. 16).

Referências

- BASAR, E. DUZGUN, A. The brain as a working syncytium and memory as a continuum in a hypertime-space: Oscillations lead to a new model. *International Journal of Psychophysiology*. v. 103, p. 199-214 maio 2016, páginas 199-214. Disponível em <www.sciencedirect.com/science/article/pii/S0167876015000501> Acesso em: 08 maio 2017.
- BERGSON, H. *Matéria e memória: Ensaio sobre a relação do corpo com o espírito*. São Paulo: Martins Fontes, 1999.
- _____. *O Pensamento e o Movente*. São Paulo: Martins Fontes, 2006.
- _____. O Cérebro e o Pensamento: Uma Ilusão Filosófica. In: *A Energia Espiritual*. São Paulo: Martins Fontes, 2009, p. 191-209.
- _____. *Ensaio sobre os dados imediatos da consciência*. Lisboa: Edições 70, 2011.
- _____. *Memória e vida*. São Paulo: Martins Fontes, 2011a.
- Eichenbaum, H. What H.M. Taught Us. *Journal of Cognitive Neuroscience* v. 25, n 1 p. 14-21, 2012. Disponível em: <<http://www.smbs.buffalo.edu/acb/neuro/readings/Memory-WhatHMTaught.pdf>> Acesso em: 07 maio 2017
- JAMES, W. *Pragmatismo*. São Paulo: Abril Cultural, 1979.
- MCNAB F, KLINGBERG, T. Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*. v. 11, p. 103-110, jan. 2008.
- ZUNINO, P. Do pragmatismo à intuição mística: uma leitura bergsoniana de William James. *Cognitio-estudos*. v. 7, n. 2, p. 221-235, jul-dez. 2010. Disponível em: <<https://revistas.pucsp.br/index.php/cognitio/article/view/3319>> Acesso em: 10 de maio 2017.

Pampsiquismo em Nietzsche? Apontamentos introdutórios

Panpsychism in Nietzsche? Introductory notes

William Mattioli¹⁰⁵

RESUMO: Nietzsche foi um filósofo muito sensível aos debates da época em torno do materialismo, e muitas de suas reflexões epistemológicas e metafísicas (incluindo sua famosa hipótese da vontade de poder) podem ser melhor compreendidas se pensadas como tentativas de se posicionar dentro deste debate. Nesse sentido, o presente trabalho tem por objetivo analisar em que medida Nietzsche se aproxima de posições que podem ser chamadas de pampsiquistas tendo em vista fornecer uma resposta a uma série de questões associadas ao problema da lacuna explicativa, tal como formulado por Du Bois-Reymond em 1872 com sua famosa declaração: *ignorabimus*, e confrontado por outros autores que lhe eram caros, como Friedrich Albert Lange. O trabalho tem como pano de fundo três teses principais: segundo Nietzsche, I) o pampsiquismo pode ser visto como alternativa mais plausível a versões clássicas do emergentismo; II) o pampsiquismo pode ser visto como uma resposta possível ao problema da lacuna explicativa; e III) o pampsiquismo pode ser visto como uma hipótese capaz de responder, de modo mais satisfatório que o fisicalismo, ao critério de acessibilidade intuitiva que estaria nas raízes da nossa noção de causalidade e da própria noção de explicação.

PALAVRAS-CHAVE: Nietzsche; pampsiquismo; consciência; causalidade; vontade de poder.

ABSTRACT: Nietzsche was a philosopher who was very sensitive to the ongoing debates on materialism, and many of his epistemological and metaphysical remarks (including his famous hypothesis of will to power) are better understood as attempts to take a stance on that debate. Having that in mind, this work aims to analyze to what extent Nietzsche comes closer to positions that can be called panpsychist in his responses to a series of a questions associated to the explanatory gap problem as they are formulated by Du Bois-Reymond, in 1872, in his famous declaration: *ignorabimus* and confronted by other authors who were very close to him, such as Friedrich Albert Lange. This work has as its background three crucial theses: according to Nietzsche I) panpsychism can be understood as a more plausible alternative to the classical versions of emergentism; II) panpsychism can be seen as a possible reply to the explanatory gap problem; III) panpsychism can be understood as a hypothesis capable of explaining, in a more satisfactory way than physicalism, the criterion of intuitive accessibility that is entrenched in the very root of our notions of causality and explanation.

KEYWORDS: Nietzsche; panpsychism; consciousness; causality; will to power.

¹⁰⁵ Bolsista de Pós-Doutorado no quadro do PNPD (CAPES) junto ao Departamento de Filosofia da Universidade Federal Fluminense.

1. Introdução

O objetivo deste trabalho é traçar o esboço inicial de um problema. Não se trata de uma exposição analítica e sistemática exaustiva, nem de uma pesquisa histórica detalhada, mesmo que em minhas reflexões em torno do problema do pampsiquismo em Nietzsche eu esteja, num sentido amplo, partindo dessas duas perspectivas (a analítico-sistemática e a histórico-crítica). Meu intuito é apresentar alguns apontamentos iniciais e levantar algumas questões no tocante a este tema tão importante e pouco trabalhado na literatura sobre Nietzsche, especialmente no Brasil (e que ainda ocupa um lugar relativamente marginal nas discussões em filosofia da mente). Não pretendo fornecer respostas conclusivas, mas hipóteses para uma interpretação da filosofia de Nietzsche a partir de uma chave de leitura específica, elaborada num confronto com algumas discussões contemporâneas do problema, em especial com a defesa do pampsiquismo realizada pelo filósofo Galen Strawson, que vê em Nietzsche igualmente um defensor desta posição. Algumas das hipóteses por trás das reflexões desenvolvidas neste trabalho são: a) o pampsiquismo pode ser visto como alternativa mais plausível a versões clássicas do emergentismo; b) o pampsiquismo pode ser visto como uma resposta possível ao problema da lacuna explicativa; e c) o pampsiquismo pode ser visto como uma hipótese capaz de responder, de modo mais satisfatório que o fisicalismo, ao critério de acessibilidade intuitiva que, segundo Nietzsche, está nas raízes da nossa noção de causalidade e da própria noção de explicação.

2. Considerações iniciais

Harold Morowitz publicou um interessante artigo na coletânea organizada por Daniel Dennett e Douglas Hofstadter, intitulada *The Minds I: Fantasies and Reflections on Self and Soul*. O artigo de Morowitz: “Rediscovering the Mind”, busca elaborar uma interpretação do problema da consciência a partir da integração de três perspectivas científicas: da psicologia, da biologia e da física. O autor toma como ponto de partida a tendência majoritária na virada do século XIX para o século XX e que dominou boa parte das abordagens do problema nas últimas décadas: a tendência reducionista, que visa explicar um determinado conjunto de fenômenos em termos de conceitos pertencentes a um nível supostamente mais básico ou fundamental. Assim, fenômenos psicológicos devem ser compreendidos em termos biológicos (uma abordagem que

remonta às origens do darwinismo), e fenômenos biológicos, por sua vez, em termos físicos (que, segundo o materialismo tradicional, deveriam ser considerados como os mais primitivos, tanto do ponto de vista explicativo quanto do ponto de vista propriamente ontológico).

Partindo desta mentalidade reducionista dominante em boa parte da ciência na segunda metade do século XIX e século XX adentro, Morowitz alude então a uma possível mudança de tendência que ele atribui sobretudo à consolidação das teorias em física quântica. Isso porque estas teorias dependem fundamentalmente de uma integração da consciência ao conjunto dos componentes que influenciam e condicionam o comportamento de certos sistemas físicos. Essa mudança teria ocasionado uma reorganização um tanto quanto peculiar na relação entre as abordagens características daquelas três ciências. Numa passagem crucial, o autor afirma o seguinte:

primeiramente, a mente humana, incluindo a consciência e o pensamento reflexivo, pode ser explicada pelas atividades do sistema nervoso central, que, por sua vez, pode ser reduzido à estrutura e à função biológicas daquele sistema fisiológico. Em segundo lugar, os fenômenos biológicos, em todos os níveis, podem ser totalmente compreendidos em termos da física atômica, isto é, através da ação e interação dos átomos de carbono, nitrogênio, oxigênio e assim por diante. Finalmente, em terceiro lugar, a física atômica, que é agora compreendida quase integralmente por meio da mecânica quântica, deve ser formulada de tal modo que a mente seja considerada como um componente primitivo do sistema. (Morowitz 1981, p. 39)

Segundo Morowitz, esse círculo, que parte da mente e retorna à mente, é resultado de uma cadeia de raciocínio coerente embasada numa conjugação dos resultados mais atuais das nossas melhores teorias científicas.

Um argumento que guarda ressonâncias bastantes interessantes em relação a este círculo apresentado por Morowitz pode ser encontrado numa obra, hoje pouco conhecida, de um filósofo do século XIX que influenciou decisivamente o pensamento de Nietzsche. O filósofo é Friedrich Albert Lange, e sua obra: *História do materialismo e crítica de seu significado para o presente*. Nesta obra, Lange defendia que os resultados mais recentes do que era na época a fisiologia da percepção, ciência baseada na observação empírica e nos pressupostos metodológicos da explicação materialista do real, representavam uma comprovação das teses de base do idealismo

kantiano, e que o materialismo, pensado de modo consequente, teria de necessariamente se converter em idealismo. Num momento importante do livro, após discutir a possibilidade de reconduzirmos todo pensamento a mecanismos físicos em atuação no corpo, ele observa:

estamos então imediatamente diante da questão: o que é um corpo? O que é a matéria? O que é o físico? [...] a fisiologia de hoje, assim como a filosofia, deve nos dar a seguinte resposta a essas questões: que isso tudo são apenas nossas representações; representações necessárias, representações resultantes de leis naturais, mas ainda assim elas não são coisas em si mesmas. A consideração materialista consequente se converte assim imediatamente em uma consideração idealista consequente. (Lange 1866, p. 496s.)

A analogia entre o argumento apresentado por Morowitz e o argumento de Lange é obviamente limitada ao fato de que, tanto num caso quanto no outro, parte-se de uma redução da mente à matéria para, ao fim, vermos a mente ser reintroduzida como elemento incontornável e inextirpável do sistema explicativo; mais ainda, como um elemento primitivo do sistema. Note-se que ainda sequer estamos falando da possibilidade ou impossibilidade de uma recondução explicativa consistente da mente à matéria. No caso em questão, mesmo que pudéssemos efetuar uma tal recondução, aquilo que entendemos por matéria ou por sistema físico observável seria, em algum nível, necessária e incontornavelmente, dependente da mente. Nesse sentido, a afirmação de que a mente é um componente primitivo do sistema anda lado a lado com a constatação de que ela é logicamente anterior à possibilidade mesma de se pensar um sistema físico qualquer. Dessa anterioridade lógica talvez possamos deduzir então sua irreducibilidade do ponto de vista explicativo.

Historicamente falando, a abordagem materialista revela suas limitações bem cedo, tão logo vem à luz sua incapacidade de explicar a ocorrência de fenômenos mentais tal como os experienciamos subjetivamente. Isso inclui a própria percepção do mundo físico, isto é, desde nossa percepção da matéria visível até a leitura dos dados nos equipamentos de medição da velocidade e posição das partículas subatômicas em um sistema quântico. Ou seja, o aspecto propriamente fenomenológico da visão, audição e interpretação não pode ser explicado a partir de uma descrição exaustiva da biofísica cerebral e corporal de um modo geral. Trata-se aqui do problema clássico dos *qualia* ou da lacuna explicativa.

Uma primeira formulação bastante polêmica deste problema, enunciada pelo cientista Emil Du Bois-Reymond, teve lugar num congresso que reunia cientistas e médicos alemães em Leipzig no ano de 1872. Ali, Du Bois-Reymond pronunciou seu histórico *ignorabimus*. A fórmula, que significa literalmente “ignoraremos”, aponta para o que seria, segundo o autor, dois limites intransponíveis da ciência da natureza: o conhecimento da essência da matéria e o conhecimento da relação existente entre estados mentais e sua suposta base material. Trata-se aqui não apenas da constatação de que, no atual estado da ciência, não sabemos essas coisas, mas da aposta mais radical de que jamais viremos a sabê-las, não importa quão desenvolvida esteja nossa ciência. No que diz respeito mais especificamente ao assim chamado “*hard problem*”, a posição de Du Bois-Reymond é clara: jamais seremos capazes de explicar a consciência a partir da estrutura material do cérebro, ou seja, estados mentais não podem ser reduzidos, do ponto de vista explicativo, a estados da matéria, tal como concebida pelo materialismo tradicional.

3. O pampsiquismo em Nietzsche

Nietzsche foi um filósofo muito sensível aos debates da época em torno do materialismo, e muitas de suas reflexões epistemológicas e metafísicas podem ser melhor compreendidas se pensadas como tentativas de se posicionar dentro deste debate. Dois importantes fragmentos póstumos do período entre 1884 e 1885 devem servir como ponto de partida para a inserção de Nietzsche na presente discussão. No primeiro deles, ele diz o seguinte:

Quando penso em minha genealogia filosófica, eu me sinto ligado ao movimento antiteleológico [...] de nossa época; [...] assim como me sinto ligado ao movimento mecanicista (redução de todas as questões morais e estéticas a questões fisiológicas, de todas as questões fisiológicas a químicas, de todas as questões químicas a mecânicas), com a diferença, entretanto, que eu não acredito em ‘matéria’ e considero Boscovich [...] um dos grandes pontos de inflexão. (N 1884, 26[432])¹⁰⁶

¹⁰⁶ As referências à obra publicada de Nietzsche seguem o padrão nacional de abreviaturas (A = *Aurora*, GC = *A gaia ciência*, ABM = *Além do bem e do mal*, GM = *Genealogia da moral*, CI = *Crepúsculo dos ídolos*) e são acompanhadas do número do aforismo, precedido, quando for o caso, do algarismo romano ou do resumo do título do capítulo. As citações dos fragmentos póstumos seguem o padrão internacional: N = *Nachlass*, seguido do ano e do número do fragmento.

Aos olhos de Nietzsche, o físico Roger Boscovich representa a superação do ponto de partida da física corpuscular, na medida em que abdica do conceito atomista de matéria e procura construir uma teoria fundada na noção de pontos inextensos compreendidos como centros dinâmicos de força. Diversos fragmentos e trechos da correspondência de Nietzsche da década de 80 dão prova da importância que as teorias de Boscovich tiveram no desenvolvimento de sua crítica ao mecanicismo e de sua visão dinâmica do mundo, baseada no conceito de força tal como pensado pelo físico dalmata.¹⁰⁷ Contudo, mesmo que Nietzsche não aceite o conceito de matéria do mecanicismo clássico, como é afirmado no texto que acabei de citar, ele se sente ligado a este movimento na medida em que pressupõe a possibilidade de se reconduzir todos os fenômenos, inclusive os fenômenos morais e estéticos, ao vocabulário fisicalista. Nietzsche entende, na verdade, que a nova física dinâmica de Boscovich nada mais é do que o atomismo levado às suas últimas consequências.

Observemos, porém, que o argumento de Nietzsche não termina neste ponto. No segundo fragmento que mencionei acima, Nietzsche escreve: “O vitorioso conceito de força, com o qual nossos físicos criaram Deus e o mundo, carece ainda de um complemento: deve ser-lhe atribuído um mundo interior, que eu designo como 'vontade de poder'.” (N 1885, 36[31])

Se entendo corretamente o que Nietzsche quer dizer com “mundo interior”, e que é designado pela já famosa expressão “vontade de poder”, acredito haver aqui um interessante ponto de virada no argumento, que em muito se assemelha à estrutura circular apresentada anteriormente a partir dos argumentos de Morowitz e de Lange: partimos de uma recondução dos fenômenos mentais aos fenômenos físicos, cuja explanação gira em torno do conceito de força, para em seguida sermos confrontados com a necessidade de se complementar esse conceito de força, essa *dynamis*, com uma qualidade mental: um “mundo interior”. Ora, a “vontade de poder”

¹⁰⁷ Para uma análise mais detalhada da relação entre Nietzsche e Boscovich, ver por exemplo Schlechta & Anders 1962, pp. 127-153, Whitlock 1996 (em especial sobre a influência de Boscovich no fragmento de 1873 conhecido como *Zeitatomlehre*), Whitlock 1997, Itaparica 2003 e Nasser 2015.

indica, antes de mais nada, o que aos olhos de Nietzsche é o elemento mais saliente de nossa vida psíquica, afetiva e pulsional, tal como nos é dado na perspectiva de primeira pessoa.¹⁰⁸

O único texto da obra publicada onde o filósofo apresenta mais claramente esta noção é o aforismo 36 de *Além do bem e do mal*.¹⁰⁹ Ali, Nietzsche escreve o seguinte:

Supondo que nada seja “dado” como real, exceto nosso mundo de desejos e paixões, e que não possamos descer ou subir a nenhuma outra “realidade”, exceto à realidade de nossos impulsos [...]: não é lícito fazer a tentativa e colocar a questão de se isso que é dado não bastaria para compreender, a partir do que lhe é igual, também o chamado mundo mecânico (ou “material”)? Afinal, não é apenas lícito fazer essa tentativa: é algo imposto pela consciência do *método*. Não admitir várias espécies de causalidade enquanto não se leva ao limite extremo (– até ao absurdo, diria mesmo) a tentativa de se contentar com uma só [...] A questão é, afinal, se reconhecemos a vontade como *atuante*, se acreditamos na causalidade da vontade: assim ocorrendo – e no fundo a crença nisso é justamente a nossa crença na causalidade mesma –, *temos* então de fazer a tentativa de hipoteticamente ver a causalidade da vontade como a única. “Vontade”, é claro, só pode atuar sobre “vontade” – e não sobre “matéria” (sobre “nervos”, por exemplo –): em suma, é preciso arriscar a hipótese de que em toda parte onde

¹⁰⁸ O restante do fragmento nos permite entrever uma premissa história importante por trás do argumento de Nietzsche: a discussão em torno da noção de 'ação à distância': “os físicos não conseguem excluir de seus princípios a 'ação à distância': tampouco a força de repulsão (ou de atração). Isso em nada ajuda: precisamos conceber todos os movimentos, todos os 'fenômenos', todas as 'leis' apenas como sintomas de um evento interno e nos servir até o fim da analogia com o humano.” (N 1885, 36[31]) Lopes (2011, p. 346) observa que “a força do modelo mecanicista clássico residia em duas virtudes: sua sobriedade e economia e o caráter intuitivamente acessível de seu modelo explicativo, baseado na noção de causação por impacto ou contato. Esta segunda virtude desaparece com Newton, na medida em que ele é obrigado a acolher a noção contra-intuitiva de uma ação à distância para tornar compreensível o modo de atuação da força gravitacional. A suposição de uma ação à distância permitiu descrever de forma unificada e econômica os fenômenos naturais, mas introduziu uma grande dúvida em relação à capacidade explicativa do modelo corpuscular clássico. Os físicos passaram paulatinamente a considerar este modelo um dispositivo heurístico e uma ficção tolerável para fins descritivos. Nietzsche argumenta [...] que o tipo de atuação que nós atribuímos à vontade (fundada numa relação de mando e obediência que não pressupõe contato entre uma instância e outra) permite uma apreensão intuitiva daquilo que seria a ação à distância, comprovando assim a superioridade (do ponto de vista da acessibilidade intuitiva) do modelo baseado nas vontades de potência.”

¹⁰⁹ Loeb (2015) argumenta que o modo como a teoria da vontade de poder é apresentada em ABM 36 não compromete Nietzsche com o pampsiquismo. Dado o escopo do presente trabalho, porém, uma crítica detida à sua leitura deverá aguardar uma outra ocasião.

se reconhecem “efeitos”, vontade atua sobre vontade – e de que todo acontecer mecânico, na medida em que nele age uma força, é justamente força de vontade, efeito de vontade. (ABM 36)

O ponto de partida de Nietzsche neste texto pode ser visto como uma versão do argumento em favor da existência de *qualia*, vistos aqui como *qualia* de natureza afetiva, desiderativa ou emotiva (Hill 2007, p. 82); em suma, como o elemento volitivo de nossa vida psíquica que constitui nossas disposições em geral, sendo capaz de alterá-las e colocar em movimento nosso corpo, nos mais diversos níveis. No que diz respeito ao seu aspecto propriamente subjetivo, Nietzsche parece assumir, por um lado, que estes elementos volitivos são suficientemente salientes para que os possamos considerar do ponto de vista de sua acessibilidade intuitiva (cf. Lopes 2011, p. 346; Constâncio 2013, p. 153ss.). Nisso o filósofo acompanha uma sugestão de Schopenhauer,¹¹⁰ mas, como veremos, vale-se de modo bem mais consistente que este último do expediente do argumento analógico; por outro lado, Nietzsche assume que há um sentido importante e originário da causalidade que é derivado da fenomenologia de nossos atos volitivos e do sentimento de nosso poder agencial. Já num fragmento do período de juventude, de 1872, encontramos um raciocínio semelhante: “A única causalidade que nos é consciente é aquela entre o querer e a ação – nós a transpomos então a todas as coisas e interpretamos assim a relação entre duas mudanças que sempre ocorrem conjuntamente.” (N 1872, 19[209]) A tese de que nosso conceito de causalidade tem sua origem psicológica na experiência que temos de nossos próprios atos volitivos é uma constante nos textos de maturidade de Nietzsche.¹¹¹ Mesmo que, na maioria das vezes, a recondução deste conceito à fenomenologia do querer tenha como pano de fundo uma crítica às suas credenciais

¹¹⁰ Cf. *O mundo como vontade e representação*, §22 (2005, p. 171): “Se, portanto, remetemos o conceito de *força* ao de *vontade*, em realidade remetemos algo desconhecido a algo infinitamente mais bem conhecido, àquilo que unicamente nos é conhecido de maneira imediata e completa e que amplia de maneira enorme nosso conhecimento.” Nietzsche, porém, mesmo que pareça pressupor que nosso conhecimento da vontade é mais intuitivo do que o conhecimento que temos das relações do mundo exterior segundo o modelo de explicação da física, questionará a afirmação schopenhaueriana de que a vontade possa ser de fato conhecida, via mera introspecção, “de maneira imediata e completa” (ABM 19; sobre isso, ver ainda Constâncio 2013).

¹¹¹ A meu ver, o texto mais ilustrativo no tocante a este ponto é o fragmento 14[81] de 1888, que leva o título: “Crítica do conceito de ‘causa’”.

epistêmicas, uma vez que Nietzsche frequentemente considera a causalidade da vontade, tal como a entendemos vulgarmente, como uma ilusão, isso em nada muda sua concepção de que aquilo que compreendemos ou acreditamos compreender quando observamos uma relação de causa e efeito no mundo externo depende fundamentalmente da compreensão intuitiva que temos do poder causal de nossa própria vontade. Isso inclui não apenas os movimentos voluntários propriamente ditos, mas também o sentimento de resistência, de obstrução, de pressão e tensão muscular, de ser afetado, etc.¹¹²

Neste ponto, devemos notar o seguinte: em que pese sua crítica recorrente à concepção vulgar que temos acerca da causalidade da vontade e, conseqüentemente, ao próprio conceito de causalidade que daí resulta, acredito que o objetivo desta crítica não seja exatamente o de negar que haja no mundo algo como relações ou processos causais e que aquilo que constitui essencialmente nosso mecanismo volitivo tenha algum poder causal. Ao meu ver, o objetivo da crítica de Nietzsche nesse contexto é denunciar uma falsa concepção de causalidade, em grande medida devida a uma série de erros comuns na compreensão que temos de nós mesmos em virtude de uma falsa auto-observação. Esses erros, potencialmente corrigíveis dentro de certos limites, nos induzem a crer, por um lado, na substancialidade, espontaneidade e no poder auto-causal da consciência; por outro, na unidade e simplicidade do querer tal como este se apresenta à consciência. A concepção de causalidade que daí emerge pressupõe que estejamos autorizados e distinguir claramente entre uma substância, entendida como causa, e sua ação ou atividade, entendida como seu efeito. É justamente essa reificação e dualização da efetividade causal que Nietzsche pretende apontar como falsa.

¹¹² Um dos argumentos de Nietzsche em favor desta tese parece ser o seguinte: a percepção que temos da eficácia causal dos objetos do mundo em geral depende da vivência que temos de nosso próprio poder agencial em termos de “performance” para ação, isto é, nossa percepção da eficácia causal dos objetos depende geneticamente de sua capacidade de agir obstrutivamente sobre nossos esforços para executar ações: “um objeto é a soma das *obstruções* vivenciadas que se nos tornaram *conscientes*.” (N 1885, 2[77]) Sua causalidade é percebida por nós, portanto, somente na medida em que tem poder inibidor sobre nossa vontade. Isso não significa necessariamente que toda percepção de uma causalidade empírica está diretamente associada à vivência concreta de inibição de uma ação atualmente visada. Seria mais correto dizer que só somos capazes de perceber a eficácia causal porque somos naturalmente predispostos para a ação, de modo que tudo aquilo que diz respeito ao nosso conhecimento da causalidade depende constitutivamente de uma experiência primária de poder agencial (cf. Poellner 2013, p. 681s.).

Por sua vez, uma correção daqueles erros, que depende não só de uma autoanálise mais perspicaz, mas também de uma investigação cuidadosa do corpo e de seus processos fisiológicos, nos revela um universo muito mais complexo de causas num nível que nos é, a princípio, inteiramente inconsciente. Nesse nível, forças agem umas sobre as outras numa relação multipolar, e sua ação não deve ser separada de sua própria essência: seu ser é seu fazer-efeito mesmo.¹¹³ Ora, o modo como Nietzsche tenta compreender então esse âmbito inconsciente das causas é ainda essencialmente dependente de um vocabulário intencional. Se tomarmos como exemplo apenas dois aforismos de *Além do bem e do mal*, os aforismos 12 e 19, veremos ali que ele fala em “subalmas”, “subvontades”, “pluralidade do sujeito”, “estrutura social de impulsos e afetos” e ainda “estrutura social de muitas almas”. Nos fragmentos póstumos do período, os termos são ainda mais ousados: o autor chega a caracterizar esses componentes fundamentais de nosso organismo como “intelectos”, “consciências”, “mentes” e “eus”.¹¹⁴

Mas isso de modo algum impede que Nietzsche assuma uma posição eliminativista e/ou epifenomenista no que diz respeito à concepção tradicional (em especial à concepção cartesiana)

¹¹³ Dentro desse contexto, e na esteira do que disse logo acima, acredito que o principal problema que Nietzsche enxerga na noção tradicional de causalidade é sua dependência com relação a um esquema substancialista, a partir do qual distinguimos uma *coisa* (sujeito / substância) que age sobre uma outra coisa: “causa e efeito – um conceito perigoso, na medida em que pensamos um *algo* que *causa* e um *algo* que *sofre o efeito*.” (N 1887, 9[91]) No meu entender, Nietzsche aceita que falemos de causalidade no âmbito daquela dinâmica volitiva caso sejamos capazes de pensá-la a partir de um modelo não substancialista: “aqueles minúsculos seres vivos que constituem nosso corpo (ou dito de forma mais precisa: de cujo trabalho conjunto aquilo que chamamos 'corpo' é a melhor alegoria) não devem ser vistos como átomos anímicos, mas como algo que cresce, luta, se multiplica e perece: de modo que seu número muda constantemente, e nossa vida, como toda vida, é ao mesmo tempo um constante morrer. Portanto, há no homem tantas 'consciências' quanto há seres – em cada momento de sua existência – que constituem seu corpo.” (N 1885, 37[4]) O modelo privilegiado por ele aqui é aquele da relação de comando e obediência. Ele teria a vantagem de partir de uma dinâmica complexa de relações de força, pensadas com base numa metáfora política e ancoradas num cenário pluralista e de interdependência entre os diversos campos de força em conflito. O próprio filósofo mantém o vocabulário causal em muitas passagens, apesar de efetuar uma reviravolta no conceito ao destacar sobretudo a noção de efeito (*Wirkung, Folge*) em detrimento da noção de causa (mais uma vez buscando evitar uma imagem substancialista da causalidade). O fragmento N 1881, 11[81] é bastante ilustrativo nesse sentido: “não há para nós causa e efeito (*Ursache und Wirkung*), mas apenas efeitos (*Folgen*), ('descargas', *Auslösungen*)”. Ainda assim, o conceito de causa continua a comparecer em suas reflexões (cf. por exemplo o famoso aforismo 36 de ABM e as diversas considerações sobre a causalidade em CI). No fragmento póstumo 34[46] de 1885, por exemplo, ao discutir a origem do sentir, querer e pensar conscientes a partir da trama dos impulsos, ele nos diz que “o verdadeiro mundo das causas nos é oculto: ele é indizivelmente mais complicado.” Sobre isso, cf. ainda Hill 2007, p. 79s., Mattioli 2014, Strawson 2008, p. 175 (nota 67) e Strawson 2015, p. 25.

¹¹⁴ Cf. por exemplo N 1884, 26[36], N 1885, 37[4] e N 1885, 40[42].

de consciência, de sujeito ou de “eu”. Para ele, a consciência ou o eu não designam uma faculdade substantiva, uma substância ou, de modo geral, uma “coisa”, uma *res*, tal como nos aparece normalmente na fenomenologia de nosso mundo interior, na percepção ordinária que o sujeito tem de si mesmo via introspecção. Nesse sentido, a consciência e a unidade do eu consciente, tal como entendidas tradicionalmente, são uma ilusão, condicionada pela crença intuitiva em certas estruturas sintáticas responsáveis por nossa apreensão do mundo e pela percepção que temos de nós mesmos.¹¹⁵ Por outro lado, devemos pensar que a consciência é, antes de mais nada, uma *propriedade de estados mentais* (cf. Katsafanas 2005),¹¹⁶ sujeita a

¹¹⁵ Cf. em especial ABM 16, 17 e 54, mas também A 115, GC 354, ABM 19, GM I 13, CI “Razão” 5; cf. ainda Katsafanas 2015, p. 119 e p. 126ss., Riccardi 2015b, p. 225 e 2015c, p. 543.

¹¹⁶ Este é um ponto importante e que deve ser devidamente considerado, pois ele desautoriza, contra um argumento comum em favor do eliminativismo, qualquer tentativa de inferir a inexistência de estados mentais ou conscientes em geral do caráter ilusório da estrutura da consciência introspectiva ordinária. O mais adequado nesse contexto seria então distinguir conceitualmente entre diferentes noções de consciência, por exemplo, entre a autoconsciência (de caráter reflexivo e linguisticamente articulada) e uma forma mais básica de consciência representacional, de caráter perceptivo e qualitativo (para uma discussão mais aprofundada deste ponto, cf. Riccardi 2015a e Welshon 2014, p. 137s.). Isso não quer dizer que sejamos obrigados a aceitar dois tipos de consciência ontologicamente distintos. E isso por duas razões: em primeiro lugar, porque a consciência unitária, substantiva, que, segundo Nietzsche, depende cognitivamente de uma articulação linguística específica, é destituída de carga ontológica própria. Nesse sentido, teríamos, no nível ontológico, apenas um mesmo tipo básico de “substância mental” a manifestar-se em diferentes formas. Em segundo lugar, estaremos nos expressando de modo mais acurado se, ao invés de falar em ‘tipos de consciência’, falarmos em certos conjuntos de propriedades de estados psicológicos conscientes, como ‘ser acessível introspectivamente’, ‘ter um caráter qualitativo’ ou ‘estar reflexivamente determinado’ (cf. Welshon 2014, p. 207, nota 2). Além disso, é importante notar o seguinte: Nietzsche é claro quanto à tese de que estados mentais não são necessariamente conscientes, ou seja, para que um determinado estado seja considerado um estado mental, ele não precisa ser consciente: “a consciência é tão somente um *accidens* da representação, *não* seu atributo necessário e essencial”; “isso que denominamos consciência constitui apenas um estado de nosso mundo espiritual e psíquico [...] e *de modo algum ele próprio*” (GC 357). Essa observação é de fundamental importância, pois permite que concebamos *estados mentais inconscientes* e impede a associação equivocada da dimensão inconsciente à dimensão física ou fisiológica (no sentido materialista), tal como ocorre em algumas versões do dualismo de propriedades. Assumir a existência de estados mentais inconscientes é, ao meu ver, um passo importante para tornar o pampsiquismo uma teoria mais plausível, uma vez que não estamos naturalmente inclinados a aceitar que coisas como árvores, pedras ou elétrons possuam consciência tal como nós possuímos. Ao separar assim o “mental” ou “psíquico” do “consciente”, podemos aceitar que a realidade seja perpassada por “mentalidade”, sem ter que aceitar ao mesmo tempo que tudo o que existe é consciente. Mais uma vez, porém, vale a seguinte observação: a depender de como entendemos a noção de consciência, podemos dizer sem problemas que árvores, pedras ou elétrons possuem *algum grau ou forma primitiva* de consciência, o que não estaria em total desacordo com a tese de que eles possuem estados mentais *inconscientes*. Há, portanto, um nível de análise no qual a distinção entre consciência e inconsciente já não é mais tão relevante. Luca Lupo (2006, p. 49 e pp. 85-132), por exemplo, prefere os termos “consciência primária” (que se refere ao nível subpessoal de inteligência corpórea, expansivo de modo gradual à totalidade do mundo físico) e “consciência secundária” (que se refere ao nível pessoal da cognição em primeira

diversos graus. Num sentido bastante amplo, quando se vale do termo de modo polêmico e subversivo, como nos fragmentos acima referidos, Nietzsche fala de “consciências” para designar os diversos sistemas intencionais em atividade no organismo e, assim, pretende reconduzir a suposta unidade da consciência a uma multiplicidade de estados mentais e disposições afetivas no nível subpessoal da dinâmica do organismo.¹¹⁷ Essa multiplicidade de estados mentais e disposições afetivas é igualmente pensada em termos de uma multiplicidade de complexos volitivos, cuja relação se daria, como já foi dito, segundo o modelo das relações de dominação, isto é, de comando e obediência.

O que fica claro aqui é o objetivo de Nietzsche não somente de reconduzir o âmbito do mental e do espiritual ao nível do orgânico, mas também, de modo reverso, de caracterizar o orgânico em certa medida por meio de qualidades espirituais (cf. Abel 2004, p. 216). Esse é justamente um dos aspectos centrais de sua hipótese das *vontades de poder*. O modelo proposto

peessoa, linguisticamente articulado e referido ao indexical “eu”). Num sentido ainda mais radical, Martin Liebscher (2010, p. 256s.) argumenta que o caráter essencialmente anti-dualista da teoria das vontades de poder conduz necessariamente a uma dissolução da dicotomia consciente/inconsciente no pensamento maduro de Nietzsche e, portanto, torna o conceito de inconsciente inútil. Apesar de não estar inteiramente de acordo com esta leitura de Liebscher, considero que ele toca num ponto importante, qual seja, o estabelecimento de um nível de análise ontológico onde a distinção entre o consciente e o inconsciente já não faz mais tanto sentido. Isso pode ser constatado pelo emprego que faz o próprio Nietzsche da noção de consciência para caracterizar certas atividades constitutivas da dinâmica pulsional do organismo claramente caracterizadas por ele, em outros momentos, como *inconscientes*. Nesse contexto, o termo inconsciente deve certamente ser usado com cautela, sob o risco de se criar uma polarização demasiadamente rígida entre esses dois níveis cognitivos, para a qual não há lugar numa teoria que aposta no princípio da continuidade e busca romper com os dualismos. Isso não quer dizer, porém, que devemos ou podemos simplesmente abandonar o conceito de inconsciente, uma vez que a distinção consciente/inconsciente desempenha uma função explicativa fundamental numa perspectiva epistemológica ou fenomenológica e ocupa uma posição importante nas teses de Nietzsche acerca dos condicionantes da experiência de um modo geral e, em especial, acerca dos condicionantes de nossas ações. O mais importante é, portanto, saber distinguir entre os contextos em que Nietzsche usa o termo “consciência” em sua acepção tradicional, como uma espécie de faculdade substantiva ou substância pensante (para então criticá-la), aqueles em que ele o utiliza para se referir a uma propriedade específica de estados mentais, e aqueles em que ele o emprega de modo mais amplo e inespecífico para designar simplesmente estados mentais ou psíquicos que, de acordo com a teoria das vontades de poder, perfazem a totalidade dos campos de força que constituem a realidade.

¹¹⁷ Para uma análise dessa tese nos termos de uma abordagem homuncular e dos problemas a ela associados, cf. Poellner 1995, p. 213ss., Clark e Dudrick 2012, p. 196ss., Lopes 2012, p. 153ss., Katsafanas 2013, p. 729ss., Riccardi 2015c, p. 439s. Um tratamento apropriado deste tópico em específico, de fundamental importância para uma interpretação pampsiquista de Nietzsche, ultrapassaria em muito o escopo deste trabalho introdutório.

aqui rompe claramente com o dualismo entre corpo e mente, apresentando uma compreensão da relação entre processos orgânicos e atividades intelectuais em termos de um *continuum*.

Nesse sentido, Nietzsche se afasta igualmente tanto do materialismo da época quanto do conceito de vontade tal como pensado por Schopenhauer. Sua crítica a este último tem por objetivo mostrar que ele substituiu o velho dualismo entre mente e corpo por um novo, o dualismo entre vontade e representação (ou intelecto), que é tão condenável quanto o primeiro (cf. Lopes 2012, p. 148). Esse dualismo resulta de um mau uso do argumento da analogia, que pretende transpor ao restante da natureza os resultados da análise de nossa vida corporal, afetiva e pulsional em sua totalidade. Schopenhauer faz um mau uso deste expediente argumentativo, uma vez que, na transposição do sentido psicológico para o sentido propriamente metafísico da vontade, ele subtraiu sua qualidade intencional, seu direcionamento a um objeto, de modo que a partir daí o que ele chama de vontade, e que funda sua concepção de uma irracionalidade que é constitutiva do real, não passa de uma “palavra vazia” (N 1888, 14[121]). Trata-se do “equívoco fundamental de Schopenhauer acerca da vontade” (N 1887, 9[169]). Essa crítica, formulada de forma clara nos textos da década de 80, já se encontra em germe numa série de apontamentos escritos entre 1867 e 1868 e reunidos sob o título *Sobre Schopenhauer* (cf. KGW I/4, p. 418-430).

Nessa série de apontamentos, Nietzsche faz a seguinte crítica ao autor do *Mundo como vontade e representação*: se o pressuposto para pensarmos todo o mundo material, segundo as teses idealistas defendidas no primeiro livro de sua obra principal, é a existência de um intelecto capaz de representação, e se este intelecto, segundo as teses naturalistas apresentadas na segunda parte da obra, só surge num estágio tardio da evolução das manifestação da vontade, isto é, num estágio tardio da história do desenvolvimento dos organismos, sendo uma função do cérebro – neste caso, como pensar a gênese do intelecto, ou seja, a gênese da própria representação, sem pressupor sua existência anterior? Notemos que não se trata aqui de um problema local, mas antes de um problema constitutivo do sistema, na medida em que o autor possui uma teoria realista da vontade como essência do mundo, segundo a qual a vontade é absolutamente alheia a

qualquer predicado mental.¹¹⁸ “De onde [provém] o aparato representacional no qual a vontade aparece?”, pergunta então Nietzsche. “Como foi possível o surgimento do intelecto?”¹¹⁹

Está claro que essa discussão parte de uma premissa idealista. Contudo, do ponto de vista do argumento naturalista avançado por Schopenhauer, o problema central não é outro senão a questão da emergência de estados mentais; da origem da representação, percepção ou experiência. Como seria possível, dentro desse contexto, que a forma da representação *emergisse* repentina e espontaneamente a partir de uma dimensão absolutamente não-representacional? A origem da representação permanece, nesse caso, um mistério. Segundo Nietzsche, nos confrontamos aqui com uma aporia: ou a vontade é aquilo mesmo que representa (isto é, que possui em si a capacidade de representação) ou não pode haver representação, pois esta jamais poderia ter origem. Uma vez que há representação, isto é, que há estados mentais (o que é imediatamente evidente do ponto de vista da fenomenologia da consciência), ele conclui que a representação deve ser um traço constitutivo daquilo que Schopenhauer chama de vontade, e que Nietzsche, em muitos momentos, chama simplesmente de natureza. Este é, ao meu ver, o primeiro esboço de um modelo do *continuum*, que tem como pressuposto a inexistência de algo como uma emergência radical do mental a partir do não mental.

Em sua defesa do pampsiquismo, Galen Strawson (2008, 2016) argumenta que, segundo as premissas comumente aceitas pelo naturalismo, não há nem pode haver algo como “emergência radical” ou um “salto qualitativo abrupto” nos processos naturais. Trata-se aqui, como ele mesmo aponta, de uma versão da antiga tese metafísica: *natura non facit saltum*. O gradualismo é visto, portanto, como uma premissa do próprio naturalismo. Partindo dessa premissa, temos o seguinte cenário: a não ser que estejamos dispostos a negar a existência de estados mentais, assumindo uma posição comprometida ontologicamente com o eliminativismo (o que, para Strawson, é bastante absurdo, já que a existência desses estados é a única coisa da qual temos um conhecimento indubitável), e a não ser que queiramos assumir uma posição dualista, não podemos senão supor que esses estados fazem parte da constituição da própria

¹¹⁸ Entendo, portanto, que Schopenhauer não é efetivamente um pampsiquista, como se afirma costumeiramente, e nisso acompanho a interpretação de Janaway (1999, p. 147).

¹¹⁹ *Zu Schopenhauer*, KGW I/4, p. 425. Cf. sobre isso Mattioli 2013.

realidade. Há duas alternativas aqui: ou aceitamos que a realidade concreta é composta essencialmente por estados e processos mentais, ou negamos o gradualismo, isto é, aceitamos que há saltos qualitativos abruptos na natureza. Strawson argumenta que uma posição naturalista consequente estará mais inclinada à primeira alternativa, haja vista seu comprometimento com o monismo e com o gradualismo.

Esta é uma posição que, segundo penso (e nisso acompanho o próprio Strawson),¹²⁰ pode ser encontrada em Nietzsche. Contudo, não se trata de uma posição elaborada por ele de modo sistemático, mas que pode ser reconstruída a partir de diversas formulações espalhadas por sua obra. Dois pontos importantes se destacam nesse contexto: a crítica ao mecanicismo e a crítica a Schopenhauer como momentos constitutivos de sua teoria das vontades de poder. Mesmo que essas críticas e a elaboração da teoria das vontades de poder façam parte sobretudo de sua filosofia de maturidade, é possível identificar alguns de seus elementos principais já nos primeiros escritos do filósofo (publicados e não publicados).¹²¹

¹²⁰ Cf. Strawson 2015, p. 11.

¹²¹ Além das notas críticas sobre Schopenhauer de 1867-68, um conjunto privilegiado de reflexões em torno deste problema se encontra nos textos póstumos do período de 1872-73, em especial no contexto das reflexões epistemológicas elaboradas em *Sobre verdade e mentira no sentido extramoral*. Neste momento, Nietzsche estava discutindo com algumas teorias da época no campo da fisiologia da percepção que buscavam responder à pergunta pelo modo como estímulos nervosos produziam imagens visuais. Os cientistas de então não haviam chegado a uma resposta para a questão acerca da natureza da relação entre estímulo e imagem, que corresponde à mediação entre o fisiológico e o mental (quanto a isso, continuamos ainda hoje basicamente na mesma situação). Assim, as teorias em voga não eram capazes de demonstrar como estados intencionais emergem a partir dos movimentos da massa cerebral. O *ignorabimus* de Du Bois-Reymond é uma posição emblemática nesse sentido. Na percepção de Nietzsche, este era talvez o exemplo mais claro dos limites do conhecimento científico, para além dos quais algo se anunciava e que escapava à possibilidade de explicação causal tradicional, isto é, mecânico-materialista. No modelo proposto pelo jovem Nietzsche, a superação da lacuna explicativa apresentada emblematicamente no *ignorabimus* de Du Bois-Reymond depende da recondução dos princípios materialistas a princípios intencionais, não o contrário. Seu modelo é centrado no conceito de metáfora, a designar não apenas certos deslocamentos semânticos no interior da linguagem, mas uma função cognitiva primitiva: é ela que está na base dos processos fisiológicos de transposição entre estímulo nervoso e imagem, por um lado, e entre imagem e som, por outro. Nietzsche parte aqui do sentido etimológico do conceito de metáfora: “transferência”, “transposição” (“*Übertragung*”, no alemão), para empregá-lo de modo produtivo, indicando uma continuidade entre a transposição de esferas no âmbito da fisiologia da percepção (do estímulo nervoso em imagem), a transposição da percepção para o processo de gênese da linguagem (da imagem ao som) e as transposições semânticas no interior da própria linguagem (do “sentido próprio” em “sentido figurado”). Nietzsche sugere que toda percepção sensível, toda a formação da linguagem e, consequentemente, todo o processo de emergência da consciência dependem daquelas operações metafóricas realizadas no nível inconsciente da cognição. Assim, a relação entre o mundo representado subjetivamente em nossa consciência e o mundo existente objetivamente é

4. Conclusão

Para concluir, gostaria de fazer uma última observação: acredito que há em Nietzsche diversas motivações para o pampsiquismo, e a possibilidade de conciliar essas diversas motivações é um problema à parte. Até onde consigo ver, podemos encontrar ao menos quatro motivações distintas, duas delas de natureza teórica e duas de natureza prática ou moral. Em especial, elas atravessam as mais diversas elaborações de Nietzsche em torno de temas centrais de sua filosofia de maturidade, como a vontade de poder e o perspectivismo. A primeira motivação de natureza teórica é resultado da combinação de duas inclinações opostas – uma combinação já presente em Schopenhauer, mas que nele permanece uma antinomia irresolúvel: o compromisso com o correlacionismo (isto é, com a tese de que todo objeto pressupõe algum tipo de contraparte cognitiva) e a tendência ao realismo. Dessa combinação emerge, como solução de compromisso, uma forma de idealismo objetivo (na qual a contraparte cognitiva é retirada do sujeito, tal como entendido no idealismo subjetivo, e como que pulverizada e infundida na própria dinâmica dos processos naturais);¹²² a segunda é uma tentativa de tornar o conceito de causalidade e, conseqüentemente, a relação mente-corpo intuitivamente acessíveis, e isso dentro de um cenário monista (tendo em vista assim o critério metodológico da economia de princípios). Esta motivação está associada ainda à sua compreensão do que significa conhecimento ou explicação: a recondução daquilo que nos é estranho a algo que nos é familiar.¹²³

Por outro lado, a primeira motivação de natureza prática ou moral é a tentativa de desconstruir o que se convencionou chamar de “exclusivismo humano” e retraduzir assim o homem de volta à natureza: as qualidades que o ser humano possui (inclusive as qualidades

definida como uma relação “retórica”. Por fim, ao entender a metáfora como função cognitiva, referindo-a à transposição de esferas no âmbito da fisiologia da percepção, Nietzsche parece conceber o caminho neuronal e energético percorrido pelo estímulo até sua configuração imagética como um processo interpretativo que só pode ser satisfatoriamente descrito mediante um vocabulário intencional. Ora, atribuir à metáfora o papel de componente integrativo da experiência, atuante no nível inconsciente e pré-predicativo dos sistemas neuronais, significa, *mutatis mutandis*, atribuir aos componentes orgânicos envolvidos nessa atividade poderes de interação que só nos são compreensíveis a partir da perspectiva de primeira pessoa. Para uma discussão mais detalhada deste conjunto de problemas, cf. Reuter 2009 e Mattioli 2016 (capítulo V).

¹²² Uma discussão exemplar deste aspecto da filosofia de Nietzsche pode ser encontrada em Hill 2008.

¹²³ Cf. por exemplo CI “Quatro” 5, N 1885, 2[83] e N 1888, 14[81].

mentais e espirituais) não são exclusividade de sua espécie, mas são compartilhadas pelos extratos mais 'baixos' da natureza, e a forma na qual se apresentam no estágio ocupado pelo homem na escala evolutiva não se dá em virtude de uma ruptura radical com fases anteriores, mas por uma gradativa adaptação de sua estrutura básica às necessidades da espécie. É provável que boa parte daquilo que subjaz ao seu clamor pela fidelidade à terra esteja diretamente associado a esta visão do lugar do homem na natureza. Por fim, a segunda motivação prática ou moral tem a ver com seu programa de superação do niilismo enquanto movimento de esvaziamento do sentido moral ou de aniquilamento da força normativa de nossos ideais mais elevados. A própria compreensão do fenômeno do niilismo, seu diagnóstico e, portanto, a estratégia pensada por Nietzsche para sua superação, que tem como uma de suas marcas o ideal do perfeccionismo e a noção de autossuperação, parecem pressupor a tese de que os valores – ou, dito de outra forma, atividades de avaliação – correspondem a processos constitutivos de toda e qualquer forma de vida e, por extensão, da natureza como um todo, na medida em que esta é composta por perspectivas valorativas.

Na impossibilidade de elaborar mais detalhadamente essas sugestões no momento, encerro com a indicação de que se trata de um trabalho a ser ainda desenvolvido.

Bibliografia

ABEL, G. *Zeichen der Wirklichkeit*. Frankfurt am Main: Suhrkamp, 2004.

CLARK, M. & DUDRICK, D. *The Soul of Nietzsche's Beyond Good and Evil*. Cambridge: Cambridge University Press, 2012.

CONSTÂNCIO, J. On Nietzsche's Conception of Philosophy in Beyond Good and Evil: Reassessing Schopenhauer's Relevance. In: BORN, M. A., PICHLER, A. (Orgs.) *Texturen des Denkens. Nietzsches Inszenierung der Philosophie in Jenseits von Gut und Böse*. Berlin/Boston: Walter de Gruyter, 2013.

HILL, K. *Nietzsche: A Guide for Perplexed*. Continuum: London / New York, 2007.

_____. From kantian temporality to Nietzschean naturalism. In: DRIES, M. (ed.). *Nietzsche on time and history*. Berlin. New York: Walter de Gruyter, 2008.

ITAPARICA, A. L. M. Nietzsche e Bosovich: dinamismo e vontade potência. In: Azeredo, V. D. (Org.) *Encontros Nietzsche*. Ijuí: Editora Unijuí, 2003.

JANAWAY, C. Will and Nature. In: JANAWAY, C. (Org.) *The Cambridge Companion to Schopenhauer*. Cambridge: Cambridge University Press, 1999.

KATSAFANAS, P. Nietzsche's Theory of Mind: Consciousness and Conceptualization. *European Journal of Philosophy* 13:1, p. 1-31, 2005.

_____. Nietzsche's Philosophical Psychology. In: RICHARDSON, J., GEMES, K (eds.) *The Oxford Handbook of Nietzsche*. Oxford University Press: Oxford, 2013.

_____. Kant and Nietzsche on Self-Knowledge. In: CONSTÂNCIO, J. et. al. (Orgs.) *Nietzsche and the Problem of Subjectivity*. Berlin/Boston: Walter de Gruyter, 2015.

LANGE, F. *Geschichte des Materialismus und Kritik seiner Bedeutung in der Gegenwart*. Iserlohn: J. Baedeker, 1866.

LIEBSCHER, M. Friedrich Nietzsche's Perspectives on the Unconscious. In: NICHOLLS, A., LIEBSCHER, M. (Orgs.) *Thinking the Unconscious. Nineteenth-Century German Thought*. Cambridge University Press, 2010.

LOEB, P. Will to Power and Panpsychism: A New Exegesis of Beyond Good and Evil 36. In: DRIES, M. e KEIL, P. (Ed.) *Nietzsche on Mind and Nature*. Oxford University Press, 2015.

LOPES, R. Das politische Triebmodell Nietzsches als Gegenmodell zu Schopenhauers Metaphysik des Willens. In: GEORG. J. e ZITTEL, K. *Nietzsches Philosophie des Unbewussten*. Berlin / Boston: Walter de Gruyter, 2012.

_____. "A ambicionada assimilação do materialismo": Nietzsche e o debate naturalista na filosofia alemã da segunda metade do século XIX. *Cadernos Nietzsche*, 29. São Paulo, p. 309-352, 2011.

LUPO, L. *Le colombe dello scetico. Riflessioni di Nietzsche sulla coscienza negli anni 1880-1888*. Pisa: Edizioni ETS, 2006.

MATTIOLI, W. Inconsciente, intencionalidade e natureza: a dialética morganática entre naturalismo e transcendentalismo na metafísica da vontade de Schopenhauer. *Revista Voluntas: Estudos sobre Schopenhauer* – v. 4, p. 66-97, 2013.

_____. A temporalidade da consciência e o problema da eficácia causal da vontade em Nietzsche. *Estudos Nietzsche*, Curitiba, v. 5, n. 2, p. 305-344, 2014.

_____. *O inconsciente no jovem Nietzsche. Da intencionalidade das formas naturais à vida da linguagem*. Tese de doutorado. Belo Horizonte: UFMG, 2016.

MOROWITZ, H. Rediscovering the Mind. In: HOFSTADTER, D. e DENNET, D. (Orgs) *The Minds I: Fantasies and Reflections on Self and Soul*. Chicago, 1981.

NASSER, E. *Nietzsche e a ontologia do vir-a-ser*. São Paulo: Loyola, 2015.

NIETZSCHE, F. *Sämtliche Werke. Kritische Studienausgabe* (KSA). Berlin/New York: Walter de Gruyter, 1988.

_____. *Werke. Kritische Gesamtausgabe* (KGW). Herausgegeben von Giorgio Colli und Mazzino Montinari. Berlin. New York: Walter de Gruyter, 1995.

_____. *A Gaia Ciência*. Trad. Paulo César de Souza. São Paulo: Companhia das Letras, 2001.

_____. *Além do Bem do Mal: Prelúdio a uma Filosofia do Futuro*. Trad. Paulo César de Souza. São Paulo: Companhia das Letras, 2005.

POELLNER, P. *Nietzsche and Metaphysics*. Oxford University Press, 1995.

_____. Nietzsche's Metaphysical Sketches: Causality and Will to Power. In: RICHARDSON, J., GEMES, K (Eds.) *The Oxford Handbook of Nietzsche*. Oxford University Press: Oxford, 2013.

REUTER, S. *An der Begräbnisstätte der Anschauung“. Nietzsches Bild- und Wahrnehmungstheorie in Ueber Wahrheit und Lüge im aussermoralischen Sinne*. Basel: Schwabe, 2009.

RICCARDI, M. Nietzsche's Pluralism about Consciousness. *British Journal for the History of Philosophy*, p. 1-23, 2015a.

_____. Inner Opacity. Nietzsche on Introspection and Agency. *Inquiry: An Interdisciplinary Journal of Philosophy*, 58 (3), p. 221-243, 2015b.

_____. Nietzsche on the Embodiment of Mind and Self. In: CONSTÂNCIO, J. et. al. (Orgs.) *Nietzsche and the Problem of Subjectivity*. Berlin/Boston: Walter de Gruyter, 2015c.

SCHLECHTA, K. & ANDERS, A. *Friedrich Nietzsche: von den verborgenen Anfängen seines Philosophierens*. Stuttgart-Bad Cannstatt: Friedrich Frommann Verlag, 1962.

SCHOPENHAUER, A. *O mundo como vontade e como representação*. Tomo I. Trad. Jair Barboza. São Paulo: Editora Unesp, 2005.

STRAWSON, G. Realistic Monism: Why Physicalism Entails Panpsychism. In: *Real Materialism and Other Essays*. Oxford: Oxford University Press, 2008.

_____. Nietzsche's Metaphysics? In: DRIES, M. e KEIL, P. (Ed.) *Nietzsche on Mind and Nature*. Oxford University Press, 2015.

_____. Mind and Being: The Primacy of Panpsychism. In: BRÜNTRUP, G. e JASKOLLA, L. (Eds.) *Panpsychism: Philosophical Essays*. Oxford, 2016.

WELSHON, R. *Nietzsche's Dynamic Metapsychology. This Uncanny Animal*. New York: Palgrave Macmillan, 2014.

WHITLOCK, G. Roger Boscovich, Benedict de Spinoza and Friedrich Nietzsche: The Untold Story. *Nietzsche-Studien*, 25, p. 200-220, 1996.

_____. Examining Nietzsche's 'Time Atom Theory' Fragment from 1873. *Nietzsche-Studien*, 26, p. 350-60, 1997.